Jean-Pierre Lecoutre

TD

STATISTIQUE ET PROBABILITÉS

70%
APPLICATIONS

30%
COURS

4e
édition

DUNOD

Statistique et probabilités

Consultez nos parutions sur dunod.com

TRAVAUX DIRIGÉS

JEAN-PIERRE LECOUTRE
Maître de conférences à l'université
Panthéon-Assas (Paris II)

Statistique et probabilités

Rappels de cours
QCM et questions de réflexion
Exercices corrigés
Sujets d'annales

4e édition

DUNOD

ISBN 978-2-10-054431-8

Sommaire

Avant-propos

Ce volume de la série TD, de la collection « Éco Sup », s'adresse aux étudiants de Licence d'Économie et de Gestion. Au début de chaque chapitre, les principales notions de cours et les résultats essentiels sont rappelés de façon succincte dans « L'essentiel du cours ». Un bref texte introductif indique les points essentiels qui vont être abordés et étudiés dans le chapitre. Il ne s'agit pas d'un résumé de cours, mais seulement d'un avant-propos où on essaie d'expliquer, dans un langage peu formalisé, le fondement et l'utilité des notions définies ensuite de façon plus formelle.

Chacun des huit chapitres présente la même structure. Un certain nombre d'affirmations constituent le paragraphe « Pouvez-vous répondre ? ». La réponse en vrai-faux permet à l'étudiant de vérifier s'il a bien compris les principaux points de cours. Il doit exercer sa vigilance face à des affirmations, parfois simples, mais qui peuvent contenir un piège.

Les « Questions de réflexion » qui sont proposées ensuite ont essentiellement pour but de mettre l'accent sur certains éléments de cours un peu délicats. Il faut être attentif aux commentaires qui figurent dans la solution de l'exercice, en fin de chapitre.

Les exercices d'« Entraînement » permettent enfin à l'étudiant de tester sa capacité à passer de la théorie à la pratique. Ils suivent l'ordre de progression du cours et sont précédés d'un titre indiquant la principale notion à laquelle ils se rapportent. Une rubrique « Analyse de l'énoncé et conseils » précise la démarche à suivre et les résultats de cours à utiliser pour résoudre l'exercice proposé.

Les solutions très détaillées sont regroupées en fin de chapitre, très souvent assorties de commentaires.

Dans le dernier chapitre, les textes récents des examens de 2e année de Licence d'Économie et Gestion de l'université Panthéon-Assas permettent de retrouver les principaux points abordés dans les chapitres précédents. L'étudiant peut ainsi évaluer le niveau de difficulté de ce qui peut être demandé à une épreuve d'examen. Les corrigés sont rassemblés après les énoncés.

Je remercie Antoine Auberger pour ses judicieuses remarques et sa contribution à cette quatrième édition.

Notion de probabilité 1

L'ESSENTIEL DU COURS

Si on veut formaliser un problème dans lequel le hasard intervient, on doit construire un modèle probabiliste, choisi en fonction du but que l'on poursuit. Ce modèle est constitué d'un ensemble fondamental, d'une tribu d'événements et d'une probabilité. Le choix de l'ensemble fondamental est très important pour le calcul ultérieur des probabilités des événements. Nous introduirons aussi les notions de probabilité conditionnelle et d'indépendance. La formule de Bayes est souvent très utile pour le calcul de probabilités conditionnelles.

1. Ensemble fondamental

Le résultat d'une expérience aléatoire s'appelle un *événement élémentaire*. L'ensemble des résultats possibles s'appelle *ensemble fondamental* (ou univers) et est noté traditionnellement Ω. Chaque élément ω de Ω représente donc un événement élémentaire, et toute partie $A \subset \Omega$ (ou $A \in \mathcal{P}(\Omega)$) sera un événement. Parfois on dit que Ω est l'ensemble des *éventualités possibles* et les événements élémentaires sont alors les singletons, c'est-à-dire les ensembles réduits à un seul élément $\{\omega\}$, qui sont effectivement en toute rigueur des événements, puisqu'appartenant à $\mathcal{P}(\Omega)$, ce qui n'est pas le cas du point ω.

2. Algèbre et tribu d'événements

Le couple $(\Omega, \mathcal{P}(\Omega))$ s'appelle un *espace probabilisable*. Même si Ω est fini, le cardinal de $\mathcal{P}(\Omega)$ peut être un nombre très grand et dans ce cas on est amené alors à ne considérer qu'une famille restreinte $\mathcal{A}$ de parties de Ω, $\mathcal{A} \subset \mathcal{P}(\Omega)$. Pour que le résultat des opérations ensemblistes (union, intersection, complémentaire) soit encore un événement, il est nécessaire que la famille d'événements qui a été retenue soit fermée, ou stable, vis-à-vis de ces opérations, c'est-à-dire qu'il soit bien un élément de la famille. De plus, les événements « certain », Ω, et « impossible », $\emptyset$, doivent également appartenir à cet ensemble. Ainsi, on associera à une épreuve aléatoire un ensemble non vide de parties de Ω, noté $\mathcal{A}$, qui vérifiera :

$\boxed{C_1}$ pour tout $A \in \mathcal{A}$, alors $\bar{A} \in \mathcal{A}$;

$\boxed{C_2}$ pour tout $A \in \mathcal{A}$ et tout $B \in \mathcal{A}$, alors $A \cup B \in \mathcal{A}$.

Il y a fermeture pour le complémentaire et l'union. Cet ensemble $\mathcal{A}$ s'appelle une *algèbre* de parties de Ω. Bien entendu, grâce aux lois de Morgan, on a une définition équivalente en remplaçant la condition C_2 par :

$\boxed{C'_2}$ pour tout $A \in \mathcal{A}$ et tout $B \in \mathcal{A}$, alors $A \cap B \in \mathcal{A}$.

► Propriétés d'une algèbre

$\boxed{P_1}$ La famille étant non vide, on en conclut que :

$$\emptyset \in \mathcal{A} \quad \text{et} \quad \Omega \in \mathcal{A}$$

$\boxed{P_2}$ Si $A_j \in \mathcal{A}$ pour $1 \leq j \leq n$, on démontre par récurrence que :

$$\bigcup_{j=1}^{n} A_j \in \mathcal{A}$$

$\boxed{P_3}$ Si $A_j \in \mathcal{A}$ pour $1 \leq j \leq n$, on démontre également par passage au complémentaire que :

$$\bigcap_{j=1}^{n} A_j \in \mathcal{A}$$

Cependant, certaines expériences peuvent se dérouler indéfiniment et on a donc besoin de renforcer la propriété P_2 de fermeture pour l'union finie par une condition de fermeture pour l'union dénombrable, soit :

$\boxed{C_3}$ Si $A_n \in \mathcal{A}$ pour tout $n \in \mathbb{N}$, alors :

$$\bigcup_{n=0}^{\infty} A_n \in \mathcal{A}$$

Cette condition exprime que toute union dénombrable d'événements est encore un événement. L'ensemble $\mathcal{A}$ auquel on impose les conditions C_1 et C_3 s'appelle alors une σ-*algèbre* ou *tribu* d'événements.

Le couple formé de l'ensemble fondamental Ω et de la tribu d'événements associée $\mathcal{A}$ s'appelle un *espace probabilisable*.

3. Probabilité

Définition. On appelle probabilité P sur $(\Omega, \mathcal{A})$ une application $P : \mathcal{A} \to [0,1]$ telle que :

(i) $P(\Omega) = 1$;

(ii) pour toute suite A_n d'événements incompatibles, soit $A_n \in \mathcal{A}$ avec $A_m \cap A_n = \emptyset$ pour $m \neq n$:

$$P\left[\bigcup_{n=0}^{\infty} A_n\right] = \sum_{n=0}^{\infty} P(A_n)$$

propriété dite de σ-additivité.

Une probabilité est donc une application qui à un événement va associer un nombre compris entre 0 et 1. Le triplet $(\Omega, \mathcal{A}, P)$ s'appelle un *espace probabilisé*.

► Propriétés

$\boxed{P_1}$ L'événement impossible est de probabilité nulle :

$$P(\emptyset) = 0$$

$\boxed{P_2}$ La probabilité de l'événement complémentaire d'un événement quelconque A s'obtient par :

$$P(\bar{A}) = 1 - P(A)$$

$\boxed{P_3}$ Si un événement en implique un autre, sa probabilité est plus petite :

$$A \subset B \Rightarrow P(A) \leq P(B)$$

$\boxed{P_4}$ La probabilité de l'union de deux événements s'obtient par la formule de Poincaré :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

4. Probabilités conditionnelles

On considère l'espace probabilisé $(\Omega, \mathcal{A}, P)$ et un événement particulier B de $\mathcal{A}$ tel que $P(B) > 0$. La connaissance de la réalisation de B modifie la probabilité de réalisation d'un événement élémentaire, puisque l'ensemble des résultats possibles est devenu B et non plus Ω. Cette nouvelle probabilité, notée $P(.|B)$, est définie sur la tribu conditionnelle :

$$\mathcal{A}|B = \{A \cap B / A \in \mathcal{A}\}$$

par :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

La formule de définition de la probabilité conditionnelle peut aussi s'écrire, si $P(A) > 0$:

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$$

Elle s'appelle parfois *formule des probabilités composées* et se généralise par récurrence :

$$\begin{aligned} P(A_1 \cap A_2 \cap \ldots \cap A_n) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots \\ &\qquad P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \\ &= P(A_1) \prod_{k=2}^{n} P\left[A_k | \bigcap_{i=1}^{k-1} A_i\right] \end{aligned}$$

5. Théorème de Bayes

Un *système complet d'événements* est une partition de Ω en événements $\{A_1, \ldots, A_n\}$ de probabilités strictement positives, $P(A_i) > 0$ pour $1 \leq i \leq n$, et incompatibles deux à deux, *i.e.* avec $A_i \cap A_j = \emptyset$ pour $i \neq j$ et $\sum_{i=1}^{n} P(A_i) = 1$. On suppose que les probabilités des événements inclus dans chacun des A_i sont connues et on va donc décomposer un événement quelconque B sur ce système :

$$B = B \cap \Omega = B \cap \left[\bigcup_{i=1}^{n} A_i\right] = \bigcup_{i=1}^{n} (A_i \cap B)$$

On aboutit ainsi à la *formule de la probabilité totale :*

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

Ceci va nous permettre de calculer les probabilités *a posteriori* $P(A_i|B)$, après réalisation d'un événement B, à partir des probabilités *a priori* $P(A_i)$, $1 \leq i \leq n$:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

résultat appelé *formule de Bayes* ou parfois *théorème de Bayes.*

6. Indépendance en probabilité

Définition. *Deux événements A et B sont dits indépendants, relativement à la probabilité P, si :*

$$P(A \cap B) = P(A)P(B)$$

La probabilité de réalisation simultanée de deux événements indépendants est égale au produit des probabilités que chacun de ces événements se produise séparément. En conséquence, si $P(B) > 0$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

La réalisation d'un événement ne modifie pas la probabilité de réalisation de l'autre.

► Indépendance mutuelle

Si l'on considère n événements A_i, avec $n > 2$, il y a lieu de distinguer l'indépendance deux à deux qui impose :

$$P(A_i \cap A_j) = P(A_i)P(A_j) \quad 1 \leq i \neq j \leq n$$

de l'indépendance mutuelle, condition plus forte qui s'écrit :

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k}) \quad 2 \leq k \leq n$$

pour tous les sous-ensembles $\{i_1, \ldots, i_k\}$ de $\{1, 2, \ldots, n\}$.

POUVEZ-VOUS RÉPONDRE ?

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

QUESTIONS DE RÉFLEXION

7. On lance deux pièces de monnaie parfaitement identiques et on note les trois résultats possibles PP, FF et PF où la lettre P (resp. F) désigne le résultat pile (resp. face). L'événement élémentaire PF est-il de probabilité $1/3$?

8. Un roi sanguinaire a imaginé le jeu suivant : il fait arrêter quelques-uns de ses sujets et les fait conduire devant un sac contenant mille jetons numérotés de un à mille ; ils doivent alors tirer trois fois un jeton dont on note le numéro avant de le remettre dans le sac. Le roi leur demande alors de choisir le cas dans lequel ils seront pendus : soit lorsque le produit des trois nombres est pair, soit dans le cas contraire. Quelle est la probabilité pour un sujet d'être pendu :

– s'il connaît le calcul des probabilités ?

– s'il ne le connaît pas ?

9. On considère deux événements quelconques A et B. Exprimer en fonction de $P(A)$, $P(B)$ et $P(A \cap B)$ les probabilités conditionnelles suivantes : $P(A|A \cup B)$, $P(A|A \cap B)$ et $P(\bar{B}|\bar{A})$; que devient cette dernière probabilité lorsque A et B sont indépendants ?

10. On suppose que dans une famille de deux enfants, les différentes répartitions ordonnées fille-garçon sont équiprobables.

a) Sachant que l'un au moins des enfants est une fille, calculer la probabilité que les deux enfants soient des filles.

b) Si vous sonnez à l'appartement de cette famille et qu'une fille vous ouvre, quelle est la probabilité que l'autre enfant soit une fille ?

ENTRAÎNEMENT

Ensemble fondamental

11. On lance simultanément deux dés numérotés de 1 à 6. Déterminer l'ensemble fondamental Ω dans les cas suivants :

a) Les deux dés sont distincts (par exemple un rouge et un bleu).

b) Les deux dés sont identiques.

c) Les deux dés sont identiques et on s'intéresse seulement à la parité du résultat.

Analyse de l'énoncé et conseils. L'ensemble fondamental est déterminé à chaque fois par la manière dont un résultat élémentaire va être noté.

12. On lance six boules dans quatre cases distinctes. Représenter l'ensemble fondamental Ω dans les deux cas suivants.

a) Les boules sont numérotées de 1 à 6.

b) Les boules sont identiques.

Analyse de l'énoncé et conseils. Comme dans l'exercice précédent, il faut se poser la question de savoir ce qui caractérise un événement élémentaire.

13. Au résultat *pile* du lancer d'une pièce on associe la valeur $+1$ et la valeur -1 à *face*. Déterminer l'ensemble fondamental Ω associé aux expériences aléatoires suivantes :

a) On effectue six lancers successifs.

b) On lance six pièces identiques simultanément.

c) On lance six pièces identiques simultanément et on ne s'intéresse qu'à la valeur du résultat total obtenu.

Analyse de l'énoncé et conseils. Chaque événement élémentaire est spécifié de façon différente selon les cas.

Atomes

14. Soit A, B et C trois événements quelconques liés à une même épreuve aléatoire. Décomposer les événements $E = (A \cup B)\Delta C$ et $F = A \cup (B\Delta C)$ en une réunion d'événements incompatibles deux à deux et indécomposables, appelés *atomes*. Dans quel cas les événements E et F sont-ils confondus ?

Analyse de l'énoncé et conseils. Si un schéma ne peut en aucun cas remplacer une démonstration, la représentation des événements E et F à partir de A, B et C sera une aide précieuse pour déterminer leur décomposition en atomes. Ces derniers sont des événements qui s'expriment sous la forme $A \cap B \cap \bar{C}$, $\bar{A} \cap \bar{B} \cap C \ldots$

Algèbre d'événements

15. Soit $\Omega = \{a, b, c, d\}$. Déterminer l'algèbre $\mathcal{A}$ contenant les événements $\{c\}$ et $\{d\}$.

Analyse de l'énoncé et conseils. Ce sont les propriétés générales d'une algèbre qui vont permettre de déterminer celle qui est demandée.

16. Soit $\Omega = \{a, b, c, d, e\}$. Déterminer l'algèbre $\mathcal{A}$ engendrée par la partition $\Pi = \big\{\{a\}, \{b, c\}, \{d\}, \{e\}\big\}$ de Ω.

Analyse de l'énoncé et conseils. L'algèbre demandée doit contenir tous les éléments de la partition et donc aussi toutes leurs réunions.

17. Déterminer toutes les algèbres d'événements définies sur les ensembles Ω suivants.

a) $\Omega = \{a\}$.

b) $\Omega = \{a, b\}$.

c) $\Omega = \{a, b, c\}$.

Analyse de l'énoncé et conseils. Les algèbres finies comportent 2^k éléments, avec k variant de 1 à card Ω.

18. À l'aide des opérations d'union, d'intersection et de passage au complémentaire, construire à partir de deux sous-ensembles quelconques A et B de Ω toutes les parties possibles de cet ensemble fondamental. Montrer qu'il s'agit d'une algèbre d'événements.

Analyse de l'énoncé et conseils. Il faut considérer la partition de Ω que l'on peut former à partir des atomes (*cf.* exercice 14, page précédente).

Tribu d'événements

19. Soit Ω un ensemble infini non dénombrable et $\mathcal{C}$ l'ensemble des parties A de Ω telles que A ou bien $\bar{A}$ est dénombrable. Montrer que $\mathcal{C}$ est une σ-algèbre ou tribu d'événements.

Analyse de l'énoncé et conseils. Il faut vérifier les propriétés caractéristiques d'une tribu et utiliser le fait qu'un ensemble inclus dans un ensemble dénombrable est lui-même dénombrable.

Calcul de probabilités

20. Soit A et B deux événements tels que $P(A) = \frac{1}{4}, P(B) = \frac{2}{3}$ et $P(A \cap B) = \frac{1}{8}$. Calculer les probabilités de $E =$ « au moins l'un de ces événements se produit » et $F =$ « un seul de ces événements se produit ».

Analyse de l'énoncé et conseils. C'est une application directe de la formule de Poincaré.

21. On considère un lot de 100 bulletins sur lesquels figurent les réponses *oui* ou *non* à trois questions. Après dépouillement, les nombres de réponses *oui* aux questions 1, 2 et 3 sont respectivement 60, 40 et 30. Les nombres de réponses *oui* associées aux questions 1 et 2, 1 et 3 et 2 et 3 sont respectivement 24, 15 et 12. Enfin, sur 10 bulletins il est répondu *oui* aux trois questions. On désigne par E_i, $1 \leq i \leq 3$, l'événement « la réponse est *oui* à la i—ème question » sur un bulletin prélevé au hasard parmi les 100. On demande d'exprimer à l'aide des E_i les trois événements suivants, puis de calculer leur probabilité : sur le bulletin prélevé, on a obtenu deux *oui* et un *non*, un *oui* et deux *non*, trois *non*.

Analyse de l'énoncé et conseils. Il faut décomposer chaque événement considéré en événements incompatibles et utiliser une formule souvent utile dans ce type d'exercice : $P(A \cap \bar{B}) = P(A) - P(A \cap B)$.

22. Un étudiant doit répondre à quatre questions à choix multiple où trois réponses sont proposées à chaque fois, une seule étant correcte.

a) Dans le cas où l'étudiant répond au hasard et de façon indépendante à chaque question, calculer la probabilité qu'il donne plus de réponses justes que fausses.

b) Que devient cette probabilité s'il n'y a que deux réponses possibles à chaque question ? et s'il y en a quatre ?

Analyse de l'énoncé et conseils. On construit l'ensemble fondamental le plus simple décrivant l'ensemble des résultats possibles, de telle sorte qu'il y ait équiprobabilité des événements élémentaires. Ayant écrit l'événement considéré à l'aide de ces événements élémentaires, le calcul de sa probabilité se fait sans difficulté.

23. Deux joueurs A et B participent à un jeu avec des probabilités respectives de victoire à chaque partie p et $q = 1 - p$. Le gagnant est celui qui le premier obtient deux victoires de plus que l'autre. Quelle est la probabilité de gain de chaque joueur ?

Analyse de l'énoncé et conseils. Il est conseillé de faire un arbre du déroulement des deux premières parties et de marquer à l'extrémité des quatre branches la nouvelle valeur de la probabilité de gain du joueur A par exemple.

Problèmes de dénombrement

24. On jette trois dés identiques numérotés de 1 à 6. Calculer la probabilité d'observer les résultats suivants.

a) Trois fois le même chiffre.

b) Deux fois le même chiffre et un autre différent.

c) Trois chiffres différents.

Analyse de l'énoncé et conseils. Ayant construit l'ensemble fondamental où il y a équiprobabilité des événements élémentaires, on écrit, en utilisant par exemple les lettres de l'alphabet, la forme générale du résultat considéré pour calculer ensuite le nombre de résultats différents de cette forme.

25. On tire au hasard et sans remise cinq cartes d'un jeu de trente deux. Calculer la probabilité d'obtenir : une paire (résultat de la forme *aabcd*), deux paires (*aabbc*), un full (*aaabb*).

Analyse de l'énoncé et conseils. Voir exercice précédent. Il s'agit de problèmes combinatoires délicats qui demandent beaucoup d'attention pour bien déterminer tous les résultats différents ayant la même forme.

26. Un tiroir contient n paires de gants différentes. En raison d'une panne d'électricité, Achille Talon prend au hasard $2r$ gants dans ce tiroir, avec $r < n$. Quelle est la probabilité qu'il n'ait obtenu aucune paire de gants appareillés ?

Analyse de l'énoncé et conseils. Il faut d'abord dénombrer tous les résultats possibles équiprobables. Ensuite, on doit examiner comment les gants doivent être choisis pour qu'il n'y ait aucune paire appareillée, en n'oubliant pas qu'une paire est constituée de deux gants, un droit et un gauche !

Probabilités conditionnelles

27. On cherche une lettre qui a la probabilité p de se trouver dans l'un des quatre tiroirs d'un secrétaire. Quelle est la probabilité qu'elle se trouve dans le quatrième tiroir, sachant qu'on ne l'a pas trouvée dans les trois premiers ?

Analyse de l'énoncé et conseils. C'est une simple application de la définition d'une probabilité conditionnelle. Il faut cependant bien noter que p est la probabilité que la lettre soit dans le secrétaire et que, si elle y est, tous les tiroirs ont la même probabilité de la contenir.

28. On considère trois urnes U_1, U_2 et U_3 qui contiennent respectivement deux boules noires, deux boules blanches et une blanche et une noire. On vous présente l'une de ces trois urnes tirée au hasard ; quelle est la probabilité que ce soit U_1 :

– si vous savez que l'urne contient au moins une boule noire ?

– si vous tirez une boule noire ?

Analyse de l'énoncé et conseils. Il s'agit encore du calcul d'une probabilité conditionnelle où il faut seulement être attentif à la définition de l'événement par lequel on conditionne.

29. Une élection a lieu au scrutin majoritaire à deux tours. Deux candidats A et B sont en présence. Au premier tour 40 % des voix vont à A et 45 % à B, le reste étant constitué d'abstentions. Aucun candidat n'ayant la majorité absolue, un second tour est organisé. Tous les électeurs ayant voté la première fois voteront à nouveau. Un sondage indique par ailleurs que 5 % des voix de A se reporteront sur B et que 10 % des voix de B iront à A. On estime de plus que les deux tiers des électeurs n'ayant pas voté au premier tour voteront, à raison de 60 % pour A et 40 % pour B.

a) Quelle est la probabilité pour qu'un abstentionniste du premier tour vote pour A ? pour B ?

b) D'après ce sondage, quel candidat a la plus forte probabilité d'être élu ?

Analyse de l'énoncé et conseils. Les résultats du premier tour se traduisent en termes de probabilités et les éléments du sondage en termes de probabilités conditionnelles au vote du premier tour. On doit ensuite calculer les probabilités conditionnelles à l'abstention au premier tour et en déduire la probabilité du vote au second tour, pour chacun des candidats.

Théorème de Bayes

30. On classe les gérants de portefeuille en deux catégories : ceux qui sont bien informés et ceux qui ne le sont pas. Lorsqu'un gérant bien informé achète une valeur boursière pour son client, la probabilité que le cours de celle-ci monte est de 0,8 ; dans le cas d'un gérant mal informé, cette probabilité ne vaut que 0,5. Si on choisit au hasard un gérant dans un annuaire professionnel, la probabilité qu'il soit bien informé est de 0,2. Calculer la probabilité que le gérant ainsi choisi soit mal informé, sachant que la valeur qu'il a achetée a monté.

Analyse de l'énoncé et conseils. C'est une application directe du théorème de Bayes.

31. Dans la coupe de France de football, une équipe E de première division estime qu'elle gagnera si elle rencontre une équipe de division inférieure et qu'elle a une chance sur deux de gagner si c'est une équipe de première division. Sachant que la probabilité de rencontrer une équipe de division inférieure est p, calculer la probabilité que E ait rencontré une équipe de division inférieure, sachant qu'elle a remporté son match.

Analyse de l'énoncé et conseils. Après avoir traduit en termes de probabilités les informations fournies dans l'énoncé, on applique la formule de Bayes.

SOLUTIONS

1 ► Vrai. L'ensemble Ω dépend évidemment de l'expérience considérée, mais aussi du choix de celui qui construit le modèle, et par là présente donc un certain arbitraire. L'ensemble fondamental naturel associé à un jet de dé est $\Omega = \{1, 2, 3, 4, 5, 6\}$, mais si on ne s'intéresse qu'à la parité du résultat, on peut retenir également $\Omega = \left\{\{1, 3, 5\}, \{2, 4, 6\}\right\}$.

2 ▸ Faux. Il suffit que $P(A \cap B) = 0$, ce qui est différent de $A \cap B = \emptyset$. Soit par exemple $\Omega = \{a, b, c\}$ avec $P(\{a\}) = P(\{b\}) = 1/2$ et $P(\{c\}) = 0$; les deux événements $A = \{a, c\}$ et $B = \{b, c\}$ ne sont pas disjoints car $\{a, c\} \cap \{b, c\} = \{c\}$ et pourtant $P(A \cup B) = P(\Omega) = 1 = P(A) + P(B)$.

3 ▸ Faux. La probabilité de tirer une boule rouge au deuxième tirage sans remise dans une urne qui contient deux noires et une rouge est $P(A) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$ alors que si l'événement $B =$ « tirer une boule rouge au premier tirage » est réalisé, on a bien sûr $P(A|B) = 0$.

4 ▸ Faux. L'incompatibilité se traduit par $P(A \cap B) = 0$ et l'indépendance par $P(A \cap B) = P(A)P(B)$, produit en général différent de 0, sauf cas particulier où l'un des deux événements est de probabilité nulle.

5 ▸ Vrai. Si les événements A et B sont indépendants, on obtient :

$$\begin{aligned} P(A \cap \bar{B}) &= P(A) - P(A \cap B) = P(A) - P(A)P(B) \\ &= P(A)\,[1 - P(B)] \;\; = P(A)P(\bar{B}) \end{aligned}$$

donc les événements A et $\bar{B}$ sont indépendants. De même :

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= P(\bar{B}) - P(A \cap \bar{B}) = P(\bar{B}) - P(A)P(\bar{B}) \\ &= P(\bar{B})\,[1 - P(A)] \;\; = P(\bar{A})P(\bar{B}) \end{aligned}$$

donc $\bar{A}$ et $\bar{B}$ sont aussi indépendants.

6 ▸ Faux. En raison de l'équiprobabilité : $P(A) = P(B) = 1/2$ et $P(A \cap B) = P(\{a\}) = 1/4 = P(A)P(B)$, donc les événements A et B sont indépendants.

7 ▸ L'événement noté PF peut être obtenu de deux façons différentes, le résultat pile par exemple pouvant apparaître sur l'une ou l'autre des deux pièces ; sa probabilité est donc $2/4$. Les événements PP et FF par contre ne peuvent être obtenus que d'une seule façon (leur probabilité est $1/4$) ; il n'y a donc pas équiprobabilité des résultats possibles retenus. Ceci peut être source d'erreur dans le calcul des probabilités et il vaut mieux retenir ici l'ensemble fondamental $\Omega = \{P, F\}^2$, c'est-à-dire l'ensemble des couples ordonnés, chaque élément du couple étant associé à une pièce particulière, même si les pièces sont identiques.

8 ▸ Pour que le produit soit impair, il faut que chacun des nombres soit impair. Le sujet qui connaît le calcul des probabilités sait donc que $P(\text{impair}) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ et il choisira donc d'être pendu (événement noté PE) si le résultat est impair. Dans le cas contraire, le choix pair (P) ou impair (I) se fait au hasard, la probabilité d'être pendu étant alors calculée par la formule de la probabilité totale :

$$P(PE) = P(P)P(PE|P) + P(I)P(PE|I) = \frac{1}{2} \times \frac{7}{8} + \frac{1}{2} \times \frac{1}{8} = \frac{1}{2}$$

D'où l'intérêt parfois de connaître le calcul des probabilités !

9 ▸ Il suffit d'appliquer la formule de Bayes et de remarquer que $A \subset A \cup B$:

$$P(A|A \cup B) = \frac{P(A)}{P(A \cup B)} = \frac{P(A)}{P(A) + P(B) - P(A \cap B)} \geq P(A)$$

Savoir que A ou B est déjà réalisé augmente la probabilité de A.

Si A et B sont réalisés, l'événement A l'est bien sûr et on peut vérifier que $P(A|A \cap B) = 1$.

En appliquant à nouveau la formule de Bayes :

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

Si A et B sont indépendants, on a vu dans l'exercice 5, page 5, que $\bar{A}$ et $\bar{B}$ sont aussi indépendants, ce que l'on retrouve ici en obtenant dans ce cas $P(\bar{B}|\bar{A}) = P(\bar{B})$.

10 ▸ Par hypothèse, il y a équiprobabilité sur $\Omega = \{FF, FG, GF, GG\}$.

a) On conditionne ici par l'événement $A = \{FF, FG, GF\}$ et la probabilité demandée se calcule par :

$$P(FF|A) = \frac{P(FF)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}$$

b) La différence avec le cas précédent peut ne pas apparaître immédiatement car l'événement $B =$ « une fille vient ouvrir » implique évidemment que l'un au moins des enfants est une fille. Mais il y a un élément supplémentaire d'information qui va renforcer cette probabilité conditionnelle : la fille s'est déplacée pour venir ouvrir. Si on admet que fille et garçon ont la même curiosité de venir ouvrir, on obtient ici :

$$P(B) = \frac{1}{2}P\,(FG) + \frac{1}{2}P\,(GF) + P\,(FF) = \frac{2}{4} < P\,(A)$$

et donc

$$P(FF|B) = \frac{P(FF)}{P(B)} = \frac{1/4}{2/4} = 1/2 > P(FF|A)$$

11 ▸ a) Les deux dés étant distincts, un événement élémentaire est représenté par un couple ordonné $\omega = (x_1, x_2)$ où x_1 (resp. x_2) représente le chiffre obtenu sur le dé rouge (resp. bleu). Ainsi $\Omega = E^2$ avec $E = \{1, 2, 3, 4, 5, 6\}$ et card $\Omega = 36$.

b) Les dés étant identiques, l'ensemble fondamental est constitué par les couples non ordonnés d'éléments de E, avec ici card $\Omega = 6^2 - \frac{6 \times 5}{2} = \frac{6 \times 7}{2} = 21$.

c) Si on note I (resp. P) un résultat impair (resp. pair) sur un dé, l'ensemble fondamental se réduit ici à $\Omega = \{IP, PP, II\}$ soit card $\Omega = 3$.

12 ▸ a) Un événement élémentaire est caractérisé par le numéro C_i, $1 \leq i \leq 4$, de la case où a été lancée chacune des boules. Il s'agit donc d'une application ω de $E = \{1, 2, 3, 4, 5, 6\}$ dans $F = \{C_1, C_2, C_3, C_4\}$ avec card $\Omega = 4^6$.

b) Si les boules sont identiques, un événement élémentaire est caractérisé cette fois par le nombre x_i, $1 \leq i \leq 4$, de boules tombées dans la case i, et :

$$\Omega = \left\{ (x_1, x_2, x_3, x_4) \,/ 0 \leq x_i \leq 6, \sum_1^4 x_i = 6 \right\}$$

Le cardinal de Ω est égal au nombre de combinaisons avec répétitions (*cf.* compléments du chapitre 1 du livre *Statistique et Probabilités*, Lecoutre, Dunod, 3^e ed, 2006) de 6 éléments choisis parmi 4, soit card $\Omega = \binom{6+4-1}{6} = \binom{9}{3} = 84$.

13 ▸ a) À chaque lancer est associé l'ensemble fondamental $E = \{P, F\}$ et aux six lancers $\Omega = E^6$.

b) Un événement élémentaire est caractérisé par le nombre n_1 de résultats P obtenus et dans ce cas $\Omega = \{0, 1, 2, 3, 4, 5, 6\}$. Le nombre de résultats F étant bien sûr $n_2 = 6 - n_1$.

c) Le total obtenu a pour valeur $n_1 - n_2 = 2n_1 - 6$, ce qui conduit à $\Omega = \{-6, -4, -2, 0, 2, 4, 6\}$.

14 ▸ Par définition de la différence symétrique :

$$E = [(A \cup B) \cap \bar{C}] \cup [(\overline{A \cup B}) \cap C] = (A \cap \bar{C}) \cup (B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)$$

L'événement $A \cap \bar{C}$ par exemple n'est pas un atome et se décompose à l'aide du troisième événement B en $A \cap B \cap \bar{C}$ et $A \cap \bar{B} \cap \bar{C}$. On obtient donc la décomposition en atomes :

$$E = (A \cap B \cap \bar{C}) \cup (A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)$$

Par le même procédé, on aboutit à :

$$F = (A \cap B \cap C) \cup (A \cap B \cap \bar{C}) \cup (A \cap \bar{B} \cap C) \cup (A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)$$

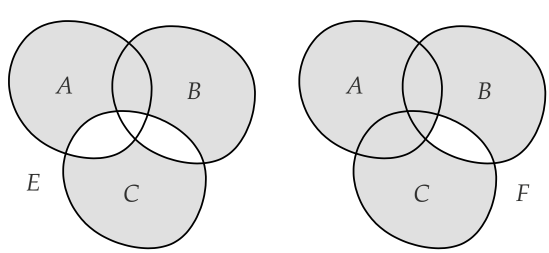


Figure 1.1

On voit ainsi que $F = E \cup (A \cap C)$, donc pour que F soit identique à E il faut que $A \cap C = \emptyset$, c'est-à-dire que A et C soient incompatibles.

15 ▸ L'algèbre $\mathcal{A}$ contenant $\{c\}$ et $\{d\}$ doit contenir aussi leur union $\{c, d\}$ et leurs complémentaires $\{a, b, d\}$ et $\{a, b, c\}$. La fermeture pour l'intersection implique qu'elle contienne aussi $\{a, b, d\} \cap \{a, b, c\} = \{a, b\}$. Au total on obtient :

$$\mathcal{A} = \big\{\emptyset, \{c\}, \{d\}, \{c, d\}, \{a, b\}, \{a, b, d\}, \{a, b, c\}, \Omega\big\}$$

On peut vérifier qu'il s'agit bien d'une algèbre à 2^8 éléments.

16 ▸ L'algèbre $\mathcal{A}$ engendrée par la partition est constituée par les 2^4 réunions des ensembles de cette partition, soit :

$$\mathcal{A} = \big\{\emptyset, \{a\}, \{b, c\}, \{d\}, \{e\}, \{a, d\}, \{a, e\}, \{d, e\}, \{a, b, c\},$$
$$\{a, d, e\}, \{b, c, d\}, \{b, c, e\}, \{a, b, c, d\}, \{a, b, c, d, e\}, \Omega\big\}$$

17 ▸ Tous les ensembles Ω étant finis, les algèbres que nous allons construire seront aussi des σ-algèbres.

a) Il n'y a que l'algèbre élémentaire $\mathcal{A} = \{\emptyset, \Omega\}$.

b) Il n'y a que l'algèbre élémentaire $\mathcal{A} = \{\emptyset, \Omega\}$ et l'algèbre la plus complète $\mathcal{P}(\Omega) = \big\{\emptyset, \{a\}, \{b\}, \Omega\big\}$.

c) On peut construire ici cinq algèbres, de la plus élémentaire $\mathcal{A}_1 = \{\emptyset, \Omega\}$ à la plus complète $\mathcal{P}(\Omega) = \big\{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \Omega\big\}$, en passant par les algèbres à quatre éléments $\mathcal{A}_2 = \big\{\emptyset, \{a\}, \{b, c\}, \Omega\big\}$, $\mathcal{A}_3 = \big\{\emptyset, \{b\}, \{a, c\}, \Omega\big\}$ et $\mathcal{A}_4 = \big\{\emptyset, \{c\}, \{a, b\}, \Omega\big\}$.

18 ▸ L'algèbre $\mathcal{A}$ est engendrée par la partition de Ω constituée des quatre atomes formés à partir de A et B, soit $\Pi = \{A \cap B, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cap \bar{B}\}$. Elle comporte $2^4 = 16$ éléments :

$$\mathcal{A} = \Big\{\emptyset, A, B, \bar{A}, \bar{B}, A \cap B, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cap \bar{B},$$
$$A \cup B, \bar{A} \cup B, A \cup \bar{B}, \bar{A} \cup \bar{B}, A \Delta B, \overline{A \Delta B}, \Omega\Big\}$$

19 ▸ L'ensemble $\mathcal{C}$ est non vide car il contient Ω, puisque $\bar{\Omega} = \emptyset$ est dénombrable (card $\emptyset = 0$). Par ailleurs, il est par définition fermé pour le passage au complémentaire. Enfin, soit $A_n \in \mathcal{C}$ pour tout entier n ; si A_n est dénombrable pour tout n, alors il en est de même de l'union dénombrable $\cup_{n \in \mathbb{N}} A_n$ qui est alors un élément de $\mathcal{C}$. S'il existe un entier n_0 tel que A_{n_0} n'est pas dénombrable, c'est donc que son complémentaire l'est, et par conséquent $\overline{\cup_{n \in \mathbb{N}} A_n} = \cap_{n \in N} \overline{A_n} \subset \overline{A_{n_0}}$ est dénombrable. Dans tous les cas il y a fermeture pour l'union dénombrable, ce qui finit de prouver que $\mathcal{C}$ est une tribu.

20 ▸ Les événements considérés sont $E = A \cup B$ et $F = (A \cap \bar{B}) \cup (\bar{A} \cap B)$ dont les probabilités se calculent par :

$$P(E) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{19}{24}$$
$$P(F) = P(A \cap \bar{B}) + P(\bar{A} \cap B) = P(E) - P(A \cap B) = \frac{2}{3}$$

21 ▸ Le premier événement s'écrit :

$$A = (E_1 \cap E_2 \cap \overline{E_3}) \cup (E_1 \cap \overline{E_2} \cap E_3) \cup (\overline{E_1} \cap E_2 \cap E_3)$$

Sa probabilité se calcule à partir de :

$$P(E_1 \cap E_2 \cap \overline{E_3}) = P(E_1 \cap E_2) - P(E_1 \cap E_2 \cap E_3)$$

Soit :

$$P(A) = \left(\frac{24}{100} - \frac{10}{100}\right) + \left(\frac{15}{100} - \frac{10}{100}\right) + \left(\frac{12}{100} - \frac{10}{100}\right) = \frac{21}{100}$$

Le deuxième événement s'écrit :

$$B = (E_1 \cap \overline{E_2} \cap \overline{E_3}) \cup (\overline{E_1} \cap \overline{E_2} \cap E_3) \cup (\overline{E_1} \cap E_2 \cap \overline{E_3})$$

Sa probabilité se calcule à partir de :

$$\begin{aligned} P(E_1 \cap \overline{E_2} \cap \overline{E_3}) &= P[E_1 \cap \overline{(E_2 \cup E_3)}] = P(E_1) - P\left[E_1 \cap (E_2 \cup E_3)\right] \\ &= P(E_1) - P(E_1 \cap E_2) - P(E_1 \cap E_3) + P(E_1 \cap E_2 \cap E_3) \end{aligned}$$

Soit :

$$P(B) = \frac{60}{100} + \frac{40}{100} + \frac{30}{100} - 2\left(\frac{24}{100} + \frac{15}{100} + \frac{12}{100}\right) + 3\frac{10}{100} = \frac{58}{100}$$

Le troisième événement s'écrit $C = \overline{E_1} \cap \overline{E_2} \cap \overline{E_3}$, de probabilité :

$$P(C) = 1 - P(A) - P(B) - P(E_1 \cap E_2 \cap E_3) = \frac{11}{100}$$

22 ▸ a) Si on note J (resp. F) l'événement « l'étudiant fournit une réponse juste (resp. fausse) », l'événement considéré correspond à trois ou quatre réponses justes, soit :

$$E = (FJJJ) \cup (JFJJ) \cup (JJFJ) \cup (JJJF) \cup (JJJJ)$$

Comme il y a équiprobabilité sur l'ensemble fondamental $\Omega = \{J, F\}^4$ et indépendance des choix à chaque question, la probabilité demandée est :

$$p = P(E) = 4\,\frac{2}{3}\left(\frac{1}{3}\right)^3 + \left(\frac{1}{3}\right)^4 = \frac{1}{9}$$

b) S'il n'y a que deux réponses possibles, cette probabilité devient :

$$p' = 4\frac{1}{2}\left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 = \frac{5}{16}$$

Avec quatre réponses possibles :

$$p'' = 4\frac{3}{4}\left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4 = \frac{13}{256}$$

On a bien sûr $p' > p > p''$.

23 ▸ Soit $P(A)$ la probabilité de gain du joueur A au début du jeu et examinons ce qu'elle devient après les deux premières parties. Si A gagne les deux parties, cette probabilité est devenue égale à 1, et à 0 s'il les a perdues. S'il gagne une seule partie elle est inchangée. On a donc la relation :

$$P(A) = 1 \times p^2 + 0 \times q^2 + P(A) \times pq + P(A) \times qp$$

D'où on déduit :

$$P(A) = \frac{p^2}{1 - 2pq} = \frac{p^2}{p^2 + q^2}$$

Par conséquent :

$$P(B) = 1 - P(A) = \frac{q^2}{p^2 + q^2}$$

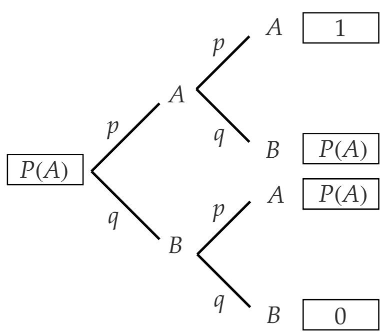


Figure 1.2

24 ▸ Il y a équiprobabilité des événements élémentaires de l'ensemble fondamental $\Omega = \{1, 2, 3, 4, 5, 6\}^3$ de cardinal 6^3. La probabilité d'un événement quelconque A se calcule donc par $P(A) = \dfrac{\text{card} A}{\text{card} \Omega}$.

a) L'événement considéré est de la forme aaa, où il y a six choix possibles pour le chiffre a. Sa probabilité est donc :

$$P(aaa) = \frac{6}{6^3} = \frac{1}{36}$$

b) Il s'agit cette fois d'un événement de la forme aab, avec six choix pour le chiffre a, cinq choix pour b et trois possibilités pour la place du résultat b, donc :

$$P(aab) = 3\,\frac{6 \times 5}{6^3} = \frac{15}{36}$$

c) Un résultat quelconque de la forme abc correspond à une permutation sans répétition de trois chiffres choisis parmi six, soit :

$$P(abc) = \frac{6 \times 5 \times 4}{6^3} = \frac{20}{36}$$

On vérifie bien que la somme de ces trois probabilités est égale à 1.

25 ▸ Il y a équiprobabilité de chaque événement élémentaire qui correspond à une *main* de cinq cartes, *i.e.* à une combinaison sans répétition de cinq cartes choisies parmi 32. Le cardinal de l'ensemble fondamental est donc $\binom{32}{5} = 201\,376$ et il va falloir maintenant dénombrer les différentes mains proposées. Une paire est déterminée par le choix de la hauteur (il y en a 8) avec $\binom{4}{2}$ choix possibles pour les couleurs. Il y a ensuite $\binom{7}{3}$ hauteurs des cartes d'accompagnement et 4^3 couleurs possibles. Le nombre total de paires est donc le produit de ces différentes valeurs, soit $8 \times \binom{4}{2} \times \binom{7}{3} \times 4^3 = 107\,520$ et :

$$P(aabcd) = \frac{107\,520}{201\,376} = \frac{480}{899} = 0{,}534$$

Il y a $\binom{8}{2}$ choix pour les hauteurs des paires avec $\binom{8}{2}$ choix de couleurs pour chacune et il reste ensuite 6 choix de la hauteur de l'autre carte, avec 4 choix de couleurs, soit un nombre total de doubles paires égal à $\binom{8}{2} \times \binom{4}{2} \times \binom{4}{2} \times 6 \times 4 =$ 24 192. Ainsi :

$$P(aabbc) = \frac{24\,192}{201\,376} = \frac{108}{899} = 0{,}120$$

Enfin, il y a 8 choix de hauteurs pour a avec $\binom{4}{3} = 4$ choix de couleurs, 7 choix de la hauteur de la paire et $\binom{4}{2} = 6$ choix pour la couleur, soit un total de $8 \times 4 \times 7 \times 6 = 1\,344$ fulls. On obtient donc :

$$P(aaabb) = \frac{1\,344}{201\,376} = \frac{6}{899} = 0{,}006\,7$$

26 ▸ Chaque ensemble de $2r$ gants pris au hasard parmi les $2n$ gants contenus dans le tiroir a la même probabilité $1 \Big/ \binom{2n}{2r}$. Pour qu'aucune paire ne soit appareillée, il faut que les $2r$ gants aient été choisis parmi les n paires diférentes, ce qui correspond à $\binom{n}{2r}$ possibilités. Mais, chacun de ces gants peut être soit le droit, soit le gauche, soit un total de $2^{2r}\binom{n}{2r}$ choix dépareillés. La probabilité demandée est donc :

$$\frac{2^{2r}\binom{n}{2r}}{\binom{2n}{2r}}$$

Pour $n = 2$ et $r = 1$ par exemple, cette probabilité vaut $\frac{2}{3}$.

27 ▸ On note A l'événement « la lettre est dans le quatrième tiroir » et B l'événement « la lettre n'est pas dans les trois premiers tiroirs ». La définition de la probabilité conditionnelle conduit à :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Chaque tiroir a la même probabilité de contenir la lettre, si celle-ci est dans l'un d'entre eux, donc $P(A \cap B) = p \times \frac{1}{4}$. On calcule ensuite $P(B) = P(A \cap B) + P(\bar{A} \cap B)$ avec bien sûr $P(\bar{A} \cap B) = 1 - p$. Ainsi :

$$P(A|B) = \frac{p/4}{p/4 + 1 - p} = \frac{p}{4 - 3p}$$

Notons que $p/4 < P(A|B) < p$ avec $P(A|B) = 0$ si $p = 0$ et $P(A|B) = 1$ si $p = 1$.

28 ▸ Avec des notations évidentes, la première probabilité s'écrit :

$$P(U_1|U_1 \cup U_3) = \frac{P(U_1)}{P(U_1) + P(U_3)} = \frac{1/3}{1/3 + 1/3} = \frac{1}{2}$$

La deuxième probabilité demandée s'écrit :

$$\begin{aligned} P(U_1|N) &= \frac{P(U_1 \cap N)}{P(N)} = \frac{P(U_1)P(N|U_1)}{P(U_1)P(N|U_1) + P(U_2)P(N|U_2)} \\ &= \frac{1 \times 1/3}{1 \times 1/3 + 1/2 \times 1/3} = \frac{2}{3} \end{aligned}$$

On voit ainsi que les deux probabilités sont distinctes. L'information du tirage d'une boule noire implique bien sûr la présence d'une boule noire mais ne se confond pas avec cette dernière information. Ceci démontre la nécessité d'une très grande rigueur dans la définition d'un événement et l'aide que peut apporter la formalisation dans le calcul d'une probabilité.

29 ▸ Soit A_i (resp. B_i), $i = 1{,}2$, l'événement « voter pour le candidat A (resp. B) au i-ème tour » et O_i l'événement « s'abstenir au i-ème tour ». Les informations fournies dans l'énoncé se traduisent par les probabilités suivantes :

$$\begin{array}{rrr} P(A_1) = 0{,}40 & P(B_1) = 0{,}45 & P(O_1) = 0{,}15 \\ P(B_2|A_1) = 0{,}05 & P(A_2|B_1) = 0{,}10 & P(\overline{O_2}|O_1) = 2/3 \\ P(A_2|O_1 \cap \overline{O_2}) = 0{,}60 & P(B_2|O_1 \cap \overline{O_2}) = 0{,}40 & \end{array}$$

a) Les événements demandés s'écrivent $A_2|O_1$ et $B_2|O_1$, leurs probabilités se calculant par :

$$P(A_2|O_1) = P(\overline{O_2}|O_1)P(A_2|O_1 \cap \overline{O_2}) = 0{,}60 \times \frac{2}{3} = 0{,}40$$
$$P(B_2|O_1) = P(\overline{O_2}|O_1)P(B_2|O_1 \cap \overline{O_2}) = 0{,}40 \times \frac{2}{3} = \frac{4}{15}$$

b) Calculons par exemple :

$$\begin{aligned} P(A_2) &= P(A_1)P(A_2|A_1) + P(B_1)P(A_2|B_1) + P(O_1)P(A_2|O_1) \\ &= 0{,}95 \times 0{,}40 + 0{,}10 \times 0{,}45 + 0{,}40 \times 0{,}15 = 0{,}485 \end{aligned}$$

On obtient de la même façon $P(B_2) = 0{,}465$, ce qui montre que les abstentionnistes apporteront la victoire au candidat A.

30 ▸ Notons MI (resp. BI) l'événement « le gérant est mal (resp. bien) informé » et M l'événement « la valeur a monté ». Les hypothèses formulées se traduisent par les probabilités $P(BI) = 0{,}2$, $P(M|BI) = 0{,}8$ et $P(M|MI) = 0{,}5$. On obtient alors par application de la formule de Bayes :

$$\begin{aligned} P(MI|M) &= \frac{P(MI \cap M)}{P(M)} = \frac{P(MI)P(M|MI)}{P(MI)P(M|MI) + P(BI)P(M|BI)} \\ &= \frac{0{,}8 \times 0{,}5}{0{,}8 \times 0{,}5 + 0{,}2 \times 0{,}8} = \frac{5}{7} \end{aligned}$$

31 ▸ Soit A l'événement « l'équipe E a rencontré une équipe de première division » et G l'événement « l'équipe E a gagné son match ». Par hypothèse $P(\bar{A}) = p$, $P(A) = 1 - p$, $P(G|A) = 1/2$ et $P(G|\bar{A}) = 1$. La formule de Bayes permet alors d'obtenir :

$$P(\bar{A}|G) = \frac{P(\bar{A} \cap G)}{P(G)} = \frac{P(\bar{A})P(G|\bar{A})}{P(A)P(G|A) + P(\bar{A})P(G|\bar{A})} = \frac{p}{(1-p)/2 + p} = \frac{2p}{p+1}$$

Le résultat est une fonction croissante de p, de la valeur 0 pour $p = 0$ à la valeur 1 pour $p = 1$.

Variable aléatoire discrète

2

L'ESSENTIEL DU COURS

Chaque résultat d'une expérience aléatoire est codé au moyen d'une application qui, si elle permet d'associer une probabilité à chacune de ses valeurs numériques discrètes, sera appelée variable aléatoire discrète. La loi de cette variable aléatoire est alors déterminée par les probabilités de toutes ses valeurs possibles. Elle est souvent caractérisée par les deux premiers moments qui sont l'espérance, caractéristique de valeur centrale, et la variance, caractéristique de dispersion autour de cette valeur centrale. Nous présenterons les principales lois de probabilité discrètes.

1. Variable aléatoire réelle discrète

1.1 Définition

On appelle v.a. discrète définie sur $(\Omega, \mathcal{A})$ une application $X : \Omega \to \mathbb{R}$ telle que $X(\Omega)$ est dénombrable, *i.e.* en correspondance bijective avec $\mathbb{N}$, et telle que pour tout x réel :

$$X^{-1}(x) = \{\omega \in \Omega / X(\omega) = x\} \in \mathcal{A}$$

ce qui exprime tout simplement que $X^{-1}(x)$ est un événement.

1.2 Loi de probabilité

La probabilité notée P_X, définie sur $\Omega' = X(\Omega)$ et définie par :

$$P_X(X = x) = P\left\{X^{-1}(x)\right\} = P\left\{\omega \in \Omega / X(\omega) = x\right\}$$

s'appelle la *probabilité image* de P par X. L'ensemble Ω' étant dénombrable, il existe une bijection permettant de représenter ses éléments par l'ensemble des x_i, $i \in \mathbb{N}$. La loi de probabilité est alors définie par les probabilités individuelles :

$$p_i = P_X(X = x_i) = P\left\{X^{-1}(x_i)\right\}, \quad i \in \mathbb{N}$$

On appelle alors *distribution* ou *loi de probabilité* de X l'ensemble des couples $(x_i, p_i)_{i \in \mathbb{N}}$.

1.3 Fonction de répartition

On appelle *fonction de répartition* de la v.a. X, la fonction F définie pour x réel par :

$$F(x) = P_X\{X < x\} = \sum \left\{p_i / x_i < x\right\}$$

Cette valeur représente la probabilité de toutes les réalisations strictement inférieures au réel x.

1.4 Moments d'une v.a. discrète

► Espérance mathématique

Définition. On appelle *espérance mathématique* de la v.a. X la quantité, si elle existe :

$$E(X) = \sum_{i \in \mathbb{N}} p_i x_i$$

Propriétés. Si on ajoute une constante à une v.a., il en est de même pour son espérance :

$$E(X + a) = E(X) + a \quad \forall a \in \mathbb{R}$$

Si on multiplie une v.a. par une constante, il en est de même pour son espérance :

$$E(aX) = aE(X) \quad \forall a \in \mathbb{R}$$

L'espérance d'une somme de deux v.a. est la somme des espérances :

$$E(X + Y) = E(X) + E(Y)$$

On résume ces trois propriétés en disant que l'opérateur espérance est linéaire :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y) \qquad \forall \lambda \in \mathbb{R}, \quad \forall \mu \in \mathbb{R}$$

► Variance

Il s'agit d'un indicateur mesurant la dispersion des valeurs x_i autour de $E(X)$:

$$V(X) = \sum_{i \in \mathbb{N}} p_i \left[x_i - E(X)\right]^2$$

lorsque cette quantité existe. Elle s'écrit aussi :

$$V(X) = E\left[X - E(X)\right]^2 = E\left(X^2\right) - E^2(X)$$

On note encore cette quantité $V(X) = \sigma_X^2$, σ_X désignant alors l'*écart type* de la v.a. X.

Propriétés. Par définition :

$$V(X) \geq 0$$

Pour tout réel a :

$$V(X+a) = V(X)$$

Pour tout réel a :

$$V(aX) = a^2 V(X)$$

Si X et Y sont deux v.a. indépendantes, alors :

$$V(X+Y) = V(X) + V(Y)$$

On définit la *covariance* de deux v.a. X et Y par $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$ et on a, dans le cas général :

$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y)$$

► **Moments non centrés et centrés**

On appelle *moment non centré* d'ordre $r \in \mathbb{N}^*$ la quantité, lorsqu'elle existe :

$$m_r(X) = \sum_{i \in \mathbb{N}} p_i x_i^r = E(X^r)$$

Le *moment centré d'ordre* $r \in \mathbb{N}^*$ est :

$$\mu_r(X) = \sum_{i \in \mathbb{N}} p_i \left[x_i - E(X)\right]^r = E\left[X - E(X)\right]^r$$

2. Lois usuelles discrètes

2.1 Loi de Dirac

Soit $a \in \mathbb{R}$ un point fixé. On appelle *loi de Dirac* la loi de la v.a. certaine X, c'est-à-dire qui est constante, avec $X(\omega) = a$ quel que soit le résultat ω de l'épreuve. Ainsi :

$$X(\Omega) = \{a\} \quad \text{et} \quad P_X(X=a) = P\{\omega \in \Omega / X(\omega) = a\} = P(\Omega) = 1$$

Bien entendu, on obtient comme moments $E(X) = a$ et $V(X) = 0$.

2.2 Loi de Bernoulli

On appelle v.a. indicatrice de l'événement A de $\mathcal{A}$, la v.a. définie par $X = \mathbf{1}_A$:

$$X(\omega) = \mathbf{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \in \bar{A} \end{cases}$$

Ainsi $X(\Omega) = \{0,1\}$ avec $P_X(X=0) = P(\bar{A}) = 1 - P(A) = q$ et $P_X(X=1) = P(A) = p$. On dit que X suit une loi de Bernoulli de paramètre p, ce qu'on écrit $X \rightsquigarrow \mathcal{B}(1,p)$, de moments $E(X) = p$ et $V(X) = pq$.

2.3 Loi binômiale

On effectue n épreuves successives indépendantes où on observe à chaque fois la réalisation ou non d'un certain événement A, et on note le nombre X de réalisations de A. On définit ainsi une v.a. X qui suit une loi binômiale de paramètres n et $p = P(A)$, caractérisée par $X(\Omega) = \{0, 1, \ldots, n\}$ et, pour $k \in X(\Omega)$, par :

$$P_X(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

On écrit $X \rightsquigarrow \mathcal{B}(n, p)$ avec $E(X) = np$ et $V(X) = npq$. Si $X_1 \rightsquigarrow \mathcal{B}(n_1, p)$ et $X_2 \rightsquigarrow \mathcal{B}(n_2, p)$, les v.a. X_1 et X_2 étant indépendantes, alors $X_1 + X_2 \rightsquigarrow \mathcal{B}(n_1 + n_2, p)$.

2.4 Loi hypergéométrique

On effectue n tirages *sans remise* dans une urne contenant N objets dont N_A objets A et on note X le nombre d'objets A tirés. Pour tout k tel que $0 \leq k \leq n$:

$$P_X(X = k) = \frac{\binom{N_A}{k}\binom{N - N_A}{n - k}}{\binom{N}{n}}$$

La loi hypergéométrique s'écrit symboliquement $X \rightsquigarrow \mathcal{H}(N, n, N_A)$, de moments :

$$E(X) = np, \quad V(X) = npq\frac{N - n}{N - 1}$$

ayant posé $p = P(A) = N_A/N$ et $q = 1 - p = 1 - N_A/N$.

2.5 Loi de Poisson

Une v.a. X suit une loi de Poisson de paramètre $\lambda > 0$ si c'est une variable à valeurs entières, $X(\Omega) = \mathbb{N}$, avec pour $k \in \mathbb{N}$:

$$P_X(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

On utilise l'écriture symbolique $X \rightsquigarrow \mathcal{P}(\lambda)$ et le calcul des moments donne $E(X) = V(X) = \lambda$. Si deux variables suivent des lois de Poisson et sont indépendantes, $X \rightsquigarrow \mathcal{P}(\lambda)$ et $Y \rightsquigarrow \mathcal{P}(\mu)$, alors leur somme suit aussi une loi de Poisson : $X + Y \rightsquigarrow \mathcal{P}(\lambda + \mu)$.

2.6 Loi géométrique ou de Pascal

On effectue des épreuves successives indépendantes jusqu'à la réalisation d'un événement particulier A et on note X le nombre d'épreuves effectuées ; pour $k \in \mathbb{N}^*$:

$$P_X(X = k) = (1 - p)^{k-1}p$$

où $p = P(A)$ avec $E(X) = \dfrac{1}{p}$ et $V(X) = \dfrac{q}{p^2}$, $q = 1 - p$.

2.7 Loi binômiale négative

On effectue des épreuves successives indépendantes jusqu'à ce que n événements A soient réalisés et on note Y le nombre d'épreuves effectuées. Pour tout entier $y \geq n$:

$$P_Y(Y = y) = \binom{y-1}{n-1} p^n (1-p)^{y-n} \quad \text{où } p = P(A)$$

avec comme moments $E(Y) = \dfrac{n}{p}$ et $V(Y) = \dfrac{nq}{p^2}, q = 1 - p \cdot$

POUVEZ-VOUS RÉPONDRE ?

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

QUESTIONS DE RÉFLEXION

7. Une urne contient trois billes numérotées 1, 2, 3, celle portant le n°1 étant la seule en métal. On retient donc sur l'ensemble fondamental $\Omega = \{1, 2, 3\}$ la tribu associée $\mathcal{A} = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$. Si on extrait une bille de cette urne au moyen d'un aimant, déterminer la probabilité P associée à cette expérience. On considère alors l'application identité X sur Ω et l'application Y définie par $Y(\omega) = 1$ pour tout ω de Ω. Établir que $P(X = Y) = 1$ et que pourtant on ne peut pas en conclure que X et Y ont la même loi de probabilité.

8. Si X et Y sont deux v.a. indépendantes qui admettent une variance, on sait que $V(X+Y) = V(X) + V(Y)$. Dans quel cas peut-on avoir la relation $V(X-Y) = V(X) - V(Y)$?

9. Un examen comporte n questions qu'un candidat tire au hasard dans le programme, avec remise. On note X la v.a. qui représente le nombre de questions tirées dont il connaît la réponse et Y celle qui représente le nombre de celles dont il ignore la réponse. Si p est la proportion de questions du programme que ce candidat connaît, calculer $V(X)$, $V(Y)$ puis $V(X+Y)$. Que remarque-t-on ?

ENTRAÎNEMENT

Loi de probabilité

10. On considère une urne contenant une boule blanche et deux boules noires identiques. On effectue deux tirages successifs dans cette urne, la première boule tirée étant remplacée par une boule de couleur différente. On demande de construire l'ensemble fondamental Ω associé à cette épreuve aléatoire, si on tient compte de l'ordre des tirages, et de déterminer la probabilité de chacun des événements élémentaires. En déduire la loi de probabilité de la v.a. X qui représente le nombre de boules noires tirées.

Analyse de l'énoncé et conseils. Dans tout problème de tirages successifs, il est important d'établir s'ils sont indépendants ou non. Pour la construction de l'ensemble fondamental associé, il est recommandé de toujours tenir compte de l'ordre des tirages, même dans le cas où cela n'a aucune influence sur les événements considérés, car cela facilite généralement le calcul des probabilités de ceux-ci. La loi de probabilité d'une v.a. discrète X se déduit des probabilités des événements élémentaires, en déterminant pour toutes ses valeurs possibles x l'ensemble $X^{-1}(x)$ des événements élémentaires auxquels correspond cette même valeur x.

11. Une urne contient six boules dont quatre blanches et deux noires. On extrait une boule de l'urne, puis on la remet et on effectue ensuite des tirages sans remise jusqu'à obtention d'une boule de même couleur. Déterminer la loi de probabilité du nombre X de tirages après remise de la boule tirée initialement.

Analyse de l'énoncé et conseils. Comme dans l'exercice précédent, on note que la composition de l'urne évolue au cours des tirages qui sont sans remise, donc dépendants. Pour déterminer la loi de X, il faut d'abord déterminer l'ensemble de ses valeurs possibles et ensuite leurs probabilités. Pour cela, il est conseillé de distinguer deux cas selon le résultat du tirage préliminaire, c'est-à-dire de déterminer la loi conditionnelle, puis de regrouper ces deux cas en appliquant la formule de la probabilité totale vue au chapitre précédent.

Espérance mathématique

12. Un joueur paie une mise de cent francs pour participer à un jeu où il doit obtenir deux résultats pile successifs en quatre lancers au plus d'une pièce de monnaie. Son gain est alors $G = 20^{\,4-X}$ pour un nombre aléatoire X de lancers. Calculer son espérance de gain.

Analyse de l'énoncé et conseils. Une espérance mathématique se calcule par rapport à une loi de probabilité. Comme G est fonction de X qui est une v.a., c'est aussi une v.a. Pour chaque valeur possible x de X, il faut calculer la valeur correspondante du gain G et multiplier par la probabilité du résultat conduisant à ce gain et qui ne sera pas la probabilité de la valeur x. En effet, l'expression de G ne correspond qu'aux résultats qui se terminent par deux pile.

13. Un garagiste commande au constructeur N voitures, le nombre aléatoire X de voitures qu'il peut vendre dans l'année étant l'un des entiers de 0 à n, tous ayant la même probabilité, où n est un entier fixé tel que $n \geq N$. Toute voiture vendue rapporte au garagiste un bénéfice a et toute voiture invendue entraîne une perte b. Calculer l'espérance du gain G du garagiste et en déduire ce que doit être sa commande optimum.

Analyse de l'énoncé et conseils. Il faut au préalable exprimer le gain en fonction de X et son espérance va ensuite être calculée par rapport à la loi de X qu'il faut déterminer en tenant compte des éléments fournis par l'énoncé. Une commande est optimum si elle rend maximale cette espérance.

Loi de probabilité et moments

14. Une v.a. X peut prendre l'une des trois valeurs 0, 1 ou 2 avec des probabilités positives. Déterminer sa loi de probabilité sachant que $E(X) = 1$ et $V(X) = 1/2$.

Analyse de l'énoncé et conseils. L'expression des deux premiers moments non centrés à l'aide des probabilités individuelles va nous fournir deux relations qui lient ces trois valeurs. La troisième permettant de les déterminer s'obtient à partir de la condition générale relative à une loi de probabilité.

15. Un jouet se trouve caché dans l'une des N boîtes fermées où un enfant le cherche. Celui-ci ouvre une boîte au hasard et recommence jusqu'à ce qu'il trouve le jouet. On suppose qu'à chaque tentative il a oublié le résultat de toutes les précédentes. Déterminer la loi de probabilité de la v.a. X égale au nombre de tentatives effectuées jusqu'à la découverte du jouet, puis calculer son espérance. Préciser le nombre N_0 de boîtes si l'enfant a environ trois chances sur quatre de trouver le jouet à l'issue de ses trois premières tentatives.

Analyse de l'énoncé et conseils. On reconnaîtra dans le descriptif de l'expérience le schéma d'une loi usuelle et il suffira d'appliquer la formule du cours pour le calcul de l'espérance. La valeur de la probabilité que X soit inférieure ou égale à 3 étant fixée à 0,75 environ, ceci va déterminer l'entier N_0 auquel est associée une valeur proche.

Loi du maximum

16. Dans une urne qui contient N boules numérotées de 1 à N, on en extrait avec remise n. Déterminer la loi de probabilité de la v.a. X égale au plus grand des n numéros tirés.

Analyse de l'énoncé et conseils. Les tirages étant avec remise sont indépendants. Il est facile de voir que X peut prendre toutes les valeurs entières de 1 à N. Pour déterminer la probabilité d'un entier k donné, il est conseillé de calculer d'abord la probabilité de l'événement $X \leq k$ puis de remplacer k par $k - 1$.

17. La v.a. X représente le chiffre obtenu après le lancer d'un dé à six faces numérotées de 1 à 6. Déterminer la loi de probabilité de la v.a. $Y = X(7 - X)$ puis calculer $E(Y)$ et $V(Y)$. On désigne par $Y_1, \ldots, Y_n$ les valeurs observées à l'issue de n lancers indépendants. Déterminer la loi de probabilité de la v.a. M_n égale à la plus grande de ces valeurs.

Analyse de l'énoncé et conseils. La loi de Y se déduit de celle de X en calculant la valeur obtenue pour chaque valeur possible de X. Pour déterminer ensuite la loi du maximum, il faut exprimer l'événement $M_n = k$ à l'aide d'événements portant sur les Y_i, en faisant parfois intervenir l'événement complémentaire et en se rappelant que toutes les v.a. Y_i sont indépendantes et de même loi. Par ailleurs, il est conseillé de commencer par les valeurs de k qui correspondent aux événements les plus simples à exprimer.

Lois usuelles

18. On désigne par X la v.a. qui représente le nombre de boules rouges obtenues après cinq tirages avec remise dans une urne qui contient deux boules rouges et six boules blanches. Déterminer sa loi de probabilité puis calculer $E(X)$ et $V(X)$.

Analyse de l'énoncé et conseils. Il est facile de reconnaître ici une loi usuelle dont les moments sont donnés par le cours.

19. On désigne par X la v.a. qui représente le nombre de boules blanches obtenues après trois tirages sans remise dans une urne qui contient deux boules noires et huit boules blanches. Déterminer sa loi de probabilité puis calculer $E(X)$ et $V(X)$.

Analyse de l'énoncé et conseils. Voir exercice précédent.

SOLUTIONS

1 ▸ Faux. Une variable aléatoire est une application qui à un événement fait correspondre un nombre.

2 ▸ Vrai. Si on prend comme tribu $\mathcal{A} = \mathcal{P}(\Omega)$, toute application X est une v.a. puisque $X^{-1}(x) \in \mathcal{P}(\Omega)$ pour tout x réel.

3 ▸ Faux. C'est simplement la moyenne en probabilité de toutes ses valeurs possibles et ne correspond d'ailleurs pas forcément à l'une des valeurs qu'elle peut prendre. Par exemple, si X est la v.a. qui code 0 le résultat pile et 1 le résultat face d'un lancer de pièce de monnaie :

$$E(X) = 0 \times P_X(X = 0) + 1 \times P_X(X = 1) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$$

La valeur moyenne en probabilité de X est $1/2$, bien que X ne prenne jamais cette valeur.

4 ▸ Vrai. La seule condition est bien sûr que l'espérance de chacune de ces deux variables existe.

5 ▸ Faux. Une espérance commune n'a aucune incidence sur la valeur de la variance. Considérons par exemple les deux distributions suivantes :

X	4	8	12
	1/4	1/4	1/2

Y	−8	6	66
	1/2	1/3	1/6

Elles ont comme valeurs moyennes :

$$E(X) = 1 + 2 + 6 = 9 \quad \text{et} \quad E(Y) = -4 + 2 + 11 = 9$$

donc même centre de distribution. Par contre :

$$E\left(X^2\right) = 4 + 16 + 72 = 92 \quad \text{et} \quad E\left(Y^2\right) = 32 + 12 + 726 = 770$$

d'où $V(X) = 11$ et $V(Y) = 689$, valeur très supérieure qui indique une dispersion de Y autour de sa moyenne beaucoup plus grande que celle de X.

6 ▸ Faux. La variance d'une somme de deux v.a. indépendantes est égale à la somme de leurs variances, mais le résultat n'est pas vrai en général pour deux variables quelconques, contrairement à ce qui se produit pour l'espérance.

7 ▸ Seule la pièce en métal est attirée par l'aimant et on a donc $P(\{1\}) = 1$ et $P(\{2, 3\}) = 0$. Par ailleurs :

$$P(X = Y) = P\big(\{\omega \in \Omega / X(\omega) = Y(\omega)\}\big) = P(\{1\}) = 1$$

donc ces deux applications sont égales avec une probabilité égale à 1 (événement certain). L'application Y est constante, avec $Y^{-1}(1) = \Omega$, donc c'est une variable aléatoire certaine : $P(Y = 1) = 1$. Cependant, on ne peut pas en conclure que la loi de X est la même, puisque X n'est pas une variable aléatoire et que cette notion n'a donc pas de sens. En effet, $X^{-1}(2) = \{2\} \notin \mathcal{A}$, donc l'application X n'est pas une v.a. et on ne peut donc pas lui associer une loi de probabilité.

8 ▸ Si on pose $Z = X - Y$, on peut écrire $X = Y + Z$ et la relation proposée s'écrit alors $V(Z) = V(X) - V(Y)$, ce qui est la même chose que $V(X) = V(Y) + V(Z)$. Cette relation sera donc vérifiée en particulier si les variables Y et Z sont indépendantes. Il ne faut surtout pas lire cette relation sous la forme « variance d'une différence égale à la différence des variances » car c'est simplement un jeu d'écriture consistant à écrire sous une autre forme la relation (qui elle est vraie) « variance d'une somme de deux v.a. indépendantes égale à la somme des variances ».

9 ▸ La v.a. X suit une loi binômiale $\mathcal{B}\left(n, p\right)$, la probabilité que le candidat connaisse une réponse tirée étant p et les tirages étant indépendants, puisque avec remise. On en déduit comme résultat du cours que $V(X) = np\left(1 - p\right)$. De même, puisqu'il y a une probabilité $1 - p$ que le candidat ne connaisse pas une réponse, la v.a. Y suit une loi binômiale $\mathcal{B}\left(n, 1 - p\right)$ et on a aussi $V\left(Y\right) = np\left(1 - p\right)$. Bien entendu on a toujours $X + Y = n$, la somme de ces deux variables n'étant pas aléatoire, avec donc $V(X + Y) = 0$. La somme des variances n'est pas égale à la variance de la somme de ces deux v.a. qui ne sont pas indépendantes, leur somme étant d'ailleurs constante, donc de variance nulle.

10 ▸ La composition initiale de l'urne est $\{B, N, N\}$ et selon que la première boule tirée sera B ou N, les compositions respectives avant le second tirage seront $\{N, N, N\}$ ou $\{B, B, N\}$. Dans le premier cas, seul le résultat (BN) est possible, alors que dans le second cas on peut avoir (NB) ou (NN). Ceci correspond donc à l'ensemble fondamental $\Omega = \{(BN), (NB), (NN)\}$. Les tirages sont bien sûr dépendants ici puisque la composition de l'urne dépend du tirage précédent. Pour le calcul des probabilités des événements élémentaires, nous allons noter B_i (resp. N_i) l'événement « avoir tiré une boule blanche (resp. noire) au i-ème tirage », avec $i = 1$ ou 2. On obtient alors aisément $P\left(BN\right) = P\left(B_1\right) P\left(N_2|B_1\right) = \frac{1}{3} \times 1 = \frac{1}{3}$, $P\left(NB\right) = P\left(N_1\right) P\left(B_2|N_1\right) = \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$ et $P\left(NN\right) = P\left(N_1\right) P\left(N_2|N_1\right) = \frac{2}{3} \times \frac{1}{3} = \frac{2}{9}$. L'ensemble fondamental nous indique que X ne peut prendre comme valeurs que 1 ou 2, avec $X^{-1}(1) = \{(BN), (NB)\}$ et $X^{-1}\left(2\right) = \{(NN)\}$ et par conséquent :

$$P\left(X = 1\right) = P\left(BN\right) + P\left(NB\right) = \frac{7}{9}$$

$$P\left(X = 2\right) = P\left(NN\right) = \frac{2}{9}$$

11 ▸ Si la boule tirée initialement est blanche, comme il n'y a que deux noires, il y aura au maximum trois tirages. Nous noterons B_i (resp. N_i) l'événement « avoir tiré une boule blanche (resp. noire) au i-ème tirage », avec $0 \leq i \leq 5$. La loi de X, conditionnellement à B_0, se détermine par :

$$P\left(X = 1|B_0\right) = P\left(B_1\right) = \frac{4}{6} = \frac{10}{15}$$

$$P\left(X = 2|B_0\right) = P\left(N_1 B_2\right) = \frac{2}{6} \times \frac{4}{5} = \frac{4}{15}$$

$$P\left(X = 3|B_0\right) = P\left(N_1 N_2 B_3\right) = \frac{2}{6} \times \frac{1}{5} \times \frac{4}{4} = \frac{1}{15}$$

Notons au passage que la somme de ces probabilités conditionnelles est bien égale à 1.

Si la boule tirée initialement est noire, il peut y avoir cette fois jusqu'à cinq tirages avant d'obtenir une noire, puisqu'il y a quatre blanches. La loi conditionnelle est définie par :

$$P(X=1|N_0) = P(N_1) = \frac{2}{6} = \frac{5}{15}$$
$$P(X=2|N_0) = P(B_1N_2) = \frac{4}{6} \times \frac{2}{5} = \frac{4}{15}$$
$$P(X=3|N_0) = P(B_1B_2N_3) = \frac{4}{6} \times \frac{3}{5} \times \frac{2}{4} = \frac{3}{15}$$
$$P(X=4|N_0) = P(B_1B_2B_3N_4) = \frac{4}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{2}{15}$$
$$P(X=5|N_0) = P(B_1B_2B_3B_4N_5) = \frac{4}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \times \frac{2}{2} = \frac{1}{15}$$

La loi de X se déduit de la formule de la probabilité totale en décomposant l'événement $(X = x)$ sur B_0 et N_0, avec pour $x \in X(\Omega) = \{1, 2, 3, 4, 5\}$:

$$P(X=x) = P(B_0)\,P(X=x|B_0) + P(N_0)\,P(X=x|N_0)$$

soit $P(X=1) = \frac{2}{3} \times \frac{10}{15} + \frac{1}{3} \times \frac{5}{15} = \frac{5}{9}$, $P(X=2) = \frac{2}{3} \times \frac{4}{15} + \frac{1}{3} \times \frac{4}{15} = \frac{4}{15}$, $P(X=3) = \frac{2}{3} \times \frac{1}{15} + \frac{1}{3} \times \frac{3}{15} = \frac{1}{9}$, $P(X=4) = \frac{1}{3} \times \frac{2}{15} = \frac{2}{45}$ et $P(X=5) = \frac{1}{3} \times \frac{1}{15} = \frac{1}{45}$. On vérifie à nouveau que la somme de ces probabilités est égale à 1.

12 ▸ Les valeurs possibles de X sont 2, 3 et 4, avec un gain qui sera associé seulement aux résultats correspondants PP, FPP et $FFPP$ ou $PFPP$ de probabilités respectives $\frac{1}{2^2}$, $\frac{1}{2^3}$ et $\frac{1}{2^4}$. On obtient donc comme espérance de gain, en tenant compte de la mise initiale de 100 francs :

$$E(G) = \frac{1}{2^2} \times 20^{4-2} + \frac{1}{2^3} \times 20^{4-3} + \frac{2}{2^4} \times 20^{4-4} - 100 = \frac{21}{8}$$

13 ▸ Si $X \geq N$, le nombre de voitures vendues est X, ce qui correspond à un gain pour le garagiste de valeur $G = aN$. Si $X < N$, le garagiste vendra X voitures et il lui en restera $N - X$, ce qui correspond à un gain $G = aX - b(N - X)$. L'expression de l'espérance du gain est donc définie à partir de la loi de X par :

$$E(G) = aNP(X \geq N) + \sum_{x=0}^{N-1} [ax - b(N-x)]\,P(X=x)$$

Tous les entiers x de $\{0, 1, \ldots, n\}$ ont la même probabilité, c'est-à-dire que X suit une loi uniforme discrète, avec $P(X = x) = 1/(n+1)$. On obtient donc :

$$\begin{aligned} E(G) &= aN \sum_{x=N}^{n} \frac{1}{n+1} + \sum_{x=0}^{N-1} \frac{(a+b)x - bN}{n+1} \\ &= \frac{aN(n-N+1)}{n+1} + (a+b)\frac{N(N-1)}{2(n+1)} - \frac{bN^2}{n+1} \\ &= \frac{N[(2n+1)a - b - (a+b)N]}{2(n+1)} \end{aligned}$$

Cette valeur est maximum quand le numérateur $h(N)$ est maximum, avec $h'(N) = (2n+1)a - b - 2(a+b)N$ qui s'annule pour $N = \dfrac{(2n+1)a - b}{2(a+b)}$ et une dérivée seconde $h''(N) = -2(a+b) < 0$. La commande optimum correspond donc bien à cette valeur de N.

14 ▸ Les probabilités $p_i = P(X = i)$ pour $i = 0, 1, 2$ sont bien sûr liées par $p_0 + p_1 + p_2 = 1$. On donne d'autre part $E(X) = p_1 + 2p_2 = 1$. Enfin, par $E(X^2) = V(X) + E^2(X) = \dfrac{3}{2} = p_1 + 4p_2$, on déduit $p_2 = 1/4$, $p_1 = 1/2$ et $p_0 = 1/4$.

15 ▸ L'absence de mémoire de l'enfant se traduit par une indépendance des tentatives successives qui ont à chaque fois la même probabilité $1/N$ de réussir. Ces tentatives s'effectuant jusqu'à la découverte du jouet, il s'agit d'une loi géométrique de paramètre $p = 1/N$, soit pour tout entier strictement positif k :

$$P(X = k) = \left(1 - \frac{1}{N}\right)^{k-1} \frac{1}{N}$$

et une espérance $E(X) = N$.

Le jouet est trouvé à l'issue de n tentatives avec une probabiité qui se calcule par :

$$P(X \leq n) = \sum_{k=1}^{n} P(X = k) = \frac{1}{N} \sum_{k=1}^{n} \left(1 - \frac{1}{N}\right)^{k-1} = 1 - \left(1 - \frac{1}{N}\right)^{n}$$

Pour $n = 3$, on doit donc trouver une valeur de N telle que $(1 - 1/N)^3$ soit proche de 0,25, ce qui correspond à N voisin de 2,7, donc le nombre de boîtes est l'entier le plus proche $N_0 = 3$.

16 ▸ L'événement $X \leq k$ exprime qu'à chaque tirage on a obtenu une boule de l'ensemble $\{1, 2, \ldots, k\}$ et, en raison de l'indépendance des tirages :

$$P(X \leq k) = \left(\frac{k}{N}\right)^{n}$$

On en déduit alors par différence :

$$P(X = k) = P(X \leq k) - P(X \leq k-1) = \left(\frac{k}{N}\right)^{n} - \left(\frac{k-1}{N}\right)^{n}$$

pour $2 \leq k \leq N$, avec $P(X = 1) = (1/N)^n$.

17 ▸ Le plus simple est de déterminer les valeurs de Y à partir du tableau des valeurs possibles de X :

X	1	2	3	4	5	6
Y	6	10	12	12	10	6

La loi de X étant la loi uniforme discrète sur les entiers $\{1, 2, 3, 4, 5, 6\}$, il en est de même de la loi de Y sur les entiers $\{6, 10, 12\}$. On obtient alors aisément :

$$E(X) = \frac{1}{3}(6 + 10 + 12) = \frac{28}{3}$$
$$E\left(X^2\right) = \frac{1}{3}(36 + 100 + 144) = \frac{280}{30} \qquad V(X) = \frac{56}{9}$$

La variable M_n peut prendre aussi les valeurs 6, 10 ou 12. Pour que M_n soit égale à 6, il faut que toutes les variables Y_i aient pris la valeur 6, d'où :

$$P(M_n = 6) = P\left\{\bigcap_{i=1}^{n}(Y_i = 6)\right\} = \left(\frac{1}{3}\right)^n$$

Si M_n est égale à 12, c'est que l'un des événements $Y_i = 12$ est réalisé, le complémentaire étant tous les événements $Y_i < 12$ sont réalisés, soit :

$$P(M_n = 12) = P\left\{\bigcup_{i=1}^{n}(Y_i = 12)\right\} = 1 - P\left\{\bigcap_{i=1}^{n}(Y_i \leq 10)\right\} = 1 - \left(\frac{2}{3}\right)^n$$

La probabilité du dernier événement $M_n = 10$ pourrait se déduire des probabilités précédentes puisque la somme de ces valeurs doit être égale à 1, mais cette méthode est vivement déconseillée car elle enlève justement la possibilité de détection d'une erreur en faisant la somme des probabilités individuelles dont on sait qu'elle doit être égale à 1. On exprime donc l'événement $M_n = 10$ qui signifie que tous les événements $Y_i \leq 10$ sont réalisés, mais pas l'événement tous les Y_i sont égaux à 6, soit :

$$\begin{aligned} P(M_n = 10) &= P\left\{\bigcap_{i=1}^{n}(Y_i \leq 10) \bigcap \overline{\bigcap_{i=1}^{n}(Y_i = 6)}\right\} \\ &= P\left\{\bigcap_{i=1}^{n}(Y_i \leq 10)\right\} - P\left\{\bigcap_{i=1}^{n}(Y_i = 6)\right\} \\ &= \left(\frac{2}{3}\right)^n - \left(\frac{1}{3}\right)^n \end{aligned}$$

18 ▸ Les tirages étant avec remise sont indépendants et donc X suit une loi binômiale de paramètres $n = 5$ et $p = 2/8$ avec $E(X) = np = 5/4$ et $V(X) = np\left(1 - p\right) = 15/16$.

19 ▸ On reconnaît le schéma de la loi hypergéométrique de paramètres $N = 10$, $n = 3$ et $N_A = 8$, soit d'après les résultats du cours :

$$P(X = 0) = 0\,, \quad P(X = 1) = \frac{1}{15}\,, \quad P(X = 2) = P(X = 3) = \frac{7}{15}$$

avec $E(X) = 3 \times \frac{8}{10} = \frac{12}{5}$ et $V(X) = 3 \times \frac{8}{10} \times \frac{2}{10} \times \frac{7}{9} = \frac{28}{75}$.

Variable aléatoire continue

3

L'ESSENTIEL DU COURS

On associe aux résultats d'une expérience aléatoire un ensemble de valeurs qui forment un (ou plusieurs) intervalle(s) réel(s). Si on peut calculer la probabilité de tous les intervalles qui sont inclus dans cet ensemble, l'application qui réalise ce codage se nomme variable aléatoire (v.a.) continue. Sa loi de probabilité est généralement définie par une densité qui sera la dérivée de la fonction de répartition. Le calcul des moments s'effectue à l'aide de cette densité par une intégration.

1. Variable aléatoire réelle continue

1.1 Définition

On appelle v.a. réelle définie sur $(\Omega, \mathcal{A})$ une application $X : \Omega \to \mathbb{R}$ telle que pour tout intervalle $I \subset \mathbb{R}$ on ait :

$$X^{-1}(I) = \{\omega \in \Omega / X(\omega) \in I\} \in \mathcal{A}$$

Cette condition est vérifiée si pour tout réel x, on a $X^{-1}\left(]-\infty, x[\right) \in \mathcal{A}$.

1.2 Loi de probabilité

Elle est déterminée par la fonction de répartition (f.r.) F, définie pour tout x réel par :

$$F(x) = P_X(X < x) = P\left\{X^{-1}\left(]-\infty, x[\right)\right\} = P\left\{\omega \in \Omega / X(\omega) < x\right\}$$

1.3 Propriétés de la fonction de répartition

Elle est croissante au sens large et prend ses valeurs entre 0 et 1 :

$$0 \leq F(x) \leq 1 \quad \text{avec} \quad \lim_{x\to-\infty} F(x) = 0 \quad \text{et} \quad \lim_{x\to+\infty} F(x) = 1$$

Elle est continue à gauche :

$$\lim_{h\to 0^+} F(x-h) = F(x)$$

1.4 Loi continue

Si la fonction F est continue, on dit que X est une *variable aléatoire réelle continue*. Dans ce cas, pour tout réel x, $P_X(X = x) = 0$, et on dit que la loi est *diffuse*.

1.5 Loi absolument continue

Dans le cas où la fonction F admet une dérivée f, celle-ci est appelée *densité de probabilité* de la v.a. X, de loi qualifiée d'*absolument continue*. La f.r. est alors déterminée par l'intégrale :

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

Une densité est positive et intégrable sur $\mathbb{R}$, d'intégrale égale à 1 :

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$$

La probabilité d'un intervalle s'obtient en intégrant la densité sur cet intervalle :

$$P_X\{X \in [x_1, x_2]\} = \int_{x_1}^{x_2} f(t)\,\mathrm{d}t$$

1.6 Moments d'une v.a. absolument continue

► Espérance mathématique

Elle est définie par :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x$$

lorsque cette intégrale généralisée existe. Pour tout réel a :

$$E(X+a) = E(X) + a \quad \text{et} \quad E(aX) = aE(X)$$

Si X et Y sont deux v.a. qui admettent une espérance :

$$E(X+Y) = E(X) + E(Y)$$

► **Variance**

Elle est définie par :

$$V(X) = E\left[X - E(X)\right]^2 = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)\,\mathrm{d}x = E\left(X^2\right) - E^2(X) = \sigma^2(X)$$

lorsque cette intégrale généralisée existe. Pour tout réel a :

$$V(X + a) = V(X) \quad \text{et} \qquad V(aX) = a^2 V(X)$$

Si X et Y sont deux v.a. indépendantes admettant une variance :

$$V(X + Y) = V(X) + V(Y)$$

► **Moments non centrés et centrés**

Le *moment non centré d'ordre* $r \in \mathbb{N}^*$ de X est la quantité, lorsqu'elle existe :

$$m_r(X) = E(X^r) = \int_{-\infty}^{+\infty} x^r f(x)\,\mathrm{d}x$$

Le *moment centré d'ordre* $r \in \mathbb{N}^*$ de X est la quantité, lorsqu'elle existe :

$$\mu_r(X) = E\left[X - E(X)\right]^r = \int_{-\infty}^{+\infty} [x - E(X)]^r f(x)\,\mathrm{d}x$$

2. Lois usuelles continues

2.1 Loi uniforme

Une v.a. X suit une loi uniforme si sa densité est constante sur un intervalle $[a, b]$:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon} \end{cases}$$

On écrit $X \rightsquigarrow \mathcal{U}([a, b])$. Sa fonction de répartition est définie par :

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \dfrac{x-a}{b-a} & \text{si } a \leq x < b \\ 1 & \text{si } b \leq x \end{cases}$$

Elle admet pour moments :

$$E(X) = \frac{b+a}{2} \quad \text{et} \qquad V(X) = \frac{(b-a)^2}{12}$$

2.2 Loi exponentielle

La loi exponentielle de paramètre $\theta > 0$ est celle d'une variable positive de densité :

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

On écrit $X \rightsquigarrow \mathcal{E}(\theta)$. Sa f.r. est nulle pour $x \leq 0$, et pour $x > 0$:

$$F(x) = \theta \int_0^x e^{-\theta t}\, dt = \left[-e^{-\theta t}\right]_0^x = 1 - e^{-\theta x}$$

Ses moments sont :

$$E(X) = \frac{1}{\theta} \quad \text{et} \quad V(X) = \frac{1}{\theta^2}$$

2.3 Loi normale ou de Laplace-Gauss

C'est la loi d'une variable aléatoire X à valeurs dans $\mathbb{R}$, de densité :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x-m)^2}{2\sigma^2}$$

On note $X \rightsquigarrow N(m, \sigma)$ avec $E(X) = m$ et $V(X) = \sigma^2$.

Si $X \rightsquigarrow N(m_1, \sigma_1)$ et $Y \rightsquigarrow N(m_2, \sigma_2)$ sont des v.a. indépendantes, alors :

$$X + Y \rightsquigarrow N\left(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right).$$

2.4 Loi gamma

Une v.a. X de loi gamma de paramètres $p > 0$ et $\theta > 0$ est positive, de densité :

$$f(x) = \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1} \quad \text{si } x \geq 0$$

La fonction gamma est définie pour tout $p > 0$ par :

$$\Gamma(p) = \int_0^{+\infty} e^{-x} x^{p-1}\, dx$$

On écrit $X \rightsquigarrow \gamma(p, \theta)$, de moments :

$$E(X) = \frac{p}{\theta} \quad \text{et} \quad V(X) = \frac{p}{\theta^2}$$

La v.a. $Y = \theta X$ est de loi $\gamma(p, 1)$, notée $\gamma(p)$, de moments $E(Y) = V(Y) = p$.

Si $X \rightsquigarrow \gamma(p, \theta)$ et $Y \rightsquigarrow \gamma(q, \theta)$ sont des v.a. indépendantes, alors $X + Y \rightsquigarrow \gamma(p+q, \theta)$.

2.5 Loi du khi-deux

La loi du khi-deux à n degrés de liberté, notée χ_n^2, est la loi $\gamma(n/2, 1/2)$ où n est un entier positif. Ses moments se déduisent de ceux de la loi gamma :

$$E\left(\chi_n^2\right) = \frac{n/2}{1/2} = n \quad \text{et} \quad V\left(\chi_n^2\right) = \frac{n/2}{1/4} = 2n$$

Si $X \rightsquigarrow \chi_n^2$ et $Y \rightsquigarrow \chi_m^2$ sont des v.a. indépendantes, alors $X + Y \rightsquigarrow \chi_{n+m}^2$.

Notons une propriété importante qui peut servir de définition de cette loi : si $X_1, \ldots, X_n$ sont des v.a. indépendantes et de même loi $N(0,1)$, alors $X_1^2 + \cdots + X_n^2$ suit une loi du khi-deux à n degrés de liberté.

2.6 Loi de Student

Soit U une v.a. de loi $N(0,1)$ et Y une autre v.a. indépendante de loi χ_n^2. Le rapport $U/\sqrt{Y/n}$ suit une loi de Student à n degrés de liberté, notée T_n. Comme le numérateur U suit une loi symétrique par rapport à 0, il en est de même de T_n, avec $E(T_n) = 0$ pour $n > 1$. On obtient aussi $V(T_n) = \dfrac{n}{n-2}$ pour $n > 2$.

2.7 Loi de Fisher-Snedecor

Si U et V sont deux v.a. indépendantes de lois respectives χ_n^2 et χ_m^2, alors le rapport $(U/n)\,/\,(V/m)$ suit une loi de Fisher-Snedecor à n et m degrés de liberté, notée $F(n,m)$.

POUVEZ-VOUS RÉPONDRE ?

	Vrai	Faux
1. La f.r. d'une loi absolument continue est strictement croissante sur l'ensemble des nombres réels x tels que $F(x) > 0$.	☐	☐
2. Si une loi est symétrique par rapport à l'origine, c'est-à-dire si $F(-x) = 1 - F(x)$ ou $f(-x) = f(x)$, alors son espérance est nulle.	☐	☐
3. Si les v.a. X et Y suivent des lois normales, alors $X + Y$ suit aussi une loi normale.	☐	☐
4. Si une loi admet une variance, alors elle admet aussi une espérance.	☐	☐
5. Si une v.a. X est telle que $P(X = x) = 0$ pour tout x réel, alors elle admet une densité.	☐	☐
6. Si on définit la v.a. $Y = g(X)$ à l'aide d'une fonction continue g, alors on peut déterminer sa loi à partir de celle de X, même si g n'est pas injective.	☐	☐

QUESTIONS DE RÉFLEXION

7. Comment déterminer si la loi de probabilité d'une v.a. réelle X est continue (ou diffuse), mixte ou absolument continue, connaissant sa fonction de répartition F ?

8. Si une loi est définie par une densité, indiquer comment on peut étudier l'existence de ses moments centrés.

ENTRAÎNEMENT

Fonction de répartition

9. Soit f la fonction définie par :

$$f(x) = \begin{cases} k(1-x) & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Déterminer la constante k pour que f soit la densité d'une loi dont on précisera la fonction de répartition.

Analyse de l'énoncé et conseils. Pour qu'une fonction soit une densité, il faut qu'elle soit positive et d'intégrale égale à 1. C'est cette dernière condition qui va permettre de calculer précisément la valeur de la constante k. On détermine ensuite la f.r. en intégrant la densité de $-\infty$ à x, en choisissant x dans chacun des intervalles où l'expression de cette densité reste la même.

Loi de Pareto

10. Une v.a. X suit une loi de Pareto de densité f définie par :

$$f(x) = \begin{cases} \dfrac{k}{x^3} & \text{si } x \geq x_0 \\ 0 & \text{sinon} \end{cases}$$

où x_0 est une constante positive. Déterminer la constante k pour que f soit bien une densité et préciser la fonction de répartition de X.

Analyse de l'énoncé et conseils. Voir exercice précédent.

Loi symétrique

11. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \frac{3x}{4}(2-x) & \text{si } 0 \leq x \leq 2 \\ 0 & \text{sinon} \end{cases}$$

Déterminer la f.r. de X, puis indiquer sans calculs les valeurs de $P(X > 1)$ et $E(X)$.

Analyse de l'énoncé et conseils. La densité se présente sous la forme usuelle d'une fonction qui est non nulle seulement sur un intervalle fini, le support de la loi. La f.r. sera donc nulle avant l'origine de cet intervalle, puisqu'il n'y a aucune masse de probabilité, et vaudra 1 au-delà de l'extrémité du support puisque toute la masse de probabilité sera située à gauche du point considéré. Si on remarque une symétrie de la densité par rapport au centre de l'intervalle, celui-ci sera centre de symétrie de la loi, c'est-à-dire sera à la fois médiane et moyenne de la loi.

Calcul de moments

12. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \frac{2x}{\theta^2} & \text{si } 0 \leq x \leq \theta \\ 0 & \text{sinon} \end{cases}$$

où θ est un nombre positif donné. Déterminer la f.r. F de X puis calculer $E(X)$ et $V(X)$.

Analyse de l'énoncé et conseils. Le calcul des moments se fait par intégration, sans difficultés particulières dans le cas d'une densité finie sur un intervalle fini. Pour le calcul de la variance, il est presque toujours préférable de calculer $E\left(X^2\right)$ puis d'utiliser la formule développée.

Loi de Laplace

13. Soit X une v.a. de densité :

$$f(x) = \frac{1}{2}e^{-|x-\theta|}$$

où θ est un nombre réel donné. Calculer $E(X)$ et $V(X)$.

Analyse de l'énoncé et conseils. Le principe est le même que dans l'exercice précédent mais il s'agit ici d'intégrales généralisées dont il faut examiner la convergence, puisqu'on intègre sur tout $\mathbb{R}$. On pourra ici simplifier le calcul par la recherche d'un centre de symétrie de la loi.

Loi définie par sa f.r.

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

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

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

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

Lecture de tables : loi normale

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

b) Calculer l'espérance et la variance d'une v.a. Y de loi normale, telle que $P(Y > -3) = 0{,}6915$ et $P(Y < 2) = 0{,}9772$.

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

Loi normale et loi du khi-deux

16. a) Soit X une v.a. de loi normale telle que $P(X < 2) = 0{,}3085$ et $P(X > 8) = 0{,}0062$. Calculer la valeur de la probabilité $P\left(X^2 - 6X < 1{,}84\right)$.

b) Soit Y une v.a. de loi normale telle que $P(Y < 2) = 0{,}0228$ et $P(Y > 3{,}5) = 0{,}1587$. Calculer la valeur du réel a tel que $P\left\{(X-3)^2 < a\right\} = 0{,}975$.

Analyse de l'énoncé et conseils. Comme dans l'exercice précédent, les probabilités fournies vont permettre de déterminer l'espérance et l'écart type de la loi. Il faudra ensuite utiliser le résultat de cours relatif à la loi du carré d'une v.a. de loi normale centrée-réduite.

Fractiles des lois de Student et de Fisher-Snedecor

17. a) Déterminer les fractiles d'ordres 0,8 et 0,2 de la loi de Student à 12 degrés de liberté.

b) Déterminer le fractile d'ordre 0,05 de la loi de Fisher-Snedecor à 30 et 10 degrés de liberté.

Analyse de l'énoncé et conseils. On utilise les tables 6, page 200, et 7, page 201, des fractiles, en utilisant les propriétés de ces lois.

Loi exponentielle tronquée

18. Soit X une v.a. de densité :

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

où θ est un nombre réel donné.

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

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

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

Loi du maximum

19. Soit X une v.a. de densité :

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

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

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

Changement de variable

20. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \dfrac{1}{\theta} x^{1/\theta - 1} & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

où θ est un nombre positif donné.

Déterminer la loi de probabilité de la v.a. $Y = -\ln X$.

Analyse de l'énoncé et conseils. La loi de Y est déterminée par sa fonction de répartition que l'on exprimera en fonction de celle de X, puis que l'on dérivera pour reconnaître une loi usuelle.

Loi de Gumbel

21. Soit X une v.a. de densité :

$$f(x) = \exp\left[-(x-\theta) - e^{-(x-\theta)}\right]$$

et $X_1, \ldots, X_n$ des v.a. indépendantes et de même loi que X. Déterminer la loi de probabilité de $Y = \exp - (X - \theta)$ et en déduire celle de $S_n = \sum_{i=1}^{n} Y_i$, ayant posé $Y_i = \exp - (X_i - \theta)$, $1 \leq i \leq n$.

Analyse de l'énoncé et conseils. La loi de Y se détermine d'abord par sa f.r. que l'on exprime à l'aide de celle de X. On calcule ensuite sa densité par dérivation. On reconnaîtra une loi usuelle, ce qui permettra d'en déduire facilement celle de S_n par application d'un résultat de cours.

SOLUTIONS

1 ▸ Faux. La densité f est une fonction positive qui peut très bien s'annuler sur certains intervalles où la f.r. sera alors constante. Considérons par exemple la densité définie par :

$$f(x) = \begin{cases} -\dfrac{x}{4} - \dfrac{1}{2} & \text{si } -4 < x < -2 \\ \dfrac{x}{4} - \dfrac{1}{2} & \text{si } 2 < x < 4 \\ 0 & \text{sinon} \end{cases}$$

Pour $-2 < x \leq 2$ la densité est nulle et $F(x) = F(-2) = 1/2$ reste constante.

2 ▸ Faux. On peut donner comme contre-exemple la loi de Cauchy de densité $1/\pi\left(1+x^2\right)$ et qui pourtant n'admet pas d'espérance. Cependant, si l'intégrale généralisée $\int_{-\infty}^{+\infty} xf(x)\, \mathrm{d}x$ existe, alors on a bien $E(X)=0$.

3 ▸ Faux. La variable $Y=-X$ suit une loi normale, comme X, et $X+Y=0$, donc cette variable suit une loi de Dirac (que l'on pourrait envisager comme cas limite d'une loi normale d'écart type nul). Bien sûr l'énoncé devient vrai si on ajoute la condition d'indépendance des variables X et Y.

4 ▸ Vrai. L'existence d'un moment d'un certain ordre implique celle de tous les moments d'ordres inférieurs.

5 ▸ Faux. La loi est continue, la f.r. est continue et strictement croissante mais pas forcément dérivable. On peut trouver des exemples assez complexes où la loi est continue mais pas absolument continue, c'est-à-dire n'admettant pas de densité.

6 ▸ Vrai. Prenons l'exemple de la fonction g définie par $g(x)=|x|$. Elle n'est pas injective et pourtant on peut déterminer la loi de $Y=|X|$. En effet, pour tout $y>0$ l'événement $Y<y$ est équivalent à $-y<X<y$, dont on peut calculer la probabilité. Cependant, si g n'est pas injective on n'a pas la certitude de pouvoir toujours déterminer la loi de Y.

7 ▸ Si la fonction F est dérivable, on peut conclure que sa dérivée est la densité de la loi, qui est alors absolument continue. Sinon, on étudie la continuité de F sur son domaine de définition. Si elle est continue partout, la loi est continue, ou diffuse, c'est-à-dire qu'il n'y a aucune masse de probabilité en aucun point. Par contre, s'il existe au moins un point x_0 où F est discontinue, on a alors $F(x_0+0)=\lim_{h\to 0^+} F(x_0+h) \neq F(x_0)$; la loi est mixte, comportant une partie continue et une partie discrète, car il y a au moins un point de masse non nulle : $P(X=x_0)=F(x_0+0)-F(x_0)$.

8 ▸ L'existence d'un moment centré d'ordre entier r est équivalente à celle du moment non centré de même ordre, qui s'établit en étudiant la convergence de l'intégrale généralisée $\int_{-\infty}^{+\infty} x^r f(x)\, \mathrm{d}x$. Si cette intégrale existe pour tout r, c'est que la loi admet des moments de tous ordres. Si on établit l'existence d'un moment d'ordre r_0, tous les moments d'ordres inférieurs à r_0 existent.

9 ▸ L'intégrale de f sur $]-\infty,+\infty[$ se réduit à :

$$k\int_0^1 (1-x)\, \mathrm{d}x = \frac{k}{2}$$

qui vaut 1 pour $k=2$.

Pour $x\leq 0$, la densité est nulle et donc $F(x)=0$. Pour $0<x\leq 1$, on décompose l'intervalle d'intégration de f en deux :

$$F(x)=\int_{-\infty}^{x} f(t)\, \mathrm{d}t=\int_{-\infty}^{0} 0\, \mathrm{d}t+2\int_0^x (1-t)\, \mathrm{d}t=2x-x^2$$

Enfin, pour $x>1$ on décompose en trois intégrales qui correspondent aux trois intervalles où f conserve la même expression :

$$F(x)=\int_{-\infty}^{x} f(t)\, \mathrm{d}t=\int_{-\infty}^{0} 0\, \mathrm{d}t+2\int_0^1 (1-t)\, \mathrm{d}t+\int_1^x 0\, \mathrm{d}t=1$$

10 ▸ L'intégrale de f sur $\mathbb{R}$ se réduit à :

$$k\int_{x_0}^{+\infty} x^{-3}\,\mathrm{d}x = \frac{k}{2x_0^2}$$

qui vaut 1 pour $k = 2x_0^2$.

Pour $x < x_0$, la densité est nulle et donc $F(x) = 0$. Pour $x \geq x_0$, on décompose l'intervalle d'intégration de f en deux :

$$F(x) = \int_{-\infty}^{x} f(t)\ \mathrm{d}t = \int_{-\infty}^{x_0} 0\,\mathrm{d}t + 2x_0^2\int_{x_0}^{x} t^{-3}\,\mathrm{d}t = 1 - \frac{x_0^2}{x^2}$$

11 ▸ Pour $x \leq 0$, la densité est nulle, donc son intégrale aussi et $F(x) = 0$. Pour $0 < x \leq 2$, le calcul de la f.r. F se réduit à :

$$F(x) = \frac{3}{4}\int_0^x \left(2t - t^2\right)\ \mathrm{d}t = \frac{3x^2}{4} - \frac{x^3}{4}$$

Pour $x \geq 2$, on intègre la densité sur tout l'intervalle où elle est non nulle, donc son intégrale vaut 1 et $F(x) = F(2) = 1$. On peut remarquer ici que $f(2-x) = f(x)$ et cette symétrie de la densité implique la symétrie de la loi par rapport au centre $x = 1$ de l'intervalle, qui est donc la médiane, avec $P(X < 1) = P(X > 1) = \frac{1}{2}$, et aussi la moyenne, avec $E(X) = 1$, puisque l'intégrale s'effectue sur un intervalle de longueur finie.

12 ▸ On obtient aisément :

$$F(x) = \begin{cases} 0 & \text{si } \ x < 0 \\ \dfrac{x^2}{\theta^2} & \text{si } \ 0 \leq x \leq \theta \\ 1 & \text{si } \ x > \theta \end{cases}$$

On calcule l'espérance :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\ \mathrm{d}x = \frac{2}{\theta^2}\int_0^\theta x^2\,\mathrm{d}x = \frac{2}{\theta^2}\left[\frac{x^3}{3}\right]_0^\theta = \frac{2}{3}\theta$$

Puis le moment non centré d'ordre deux :

$$E\left(X^2\right) = \int_{-\infty}^{+\infty} x^2 f(x)\ \mathrm{d}x = \frac{2}{\theta^2}\int_0^\theta x^3\,\mathrm{d}x = \frac{2}{\theta^2}\left[\frac{x^4}{4}\right]_0^\theta = \frac{\theta^2}{2}$$

Et enfin $V(X) = E\left(X^2\right) - E^2(X) = \frac{\theta^2}{18}\cdot$

13 ▸ Tous les moments de cette loi existent, d'après le théorème des croissances comparées, en raison de la présence de l'exponentielle qui l'emporte sur toutes les puissances de x. Cela assure donc la convergence des intégrales $\int_{-\infty}^{+\infty} x^r f(x)\, dx$ pour tous les entiers r. Si on remarque que $f(2x - \theta) = f(x)$, on en conclut que θ est centre de symétrie de cette loi, donc la médiane, mais aussi la moyenne puisque nous savons que celle-ci existe : $E(X) = \theta$. Pour le calcul du moment d'ordre deux, il sera donc judicieux de faire le changement $u = x - \theta$:

$$E\left(X^2\right) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx = \frac{1}{2}\int_{-\infty}^{+\infty} (u+\theta)^2\, e^{-|u|}\, du$$
$$= \int_0^{+\infty} u^2 e^{-u}\, du + \theta^2 \int_0^{+\infty} e^{-u}\, du = 2 + \theta^2$$

et $V(X) = E\left(X^2\right) - E^2(X) = 2$.

14 ▸ Il s'agit bien d'une fonction de répartition, strictement croissante de $F(1) = 0$ à $F(e) = 1$, de dérivée nulle en dehors de l'intervalle $[1, e]$, avec $F'(x) = f(x) = 1/x$ pour $1 < x \leq e$. On calcule donc :

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx = \int_1^e dx = e - 1$$
$$E\left(X^2\right) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx = \int_1^e x\, dx = \frac{1}{2}\left(e^2 - 1\right)$$

et $V(X) = E\left(X^2\right) - E^2(X) = \frac{1}{2}\left(-e^2 + 4e - 3\right)$.

15 ▸ **a)** On retranche 35 puis on divise par 5 pour pouvoir ensuite utiliser la table 1, page 194 :

$$P(X < 25) = P\left(\frac{X - 35}{5} < -2\right) = \Phi(-2) = 1 - \Phi(2) = 0{,}022\,8$$

où Φ est la f.r. de la loi $N(0, 1)$. On obtient par le même procédé :

$$P(37{,}5 < X < 40) = \Phi(1) - \Phi(0{,}5) = 0{,}149\,8$$

et :

$$P(32{,}5 < X < 37{,}5) = \Phi(0{,}5) - \Phi(-0{,}5) = 2\Phi(0{,}5) - 1 = 0{,}383\,0$$

b) Les valeurs fournies pour les probabilités figurent dans la table 1, page 194, associées aux nombres 0,5 et 2. En centrant sur $m = E(Y)$ puis divisant par $\sigma = \sqrt{V(Y)}$, les hypothèses se traduisent donc par :

$$P(Y > -3) = P\left(\frac{Y - m}{\sigma} > \frac{-3 - m}{\sigma}\right) = 1 - \Phi\left(\frac{-3 - m}{\sigma}\right) = \Phi(0{,}5)$$
$$P(Y < 2) = P\left(\frac{Y - m}{\sigma} < \frac{2 - m}{\sigma}\right) = \Phi\left(\frac{2 - m}{\sigma}\right) = \Phi(2)$$

Ce qui conduit au système :

$$\frac{-3-m}{\sigma} = -0{,}5 \quad \text{et} \quad \frac{2-m}{\sigma} = 2$$

de solution $m = -2$ et $\sigma = 2$.

16 ▸ **a)** La lecture de la table 1, page 194, permet d'obtenir :

$$P\left(\frac{X-m}{\sigma} < \frac{2-m}{\sigma}\right) = 0{,}308\,5 = 1 - 0{,}691\,5 = 1 - \Phi\,(0{,}5) = \Phi\,(-0{,}5)$$

où Φ est la f.r. de la loi $N\,(0,1)$ et ayant posé $m = E\,(X)$ et $\sigma = \sqrt{V(X)}$. De même, $P\left(\frac{X-m}{\sigma} < \frac{8-m}{\sigma}\right) = 0{,}993\,8 = \Phi\,(2{,}5)$, ce qui donne les équations :

$$\frac{2-m}{\sigma} = -0{,}5 \quad \text{et} \quad \frac{8-m}{\sigma} = 2{,}5$$

de solution $m = 3$ et $\sigma = 2$. On remarque ensuite que $X^2 - 6X = (X-3)^2 - 9$, donc on doit calculer la probabilité $p = P\left\{(X-3)^2 < 10{,}84\right\}$. On sait par ailleurs que $(X-3)^2/4$ suit une loi du khi-deux à un degré de liberté, donc on calcule $p = P\left\{(X-3)^2/4 < 2{,}71\right\} = 0{,}90$, valeur lue dans la table 5, page 199.

b) Avec $m = E\,(Y)$ et $\sigma = \sqrt{V(Y)}$, on obtient :

$$P\left(\frac{Y-m}{\sigma} < \frac{2-m}{\sigma}\right) = 1 - 0{,}977\,2 = \Phi\left(\frac{2-m}{\sigma}\right) = 1 - \Phi\,(2) = \Phi\,(-2)$$

et :

$$P\left(\frac{Y-m}{\sigma} < \frac{3{,}5-m}{\sigma}\right) = 0{,}841\,3 = \Phi\,(1)$$

soit :

$$\frac{2-m}{\sigma} = -2 \quad \text{et} \quad \frac{3,5-m}{\sigma} = 1$$

de solution $m = 3$ et $\sigma = 1/2$. La loi de $4(X-3)^2$ est une loi du khi-deux à un degré de liberté, donc $4a$ est le fractile d'ordre 0,975 de cette loi, lu dans la table 5, soit $4a = 5{,}02$ et $a = 1{,}255$.

17 ▸ **a)** La lecture de la table 6, page 200, fournit la valeur 0,873 du fractile d'ordre 0,8. En raison de la symétrie de cette loi par rapport à l'origine, le fractile d'ordre 0,2 est donc $-0{,}873$.

b) La table 7, page 201, ne comporte pas le couple de degrés de liberté $(30,10)$, mais le couple $(10,30)$, qui fournit le fractile 2,16, soit $P(X < 2{,}16) = 0{,}95$ avec $X \rightsquigarrow F(10,30)$. L'inverse de X suit une loi $F(30,10)$, donc $P(1/X > 1/2{,}16) = 0{,}95$ et le fractile d'ordre 0,05 de la loi $F(30,10)$ est donc $1/2{,}16 = 0{,}46$.

18 ▸ **a)** La f.r. est nulle pour $x \leq \theta$. Son expression pour $x > \theta$ est :

$$F(x) = \int_0^x e^{-(t-\theta)}\,dt = 1 - e^{-(x-\theta)}$$

La médiane vérifie donc $e^{-(Md-\theta)} = 1/2$, soit $Md = \theta + \ln 2$.

b) L'événement $(m_n > x)$ est équivalent à l'événement « tous les éléments de l'échantillon sont plus grands que x », ce qui s'écrit :

$$(m_n > x) = \bigcap_{i=1}^{n} (X_i > x)$$

Du fait de l'indépendance des variables X_i et de leur identité de loi avec X :

$$P(m_n > x) = P\left\{\bigcap_{i=1}^{n} (X_i > x)\right\} = \prod_{i=1}^{n} P(X_i > x) = P^n(X > x)$$

Ainsi :

$$G(x) = P(m_n < x) = 1 - P^n(X > x) = 1 - [1 - F(x)]^n$$

soit :

$$G(x) = \begin{cases} 0 & \text{si } x \leq \theta \\ 1 - e^{-n(x-\theta)} & \text{si } x > \theta \end{cases}$$

Et par dérivation :

$$g(x) = G'(x) = nf(x)[1 - F(x)]^{n-1} = \begin{cases} 0 & \text{si } x \leq \theta \\ ne^{-n(x-\theta)} & \text{si } x > \theta \end{cases}$$

19 ▸ La f.r. de X est nulle pour $x \leq -3$ et vaut un pour $x > 3$. Si $-3 \leq x \leq 3$, on obtient :

$$F(x) = \frac{1}{36}\int_{-3}^{x} \left(9 - t^2\right) dt = -\frac{x^3}{108} + \frac{x}{4} + \frac{1}{2}$$

L'événement $(M_n < x)$ est équivalent à « toutes les variables de l'échantillon sont inférieures à x » et comme toutes les v.a. sont indépendantes et de même loi que X, on obtient :

$$G(x) = P(M_n < x) = P\left\{\bigcap_{i=1}^{n} (X_i < x)\right\} = \prod_{i=1}^{n} P(X_i < x) = F^n(x)$$

La densité étant :

$$g(x) = nf(x)F^{n-1}(x)$$

20 ▸ Puisque $0 < X < 1$, la variable $Y = -\ln X$ est une variable positive dont la f.r. est définie pour $y > 0$ par :

$$G(y) = P\left(Y < y\right) = P\left(-\ln X < y\right) = P\left(X > \mathrm{e}^{-y}\right) = 1 - F\left(\mathrm{e}^{-y}\right)$$

Sa densité est obtenue par dérivation :

$$g(y) = G'(y) = \mathrm{e}^{-y} f\left(\mathrm{e}^{-y}\right) = \frac{1}{\theta}\mathrm{e}^{-y/\theta}$$

On reconnaît la loi exponentielle de paramètre $1/\theta$.

21 ▸ La v.a. Y est positive et pour $y > 0$ sa f.r. est définie par :

$$\begin{aligned} G(y) &= P\left(Y < y\right) = P\left\{\exp - (X - \theta) < y\right\} = P\left\{-(X - \theta) < \ln y\right\} \\ &= P\left\{X - \theta > -\ln y\right\} = P\left\{X > \theta - \ln y\right\} = 1 - F\left(\theta - \ln y\right) \end{aligned}$$

Par dérivation, on obtient comme densité pour $y > 0$:

$$g(y) = \frac{1}{y} f\left(\theta - \ln y\right) = \mathrm{e}^{-y}$$

qui est la densité d'une loi exponentielle ou loi $\gamma(1)$. La variable S_n est donc la somme de v.a. indépendantes qui suivent la même loi $\gamma(1)$ et suit donc une loi $\gamma(n)$.

Couple et vecteur aléatoires

4

L'ESSENTIEL DU COURS

Ce chapitre généralise à plusieurs dimensions la notion de variable aléatoire. Un vecteur aléatoire a comme composantes des variables aléatoires réelles. Le vecteur des espérances va définir l'espérance de ce vecteur. La généralisation de la variance est la matrice de variances-covariances qui contient les variances des composantes sur la diagonale principale et les covariances comme autres éléments. Si toute combinaison linéaire des composantes d'un vecteur aléatoire suit une loi normale, alors ce vecteur suit une loi normale multidimensionnelle.

1. Couple de v.a. discrètes

1.1 Loi d'un couple

Un couple de v.a. discrètes est constitué de deux v.a. discrètes X et Y dont les ensembles de valeurs possibles sont $\{x_i\}_{i\in I}$ et $\{y_j\}_{j\in J}$, avec $I, J \subset \mathbb{N}$, de loi définie par :

$$p_{ij} = P\left(X = x_i, Y = y_j\right)$$

1.2 Lois marginales

À la loi d'un couple sont associées deux lois marginales :

$$P\left(X = x_i\right) = \sum_{j\in J} P\left(X = x_i, Y = y_j\right) = \sum_{j\in J} p_{ij} = p_{i.}$$

$$P\left(Y = y_j\right) = \sum_{i\in I} P\left(X = x_i, Y = y_j\right) = \sum_{i\in I} p_{ij} = p_{.j}$$

1.3 Lois conditionnelles

Pour $Y = y_j$ fixé, la loi conditionnelle de X est définie par :

$$P\left(X = x_i | Y = y_j\right) = \frac{P\left(X = x_i, Y = y_j\right)}{P\left(Y = y_j\right)} = \frac{p_{ij}}{p_{\cdot j}} = p_i^j$$

1.4 Indépendance

Les v.a. X et Y sont indépendantes si pour tout $i \in I$ et tout $j \in J$:

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

1.5 Moments associés à un couple

Si $h : \mathbb{R}^2 \to \mathbb{R}$ est une application continue, elle définit une v.a. $h(X, Y)$ d'espérance qui se calcule par :

$$E\left[h(X, Y)\right] = \sum_{i \in I} \sum_{j \in J} p_{ij} h\left(x_i, y_j\right)$$

Pour $h(X, Y) = [X - E(X)]\,[Y - E(Y)]$, on définit le moment appelé *covariance* de X et Y :

$$\text{Cov}\,(X, Y) = E\left\{[X - E(X)]\,[Y - E(Y)]\right\} = E(XY) - E(X)E(Y)$$

On appelle *coefficient de corrélation linéaire* de deux v.a. X et Y le nombre réel :

$$\rho = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

1.6 Loi d'une somme

La loi de probabilité de la v.a. $Z = X + Y$ est définie par :

$$P\left(Z = z_k\right) = \sum \left\{P\left(X = x_i, Y = y_j\right) / x_i + y_j = z_k\right\}$$

Si X et Y sont indépendantes, on parle alors de *convolution* des lois de X et Y et :

$$\begin{aligned} P\left(Z = z_k\right) &= \sum_{i \in I} P(X = x_i) P(Y = z_k - x_i) \\ &= \sum_{j \in J} P\left(Y = y_j\right) P\left(X = z_k - y_j\right) \end{aligned}$$

2. Couple de v.a. continues

2.1 Loi du couple

Si X et Y sont deux v.a. continues, la loi de (X, Y) est déterminée par sa f.r. :

$$F\left(x, y\right) = P\left(X < x, \quad Y < y\right)$$

Si F est dérivable par rapport à x et y, la loi de (X, Y) admet une densité f :

$$f\left(x, y\right) = \frac{\partial^2 F\left(x, y\right)}{\partial x \partial y}$$

2.2 Lois marginales

Les fonctions de répartition marginales de X et Y sont définies par :

$$F_X(x) = P(X < x) = F(x, +\infty) \quad \text{et} \quad F_Y(y) = P(Y < y) = F(+\infty, y)$$

Si la loi du couple est définie par sa densité, les densités marginales sont :

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}y \quad \text{et} \quad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}x$$

2.3 Lois conditionnelles

Pour une loi de densité f, les lois conditionnelles sont définies par les densités :

$$f_X\left(x | Y = y\right) = \frac{f(x, y)}{f_Y(y)} \quad \text{et} \quad f_Y\left(y | X = x\right) = \frac{f(x, y)}{f_X(x)}$$

2.4 Indépendance

Les v.a. X et Y sont dites *indépendantes*, si pour tous les réels x et y :

$$f(x, y) = f_X(x) f_Y(y)$$

2.5 Moments associés à un couple

Si $h : \mathbb{R}^2 \to \mathbb{R}$ est une application continue, elle définit une v.a. $h(X, Y)$ dont l'espérance se calcule par l'intégrale double :

$$E\left[h(X, Y)\right] = \iint_{\mathbb{R}^2} h(x, y) f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

2.6 Régression

La fonction : $x \mapsto E(Y|X=x)$, s'appelle *fonction de régression* de Y en X, avec :

$$E(Y|X=x) = \int_{-\infty}^{+\infty} y f_Y(y|X=x)\,\mathrm{d}y = \int_{-\infty}^{+\infty} y \frac{f(x,y)}{f_X(x)}\,\mathrm{d}y$$

2.7 Loi d'une somme

La loi de la v.a. $Z = X + Y$ se détermine par sa f.r. G, définie par :

$$G(z) = P(Z < z) = P(X + Y < z)$$

Elle peut se calculer, dans le cas où ce couple admet une densité f, par :

$$G(z) = \iint_D f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

où $D = \{(x,y)\,/x + y < z\}$. Dans le cas particulier où X et Y sont indépendantes :

$$G(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z} f_X(x) f_Y(s-x)\,\mathrm{d}s\,\mathrm{d}x = \int_{-\infty}^{z} g(s)\,\mathrm{d}s$$

où g est la densité de Z, définie par :

$$g(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,\mathrm{d}y$$

3. Vecteur aléatoire

Un vecteur aléatoire X de $\mathbb{R}^n$ est une application de Ω dans $\mathbb{R}^n$ dont toutes les composantes X_i, $1 \leq i \leq n$, sont des v.a. réelles. On définit son espérance par le vecteur :

$$E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{pmatrix}$$

La généralisation de la variance est la *matrice de variances-covariances* :

$$V(X) = E\left\{[X - E(X)]\ {}^t[X - E(X)]\right\}$$

Si A est une matrice de format (m, n) et b un vecteur de $\mathbb{R}^m$:

$$E(AX + b) = AE(X) + b \quad \text{et} \quad V(AX + b) = V(AX) = AV(X)\,{}^tA$$

3.1 Loi multinomiale

Le vecteur aléatoire N, de composantes $N_1, \ldots, N_k$, suit une *loi multinomiale* de paramètres $n, p_1, \ldots, p_k$ si :

$$P(N_1 = n_1, \ldots, N_k = n_k) = \frac{n!}{n_1! \ldots n_k!} p_1^{n_1} \ldots p_k^{n_k}$$

où $\sum_{j=1}^{k} p_j = 1$. Si on pose ${}^tp = \left(p_1, \ldots, p_k\right)$, on écrit $N \rightsquigarrow \mathcal{M}(n; p)$ d'espérance $E(N) = np$ et de matrice de variances-covariances :

$$V(N) = n\left[p_i\left(\delta_j^i - p_j\right)\right]_{1 \leq i,j \leq k}$$

où δ_j^i est le symbole de Kronecker qui vaut 1 quand $i = j$ et 0 sinon.

3.2 Loi normale vectorielle

Définition. Un vecteur aléatoire X à valeurs dans $\mathbb{R}^n$ suit une *loi normale* si toute combinaison linéaire de ses composantes suit une loi normale dans $\mathbb{R}$:

$$X \rightsquigarrow \mathcal{N}_n \quad \Longleftrightarrow \quad \forall a \in \mathbb{R}^n, \quad {}^taX = \sum_{i=1}^{n} a_i X_i \rightsquigarrow \mathcal{N}_1$$

La loi normale est déterminée par le vecteur espérance $\mu = E(X)$ et la matrice variances-covariances $\Sigma = V(X)$, la densité au point $x = (x_1, \ldots, x_n)$ étant :

$$f(x) = \frac{1}{\left(\sqrt{2\pi}\right)^n \sqrt{\det \Sigma}} \exp -\frac{1}{2}\,{}^t(x - \mu)\,\Sigma^{-1}\,(x - \mu)$$

Si $X_1, \ldots, X_n$ sont des v.a. normales **indépendantes**, de lois respectives $N(\mu_i, \sigma_i)$, $1 \leq i \leq n$, alors elles constituent les composantes d'un vecteur normal X de densité :

$$f(x) = \frac{1}{\left(\sqrt{2\pi}\right)^n \left[\prod_{i=1}^{n} \sigma_i\right]} \exp -\frac{1}{2} \sum_{i=1}^{n} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2$$

POUVEZ-VOUS RÉPONDRE ?

	Vrai	Faux
1. Les lois de X et Y permettent de déterminer celle du couple (X, Y).	☐	☐
2. L'expression $E(Y\|X)$ désigne une variable aléatoire.	☐	☐
3. Deux variables aléatoires indépendantes ont une covariance nulle.	☐	☐
4. La somme de deux lois binômiales est une loi binômiale.	☐	☐
5. Si on connaît les densités des deux v.a. continues X et Y, on peut calculer celle de $X + Y$.	☐	☐
6. Dans le cas multidimensionnel, l'opérateur espérance est linéaire.	☐	☐
7. Si la densité f d'un couple (X, Y) s'écrit sous la forme $f(x, y) = g(x)h(y)$, où g et h sont deux fonctions positives, alors les v.a. X et Y sont indépendantes.	☐	☐
8. Le vecteur dont les composantes $X_1, \ldots, X_n$ sont des v.a. normales est un vecteur normal.	☐	☐
9. Si X et Y sont deux variables aléatoires réelles normales de covariance nulle, elles sont indépendantes.	☐	☐

QUESTIONS DE RÉFLEXION

10. Comment l'examen du tableau donnant la loi d'un couple de v.a. discrètes permet-il de conclure facilement à la dépendance de ces deux variables ?

11. Indiquer le moyen généralement le plus simple pour obtenir les densités marginales d'un couple de v.a. continues dont on connaît la fonction de répartition.

12. Si X est une v.a. réelle de loi $N(\mu, \sigma)$, on lui associe la variable centrée-réduite $U = \sigma^{-1}(X - \mu)$, notamment pour pouvoir utiliser les tables. Dans le cas d'un vecteur aléatoire X de loi $\mathcal{N}_n(\mu, \Sigma)$, indiquer la transformation qui permet d'obtenir un vecteur de loi normale centrée-réduite $\mathcal{N}_n(0, I)$.

ENTRAÎNEMENT

Variance d'une somme

13. On lance un dé à six faces numérotées de 1 à 6 et on note X la v.a. qui représente le chiffre obtenu. Si celui-ci est pair, la v.a. Y prend la valeur 1 et 0 sinon. Calculer la variance $V(X+Y)$.

Analyse de l'énoncé et conseils. La valeur de Y dépend du résultat X ; donc on peut penser intuitivement que ces variables sont dépendantes. Pour pouvoir calculer la variance de la somme, il faut donc déterminer sa loi de probabilité, ce qui nécessite au préalable d'obtenir la loi du couple (X, Y).

Somme de variables dépendantes

14. Une urne contient quatre boules numérotées de 1 à 4 et l'on effectue deux tirages successifs sans remise, notant X (resp. Y) la v.a. indiquant le chiffre marqué sur la boule tirée au premier (resp. second) tirage. Déterminer la loi de probabilité de la v.a. $X+Y$.

Analyse de l'énoncé et conseils. La loi de la somme de ces variables ne peut être déterminée qu'après avoir établi celle du couple, car elles ne sont pas indépendantes. Les tirages étant sans remise, on ne peut pas tirer deux fois le même numéro. Pour obtenir facilement la loi de la somme, il est conseillé de faire figurer sa valeur dans le tableau donnant la loi du couple.

Loi conditionnelle discrète

15. On jette deux dés à six faces numérotées de 1 à 6 et on note X la v.a. qui représente la somme des deux chiffres obtenus et Y la v.a. qui est égale au plus grand de ces deux chiffres.

a) Déterminer la loi de probabilité du couple (X, Y) puis calculer $E(X)$. Les v.a. X et Y sont-elles indépendantes ?

b) Déterminer la loi conditionnelle de Y pour $X = 6$, puis son espérance $E(Y|X=6)$.

Analyse de l'énoncé et conseils. Les couples ordonnés (x_1, x_2) de résultats associés aux deux dés ont la même probabilité $1/36$. Pour chacun d'eux il faut calculer la somme x_1+x_2, qui donnera la valeur de X, et $\max\{x_1, x_2\}$ qui sera celle de Y, puis regrouper les couples de valeurs identiques ; leur nombre sera la probabilité, en $1/36$-ème du couple obtenu. La loi conditionnelle est obtenue à partir de la colonne du tableau associée à $X = 6$, en divisant les probabilités de cette colonne par la probabilité marginale $P(X=6)$.

Espérance conditionnelle

16. La loi de probabilité du couple (X, Y) est indiquée dans le tableau suivant, les valeurs étant exprimées en dizièmes :

$Y \backslash X$	1	2	3	4
1	1	1	2	1
2	0	1	1	0
3	1	0	0	2

Déterminer la loi de probabilité de la v.a. $E(Y|X)$, puis calculer son espérance et la comparer avec la valeur de $E(Y)$.

Analyse de l'énoncé et conseils. L'expression $E(Y|X)$ est bien celle d'une variable aléatoire puisqu'il s'agit d'une fonction de X, qui est elle-même une v.a. Pour une réalisation ω telle que X prend la valeur $X(\omega) = x$, cette variable prend la valeur de la régression en x, c'est-à-dire que $E(Y|X)(\omega) = E(Y|X = x)$. Il faut donc calculer les différentes valeurs de cette fonction de régression, qui seront les valeurs possibles de cette v.a. On peut ensuite calculer l'espérance de cette v.a. par rapport à la loi de X.

Loi conditionnelle binômiale

17. Le montant de l'indemnité versée à une entreprise qui a souscrit une police d'assurance incendie, après un sinistre, est une v.a. X de loi connue. Le nombre annuel d'incendies est une v.a. N qui suit une loi de Poisson de paramètre λ. Déterminer la loi de probabilité de la v.a. K_x qui représente le nombre annuel d'incendies d'un montant supérieur à une valeur x positive donnée. Déterminer ensuite la loi conditionnelle de N pour $K_x = k$.

Analyse de l'énoncé et conseils. Il faut envisager toutes les valeurs possibles n de N et ensuite reconnaître la loi de probabilité de K_x, pour chaque valeur fixée $N = n$. La loi de K_x s'obtiendra comme loi marginale du couple (K_x, N). Pour obtenir la loi conditionnelle de N, il suffit ensuite d'appliquer la définition d'une probabilité conditionnelle.

Couple de v.a. indépendantes

18. Soit (X, Y) un couple de v.a. dont la loi est définie par la densité :

$$f(x,y) = \begin{cases} 4x(1-y) & \text{si } 0 \leq x \leq 1, \quad 0 \leq y \leq 1 \\ 0 & \text{sinon} \end{cases}$$

a) Calculer la probabilité de l'événement $(0 \leq X < 1/3,\ 0 \leq Y < 1/3)$.

b) Déterminer les fonctions de répartition marginales des v.a. X et Y et établir qu'elles sont indépendantes.

Analyse de l'énoncé et conseils. Pour calculer la probabilité d'un événement déterminé par deux v.a. continues, on intègre la densité sur l'ensemble des valeurs qui réalisent cet événement. Les f.r. marginales vont se déduire des densités marginales obtenues par intégration de la densité du couple par rapport à l'autre variable. Si le produit de ces densités est égal à la densité donnée, c'est que les variables du couple sont indépendantes.

Couple de v.a. dépendantes

19. Soit (X, Y) un couple de v.a. dont la loi est définie par la densité :

$$f(x, y) = \begin{cases} 2e^{-(x+y)} & \text{si } 0 \leq y \leq x \\ 0 & \text{sinon} \end{cases}$$

a) Calculer la probabilité de l'événement $(X + Y < 1)$.

b) Déterminer la fonction de répartition du couple (X, Y).

c) Déterminer les densités marginales de X et Y. Ces variables sont-elles indépendantes ?

Analyse de l'énoncé et conseils. Pour déterminer la f.r. du couple en un point $M(x_0, y_0)$, il faut intégrer la densité sur le domaine $E =]-\infty, x_0[\times]-\infty, y_0[$. Il faudra alors distinguer plusieurs cas selon la position du point M par rapport au domaine D où cette densité est non nulle. Ayant déterminé cette f.r., on en déduit aisément les f.r. marginales en faisant tendre successivement y_0 et x_0 vers plus l'infini. C'est la méthode ici la plus simple pour obtenir ensuite les densités marginales par dérivation de ces f.r. marginales.

Valeurs extrêmes d'une loi uniforme

20. Soit $X_1, \ldots, X_n$ des v.a. indépendantes et de même loi uniforme sur $[0, 1]$. Déterminer la loi de probabilité du couple de valeurs extrêmes (m_n, M_n), ayant posé $m_n = \min\{X_1, \ldots, X_n\}$ et $M_n = \max\{X_1, \ldots, X_n\}$.

Analyse de l'énoncé et conseils. Il faut d'abord exprimer la f.r. du couple à l'aide de l'événement $(m_n \geq x) \cap (M_n < y)$ qui s'écrit facilement à l'aide d'événements associés aux variables $X_1, \ldots, X_n$ et dont la probabilité sera facile à calculer. La densité de ce couple de variables continues s'en déduira alors par dérivation.

Fonction de régression

21. Soit (X, Y) un couple de v.a. dont la loi est définie par la densité :

$$f(x, y) = \begin{cases} 3y & \text{si} \quad (x, y) \in D \\ 0 & \text{sinon} \end{cases}$$

où D est l'intérieur du triangle de sommets $(-1, 0), (0, 1)$ et $(1, 0)$. Déterminer l'expression de la fonction de régression de Y en X.

Analyse de l'énoncé et conseils. La fonction de régression est définie au point x par la valeur de l'espérance conditionnelle $E(Y|X = x)$. La loi étant continue, cette espérance va se calculer après détermination de la densité conditionnelle, obtenue en divisant la densité du couple par la densité marginale de X.

Loi multinomiale

22. On effectue n tirages avec remise dans une urne qui contient une boule noire, deux blanches et trois vertes. Déterminer la loi de probabilité du vecteur N dont les composantes N_1, N_2 et N_3 désignent respectivement le nombre de boules noires, blanches et vertes tirées. Calculer ensuite $E(N)$ et $V(N)$.

Analyse de l'énoncé et conseils. Les résultats du cours permettent d'obtenir sans difficultés les réponses.

Transformation linéaire d'un vecteur gaussien

23. Si X_1, X_2 et X_3 sont trois v.a. indépendantes de même loi $N(0, 1)$, déterminer la loi de probabilité du vecteur Y dont les composantes sont définies par :

$$Y_1 = \frac{1}{2}(X_1 + X_2), \quad Y_2 = -\frac{1}{3}(X_1 + X_2 - X_3), \quad Y_3 = \frac{1}{6}(-X_1 + X_2 + 2X_3)$$

Analyse de l'énoncé et conseils. Le vecteur Y est le transformé linéaire du vecteur X de composantes X_1, X_2 et X_3. On établira facilement que X est un vecteur gaussien, ce qui permettra d'en déduire par application du cours qu'il en est de même de Y.

SOLUTIONS

1 ▸ Faux. Dans le cas général, la connaissance des lois marginales ne permet pas de déterminer la loi du couple. Par contre, dans le cas particulier où les variables sont indépendantes, la loi du couple s'obtient comme produit des lois marginales.

2 ▸ Vrai. Il s'agit d'une variable aléatoire car c'est une fonction de X, qui est une variable aléatoire. Pour un événement élémentaire ω tel que X prend la valeur $X(\omega) = x$, cette v.a. aura comme réalisation la valeur de la fonction de régression au point x, c'est-à-dire $E(Y|X)(\omega) = E(Y|X = x)$. On peut donc calculer l'espérance de cette variable par rapport à la loi de X et on établit que $E[E(Y|X)] = E(Y)$.

3 ▸ Vrai. Si deux v.a. X et Y sont indépendantes, alors $E(XY) = E(X)E(Y)$, ce qui implique que $\text{Cov}(X, Y) = 0$. Par contre, la réciproque est fausse dans le cas général, c'est-à-dire que deux variables peuvent avoir une covariance nulle et être dépendantes.

4 ▸ Faux. Pour que l'énoncé devienne vrai en général, il faut ajouter la condition d'indépendance des deux lois. Si par exemple X suit une loi binômiale, $X + X = 2X$ ne suit pas une loi binômiale.

5 ▸ Faux. Il est nécessaire, dans le cas général, de connaître la loi du couple (X, Y) pour pouvoir déterminer celle de $X + Y$, puisqu'elle implique ces deux variables simultanément. Cependant, si les deux variables sont indépendantes, l'énoncé devient vrai.

6 ▸ Vrai. Cette propriété, qui était vraie en unidimensionnel, subsiste dans le cas multidimensionnel, c'est-à-dire que si X et Y sont deux vecteurs aléatoires à n composantes et A, B deux matrices de format (p, n), alors :

$$E(AX + BY) = AE(X) + BE(Y)$$

7 ▸ Vrai. Cependant, g et h ne sont pas forcément les densités marginales de X et Y car il y a plusieurs factorisations possibles. Par contre, si ce sont des densités, ce sont les densités marginales.

8 ▸ Faux. Les composantes d'un vecteur normal sont bien des v.a. normales, mais la réciproque est faussse en général : des composantes normales ne constituent pas toujours un vecteur normal. Il faut pour cela qu'elles soient indépendantes.

9 ▸ Faux. Pour qu'une covariance nulle implique l'indépendance de deux v.a., il ne suffit pas qu'elles soient normales, il faut qu'elles soient les composantes d'un vecteur normal.

10 ▸ Il suffit de trouver dans les marges une valeur $p_{i.}$ en ligne et une valeur $p_{.j}$ en colonne dont le produit soit différent de la valeur p_{ij} du tableau pour conclure à la dépendance de ces deux variables. Par contre, pour établir l'indépendance des variables d'un couple, il faut effectuer tous les produits des probabilités marginales et retrouver les valeurs données dans le tableau.

11 ▸ Connaissant les valeurs $F(x,y)$ de la f.r. du couple, en faisant tendre successivement y et x vers plus l'infini, on obtient les f.r. marginales $F_X(x)$ et $F_Y(y)$. Il suffit alors de dériver pour obtenir les densités marginales :

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} \quad \text{et} \quad f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y}$$

12 ▸ Le vecteur $X-\mu$ est bien sûr centré. Pour réduire, on utilise la propriété, liée aux matrices symétriques et définies-positives, d'existence d'une matrice symétrique S telle que $S^2 = \Sigma^{-1}$ et que nous noterons $S = \Sigma^{-1/2}$. La transformation $Y = \Sigma^{-1/2}(X-\mu)$ est celle demandée puisque :

$$V(Y) = \Sigma^{-1/2}V(X)\Sigma^{-1/2} = S(\Sigma S) = SS^{-1} = I$$

car $\Sigma S^2 = I$.

13 ▸ La loi du couple (X,Y) s'obtient sans difficultés. Elle est indiquée dans le tableau suivant :

Y \ X	1	2	3	4	5	6	
0	1/6	0	1/6	0	1/6	0	1/2
1	0	1/6	0	1/6	0	1/6	1/2
	1/6	1/6	1/6	1/6	1/6	1/6	1

On peut voir par exemple que $P(X=1, Y=1) = 0 \neq P(X=1)P(Y=1)$ et donc que les variables X et Y sont bien dépendantes. On déduit de ce tableau la loi de la somme :

$X+Y$	1	3	5	7
	1/6	1/3	1/3	1/6

Puis on calcule :

$$E(X+Y) = \frac{1}{6}\times 1 + \frac{1}{3}\times 3 + \frac{1}{3}\times 5 + \frac{1}{6}\times 7 = 4$$

et :

$$E(X+Y)^2 = \frac{1}{6}\times 1 + \frac{1}{3}\times 3^2 + \frac{1}{3}\times 5^2 + \frac{1}{6}\times 7^2 = \frac{59}{3}$$

soit enfin $V(X+Y) = \frac{11}{3}$. Le calcul de $V(X) = 35/12$ et $V(Y) = 1/4$ confirme que $V(X) + V(Y) \neq V(X+Y)$ et donc que X et Y ne peuvent pas être indépendantes.

14 ▸ Le premier chiffre x, tiré dans l'ensemble $E = \{1, 2, 3, 4\}$, a la probabilité $1/4$ et le second y, tiré dans $E - \{x\}$, la probabilité $1/3$ puisqu'il ne reste plus que trois boules. Chaque couple (x, y) tel que $x \neq y$ est donc de probabilité $\frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$. Cela permet de construire le tableau de la loi de probabilité du couple (X, Y), où les valeurs numériques sont exprimées en douzièmes et où la valeur entre parenthèses est celle de $x + y$.

Y \ X	1	2	3	4	
1	0(2)	1(3)	1(4)	1(5)	3
2	1(3)	0(4)	1(5)	1(6)	3
3	1(4)	1(5)	0(6)	1(7)	3
4	1(5)	1(6)	1(7)	0(8)	3
	3	3	3	3	12

De ce tableau se déduit aisément la loi de la somme :

$X + Y$	3	4	5	6	7
	$1/6$	$1/6$	$1/3$	$1/6$	$1/6$

15 ▸ **a)** La loi de probabilité du couple (X, Y) figure dans le tableau ci-dessous, les probabilités étant exprimées en $1/36$. La somme $X = 6$ est obtenue par exemple pour les deux résultats de dés $(1, 5)$ et $(5, 1)$ pour lesquels $Y = 5$, pour $(2, 4)$ et $(4, 2)$ avec $Y = 4$ et pour l'unique couple $(3, 3)$ qui donne $Y = 3$. Le couple (X, Y) prend donc les valeurs $(6, 5)$, $(6, 4)$ et $(6, 3)$ avec les probabilités respectives $2/36$, $2/36$ et $1/36$.

Y \ X	2	3	4	5	6	7	8	9	10	11	12	
1	1	0	0	0	0	0	0	0	0	0	0	1
2	0	2	1	0	0	0	0	0	0	0	0	3
3	0	0	2	2	1	0	0	0	0	0	0	5
4	0	0	0	2	2	2	1	0	0	0	0	7
5	0	0	0	0	2	2	2	2	1	0	0	9
6	0	0	0	0	0	2	2	2	2	2	1	11
	1	2	3	4	5	6	5	4	3	2	1	36

La dernière ligne du tableau donne la loi marginale de X. On constate qu'elle est symétrique par rapport à 7, donc $E(X) = 7$. Le produit des éléments des marges n'est pas toujours égal aux éléments du tableau ; par exemple $P(X = 3, Y = 1) = 0 \neq P(X = 3)\,P(Y = 1)$, donc les variables X et Y ne sont pas indépendantes.

b) En divisant la colonne $X = 6$ par $P(X = 6) = 5/36$, on obtient la loi conditionnelle :

$Y \mid X = 6$	3	4	5
	1/5	2/5	2/5

L'espérance de cette loi est :

$$E(Y|X = 6) = \frac{1}{5} \times 3 + \frac{2}{5} \times 4 + \frac{2}{5} \times 5 = \frac{21}{5}$$

16 ▸ Pour les différentes valeurs possibles x de X, on calcule $E(Y|X = x)$ et on obtiendra ainsi les différentes valeurs possibles de la variable aléatoire discrète $E(Y|X)$, avec les probabiliés associées $P(X = x)$. Par exemple :

$$E(Y|X = 1) = \frac{1}{2} \times 1 + \frac{1}{2} \times 3 = 2$$

valeur dont la probabilité est $P(X = 1) = 2/10$. On construit ainsi le tableau donnant la loi de probabilité de la v.a. $E(Y|X)$:

$E(Y \mid X)$	2	3/2	4/3	7/3
	2/10	2/10	3/10	3/10

Son espérance est donc :

$$E[E(Y|X)] = \frac{2}{10} \times 2 + \frac{2}{10} \times \frac{3}{2} + \frac{3}{10} \times \frac{4}{3} + \frac{3}{10} \times \frac{7}{3} = \frac{9}{5}$$

La loi marginale de Y, obtenue en additionnant les colonnes, permet d'obtenir :

$$E(Y) = \frac{1}{2} \times 1 + \frac{1}{5} \times 2 + \frac{3}{10} \times 3 = \frac{9}{5}$$

On remarque ainsi que $E[E(Y|X)] = E(Y)$, résultat qui est d'ailleurs toujours vérifié.

17 ▸ Pour tout entier k, l'événement $(K_x = k)$ peut se décomposer à l'aide des événements disjoints $\{(K_x = k) \cap (N = n)\}$ avec n qui peut varier de l'entier k à l'infini. Ceci permet d'écrire :

$$P(K_x = k) = \sum_{n=k}^{\infty} P\{(K_x = k) \cap (N = n)\} = \sum_{n=k}^{\infty} P(N = n) P(K_x = k|N = n)$$

Lorsque le nombre annuel d'incendies est fixé à n, chacun d'eux est d'un montant supérieur à x avec une probabilité connue $p_x = P(X > x)$. La loi conditionnelle de $K_x|N = n$ est donc une loi binômiale de paramètres n et p_x. Ainsi :

$$P(K_x = k) = \sum_{n=k}^{\infty} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} \binom{n}{k} p_x^k (1 - p_x)^{n-k} = \mathrm{e}^{-\lambda} \frac{(\lambda p_x)^k}{k!} \sum_{n=k}^{\infty} \frac{[\lambda(1 - p_x)]^{n-k}}{(n - k)!}$$

$$= \mathrm{e}^{-\lambda} \frac{(\lambda p_x)^k}{k!} \mathrm{e}^{\lambda(1-p_x)} = \mathrm{e}^{-\lambda p_x} \frac{(\lambda p_x)^k}{k!}$$

ayant utilisé le résultat $\sum_{j=0}^{\infty} \frac{z^j}{j!} = \mathrm{e}^z$. L'expression de cette probabilité montre que K_x suit une loi de Poisson de paramètre λp_x. Pour $K_x = k$ fixé, la v.a. N ne peut prendre que les valeurs entières $n \geq k$, avec :

$$P(N = n|K_x = k) = \frac{P\{(N = n) \cap (K_x = k)\}}{P(K_x = k)} = \frac{\mathrm{e}^{-\lambda} \lambda^n p_x^k (1 - p_x)^{n-k}}{(n - k)! \mathrm{e}^{-\lambda p_x} (\lambda p_x)^k}$$

$$= \mathrm{e}^{-\lambda(1-p_x)} \frac{[\lambda(1 - p_x)]^{n-k}}{(n - k)!}$$

18 ▸ **a)** On intègre la densité f sur le domaine de $\mathbb{R}^2$ où $0 \leq x < 1/3$ et $0 \leq y < 1/3$ pour obtenir la probabilité demandée, soit :

$$P(0 \leq X < 1/3, 0 \leq Y < 1/3) = 4 \int_0^{1/3} x \, \mathrm{d}x \int_0^{1/3} (1 - y) \, \mathrm{d}y = \frac{5}{81}$$

b) Les densités marginales s'obtiennent par intégration de la densité du couple par rapport à l'autre variable. Pour $0 \leq x \leq 1$, la densité de X s'obtient donc par :

$$g(x) = 4x \int_0^1 (1 - y) \, \mathrm{d}y = 2x$$

De la même façon, pour $0 \leq y \leq 1$:

$$h(y) = 4(1 - y) \int_0^1 x \, \mathrm{d}x = 2(1 - y)$$

En intégrant ces densités, on obtient les f.r. marginales demandées :

$$G(x) = \begin{cases} 0 & \text{si } x < 0 \\ x^2 & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } 1 < x \end{cases}$$

et :

$$H(y) = \begin{cases} 0 & \text{si } y < 0 \\ 2y - y^2 & \text{si } 0 \leq y \leq 1 \\ 1 & \text{si } 1 < y \end{cases}$$

Nous pouvions remarquer que la densité de (X, Y) peut s'écrire comme produit de deux fonctions positives, par exemple $4x\mathbf{1}_{[0,1]}(x)$ et $(1-y)\mathbf{1}_{[0,1]}(y)$, et donc que les v.a. X et Y sont indépendantes. Cependant, les deux fonctions retenues ici ne sont pas les densités marginales, qui bien sûr factorisent aussi f sous la forme $f(x,y) = g(x)h(y)$, ce qui confirme l'indépendance de ces deux variables.

19 ▸ a) On intègre la densité f sur la partie A de $\mathbb{R}^2$ où $x + y < 1$ pour obtenir :

$$P(X + Y < 1) = \iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 2\iint_{A\cap D} \mathrm{e}^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y$$

où D est le domaine où f est non nulle. L'ensemble $A \cap D$ est le triangle rectangle représenté dans la figure 4.1.

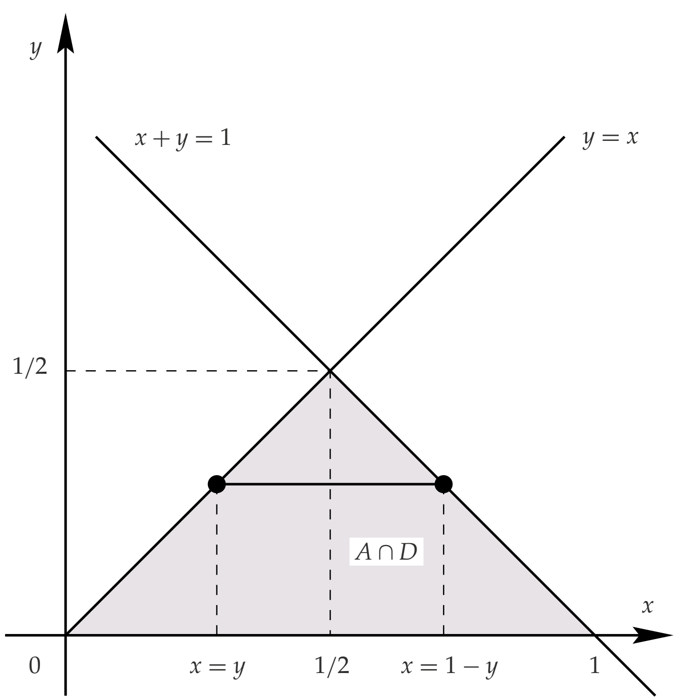


Figure 4.1

On voit donc qu'il est préférable d'intégrer d'abord par rapport à x sur une horizontale, pour ne pas avoir à distinguer deux cas, soit :

$$P(X+Y<1) = 2\int_0^{1/2}\left(\int_y^{1-y} \mathrm{e}^{-x}\,\mathrm{d}x\right)\mathrm{e}^{-y}\,\mathrm{d}y = 2\int_0^{1/2}\left(\mathrm{e}^{-y}-\mathrm{e}^{y-1}\right)\mathrm{e}^{-y}\,\mathrm{d}y$$
$$= \left[-\mathrm{e}^{-2y} - 2\mathrm{e}^{-1}y\right]_0^{1/2} = 1-2\mathrm{e}^{-1}$$

b) La f.r. F du couple (X,Y) est définie au point $M(x_0,y_0)$ par :

$$F(x_0,y_0) = P\left\{(X<x_0)\cap(Y<y_0)\right\} = \iint_E f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

avec $E =]-\infty, x_0[\times]-\infty, y_0[$. Pour $x_0 \leq 0$ ou $y_0 \leq 0$, l'ensemble E n'a aucune partie commune avec D et par conséquent $F(x_0,y_0) = 0$. Si le point M est situé à l'intérieur de D, on intègre sur le domaine grisé $D\cap E$ représenté sur la figure 4.2 où on voit qu'il est préférable d'intégrer sur une horizontale, c'est-à-dire d'abord par rapport à x. Ainsi, pour $x_0 \geq y_0 \geq 0$ on obtient :

$$F(x_0,y_0) = 2\int_0^{y_0}\left(\int_y^{x_0}\mathrm{e}^{-x}\,\mathrm{d}x\right)\mathrm{e}^{-y}\,\mathrm{d}y = 2\int_0^{y_0}\left(\mathrm{e}^{-y}-\mathrm{e}^{-x_0}\right)\mathrm{e}^{-y}\,\mathrm{d}y$$
$$= \left[-\mathrm{e}^{-2y} + 2\mathrm{e}^{-x_0}\mathrm{e}^{-y}\right]_0^{y_0} = 1 - 2\mathrm{e}^{-x_0}\left(1-\mathrm{e}^{-y_0}\right) - \mathrm{e}^{-2y_0}$$

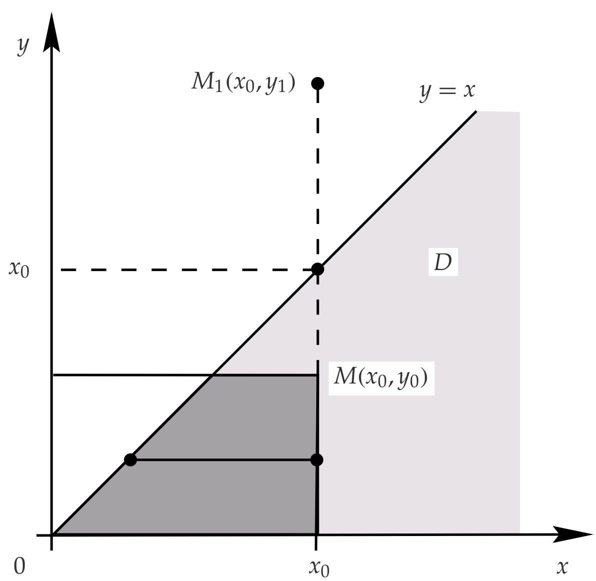


Figure 4.2

Si on sort du domaine D par une verticale de M, l'intersection avec l'ensemble E défini par le point $M_1(x_0,y_1)$, où $y_1 > x_0$, est la même que celle obtenue pour le point situé sur la première bissectrice, qui est la frontière de D. Cela veut dire que l'on peut calculer la valeur de F au point M_1 à partir de l'expression précédente qui est encore valable sur la frontière, soit :

$$F\left(x_0, y_1\right) = F\left(x_0, x_0\right) = 1 - 2\mathrm{e}^{-x_0} + \mathrm{e}^{-2x_0}$$

c) On pourrait avoir ici la tentation d'écrire $f\left(x, y\right) = 2\mathrm{e}^{-x} \times \mathrm{e}^{-y}$ et d'en conclure que les variables X et Y sont indépendantes, la densité s'écrivant comme produit de deux fonctions positives. Ceci est faux, car l'expression de f n'est celle-ci que sur le domaine D où $0 \leq y \leq x$ et il faudrait donc ajouter le terme $\mathbf{1}_D$, qui lui ne peut pas se factoriser. Cette impossibilité de factorisation montre donc, au contraire, que ces variables sont dépendantes. Nous allons le vérifier à partir de l'expression des densités marginales, que nous allons obtenir à partir des f.r. marginales. Pour y tendant vers plus l'infini, on obtient celle de X, soit pour $x > 0$:

$$F_X\left(x\right) = F\left(x, +\infty\right) = 1 - 2\mathrm{e}^{-x} + \mathrm{e}^{-2x}$$

La densité marginale est donc, par dérivation :

$$f_X\left(x\right) = 2\mathrm{e}^{-x}\left(1 - \mathrm{e}^{-x}\right)$$

De même, on obtient les f.r. et densité marginales de Y, soit pour $y > 0$:

$$F_Y\left(y\right) = F\left(+\infty, y\right) = 1 - \mathrm{e}^{-2y}$$
$$f_Y\left(y\right) = 2\mathrm{e}^{-2y}$$

Le produit de f_X et f_Y ne redonne pas f, donc on retrouve bien le fait que X et Y ne sont pas indépendantes.

20 ▸ La f.r. du couple (m_n, M_n) est définie par :

$$F\left(x, y\right) = P\left\{\left(m_n < x\right) \cap \left(M_n < y\right)\right\} = P\left(M_n < y\right) - P\left\{\left(m_n \geq x\right) \cap \left(M_n < y\right)\right\}$$

Les événements qui apparaissent au second membre peuvent s'exprimer facilement à l'aide des v.a. $X_1, \ldots, X_n$:

$$\left(M_n < y\right) = \bigcap_{i=1}^{n}\left(X_i < y\right) \quad \text{et} \quad \left(m_n \geq x\right) \cap \left(M_n < y\right) = \bigcap_{i=1}^{n}\left(x \leq X_i < y\right)$$

Les v.a. X_i étant indépendantes, il en est de même des événements qu'elles définissent, donc :

$$F\left(x, y\right) = \prod_{i=1}^{n} P\left(X_i < y\right) - \prod_{i=1}^{n} P\left(x \leq X_i < y\right)$$

D'autre part, toutes ces variables ont la même loi, de f.r. définie pour $0 \leq u \leq 1$ par $P\left(X_1 < u\right) = u$, d'où :

$$F\left(x, y\right) = y^n - \left(y - x\right)^n$$

pour $0 \leq x \leq y \leq 1$. En dérivant deux fois par rapport à x et à y, on en déduit la densité du couple :

$$f\left(x, y\right) = n\left(n - 1\right)\left(y - x\right)^{n-2}$$

21 ▸ La densité marginale de X s'obtient en intégrant la densité du couple sur D par rapport à y. Le triangle D, représenté dans la figure 4.3, est délimité par les droites $y = 1 + x$ et $y = 1 - x$. Il est symétrique par rapport à l'axe des ordonnées et comme l'expression de $f(x, y)$ est symétrique en x, la densité g de X est une fonction paire, c'est-à-dire que la loi est symétrique par rapport à l'origine. Pour $0 \leq x \leq 1$, on obtient :

$$g(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy = 3 \int_0^{1-x} y\, dy = \frac{3}{2}(1 - x)^2$$

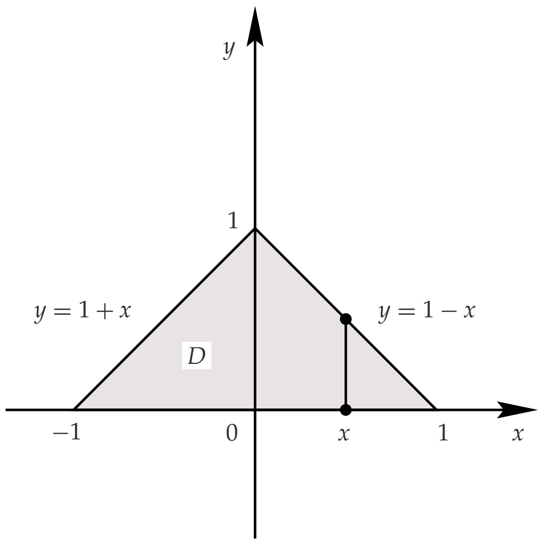


Figure 4.3

On peut donc écrire la densité de X sous la forme $g(x) = \frac{3}{2}(1 - |x|)^2$ pour $|x| \leq 1$ et la densité conditionnelle de $Y|X = x$ s'écrit donc, pour $0 < y < 1 - |x|$:

$$h(y|X = x) = \frac{2y}{(1 - |x|)^2}$$

L'expression de la régression est donc, pour $|x| < 1$:

$$E(Y|X = x) = \int_{-\infty}^{+\infty} yh(y|X = x)\, dy = \frac{2}{(1 - |x|)^2} \int_0^{1-|x|} y^2\, dy = \frac{2}{3}(1 - |x|)$$

22 ▸ Le vecteur N suit une loi multinomiale de paramètres n et $p_1 = 1/6, p_2 = 2/6, p_3 = 3/6$, soit :

$$P(N_1 = n_1, N_2 = n_2, N_3 = n_3) = \frac{n!}{n_1!n_2!n_3!} \frac{2^{n_2} 3^{n_3}}{6^n}$$

les n_i, $1 \leq i \leq 3$, étant des entiers positifs tels que $n_1 + n_2 + n_3 = n$. Le vecteur espérance est :

$$E(N) = \frac{n}{6}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

et la matrice de variances-covariances :

$$V(N) = \frac{n}{36}\begin{pmatrix} 5 & -2 & -3 \\ -2 & 8 & -6 \\ -3 & -6 & 9 \end{pmatrix}$$

23 ▸ Les v.a. X_1, X_2, X_3 étant indépendantes et normales constituent les composantes d'un vecteur normal X, de loi $\mathcal{N}_3(0, I)$ et de densité :

$$f(x_1, x_2, x_3) = \frac{1}{(2\pi)^{3/2}} \exp -\frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$$

Le vecteur Y se déduit de X par une application linéaire qui s'écrit $Y = AX$, où A est la matrice :

$$A = \frac{1}{6}\begin{pmatrix} 3 & 3 & 0 \\ -2 & -2 & 2 \\ -1 & 1 & 2 \end{pmatrix}$$

Toute combinaison linéaire des composantes de Y est une combinaison linéaire des composantes de X, qui par définition est une variable aléatoire réelle. Le vecteur Y est donc normal et ses caractéristiques vont se déduire de celles de X par :

$$E(Y) = E(AX) = AE(X) = 0$$

$$V(Y) = V(AX) = AV(X)\,{}^tA = A\,{}^tA = \frac{1}{18}\begin{pmatrix} 9 & 0 & 3 \\ 0 & 6 & 2 \\ 3 & 2 & 3 \end{pmatrix}$$

Notions de convergence 5

L'ESSENTIEL DU COURS

En statistique, on est amené à étudier le comportement de suites de v.a. lorsque la taille de l'échantillon devient infinie. Pour cela nous allons définir deux types de convergence. La convergence en probabilité d'une suite de v.a. (X_n) vers une v.a. X signifie que ces variables sont d'autant plus proches de X que la valeur de n est élevée, au sens où la probabilité que ces variables soient éloignées de plus d'une valeur positive ε tend vers zéro. La convergence en loi correspond à la convergence des fonctions de répartition. Ces deux convergences sont conservées par une application continue. La suite des moyennes de v.a. indépendantes et de même loi converge, en probabilité, vers la moyenne commune, et en loi, vers la loi normale, après avoir été centrée et réduite. Ce sont les deux théorèmes fondamentaux de la statistique asymptotique, connus sous les noms de loi des grands nombres et théorème central limite.

1. Convergence en probabilité

1.1 Inégalité de Markov

Si X est une v.a. positive dont l'espérance existe, l'*inégalité de Markov* établit que pour tout $\lambda > 0$:

$$P\{X \geq \lambda E(X)\} \leq \frac{1}{\lambda} \quad \text{ou} \quad P(X \geq \lambda) \leq \frac{E(X)}{\lambda}$$

1.2 Inégalité de Bienaymé-Tchebychev

Si X est une v.a. dont la variance existe, pour tout $\varepsilon > 0$ fixé :

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

1.3 Convergence en probabilité

On dit que la suite de v.a. (X_n) *converge en probabilité* vers une v.a. X si, pour tout $\varepsilon > 0$:

$$P(|X_n - X| < \varepsilon) \to 1 \quad \text{quand} \quad n \to \infty$$

ou, de façon équivalente :

$$P(|X_n - X| > \varepsilon) \to 0 \quad \text{quand} \quad n \to \infty$$

On écrit :

$$X_n \underset{p}{\to} X$$

1.4 Théorème de Slutsky

Si f est une application réelle continue, alors :

$$X_n \underset{p}{\to} X \quad \Longrightarrow \quad f(X_n) \underset{p}{\to} f(X)$$

1.5 Théorème

Si f est une application de $\mathbb{R}^2$ dans $\mathbb{R}$ uniformément continue et si (X_n) et (Y_n) sont deux suites de v.a. qui convergent en probabilité respectivement vers les v.a. X et Y, alors :

$$f(X_n, Y_n) \underset{p}{\to} f(X, Y)$$

Si on applique ce théorème aux fonctions f définies respectivement par $f(u,v) = u+v, f(u,v) = uv$ et $f(u,v) = \dfrac{u}{v}$, de $X_n \underset{p}{\to} X$ et $Y_n \underset{p}{\to} Y$, on déduit respectivement :

$$\begin{aligned} X_n + Y_n &\underset{p}{\to} X + Y \\ X_n Y_n &\underset{p}{\to} XY \\ \frac{X_n}{Y_n} &\underset{p}{\to} \frac{X}{Y} \quad \text{à condition que} \quad P(Y = 0) = 0 \end{aligned}$$

1.6 Loi des grands nombres

Si (X_n) est une suite de v.a. mutuellement indépendantes qui admettent les mêmes moments d'ordres un et deux, c'est-à-dire avec pour tout entier n, $E(X_n) = m$ et $V(X_n) = \sigma^2$, alors quand $n \to \infty$:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \underset{p}{\to} m$$

2. Convergence en loi

On dit que la suite de v.a. (X_n), de f.r. F_n, *converge en loi* vers une v.a. X de f.r. F si la suite $\{F_n(x)\}$ converge vers $F(x)$ en tout point x où F est continue. On écrit alors :

$$X_n \underset{\text{loi}}{\rightarrow} X$$

2.1 Théorème

La convergence en probabilité d'une suite (X_n) implique sa convergence en loi :

$$X_n \underset{p}{\rightarrow} X \quad \Longrightarrow \quad X_n \underset{\text{loi}}{\rightarrow} X$$

2.2 Propriété

Si (X_n) et (Y_n) sont deux suites de v.a. telles que pour $n \to \infty$:

$$X_n \underset{\text{loi}}{\rightarrow} X \quad \text{et} \quad Y_n \underset{p}{\rightarrow} a$$

où a est un nombre réel, alors :

$$X_n + Y_n \underset{\text{loi}}{\rightarrow} X + a$$
$$X_n Y_n \underset{\text{loi}}{\rightarrow} aX$$
$$\frac{X_n}{Y_n} \underset{\text{loi}}{\rightarrow} \frac{X}{a} \quad \text{si } a \neq 0$$

2.3 Théorème de Slutsky

Si g est une application réelle continue, alors :

$$X_n \underset{\text{loi}}{\rightarrow} X \quad \Longrightarrow \quad g(X_n) \underset{\text{loi}}{\rightarrow} g(X)$$

2.4 Conditions suffisantes de convergence en loi

Dans le cas discret où $X_n(\Omega) = X(\Omega) = \{a_i / i \in I\}$:

$$\forall i \in I, \quad P(X_n = a_i) \to P(X = a_i) \quad \Longrightarrow \quad X_n \underset{\text{loi}}{\rightarrow} X, \quad n \to \infty$$

Dans le cas où X_n et X admettent pour densités respectives f_n et f :

$$\forall x \in \mathbb{R}, \quad f_n(x) \to f(x) \quad \Longrightarrow \quad X_n \underset{\text{loi}}{\rightarrow} X, \quad n \to \infty$$

2.5 Théorème central limite

Si (X_n) est une suite de v.a. indépendantes et de même loi, admettant des moments d'ordres un et deux notés $m = E(X_n)$ et $\sigma^2 = V(X_n)$, alors :

$$\sqrt{n}\,\frac{\bar{X}_n - m}{\sigma} \underset{\text{loi}}{\rightarrow} N(0,1)$$

2.6 Limite d'une suite image

Si la suite (X_n) est telle que :

$$a_n(X_n - m) \underset{\text{loi}}{\rightarrow} N(0,\sigma), \quad n \rightarrow \infty$$

avec $a_n \rightarrow \infty$ et $\sigma > 0$, alors si g est une application réelle dérivable, la suite $\big(g(X_n)\big)$ converge aussi en loi avec :

$$a_n\left[g(X_n) - g(m)\right] \underset{\text{loi}}{\rightarrow} N\left(0, \sigma\left|g'(m)\right|\right), \quad n \rightarrow \infty$$

3. Convergence en moyenne d'ordre p

3.1 Définition

On dit que la suite de v.a. (X_n) converge *en moyenne d'ordre* p, avec $0 < p < \infty$, vers la v.a. X si :

$$E\,|X_n - X|^p \rightarrow 0 \quad \text{quand} \ \ n \rightarrow \infty$$

On écrit :

$$X_n \underset{\mathcal{M}_p}{\rightarrow} X$$

Dans le cas particulier $p = 2$, la convergence en moyenne d'ordre 2 s'appelle *convergence en moyenne quadratique* (m.q.). En écrivant :

$$E(X_n - X)^2 = V(X_n - X) + E^2(X_n - X)$$

qui est la somme de deux termes positifs, on retrouve les deux conditions suffisantes de convergence en probabilité (*cf.* exercice 7, page 76) comme conditions nécessaires de convergence en moyenne quadratique :

$$X_n \underset{m.q.}{\rightarrow} X \quad \Longleftrightarrow \quad \begin{cases} E(X_n - X) \rightarrow 0 \\ V(X_n - X) \rightarrow 0 \end{cases}$$

3.2 Propriété

Pour toute valeur de $p > 0$, la convergence en moyenne d'ordre p est plus forte que la convergence en probabilité, au sens où :

$$X_n \underset{\mathcal{M}_p}{\rightarrow} X \quad \Longrightarrow \quad X_n \underset{p}{\rightarrow} X$$

Donc, en particulier :

$$X_n \underset{m.q.}{\rightarrow} X \quad \Longrightarrow \quad X_n \underset{p}{\rightarrow} X$$

POUVEZ-VOUS RÉPONDRE ?

	Vrai	Faux
1. Les majorants des probabilités dans les inégalités de Markov et de Bienaymé-Tchebychev ont des valeurs toujours inférieures à 1.	☐	☐
2. Si (X_n) est une suite de v.a. d'espérance constante, dont la variance tend vers zéro, alors elle converge en probabilité vers cette constante.	☐	☐
3. Si la suite (X_n) converge en probabilité vers la v.a. X, alors les suites $\ln X_n$ et e^{X_n} convergent en probabilité, respectivement vers $\ln X$ et e^X.	☐	☐
4. Si la suite (X_n) converge en loi vers le nombre réel a, alors elle converge aussi en probabilité vers a.	☐	☐
5. La convergence en moyenne quadratique implique la convergence en loi.	☐	☐

QUESTIONS DE RÉFLEXION

6. Si X est une v.a. de signe quelconque, telle que $E\,|X|^k$ existe pour k entier positif fixé, montrer que l'inégalité de Markov peut se généraliser sous la forme suivante :

$$P\left(|X| \geq \varepsilon\right) \leq \frac{E|X|^k}{\varepsilon^k}$$

7. Si (X_n) est une suite de v.a. telle que $E(X_n) \to a$ et $V(X_n) \to 0$ quand $n \to \infty$, alors on en déduit aisément que cette suite converge en probabilité vers a, par application de l'inégalité de Bienaymé-Tchebychev. Montrer qu'en fait, on peut énoncer un résultat plus général que celui-ci, qui établit que $E(X_n - X) \to 0$ et $V(X_n - X) \to 0$ sont des conditions suffisantes de convergence en probabilité de la suite (X_n) vers la v.a. X.

8. Si X suit une loi de Poisson de paramètre λ, tel que λ tend vers l'infini, on sait que $(X - \lambda)/\sqrt{\lambda}$ converge en loi vers une v.a. U de loi $N(0, 1)$. On peut donc approximer la probabilité $P(X \leq k)$ par $P(Y \leq k)$ où Y suit une loi $N\left(\lambda, \sqrt{\lambda}\right)$, pour tout entier k. Peut-on de la même façon approximer $P(X = k)$ par $P(Y = k)$?

9. Soit (X_n) une suite de v.a. indépendantes qui suivent la même loi de Cauchy, symétrique par rapport à l'origine, de densité $f(x) = \dfrac{1}{\pi\left(1 + x^2\right)}$. Sachant que $\overline{X_n} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$ suit aussi une loi de Cauchy, peut-on en déduire de la loi des grands nombres la convergence en probabilité de la suite $\left(\overline{X_n}\right)$?

ENTRAÎNEMENT

Inégalité de Bienaymé-Tchebychev

10. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \dfrac{1}{4}\left(1 + 3x^2\right) & \text{si} \quad -1 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Déterminer un intervalle de la forme $[-l, l]$ qui contienne X avec une probabilité supérieure à 0,75 puis calculer la probabilité exacte de cet intervalle.

Analyse de l'énoncé et conseils. On transforme la condition imposée par la majoration d'une probabilité. L'intervalle demandé peut alors être obtenu par application de l'inégalité de Bienaymé-Tchebychev, ayant remarqué que la loi était symétrique par rapport à l'origine.

Inégalité de Markov

11. Soit $X_1, \ldots, X_n$ des v.a. indépendantes et de même loi que la v.a. X de densité :

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

où θ est un nombre réel positif donné. On définit la v.a. $m_n = \min\{X_1, \ldots, X_n\}$. Pour n tendant vers l'infini, étudier la convergence en moyenne et en probabilité de la suite (m_n).

Analyse de l'énoncé et conseils. La convergence en moyenne, sous-entendu d'ordre 1, s'étudie à partir du comportement de la suite numérique $E(m_n - \theta)$ car toutes les variables sont supérieures à θ et on peut penser intuitivement que la plus petite d'entre elles va tendre vers cette valeur minimale. L'étude de la convergence en probabilité peut se faire directement en utilisant l'inégalité de Markov.

Convergence en probabilité

12. La durée d'une communication téléphonique urbaine est représentée par une v.a. D uniformément distribuée sur l'intervalle $[0, t]$, où t est un nombre réel positif donné. On souhaite étudier le comportement de la plus longue durée de n communications, définie par $M_n = \max\{D_1, \ldots, D_n\}$, lorsque n devient infini, les v.a. D_i étant supposées indépendantes et de même loi que D. Montrer que M_n converge en probabilité vers t.

Analyse de l'énoncé et conseils. Pour démontrer cette convergence à partir de la définition, on calcule pour un $\varepsilon > 0$ fixé la probabilité $P(|M_n - t| < \varepsilon)$, après avoir déterminé la f.r. de M_n et on montre que cette suite numérique converge vers 1.

Convergence en moyenne quadratique

13. Soit $X_1, \ldots, X_n$ des v.a. indépendantes et de même loi que la v.a. X de loi normale $N(0, \sigma)$, où σ est un nombre réel positif donné. Montrer que la suite (T_n) définie par :

$$T_n = \frac{1}{n} \sum_{i=1}^{n} |X_i|$$

converge en moyenne quadratique vers une limite que l'on déterminera.

Analyse de l'énoncé et conseils. La v.a. T_n est la moyenne de v.a. indépendantes et de même loi à laquelle on peut appliquer la loi des grands nombres. Elle converge donc en probabilité vers $E(|X|)$. Ce sera donc aussi la limite pour la convergence en moyenne quadratique, puisque cette dernière implique la convergence en probabilité vers la même limite.

Loi des grands nombres

14. Soit $X_1, \dots, X_n$ des v.a. indépendantes et de même loi que X, de densité :

$$f(x) = \frac{\theta}{2} e^{-\theta|x|}$$

où θ est un nombre réel positif donné. Étudier la convergence en probabilité de la suite de v.a. définie par :

$$T_n = \frac{1}{n} \sum_{i=1}^{n} |X_i|$$

Analyse de l'énoncé et conseils. La variable T_n étant la moyenne de v.a. indépendantes et de même loi, on peut lui appliquer la loi des grands nombres.

Limite de la loi de Student

15. Soit $X_1, \dots, X_n$ des v.a. indépendantes et de même loi $N(0, 1)$. Étudier, quand n devient infini, la convergence en probabilité de la suite :

$$S_n^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

et en déduire la loi limite d'une v.a. qui suit une loi de Student à n degrés de liberté.

Analyse de l'énoncé et conseils. La loi des grands nombres s'applique à la suite (S_n^2). En utilisant ce résultat et la définition de la loi de Student on peut établir qu'elle converge vers la loi normale.

Théorème central limite

16. Soit $X_1, \dots, X_n$ des v.a. indépendantes et de même loi que X, de densité :

$$f(x) = \exp\left[-(x-\theta) - e^{-(x-\theta)}\right]$$

où θ est un nombre réel positif donné. Étudier la convergence en loi de la suite (T_n) de v.a. définie par :

$$T_n = \frac{1}{n} \sum_{i=1}^{n} e^{-(X_i-\theta)}$$

Analyse de l'énoncé et conseils. La v.a. T_n est définie comme la moyenne de variables indépendantes et de même loi ; donc on va pouvoir lui appliquer le théorème central limite.

Variance infinie

17. Soit (X_n) une suite de v.a. mutuellement indépendantes dont la loi est définie par :

$$P(X_n = 0) = 1 - \frac{1}{n} \quad \text{et} \quad P(X_n = n) = P(X_n = -n) = \frac{1}{2n}$$

Étudier la convergence en probabilité, en moyenne et en moyenne quadratique de cette suite. Peut-on appliquer la loi des grands nombres ou le théorème central limite ?

Analyse de l'énoncé et conseils. Pour étudier la convergence en probabilité, on pourra calculer la probabilité de l'événement $|X_n| < \varepsilon$, pour un $\varepsilon > 0$ fixé. L'étude des convergences en moyenne d'ordres 1 et 2 ne présente pas de difficulté particulière. Pour savoir si on peut appliquer les deux théorèmes de convergence à la moyenne empirique, il faut bien examiner si les hypothèses des énoncés sont vérifiées.

Limite en loi d'une suite image

18. Soit $X_1, \ldots, X_n$ des v.a. indépendantes et qui suivent la même loi exponentielle de paramètre $\theta > 0$. Étudier la convergence en loi de la suite (T_n) de v.a. définie par :

$$T_n = \frac{n}{\sum_{i=1}^{n} X_i}$$

Analyse de l'énoncé et conseils. La v.a. T_n s'écrit sous la forme $1/\bar{X}_n$ et on peut appliquer le théorème central limite à la suite $\left(\bar{X}_n\right)$. La propriété relative à la suite image $T_n = g\left(\bar{X}_n\right)$ permet alors d'obtenir sa loi limite.

SOLUTIONS

1 ▸ Faux. Si on choisit $\lambda \leq 1$ dans l'inégalité de Markov, ou $\varepsilon \leq \sigma(X)$ dans l'inégalité de Bienaymé-Tchebychev, le majorant devient supérieur à un. L'inégalité est donc bien évidemment vérifiée, mais ne présente aucun intérêt.

2 ▸ Vrai. L'inégalité de Bienaymé-Tchebychev montre que si $E(X_n) = a$ et $V(X_n) \to 0$, alors pour tout $\varepsilon > 0$, on a $P\big(|X_n - E(X_n)| \geq \varepsilon\big) \to 0$, ce qui exprime que la suite (X_n) converge en probabilité vers a. Le résultat reste d'ailleurs vrai si on a seulement la condition $E(X_n) \to a$ au lieu de $E(X_n) = a$.

3 ▸ Vrai. L'écriture $\ln X_n$ suppose bien sûr que $X_n > 0$; les fonctions logarithme et exponentielle étant continues, le résultat découle du théorème de Slutsky.

4 ▸ Vrai. La convergence en probabilité implique toujours la convergence en loi, la réciproque étant fausse en général. Cependant, dans le cas particulier où la limite est une constante, les deux convergences sont équivalentes.

5 ▸ Vrai. La convergence en moyenne quadratique implique la convergence en probabilité, qui à son tour implique la convergence en loi.

6 ▸ On adapte la seconde forme de l'inégalité de Markov en l'appliquant à $|X|^k$:

$$P\left(|X|^k \geq \lambda\right) \leq \frac{E|X|^k}{\lambda}$$

On introduit alors un nombre $\varepsilon > 0$ tel que $\varepsilon^k = \lambda$ et on en déduit alors l'inégalité demandée puisque $|X|^k \geq \varepsilon^k$ est équivalent à $|X| \geq \varepsilon$.

7 ▸ Il suffit d'appliquer le résultat énoncé à la suite (Y_n) définie par $Y_n = X_n - X$, puisque par hypothèse $E(Y_n) \to 0$ et $V(Y_n) \to 0$ et que, par ailleurs :

$$Y_n \underset{p}{\to} 0 \iff X_n \underset{p}{\to} X$$

8 ▸ La réponse est évidemment négative, car la loi de Y est continue et par conséquent $P(Y = k) = 0$ pour tout entier k. Lorsqu'une loi discrète est approximée par une loi continue, la probabilité d'un point est alors approximée par la probabilité d'un intervalle qui contient ce point. Ainsi $P(X \leq k)$ sera approchée par $P(k - 1/2 < Y < k + 1/2)$.

9 ▸ Il n'est pas possible d'appliquer ici la loi des grands nombres, puisque nous sommes dans le cas d'une loi qui n'admet aucun moment, même pas d'espérance. On a seulement ici la convergence en loi de la suite $\left(\overline{X_n}\right)$ vers une v.a. X de loi de Cauchy, puisque c'est la loi exacte de $\overline{X_n}$.

10 ▸ La densité est paire et l'espérance existe puisqu'on intègre sur l'intervalle fini $[-1, 1]$, donc $E(X) = 0$. On calcule alors la variance :

$$V(X) = E\left(X^2\right) = \frac{1}{4}\int_{-1}^{1} x^2\left(1 + 3x^2\right) \mathrm{d}x = \frac{1}{2}\left[\frac{x^3}{3} + \frac{3}{5}x^5\right]_0^1 = \frac{7}{15}$$

La condition imposée, qui s'écrit $P(|X| \leq l) \geq 0{,}75$, est équivalente à $P(|X| > l) \leq 0{,}25$ et l'inégalité de Bienaymé-Tchebychev permet d'obtenir pour tout $l > 0$ fixé :

$$P(|X| > l) \leq \frac{V(X)}{l^2}$$

En choisisssant l tel que $V(X) \leq 0{,}25l^2$, la condition sera bien vérifiée. Cela donne ici $l^2 \geq 28/15$, soit $l \geq 1{,}37$. Comme X prend ses valeurs dans $[-1, 1]$, c'est l'intervalle que l'on retient, qui est de probabilité exacte égale à 1, valeur qui est bien supérieure à 0,75. L'application de l'inégalité de Bienaymé-Tchebychev n'a donc pas permis ici de réduire l'intervalle demandé, par rapport à l'ensemble des valeurs possibles pour X. Connaissant la loi de probabilité, on peut déterminer ici la valeur exacte de l telle que $P(-l \leq X \leq l) = 0{,}75$ et qui sera solution de l'équation :

$$0{,}75 = \frac{1}{4}\int_{-l}^{l}\left(1+3x^2\right)\,\mathrm{d}x = \frac{1}{2}\left[x+x^3\right]_0^l = \frac{1}{2}\left(l+l^3\right)$$

équivalente à $l^3 + l - 3/2 = 0$, de solution $l = 0{,}86$.

11 ▸ Nous allons déterminer la loi de probabilité de m_n pour pouvoir ensuite calculer son espérance. Comme il s'agit de la plus petite observation parmi n, il vaut mieux considérer l'événement $(m_n > x)$ qui est équivalent à la réalisation de tous les événements $(X_i > x)$, qui sont d'ailleurs indépendants en raison de l'indépendance des variables X_i. Ainsi :

$$P(m_n > x) = P\left\{\bigcap_{i=1}^{n}(X_i > x)\right\} = \prod_{i=1}^{n} P(X_i > x) = [1 - F(x)]^n$$

ayant noté F la f.r. commune des v.a. X_i. La f.r. de la v.a. m_n est donc définie par :

$$G(x) = P(m_n < x) = 1 - [1 - F(x)]^n$$

et sa densité par :

$$g(x) = G'(x) = nf(x)\,[1 - F(x)]^{n-1}$$

Il faut donc déterminer aussi la f.r. de X, qui est nulle pour $x < \theta$, et qui a comme valeur pour $x \geq \theta$:

$$F(x) = \int_{\theta}^{x} \mathrm{e}^{-(u-\theta)}\,\mathrm{d}u = \left[-\mathrm{e}^{-(u-\theta)}\right]_{\theta}^{x} = 1 - \mathrm{e}^{-(x-\theta)}$$

Par conséquent, pour $x \geq \theta$:

$$g(x) = n\mathrm{e}^{-n(x-\theta)}$$

L'espérance de m_n se calcule donc par :

$$E(m_n) = n\int_{\theta}^{+\infty} x\mathrm{e}^{-n(x-\theta)}\,\mathrm{d}x = n\int_{\theta}^{+\infty}(x-\theta+\theta)\,\mathrm{e}^{-n(x-\theta)}\,\mathrm{d}x$$
$$= \frac{1}{n}\int_{0}^{+\infty} u\mathrm{e}^{-u}\,\mathrm{d}u + \theta\int_{\theta}^{+\infty} n\mathrm{e}^{-n(x-\theta)}\,\mathrm{d}x$$

ayant posé $u = n(x-\theta)$ dans la première intégrale que l'on calcule par une intégration par parties. La seconde intégrale vaut 1 (on intègre la densité) et donc :

$$E(m_n) = \frac{1}{n} + \theta$$

On en conclut donc que :

$$E(|m_n - \theta|) = E(m_n - \theta) = \frac{1}{n} \to 0$$

quand $n \to \infty$, condition qui exprime que m_n converge en moyenne d'ordre 1 vers θ. On peut d'ailleurs en déduire la convergence en probabilité qui est plus faible. Cependant, le résultat s'obtient directement par application de l'inégalité de Markov à la variable positive $m_n - \theta = |m_n - \theta|$:

$$P(|m_n - \theta| \geq \varepsilon) = P(m_n - \theta \geq \varepsilon) \leq \frac{E(m_n - \theta)}{\varepsilon} = \frac{1}{n\varepsilon} \to 0$$

pour tout $\varepsilon > 0$, condition qui se traduit par :

$$m_n \underset{p}{\to} \theta$$

12 ▸ La f.r. de M_n est définie par :

$$G(x) = P(M_n < x) = P\left\{\bigcap_{i=1}^{n} (D_i < x)\right\} = \prod_{i=1}^{n} P(D_i < x) = F^n(x)$$

car les v.a. D_i sont indépendantes et de même loi, de f.r. F définie par $F(x) = 0$ pour $x \leq 0$, et pour $0 \leq x \leq t$:

$$F(x) = \int_0^x \frac{du}{t} = \frac{x}{t}$$

avec $F(x) = 1$ pour $x \geq t$. On peut donc calculer, pour tout $\varepsilon > 0$ fixé, la probabilité :

$$\begin{aligned} P(|M_n - t| < \varepsilon) &= P(t - \varepsilon < M_n < t + \varepsilon) \\ &= G(t+\varepsilon) - G(t-\varepsilon) = 1 - \left(1 - \frac{\varepsilon}{t}\right)^n \end{aligned}$$

si $\varepsilon < t$ (pour $\varepsilon \geq t$ cette probabilité est égale à 1). Comme $1 - \varepsilon/t < 1$, cette quantité tend vers 1 et ceci montre donc que :

$$M_n \underset{p}{\to} t$$

13 ▸ L'opérateur espérance étant linéaire et les v.a. X_i de même loi :

$$E(T_n) = \frac{1}{n}\sum_{i=1}^{n} E(|X_i|) = E(|X|)$$

Nous calculons cette valeur en faisant le changement de variable $x^2 = 2\sigma^2 u$:

$$E(|X|) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} |x|\, e^{-x^2/2\sigma^2}\, dx = \frac{2\sigma}{\sqrt{2\pi}}\int_0^{+\infty} e^{-u}\, du = \sigma\sqrt{\frac{2}{\pi}}$$

La convergence en moyenne quadratique vers cette valeur, qui est l'espérance de T_n, est équivalente à la convergence vers zéro de la variance de T_n qui vaut :

$$V(T_n) = \frac{1}{n^2}\sum_{i=1}^{n} V(|X_i|) = \frac{1}{n} V(|X|)$$

en raison de l'indépendance des v.a. X_i. On obtient :

$$V(|X|) = E\left(|X|^2\right) - E^2(|X|) = \sigma^2 - \frac{2}{\pi}\sigma^2$$

car $E\left(|X|^2\right) = E\left(X^2\right) = V(X) = \sigma^2$. Ainsi :

$$V(T_n) = \left(1 - \frac{2}{\pi}\right)\frac{\sigma^2}{n} \to 0$$

quand $n \to \infty$, d'où on peut conclure :

$$T_n \underset{m.q.}{\to} \sigma\sqrt{\frac{2}{\pi}}$$

14 ▸ Nous allons calculer les deux premiers moments de la v.a. $|X|$:

$$E(|X|) = \int_{-\infty}^{+\infty} |x| f(x)\, \mathrm{d}x = \theta \int_0^{+\infty} x\mathrm{e}^{-\theta x}\, \mathrm{d}x = \frac{1}{\theta}$$

valeur obtenue en intégrant par parties. De la même manière :

$$E\left(X^2\right) = \theta \int_0^{+\infty} x^2\mathrm{e}^{-\theta x}\, \mathrm{d}x = \frac{2}{\theta^2}$$

et $V(|X|) = 1/\theta^2$. La v.a. T_n étant la moyenne des v.a. $|X_i|$ indépendantes, la loi des grands nombres permet d'obtenir :

$$T_n \underset{p}{\to} \frac{1}{\theta}$$

15 ▸ La loi des grands nombres s'applique à la moyenne des v.a. X_i^2 qui converge vers la moyenne commune $E\left(X_1^2\right) = V(X_1) = 1$, soit :

$$S_n^2 \underset{p}{\to} 1$$

On sait par ailleurs que la v.a. $Y_n = nS_n^2$ suit une loi du khi-deux à n degrés de liberté, comme somme de n carrés de v.a. indépendantes, de même loi normale centrée-réduite. Si U est une v.a. de loi $N(0,1)$ indépendante de Y_n, on sait également par définition de la loi de Student que la v.a. $T_n = U/\sqrt{Y_n/n}$ suit une loi de Student à n degrés de liberté. Le numérateur est indépendant de n et nous venons de voir que le terme sous le radical converge en probabilité vers 1. D'après une propriété du cours, le rapport se comporte donc comme le numérateur, c'est-à-dire que cette loi de Student converge vers la loi $N(0,1)$, ce que l'on peut écrire sous la forme :

$$T_n \underset{\text{loi}}{\to} N(0,1)$$

16 ▸ La v.a. T_n est la moyenne des v.a. indépendantes $Y_i = \exp-(X_i - \theta)$, de même loi que la v.a. $Y = \exp-(X - \theta)$, de f.r. définie par :

$$G(y) = P(Y < y) = P\{\exp-(X - \theta) < y\}$$

probabilité nulle pour $y \leq 0$ et qui a pour valeur, obtenue par passage au logarithme pour $y > 0$:

$$G(y) = P\{-(X - \theta) < \ln y\} = P(X > \theta - \ln y) = 1 - F(\theta - \ln y)$$

où F est la f.r. de X. La densité est donc nulle pour $y \leq 0$ et a pour valeur, pour $y > 0$:

$$g(y) = \frac{1}{y} f(\theta - \ln y) = \mathrm{e}^{-y}$$

qui est la densité d'une loi exponentielle ou loi $\gamma(1)$. D'après le cours, ses moments sont donc $E(Y) = V(Y) = 1$ et le théorème central limite permet donc d'obtenir :

$$\sqrt{n}\,(T_n - 1) \underset{\text{loi}}{\rightarrow} N(0,1)$$

17 ▸ Pour $\varepsilon > 0$ fixé, on obtient $P(|X_n| < \varepsilon) = P(X_n = 0) = 1 - 1/n \rightarrow 1$ quand $n \rightarrow \infty$ et par conséquent la suite (X_n) converge en probabilité vers zéro. Pour la convergence en moyenne, on calcule $E(|X_n|) = 1$ et donc il n'y a pas convergence en moyenne d'ordre 1 vers zéro. On sait, *a fortiori*, qu'il ne peut pas y avoir convergence en moyenne quadratique. On le vérifie en calculant $E(X_n^2) = n$ qui devient infini avec n. Les variables X_n ont des lois différentes, donc on ne peut pas appliquer le théorème central limite. Comme d'autre part les moments d'ordre 2 sont aussi différents, puisque $V(X_n) = n$, on ne peut pas appliquer non plus la loi des grands nombres. Notons cependant que l'on peut démontrer que la suite $(\bar{X}_n)$ converge en loi vers la loi $N\left(0, 1/\sqrt{2}\right)$.

18 ▸ D'après le cours, on sait que les v.a. X_i de loi exponentielle de paramètre θ ont pour caractéristiques communes $E(X_1) = 1/\theta$ et $V(X_1) = 1/\theta^2$. Le théorème central limite permet donc d'obtenir :

$$\sqrt{n}\frac{\bar{X}_n - 1/\theta}{1/\theta} \underset{\text{loi}}{\rightarrow} N(0,1)$$

quand $n \rightarrow \infty$. La fonction g définie par $g(x) = 1/x$ est continue et dérivable pour $x > 0$, avec $g'(x) = -1/x^2$. La convergence en loi de la suite $\sqrt{n}\left(\bar{X}_n - 1/\theta\right)$ vers une loi $N(0, 1/\theta)$ se transmet donc par cette application à la suite $\sqrt{n}\left[g\left(\bar{X}_n\right) - g(1/\theta)\right]$ qui converge vers la loi normale $N\left(0, \left|g'(1/\theta)\right|/\theta\right)$, ce qui donne donc comme résultat :

$$\sqrt{n}\,(T_n - \theta) \underset{\text{loi}}{\rightarrow} N(0,\theta)$$

Estimation ponctuelle 6

L'ESSENTIEL DU COURS

Au phénomène étudié, nous associons maintenant un modèle statistique où la v.a. X considérée, à valeurs dans E, suit une loi de probabilité, notée P_θ, car elle dépend d'un paramètre θ qui appartient à un sous-ensemble donné Θ, appelé espace des paramètres, qui sera généralement un sous-ensemble de $\mathbb{R}$, mais qui peut aussi parfois être inclus dans $\mathbb{R}^k$. Pour se faire une idée de la valeur inconnue du paramètre θ qui détermine cette loi de probabilité, on utilise un échantillon de cette loi, c'est-à-dire un ensemble $(X_1, ..., X_n)$ de n v.a. indépendantes et qui suivent la même loi que X. À partir des valeurs observées $x_1, ..., x_n$ on calcule ensuite une certaine valeur numérique que l'on considérera comme une valeur approchée de θ et qu'on appellera une estimation de θ. La règle qui permettra d'effectuer ce calcul est un estimateur *auquel on demande de posséder certaines propriétés. Il est* sans biais *si son espérance est égale à la valeur du paramètre et* convergent *s'il converge en probabilité vers θ, quand la taille de l'échantillon tend vers l'infini. Un estimateur optimal dans la classe des estimateurs sans biais est dit* efficace. *Les paramètres les plus fréquents sont la moyenne et la variance de la loi, qui sont respectivement estimés par la moyenne et la variance empiriques. Dans le cas où il n'y a pas d'estimateur évident, on en cherche un par la méthode du* maximum de vraisemblance, *qui consiste à maximiser la densité de l'échantillon, ou par la* méthode des moments, *qui consiste à égaler moment théorique et moment empirique correspondant.*

1. Définition et propriétés d'un estimateur

Définition. Un *estimateur* de θ est une application T_n de E^n dans F, qui à un échantillon $(X_1, ..., X_n)$ de la loi P_θ associe une variable aléatoire réelle (ou plusieurs dans le cas d'un paramètre multidimensionnel) dont on peut déterminer la loi de probabilité.

La loi de la v.a. $T_n(X_1, \dots, X_n)$ dépend de celle de X, et donc de θ, et chaque réalisation $T_n(x_1, \dots, x_n)$ sera une estimation de θ. Un estimateur est une statistique particulière, en tant que fonction de l'échantillon, qui permet d'attribuer une valeur au paramètre θ à estimer. On pourrait donc le définir comme une statistique à valeurs dans Θ. Comme nous n'avons imposé ici aucune condition à l'ensemble F, nous allons définir une propriété, que l'on pourrait qualifier de minimale, pour un estimateur, c'est-à-dire de prendre ses valeurs dans le même ensemble que le paramètre. On dira qu'un estimateur est *strict* si $T_n(E^n) \subset \Theta$.

1.1 Biais d'un estimateur

Pour pouvoir considérer $T_n(x_1, \dots, x_n)$ comme une valeur approchée de θ, il faut que les valeurs prises par la v.a. T_n ne s'écartent pas trop de la valeur, fixe, de θ. Comme T_n est une v.a., on ne peut imposer une condition qu'à sa valeur moyenne, ce qui nous amène à définir le *biais* d'un estimateur comme l'écart entre sa moyenne et la vraie valeur du paramètre :

$$b_n(\theta) = E_\theta(T_n) - \theta$$

Définitions. Un estimateur T_n de θ est dit *sans biais* si pour tout θ de Θ et tout entier positif n :

$$E_\theta(T_n) = \theta$$

Un estimateur T_n de θ est dit *asymptotiquement sans biais* si pour tout θ de Θ :

$$E_\theta(T_n) \to \theta \quad \text{quand} \quad n \to \infty$$

1.2 Convergence d'un estimateur

Définition. Un estimateur T_n est *convergent* si la suite de v.a. (T_n) converge en probabilité vers la valeur du paramètre, c'est-à-dire pour tout θ de Θ :

$$\begin{aligned} T_n \underset{p}{\to} \theta \quad &\Longleftrightarrow \quad P_\theta(|T_n - \theta| < \varepsilon) \to 1, \quad \forall \varepsilon > 0, \quad n \to \infty \\ &\Longleftrightarrow \quad P_\theta(|T_n - \theta| > \varepsilon) \to 0 \end{aligned}$$

Théorème. Tout estimateur sans biais, ou asymptotiquement sans biais, dont la variance tend vers zéro, quand n tend vers l'infini, est convergent.

1.3 Estimateur optimal

► Qualité d'un estimateur

La qualité d'un estimateur se mesure par l'*erreur quadratique moyenne*, définie pour tout θ par :

$$EQ(T_n) = E_\theta(T_n - \theta)^2 = V_\theta(T_n) + b_n^2(\theta)$$

Dans le cas particulier d'un estimateur sans biais, cette erreur quadratique se confond avec la variance de l'estimateur. Si dans l'erreur totale d'estimation on privilégie l'erreur structurelle, mesurée par $b_n^2(\theta)$, on fera le choix d'un estimateur sans biais et l'erreur d'estimation se réduira à l'erreur statistique mesurée par la variance de l'estimateur. Si on se place donc dorénavant dans la classe des estimateurs sans biais, on pourra comparer deux estimateurs T_n et T'_n de cette classe par leur variance, qui mesure alors leur dispersion par rapport au paramètre, qui est leur espérance commune. Nous dirons que l'estimateur T_n est plus *efficace* que T'_n si pour tout θ de Θ et pour une taille d'échantillon $n > N$:

$$V_\theta (T_n) \leq V_\theta \left(T'_n\right)$$

▶ Inégalité de Fréchet-Darmois-Cramer-Rao

Définitions. On appelle *vraisemblance* (likelihood) de l'échantillon $(X_1, \ldots, X_n)$ la loi de probabilité de ce n-uple, notée $L(x_1, \ldots, x_n; \theta)$, et définie par :

$$L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} P(X_i = x_i|\theta)$$

si X est une v.a. discrète, et par :

$$L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

si X est une v.a. continue de densité $f(x; \theta)$.

La *quantité d'information de Fisher* est définie par :

$$I_n(\theta) = E_\theta \left(\frac{\partial \ln L}{\partial \theta}\right)^2$$

Théorème. Sous les hypothèses de Cramer-Rao, en particulier si $E = X(\Omega)$ est indépendant du paramètre à estimer θ, pour tout estimateur sans biais T_n de θ, on a :

$$V_\theta(T_n) \geq \frac{1}{I_n(\theta)} = B_F(\theta)$$

La quantité $B_F(\theta)$ est la borne inférieure de Fréchet-Darmois-Cramer-Rao (FDCR en abrégé). Notons que dans les conditions d'application de ce théorème, en particulier si $E = X(\Omega)$ est indépendant du paramètre à estimer θ, on obtient une expression équivalente de la quantité d'information de Fisher, qui est généralement plus simple à calculer :

$$I_n(\theta) = E_\theta \left(-\frac{\partial^2 \ln L}{\partial \theta^2}\right)$$

▶ Estimateur efficace

Définition. Un estimateur sans biais T_n est dit *efficace* si sa variance est égale à la borne inférieure de FDCR :

$$V_\theta(T_n) = \frac{1}{I_n(\theta)}$$

2. Méthodes de construction d'un estimateur

2.1 Méthode du maximum de vraisemblance

La vraisemblance $L(x_1, \dots, x_n; \theta)$ représente la probabilité d'observer le n-uple $(x_1, \dots, x_n)$ pour une valeur fixée de θ. Dans la situation inverse ici où on a observé $(x_1, \dots, x_n)$ sans connaître la valeur de θ, on va attribuer à θ la valeur qui paraît la plus vraisemblable, compte tenu de l'observation dont on dispose, c'est-à-dire celle qui va lui attribuer la plus forte probabilité. On se fixe donc la règle suivante : à $(x_1, \dots, x_n)$ fixé, on considère la vraisemblance L comme une fonction de θ et on attribue à θ la valeur qui maximise cette fonction. D'où la définition suivante.

Définition. On appelle *estimateur du maximum de vraisemblance* (emv) toute fonction $\widehat{\theta}_n$ de $(x_1, \dots, x_n)$ qui vérifie :

$$L\left(x_1, \dots, x_n; \widehat{\theta}_n\right) = \max_{\theta \in \Theta} L(x_1, \dots, x_n; \theta)$$

Cette définition ne renseigne en aucune façon, ni sur l'existence, ni sur l'unicité, d'un tel estimateur. La recherche de l'emv peut se faire directement par recherche du maximum de L, ou dans le cas particulier où la fonction L est deux fois dérivable par rapport à θ, comme solution de l'équation :

$$\frac{\partial L}{\partial \theta} = 0$$

qui vérifie aussi $\dfrac{\partial^2 L}{\partial \theta^2} < 0$. Cependant, la vraisemblance se calculant à partir d'un produit, on préfère remplacer ce dernier problème par le problème équivalent pour la log-vraisemblance :

$$\frac{\partial \ln L}{\partial \theta} = 0$$

avec $\dfrac{\partial^2 \ln L}{\partial \theta^2} < 0$.

2.2 Méthode des moments

Dans le cas où le paramètre à estimer est $\theta = E_\theta(X)$, moyenne théorique de la loi, l'estimateur naturel est la moyenne empirique, ou moyenne de l'échantillon :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

De même, pour estimer le paramètre $\theta = V_\theta(X)$, variance de la loi, nous retenons logiquement comme estimateur la variance empirique :

$$S_n'^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2$$

(la variance empirique modifiée S_n^2 est obtenue en divisant par $n-1$ au lieu de n).

Plus généralement, si l'un des moments d'ordre $k \in \mathbb{N}^*$, non centré $m_k = E_\theta\left(X^k\right) = m_k(\theta)$, ou centré $\mu_k = E_\theta(X - m_1)^k = \mu_k(\theta)$, dépend de θ, nous allons chercher un estimateur par résolution de l'équation en θ obtenue en égalant moment théorique et moment empirique correspondant, soit :

$$m_{kn} = \frac{1}{n}\sum_{i=1}^{n} X_i^k = m_k(\theta) \quad \text{ou} \quad \mu_{kn} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^k = \mu_k(\theta)$$

La solution de l'équation retenue, si elle existe et est unique, sera appelée estimateur obtenu par la *méthode des moments*.

POUVEZ-VOUS RÉPONDRE ?

	Vrai	Faux
1. La moyenne empirique est un estimateur sans biais de la moyenne d'une loi.	☐	☐
2. Toute combinaison linéaire de deux estimateurs sans biais est un estimateur sans biais.	☐	☐
3. La variance empirique est un estimateur sans biais de la variance d'une loi.	☐	☐
4. Tout estimateur dont la variance tend vers zéro est convergent.	☐	☐
5. Un estimateur efficace pour un paramètre est le meilleur estimateur de celui-ci.	☐	☐
6. On peut trouver des estimateurs sans biais d'un paramètre θ qui ont une variance plus petite que la quantité $1/I_n(\theta)$.	☐	☐
7. Chercher le maximum de la vraisemblance L est équivalent à chercher le maximum de la log-vraisemblance $\ln L$.	☐	☐
8. Si $\widehat{\theta}_n$ est un emv pour un paramètre θ, alors $g\left(\widehat{\theta}_n\right)$ est un emv du paramètre $g(\theta)$ pour toute fonction g.	☐	☐

QUESTIONS DE RÉFLEXION

9. Indiquer pourquoi l'écart type empirique modifié S_n, où :

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

ne peut pas être un estimateur sans biais de l'écart type d'une loi.

10. À partir de deux estimateurs T_1 et T_2 du paramètre θ, sans biais et indépendants, on construit l'estimateur $T_3 = \alpha T_1 + (1-\alpha)\, T_2$, où $\alpha \in \,]0,1[$. Déterminer α, en fonction des variances $V_1 = V(T_1)$ et $V_2 = V(T_2)$, de façon que T_3 soit un estimateur sans biais de variance minimale. Les estimateurs T_1 et T_2 peuvent-ils être efficaces tous les deux ?

ENTRAÎNEMENT

Estimation d'une proportion

11. Pour estimer la proportion p de foyers équipés d'un micro-ordinateur, on tire au sort un échantillon de n foyers dans la population totale, de taille supposée infinie. À chaque foyer i, $1 \leq i \leq n$, on associe une v.a. de Bernoulli :

$$X_i = \begin{cases} 1 & \text{si le foyer } i \text{ possède un micro-ordinateur} \\ 0 & \text{sinon} \end{cases}$$

a) Étudier les propriétés de l'estimateur naturel F_n du paramètre p et vérifier que c'est aussi l'estimateur du maximum de vraisemblance.

b) Soit maintenant G_m l'estimateur construit de la même façon sur un second échantillon de taille m, indépendant du précédent. Déterminer les coefficients réels α et β de façon que $\widehat{p}_{n+m} = \alpha F_n + \beta G_m$ soit un estimateur de p sans biais et de variance minimum.

Analyse de l'énoncé et conseils. Le paramètre p est une proportion dans la population totale, donc son estimateur naturel est la proportion observée sur l'échantillon, c'est-à-dire la fréquence empirique. On établira que cet estimateur est sans biais, convergent et efficace. L'estimateur $\widehat{p}_{n+m}$ est construit en fait sur un échantillon unique de taille $n+m$, réunissant les deux échantillons séparés. Ayant établi que la fréquence empirique était le meilleur estimateur sans biais d'une proportion, les valeurs de α et β doivent donc correspondre à cet estimateur construit sur la réunion des deux échantillons.

Loi de Poisson

12. Le nombre d'accidents mortels par mois, à un carrefour dangereux, est une v.a. X qui suit une loi de Poisson de paramètre λ, que l'on cherche à estimer à partir d'un échantillon $X_1, \ldots, X_n$ de cette loi. Étudier les propriétés de l'estimateur naturel de ce paramètre et vérifier que c'est aussi l'estimateur du maximum de vraisemblance.

Analyse de l'énoncé et conseils. Le paramètre λ d'une loi de Poisson est la moyenne de cette loi, donc l'estimateur naturel est la moyenne empirique, dont on sait *a priori* qu'il est sans biais et convergent. On établira de plus qu'il est efficace et l'expression de la log-vraisemblance permettra de vérifier que c'est l'emv.

Loi conditionnelle binômiale

13. Chaque jour ouvrable, le nombre de clients qui entrent dans un grand magasin est une v.a. X qui suit une loi de Poisson de paramètre λ connu. Une fois entré, il y a une probabilité p qu'un client fasse des achats pour un montant supérieur à 1 000 francs. Pour pouvoir estimer p on dispose d'un échantillon sur n jours ouvrables de v.a. $Y_1, \ldots, Y_n$ de même loi que la v.a. Y, qui représente le nombre journalier de clients ayant effectué plus de 1 000 francs d'achats. Proposer un estimateur de p, étudier ses propriétés et donner une estimation à partir de l'observation $\sum_{i=1}^{100} y_i = 360$, sachant que $\lambda = 18$.

Analyse de l'énoncé et conseils. Il faut déterminer au préalable la loi conditionnelle de $Y|X = x$, pour en déduire ensuite la loi du couple (X, Y) et enfin, celle de Y. L'interprétation du paramètre p pour cette loi conduit à un estimateur naturel qui posssède toutes les propriétés classiques.

Emv pour une loi discrète

14. Soit $X_1, \ldots, X_n$ un échantillon d'une v.a. X à valeurs entières, de loi définie pour $k \in \mathbb{N}$ par :

$$P(X = k) = \frac{\theta^k}{(1+\theta)^{k+1}}$$

où θ est un paramètre positif. Déterminer l'estimateur du maximum de vraisemblance de θ et étudier ses propriétés.

Analyse de l'énoncé et conseils. On cherche le maximum de la vraisemblance considérée comme une fonction de θ et on établira que l'estimateur obtenu est sans biais, convergent et efficace. Le calcul des moments s'effectue à partir de la somme de la série géométrique et de ses deux premières dérivées.

Loi de Pareto

15. Soit X une v.a. qui suit une loi de Pareto de densité :

$$f(x) = \begin{cases} \dfrac{\alpha - 1}{x^{\alpha}} & \text{si } x \geq 1 \\ 0 & \text{sinon} \end{cases}$$

où α est un paramètre inconnu. Déterminer l'estimateur du maximum de vraisemblance de α construit à partir d'un échantillon $X_1, \ldots, X_n$ de cette loi et montrer qu'il est convergent.

Analyse de l'énoncé et conseils. Le calcul des moments de X en fonction de α ne fait pas apparaître d'estimateur naturel de ce paramètre et il est donc nécessaire ici de recourir à une méthode de construction. L'expression obtenue de l'emv permettra de déduire sa convergence de la loi des grands nombres.

Loi uniforme

16. Soit $X_1, \ldots, X_n$ un échantillon d'une v.a. X qui suit une loi uniforme sur l'intervalle $[0, \theta]$, où θ est un paramètre positif inconnu. On demande de déterminer un estimateur sans biais de θ construit à partir de l'emv et de comparer sa variance à la quantité $1/I_n(\theta)$.

Analyse de l'énoncé et conseils. On remarque ici que l'ensemble des valeurs possibles pour la variable dépend du paramètre à estimer et donc que les conditions de l'inégalité FDCR ne sont pas remplies. Par ailleurs, l'expression de la vraisemblance ne permet pas de déterminer l'emv par l'étude des dérivées et il faut donc résoudre directement ce problème de la recherche du maximum.

Comparaison d'estimateurs

17. Soit X une v.a. qui suit une loi normale d'espérance θ et de variance $\theta(1-\theta)$ où θ est un paramètre inconnu tel que $0 < \theta < 1$. À partir d'un échantillon $X_1, \ldots, X_n$ de cette loi, on construit les estimateurs :

$$T_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad S_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Indiquer l'estimateur que l'on doit choisir.

Analyse de l'énoncé et conseils. Pour effectuer cette comparaison, il faut se donner un critère. Ayant vérifié que ces estimateurs sont sans biais, ils peuvent être comparés par leurs variances, le plus efficace, donc celui qu'il faut choisir étant de variance la plus petite. On pourra admettre que $V\left(X^2\right) = 2\theta^2\left(1 - \theta^2\right)$.

Méthode des moments

18. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \frac{2}{\theta}\left(1 - \frac{x}{\theta}\right) & \text{si } 0 \leq x \leq \theta \\ 0 & \text{sinon} \end{cases}$$

où θ est un paramètre strictement positif. Déterminer par la méthode des moments un estimateur du paramètre θ construit à partir d'un échantillon $X_1, \ldots, X_n$ de cette loi et étudier ses propriétés.

Analyse de l'énoncé et conseils. Pour utiliser la méthode des moments, on calcule l'espérance de cette loi, qui va s'exprimer en fonction du paramètre, et on écrit l'égalité avec la moyenne empirique. On résout ensuite cette équation par rapport à θ, la solution fournissant l'estimateur cherché. Avant de répondre à la question sur l'efficacité, il est nécessaire de bien examiner la définition de la loi de probabilité.

SOLUTIONS

1 ▸ Vrai. Pour toute loi qui admet une espérance, la moyenne $\bar{X}_n$ de l'échantillon est un estimateur sans biais de $E(X)$.

2 ▸ Faux. Si S et T sont deux estimateurs sans biais d'un même paramètre, $\alpha S + \beta T$ ne sera aussi sans biais que si $\alpha + \beta = 1$.

3 ▸ Faux. La variance empirique $S_n'^2$ est seulement un estimateur asymptotiquement sans biais de la variance d'une loi. Pour obtenir un estimateur sans biais, il faut diviser $\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2$ par le nombre de variables indépendantes, soit $n-1$, et non pas par le nombre n de termes. En effet, ces variables sont liées par la relation $\sum_{i=1}^n \left(X_i - \bar{X}_n\right) = 0$.

4 ▸ Faux. Par exemple, si l'on retient comme estimateur du paramètre θ d'une loi uniforme sur $[0, \theta]$ la moyenne $\bar{X}_n$ d'un échantillon de cette loi, on a bien $V(\bar{X}_n) = \theta^2/12n$ qui tend vers zéro, mais cet estimateur ne converge pas vers θ puisque $E\left(\bar{X}_n\right) = \theta/2$. Pour que l'énoncé soit vrai, il faut ajouter la condition d'être un estimateur sans biais, ou asymptotiquement sans biais.

5 ▸ Faux. L'énoncé est vrai si l'on se restreint à la classe des estimateurs sans biais. Mais si l'on utilise comme critère général de comparaison entre estimateurs l'erreur quadratique moyenne, on peut trouver un estimateur avec biais meilleur qu'un estimateur efficace. Par exemple, pour la loi $N(0, \sigma)$, l'estimateur $\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$ est efficace, avec $EQ\left(\sigma_n^2\right) = V\left(\sigma_n^2\right) = 2\sigma^4/n$. Mais l'estimateur biaisé $T_n = \frac{1}{n+2} \sum_{i=1}^n X_i^2$ a une erreur quadratique $EQ(T_n) = \frac{2\sigma^4}{n+2}$ plus petite que celle de σ_n^2.

6 ▸ Vrai. Si les hypothèses de Cramer-Rao ne sont pas vérifiées, on peut toujours calculer la quantité $B_F(\theta) = 1/I_n(\theta)$, mais elle n'est pas une borne inférieure pour la variance des estimateurs sans biais de θ. On peut donc dans certains cas trouver des estimateurs sans biais de variance plus petite que $B_F(\theta)$ (*cf.* exercice 16, page 92).

7 ▸ Vrai. La fonction ln est strictement croissante ; donc maximiser L est équivalent à maximiser $\ln L$.

8 ▸ Vrai. L'image par une fonction quelconque g de l'emv d'un paramètre θ est emv de l'image $g(\theta)$ de ce paramètre.

9 ▸ On sait que S_n^2 est un estimateur sans biais de la variance σ^2 d'une loi, c'est-à-dire que $E\left(S_n^2\right) = \sigma^2$. Si S_n était un estimateur sans biais de σ, on aurait donc $E(S_n) = \sigma$. La variance de cet estimateur serait alors $V(S_n) = E\left(S_n^2\right) - E^2(S_n) = \sigma^2 - \sigma^2 = 0$, ce qui est impossible.

10 ▸ Il est facile de vérifier que l'estimateur T_3 est sans biais. Les estimateurs T_1 et T_2 étant indépendants, sa variance est :

$$V_3 = V(T_3) = \alpha^2 V_1 + (1-\alpha)^2 V_2$$

Pour déterminer le minimum de cette fonction de α, on calcule sa dérivée $V_3'(\alpha) = 2\alpha V_1 - 2(1-\alpha) V_2$, qui s'annule pour $\alpha = V_2/(V_1 + V_2)$, avec une dérivée seconde qui est positive. C'est donc la valeur qui permet d'obtenir l'estimateur sans biais de variance minimale $V_3 = V_1 V_2/(V_1 + V_2)$. Si les deux estimateurs étaient efficaces, leurs variances seraient égales à la borne FDCR $B_F(\theta)$ et on aurait alors $V_3 = V_1/2 = B_F(\theta)/2$. Ceci est impossible car aucun estimateur sans biais ne peut avoir une variance inférieure à cette borne FDCR.

11 ▸ **a)** La fréquence théorique p est estimée par la fréquence empirique, c'est-à-dire la proportion dans l'échantillon de foyers possédant un micro-ordinateur, soit :

$$F_n = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

On retrouve les propriétés générales de la moyenne empirique qui est un estimateur sans biais de la moyenne théorique :

$$E(F_n) = E\left(\bar{X}_n\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} p = p$$

et qui est convergent d'après la loi des grands nombres :

$$\bar{X}_n \underset{p}{\to} E(X) = p$$

Pour démontrer qu'il est également efficace pour ce paramètre, nous allons déterminer la vraisemblance à partir de la densité de la loi de Bernoulli, qui peut s'écrire $P(X = x) = p^x\left(1-p\right)^{1-x}$ avec $x = 0$ ou 1, soit :

$$L\left(x_1, \ldots, x_n; p\right) = \prod_{i=1}^{n} P(X_i = x_i) = p^{\sum_{i=1}^{n} x_i}\left(1-p\right)^{n-\sum_{i=1}^{n} x_i}$$

L'expression de la log-vraisemblance est donc :

$$\ln L\left(x_1,\ldots,x_n;p\right)=\left[\sum_{i=1}^{n}x_i\right]\ln p+\left[n-\sum_{i=1}^{n}x_i\right]\ln\left(1-p\right)$$

De dérivée :

$$\frac{\partial \ln L}{\partial p}=\frac{1}{p}\sum_{i=1}^{n}x_i-\frac{n-\sum_{i=1}^{n}x_i}{1-p}$$

Avec pour dérivée seconde :

$$\frac{\partial^2 \ln L}{\partial p^2}=-\frac{1}{p^2}\sum_{i=1}^{n}x_i-\frac{n-\sum_{i=1}^{n}x_i}{\left(1-p\right)^2}$$

Ce qui permet de calculer la quantité d'information de Fisher :

$$\begin{aligned}I_n\left(p\right)&=E\left[-\frac{\partial^2 \ln L}{\partial p^2}\right]=\frac{1}{p^2}\sum_{i=1}^{n}E\left(X_i\right)+\frac{n-\sum_{i=1}^{n}E\left(X_i\right)}{\left(1-p\right)^2}\\&=\frac{np}{p^2}+\frac{n-np}{\left(1-p\right)^2}=\frac{n}{p\left(1-p\right)}\end{aligned}$$

L'estimateur a pour variance :

$$V(F_n)=V(\bar{X}_n)=\frac{1}{n^2}\sum_{i=1}^{n}V(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}p\left(1-p\right)=\frac{p\left(1-p\right)}{n}$$

qui est égale à la borne FDCR, donc il est efficace. On peut noter au passage que la dérivée de la log-vraisemblance s'annule pour $p=\sum_{i=1}^{n}x_i/n$. Comme la dérivée seconde est toujours négative, il y a un maximum en ce point, donc l'estimateur précédent est l'emv.

b) Les estimateurs F_n et G_m étant sans biais, pour que $\widehat{p}_{n+m}=\alpha F_n+\beta G_m$ le soit aussi, il faut que :

$$E\left(\widehat{p}_{n+m}\right)=\alpha p+\beta p=p$$

soit la condition $\alpha+\beta=1$. Les deux échantillons étant indépendants, la variance de cet estimateur est :

$$V(\widehat{p}_{n+m})=\alpha^2V(F_n)+\beta^2V(G_m)=\alpha^2\frac{p\left(1-p\right)}{n}+\beta^2\frac{p\left(1-p\right)}{m}$$

Pour que cette quantité soit minimale, on doit résoudre le problème de minimisation de :

$$\frac{\alpha^2}{n}+\frac{\beta^2}{m}$$

sous la contrainte $\alpha+\beta = 1$. En utilisant le multiplicateur de Lagrange λ, on obtient comme conditions nécessaires du premier ordre le système :

$$\begin{cases} 2\frac{\alpha}{n} - \lambda = 0 \\ 2\frac{\beta}{m} - \lambda = 0 \end{cases}$$

d'où :

$$\frac{\alpha}{n} = \frac{\beta}{m} = \frac{\alpha+\beta}{n+m} = \frac{1}{N}$$

ayant posé $N = n + m$. Les conditions du second ordre montrent qu'il s'agit effectivement d'un minimum et l'estimateur de variance minimale est construit avec des coefficients proportionnels aux tailles d'échantillon :

$$\widehat{p}_{n+m} = \frac{n}{N}F_n + \frac{m}{N}G_m$$

Ce n'est autre que la fréquence empirique de l'échantillon global, obtenu par réunion des deux échantillons partiels indépendants. C'est l'estimateur efficace, de variance minimale :

$$V(\widehat{p}_{n+m}) = \frac{p\left(1-p\right)}{N}$$

12 ▸ On estime le paramètre $\lambda = E(X)$, moyenne théorique de la loi, par la moyenne de l'échantillon, ou moyenne empirique :

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

On vérifie que cet estimateur est sans biais, d'après les propriétés de linéarité de l'espérance :

$$E\left(\bar{X}_n\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \lambda = \lambda$$

et qu'il est convergent, d'après la loi des grands nombres :

$$\bar{X}_n \underset{p}{\rightarrow} \lambda = E(X)$$

Il converge d'ailleurs en moyenne quadratique puisqu'il est sans biais et que sa variance tend vers zéro :

$$V(\bar{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \lambda = \frac{\lambda}{n} \rightarrow 0$$

Nous allons maintenant établir qu'il est efficace en calculant la quantité d'information de Fisher à partir de la vraisemblance :

$$L(x_1,\ldots,x_n;\lambda)=\mathrm{e}^{-n\lambda}\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

D'où la log-vraisemblance :

$$\ln L(x_1,\ldots,x_n;\lambda)=-n\lambda+\left[\sum_{i=1}^n x_i\right]\ln\lambda-\sum_{i=1}^n\ln(x_i!)$$

et sa dérivée première :

$$\frac{\partial\ln L}{\partial\lambda}=-n+\frac{1}{\lambda}\sum_{i=1}^n x_i$$

Les hypothèses de Cramer-Rao étant vérifiées, on peut utiliser la dérivée seconde :

$$\frac{\partial^2\ln L}{\partial\lambda^2}=-\frac{1}{\lambda^2}\sum_{i=1}^n x_i$$

pour calculer :

$$I_n(\lambda)=E\left[-\frac{\partial^2\ln L}{\partial\lambda^2}\right]=\frac{1}{\lambda^2}\sum_{i=1}^n E(X_i)=\frac{n\lambda}{\lambda^2}=\frac{n}{\lambda}$$

Son inverse est la borne FDCR, égale ici à la variance de $\bar{X}_n$ qui est donc un estimateur efficace du paramètre λ. La dérivée de $\ln L$ s'annule pour $\lambda=\bar{X}_n$ et comme d'autre part la dérivée seconde est toujours négative, on établit ainsi qu'il s'agit de l'expression de l'estimateur du maximum de vraisemblance de λ.

13 ▸ Lorsque $X=x$, chacun des x clients entrés effectue des achats d'un montant supérieur à 1 000 F avec une probabilité p, donc la loi de Y, conditionnellement à $X=x$, est une loi binômiale de paramètres x et p. La loi du couple (X,Y) se détermine donc par :

$$\begin{aligned}P(X=x,Y=y)&=P(X=x)\,P(Y=y|X=x)=\mathrm{e}^{-\lambda}\frac{\lambda^x}{x!}\binom{x}{y}p^y(1-p)^{x-y}\\&=\mathrm{e}^{-\lambda}\frac{(\lambda p)^y\left[\lambda(1-p)\right]^{x-y}}{y!\,(x-y)!}\end{aligned}$$

pour x et y entiers tels que $0\leq y\leq x$. On obtient la loi marginale de Y par sommation de cette loi du couple :

$$\begin{aligned}P(Y=y)&=\sum_{x=y}^{\infty}P(X=x,Y=y)=\mathrm{e}^{-\lambda}\frac{(\lambda p)^y}{y!}\sum_{x=y}^{\infty}\frac{\left[\lambda(1-p)\right]^{x-y}}{(x-y)!}\\&=\mathrm{e}^{-\lambda}\frac{(\lambda p)^y}{y!}\mathrm{e}^{\lambda(1-p)}=\mathrm{e}^{-\lambda p}\frac{(\lambda p)^y}{y!}\end{aligned}$$

Ce qui montre que Y suit une loi de Poisson de paramètre λp. Les résultats de l'exercice précédent permettent d'affirmer que $\widehat{p}_n = \bar{Y}_n/\lambda$ est un estimateur sans biais, convergent et efficace, de variance $V(\widehat{p}_n) = p/n\lambda$. On obtient comme estimation $\widehat{p}_{100} = 3{,}6/18 = 0{,}2$.

14 ▸ La vraisemblance a pour expression :

$$L\left(x_1, \ldots, x_n; \theta\right) = \prod_{i=1}^{n} P\left(X_i = x_i\right) = \frac{\theta^{\sum_{i=1}^{n} x_i}}{(1+\theta)^{n+\sum_{i=1}^{n} x_i}}$$

et la log-vraisemblance :

$$\ln L\left(x_1, \ldots, x_n; \theta\right) = \left[\sum_{i=1}^{n} x_i\right] \ln \theta - \left[n + \sum_{i=1}^{n} x_i\right] \ln (1+\theta)$$

Sa dérivée est :

$$\frac{\partial \ln L}{\partial \theta} = \frac{1}{\theta} \sum_{i=1}^{n} x_i - \frac{n + \sum_{i=1}^{n} x_i}{1+\theta}$$

qui s'annule pour $\theta = \sum_{i=1}^{n} x_i/n = \bar{x}$. Étudions la dérivée seconde :

$$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{1}{\theta^2} \sum_{i=1}^{n} x_i + \frac{n + \sum_{i=1}^{n} x_i}{(1+\theta)^2}$$

qui a comme valeur $-n/\bar{x}\,(1+\bar{x}) < 0$ pour $\theta = \bar{x}$ qui correspond donc bien à un maximum. L'emv est par conséquent la moyenne empirique :

$$\widehat{\theta}_n = \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Nous allons d'ailleurs voir que le paramètre est ici la moyenne de la loi :

$$E(X) = \sum_{k=0}^{\infty} k \frac{\theta^k}{(1+\theta)^{k+1}} = \frac{\theta}{(1+\theta)^2} \sum_{k=1}^{\infty} k \left(\frac{\theta}{1+\theta}\right)^{k-1}$$

quantité qui se calcule à partir de la somme de la série géométrique $\sum_{k=0}^{\infty} x^k = 1/\left(1-x\right)$ pour $|x| < 1$, et de sa dérivée $\sum_{k=1}^{\infty} k x^{k-1} = 1/\left(1-x\right)^2$. La valeur de la somme précédente est obtenue pour $x = \theta/\left(1+\theta\right)$, soit :

$$E(X) = \frac{\theta}{(1+\theta)^2} \times (1+\theta)^2 = \theta$$

Le calcul préalable de l'espérance aurait donc permis de retenir la moyenne empirique comme estimateur naturel, sans avoir à recourir à une méthode de construction comme celle du maximum de vraisemblance. On sait que cet estimateur est sans biais, convergent d'après la loi des grands nombres et de variance $V(X)/n$. Pour le calcul de la variance, nous allons utiliser la seconde dérivée de la somme de la série géométrique $\sum_{k=2}^{\infty} k(k-1)x^{k-2} = 2/(1-x)^3$, qui va nous permettre d'obtenir le moment factoriel :

$$E[X(X-1)] = \sum_{k=0}^{\infty} k(k-1)P(X=k) = \frac{\theta^2}{(1+\theta)^3}\sum_{k=2}^{\infty} k(k-1)\left(\frac{\theta}{1+\theta}\right)^{k-2} = 2\theta^2$$

ayant remplacé x par $\theta/(1+\theta)$. On en déduit ensuite :

$$V(X) = E\left(X^2\right) - E^2(X) = E[X(X-1)] + E(X) - E^2(X) = \theta^2 + \theta$$

Calculons maintenant la quantité d'information de Fisher :

$$I_n(\theta) = E\left(-\frac{\partial^2 \ln L}{\partial \theta^2}\right) = \frac{1}{\theta^2}\sum_{i=1}^{n} E(X_i) - \frac{n + \sum_{i=1}^{n} E(X_i)}{(1+\theta)^2} = \frac{n}{\theta} - \frac{n}{1+\theta} = \frac{n}{\theta(1+\theta)}$$

La variance de l'estimateur est :

$$V\left(\widehat{\theta}_n\right) = \frac{V(X)}{n} = \frac{\theta(1+\theta)}{n} = \frac{1}{I_n(\theta)}$$

et cet estimateur est donc aussi efficace.

15 ▸ La densité n'est définie que pour $\alpha > 1$. L'espérance a pour valeur :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x = \int_{1}^{+\infty} (\alpha-1)x^{-\alpha+1}\,\mathrm{d}x = (\alpha-1)\left[\frac{x^{-\alpha+2}}{-\alpha+2}\right]_1^{+\infty}$$

quantité qui n'est définie que pour $\alpha > 2$. Si cette condition est réalisée, on obtient :

$$E(X) = \frac{\alpha-1}{\alpha-2}$$

On voit ainsi que le paramètre n'a pas d'interprétation simple en terme de moment. On calcule donc la vraisemblance pour déterminer l'emv :

$$L(x_1,\ldots,x_n;\alpha) = (\alpha-1)^n \prod_{i=1}^{n} x_i^{-\alpha} \quad \text{avec} \quad \min x_i \geq 1$$

Les expressions de la log-vraisemblance et de ses dérivées sont donc :

$$\ln L(x_1,\ldots,x_n;\alpha) = n\ln(\alpha-1) - \alpha\sum_{i=1}^{n}\ln x_i$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{n}{\alpha-1} - \sum_{i=1}^{n}\ln x_i$$

$$\frac{\partial^2 \ln L}{\partial \alpha^2} = -\frac{n}{(\alpha-1)^2}$$

La dérivée première s'annule pour $\alpha = 1 + n / \sum_{i=1}^{n} \ln x_i$, avec une dérivée seconde négative, et il s'agit donc bien d'un maximum. L'emv est donc défini par :

$$\widehat{\alpha}_n = 1 + \frac{n}{\sum_{i=1}^{n} \ln X_i}$$

Il est de la forme $\widehat{\alpha}_n = 1 + 1/\bar{Y}_n$, ayant posé $Y_i = \ln X_i$ pour $1 \leq i \leq n$. Pour savoir si on peut appliquer la loi des grands nombres à $\bar{Y}_n$, nous allons calculer l'espérance de $Y = \ln X$:

$$E\left(\ln X\right) = (\alpha - 1) \int_1^{+\infty} x^{-\alpha} \ln x \, dx = \left[-x^{-\alpha+1} \ln x\right]_1^{+\infty} + \int_1^{+\infty} x^{-\alpha} \, dx = \frac{1}{\alpha - 1}$$

On aurait pu également obtenir ce résultat à partir de la loi de Y définie par sa f.r. $G(y) = P\left(Y < y\right) = P\left(\ln X < y\right) = P\left(X < e^y\right) = F\left(e^y\right)$, où F désigne la f.r. de X. Pour $x \geq 1$:

$$F(x) = \int_{-\infty}^{x} f(u) \, du = (\alpha - 1) \int_1^{x} u^{-\alpha} \, du = 1 - x^{1-\alpha}$$

avec $F(x) = 0$ pour $x < 1$. Ainsi, pour $y < 0$, on a $G\left(y\right) = 0$ et pour $y \geq 0$: $G(y) = 1 - e^{-(\alpha-1)y}$. La v.a. Y est donc positive, de densité $(\alpha - 1) e^{-(\alpha-1)y}$, et on reconnaît une loi exponentielle de paramètre $\alpha - 1$. On retrouve bien la valeur de l'espérance, la variance étant :

$$V(Y) = \frac{1}{(\alpha - 1)^2}$$

Par conséquent, la loi des grands nombres permet d'obtenir :

$$\bar{Y}_n \underset{p}{\rightarrow} E(Y) = \frac{1}{\alpha - 1}$$

On en déduit, du théorème de Slutsky, la convergence de l'estimateur :

$$\widehat{\alpha}_n = 1 + \frac{1}{\bar{Y}_n} \underset{p}{\rightarrow} 1 + (\alpha - 1) = \alpha$$

Notons que cet estimateur aurait pu aussi être obtenu par la méthode des moments, appliquée à la loi de Y. En effet, l'équation :

$$E(Y) = \frac{1}{\alpha - 1} = \bar{Y}_n$$

admet comme solution unique $\alpha = 1 + 1/\bar{Y}_n$.

16 ▸ La densité de cette loi uniforme est $1/\theta$ sur l'intervalle $[0, \theta]$. L'expression de la vraisemblance est donc :

$$L(x_1, \ldots, x_n; \theta) = \frac{1}{\theta^n}$$

si $0 \leq x_i \leq \theta$ pour $1 \leq i \leq n$. Elle est nulle si ces conditions ne sont pas remplies, c'est-à-dire que $L = L(\theta) = 0$ pour $\theta < \max\{x_1, \ldots, x_n\}$. Au-delà de cette valeur, la fonction L varie comme $1/\theta$, donc est décroissante (voir figure 6.1). Par conséquent, il y a un maximum en ce point, où la vraisemblance n'est pas dérivable. L'emv est donc :

$$M_n = \max\{X_1, \ldots, X_n\}$$

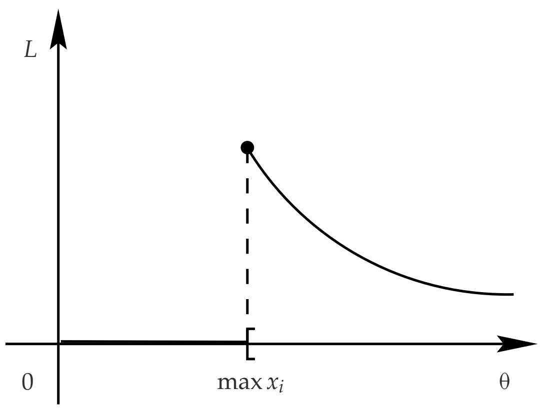


Figure 6.1

La valeur du paramètre étant la plus grande valeur possible de la variable, on l'estime logiquement par la plus grande observation de l'échantillon. Pour étudier les propriétés de cet estimateur, il est nécessaire de déterminer sa loi de probabilité, par l'intermédiaire de sa fonction de répartition :

$$G(x) = P(M_n < x) = P\left\{\bigcap_{i=1}^{n} (X_i < x)\right\} = \prod_{i=1}^{n} P(X_i < x) = F^n(x)$$

car les variables de l'échantillon sont indépendantes et de même f.r. F. On obtient la densité g en dérivant :

$$g(x) = nF^{n-1}(x) f(x)$$

où $f(x) = 1/\theta$ pour $x \in [0, \theta]$ et $f(x) = 0$ en dehors de cet intervalle. La f.r. de X est définie par $F(x) = x$ sur l'intervalle $[0, \theta]$. Ainsi :

$$E(M_n) = \int_{-\infty}^{+\infty} xg(x)\, dx = n \int_0^{\theta} \frac{x^n}{\theta^n}\, dx = \frac{n}{n+1}\theta$$

donc cet estimateur est seulement asymptotiquement sans biais. L'estimateur sans biais est :

$$T_n = \frac{n+1}{n} M_n$$

Pour calculer sa variance, nous avons besoin de :

$$E\left(M_n^2\right) = \int_{-\infty}^{+\infty} x^2 g(x)\, \mathrm{d}x = n \int_0^{\theta} \frac{x^{n+1}}{\theta^n}\, \mathrm{d}x = \frac{n}{n+2}\theta^2$$

D'où on déduit :

$$V(M_n) = E\left(M_n^2\right) - E^2(M_n) = \frac{n}{(n+2)(n+1)^2}\theta^2$$

et enfin :

$$V(T_n) = \frac{\theta^2}{n(n+2)}$$

L'ensemble des valeurs prises par la variable est ici $E = X(\Omega) = [0, \theta]$, qui dépend du paramètre ; donc pour calculer la quantité d'information de Fisher, on doit utiliser la première formule fournie par la définition :

$$I_n(\theta) = E\left(\frac{\partial \ln L}{\partial \theta}\right)^2$$

La log-vraisemblance a ici comme expression $\ln L = -n \ln \theta$, de dérivée $-n/\theta$ pour $M_n \leq \theta$, d'où :

$$I_n(\theta) = \frac{n^2}{\theta^2}$$

Si on compare la variance de l'estimateur sans biais T_n et la quantité $1/I_n(\theta)$, on observe que :

$$V(T_n) = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n^2} = \frac{1}{I_n(\theta)}$$

Cependant, ce majorant n'est pas une borne inférieure pour la variance des estimateurs sans biais, puisque les conditions de l'inégalité FDCR ne sont pas remplies. Il n'est donc pas anormal de trouver un estimateur sans biais de variance plus petite que cette quantité.

17 ▸ L'estimateur naturel du paramètre $\theta = E(X)$ est la moyenne empirique $T_n = \bar{X}_n$, qui est sans biais et convergent d'après la loi des grands nombres. Par ailleurs, on peut remarquer que $E\left(X^2\right) = V(X) + E^2(X) = \theta$, donc le paramètre θ s'interprète aussi comme la moyenne théorique de la loi de $Y = X^2$. L'estimateur $S_n = \bar{Y}_n$ est la moyenne empirique de l'échantillon de cette loi ; il est donc aussi sans biais et convergent d'après la loi des grands nombres. Ces deux estimateurs sans biais vont être comparés par leurs variances :

$$V(T_n) = V(\bar{X}_n) = \frac{V(X)}{n} = \frac{\theta(1-\theta)}{n} \quad \text{et} \quad V(S_n) = \frac{V(Y)}{n}$$

Pour pouvoir obtenir facilement la variance de $Y = X^2$, on peut utiliser le fait que $(X - \theta)^2 / \theta (1 - \theta)$ suit une loi du khi-deux à un degré de liberté, comme carré d'une loi normale centrée-réduite. On en déduit :

$$V(X - \theta)^2 = E (X - \theta)^4 - E^2 (X - \theta)^2 = 2\theta^2 (1 - \theta)^2$$

car on sait par ailleurs que $E (X - \theta)^4 = 3\theta^2 (1 - \theta)^2$.

On écrit enfin $X^2 = [(X - \theta) + \theta]^2$ pour obtenir :

$$\begin{aligned} V\left(X^2\right) &= V(X - \theta)^2 + 4\theta^2 V(X - \theta) + 4\theta \operatorname{Cov}\left[X - \theta, (X - \theta)^2\right] \\ &= V(X - \theta)^2 + 4\theta^2 V(X) = 2\theta^2 \left(1 - \theta^2\right) \end{aligned}$$

On forme le rapport des variances :

$$\frac{V(S_n)}{V(T_n)} = 2\theta (1 + \theta)$$

Comme ce rapport dépend du paramètre à estimer, étant de valeur supérieure à 1 pour $\frac{\sqrt{3} - 1}{2} < \theta < 1$, la comparaison entre ces deux estimateurs n'est pas possible.

18 ▸ Le calcul de l'espérance donne :

$$E (X) = \frac{2}{\theta} \int_0^\theta x \left(1 - \frac{x}{\theta}\right) \mathrm{d}x = \frac{2}{\theta} \left[\frac{x^2}{2} - \frac{x^3}{3\theta}\right]_0^\theta = \frac{\theta}{3}$$

On cherche donc à résoudre l'équation $\bar{X}_n = \theta/3$, de solution immédiate $\theta = 3\bar{X}_n$, ce qui fournit l'estimateur de la méthode des moments $T_n = 3\bar{X}_n$, qui va être sans biais par construction :

$$E (T_n) = 3E \left(\bar{X}_n\right) = 3E (X) = \theta$$

et convergent d'après la loi des grands nombres :

$$T_n = 3\bar{X}_n \underset{p}{\rightarrow} 3E (X) = \theta$$

La question de l'efficacité ne se pose pas ici car X prend ses valeurs dans l'intervalle $[0, \theta]$, qui dépend du paramètre de la loi ; donc les conditions de l'inégalité FDCR ne sont pas remplies.

Estimation par intervalle de confiance 7

L'ESSENTIEL DU COURS

Afin de juger de la précision du résultat d'un problème d'estimation, nous allons associer à un échantillon d'une loi inconnue un ensemble de valeurs, et non plus une valeur unique. Ces valeurs constituent généralement un intervalle inclus dans l'ensemble Θ des valeurs possibles pour le paramètre. La détermination de cet intervalle, dit de confiance, s'effectue par la donnée de la probabilité, dénommée niveau de confiance, que la vraie valeur du paramètre appartienne à cet intervalle.

1. Définition et principe de construction

Définition. Un *intervalle de confiance* pour le paramètre θ, de *niveau de confiance* $1-\alpha \in]0,1[$, est un intervalle qui a la probabilité $1-\alpha$ de contenir la vraie valeur du paramètre θ.

Pour obtenir cet intervalle, on utilise un estimateur T_n dont on connaît la loi en fonction de θ, ce qui permet de déterminer les valeurs $t_1 = t_1(\theta)$ et $t_2 = t_2(\theta)$ telles que :

$$P_\theta \{t_1(\theta) \leq T_n \leq t_2(\theta)\} = 1-\alpha$$

Il faut ensuite inverser cet intervalle pour T_n afin d'obtenir un intervalle pour θ, dont les bornes vont dépendre de l'estimateur T_n, c'est-à-dire déterminer les valeurs $a = a(T_n)$ et $b = b(T_n)$ telles que :

$$P_\theta \{a(T_n) \leq \theta \leq b(T_n)\} = 1-\alpha$$

Ainsi, $[a(T_n), b(T_n)]$ est un intervalle de confiance de niveau $1-\alpha$ pour le paramètre θ. Il s'obtient en déterminant l'ensemble des valeurs de θ pour lesquelles on a simultanément, pour T_n fixé :

$$t_1(\theta) \leq T_n \quad \text{et} \quad T_n \leq t_2(\theta)$$

Si par exemple ces deux fonctions sont croissantes, on voit sur la figure 7.1 que :

$$t_1(\theta) \leq T_n \Longleftrightarrow \theta \leq t_1^{-1}(T_n) \quad \text{et} \quad T_n \leq t_2(\theta) \Longleftrightarrow \theta \geq t_2^{-1}(T_n)$$

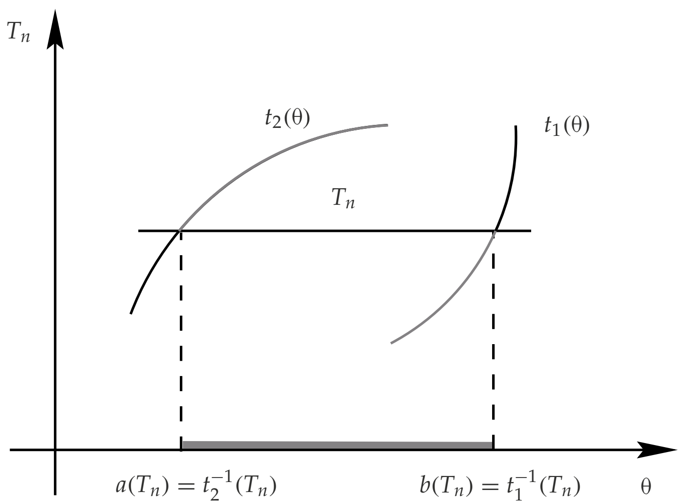


Figure 7.1

L'intervalle est facile à obtenir ici car les fonctions t_1 et t_2 sont inversibles :

$$t_1(\theta) \leq T_n \leq t_2(\theta) \quad \Longleftrightarrow \quad a(T_n) = t_2^{-1}(T_n) \leq \theta \leq t_1^{-1}(T_n) = b(T_n)$$

Le risque total $\alpha = P_\theta(T_n < t_1) + P_\theta(T_n > t_2)$ peut être *a priori* réparti de multiples façons. Posons $\alpha_1 = P_\theta\{\theta > b(T_n)\}$ et $\alpha_2 = P_\theta\{\theta < a(T_n)\}$; les différents choix possibles sont les suivants.

► **Intervalle bilatéral** $(\alpha_1\alpha_2 > 0)$

• *Symétrique :* $\alpha_1 = \alpha_2 = \alpha/2$

C'est le choix que l'on fait si la loi de T_n est symétrique, ou si on n'a aucune information particulière, ce choix étant le moins arbitraire.

• *Dissymétrique :* $\alpha_1 \neq \alpha_2$

Seules des raisons très particulières peuvent permettre de fixer les valeurs de α_1 et α_2 telles que $\alpha_1 + \alpha_2 = \alpha$.

► **Intervalle unilatéral** $(\alpha_1\alpha_2 = 0)$

• *À droite :* $\alpha_1 = 0, \alpha_2 = \alpha$

C'est l'interprétation donnée au paramètre θ qui conduit à un intervalle de la forme $\theta > a(T_n)$.

• *À gauche :* $\alpha_1 = \alpha, \alpha_2 = 0$

L'interprétation du paramètre θ peut également conduire à un intervalle de la forme $\theta < b(T_n)$.

2. Intervalle pour une proportion

Soit $(X_1, \dots, X_n)$ un échantillon d'une v.a. X qui suit une loi de Bernoulli de paramètre p, estimé par $\widehat{p}_n = \bar{X}_n$. On utilise la loi exacte de $n\widehat{p}_n = S_n$ pour déterminer les solutions $t_1 = t_1(p)$ et $t_2 = t_2(p)$ des inéquations $P(n\widehat{p}_n \leq nt_1) \leq \alpha/2$ et $P(n\widehat{p}_n \geq nt_2) \leq \alpha/2$. On trace le graphe de ces deux fonctions de p. Pour une valeur observée $\widehat{p}_n$, l'intersection de l'horizontale correspondante avec ces deux courbes permet de lire sur l'axe des abscisses l'intervalle des valeurs de p pour lesquelles on a simultanément $t_1(p) \leq \widehat{p}_n$ et $t_2(p) \geq \widehat{p}_n$ (voir figure 7.2). Ainsi, l'intervalle de confiance s'obtient par simple lecture de l'*abaque*.

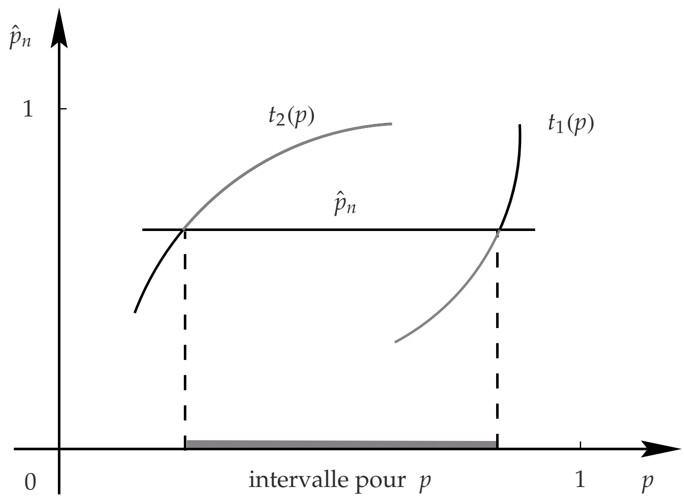


Figure 7.2

Si on ne dispose pas d'abaque, on utilise la loi asymptotique de $\widehat{p}_n$. On retient alors un intervalle symétrique, à partir de la valeur de a lue dans la table 2, page 203, de la loi normale, telle que :

$$P\left(-a < \sqrt{n}\,\frac{\widehat{p}_n - p}{\sqrt{pq}} < a\right) = 1 - \alpha$$

On fait une seconde approximation en remplaçant p par $\widehat{p}_n$ dans les bornes de l'intervalle, ce qui conduit à l'intervalle de confiance pour p :

$$\widehat{p}_n - a\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} < p < \widehat{p}_n + a\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}$$

où a est le fractile d'ordre $1 - \alpha/2$ de la loi $N(0,1)$. Cette approximation est acceptable pour $np(1-p) \geq 3$. On peut également remplacer $p(1-p)$ par sa borne supérieure $1/4$, et obtenir ainsi l'intervalle approché le plus grand :

$$\widehat{p}_n - \frac{a}{2\sqrt{n}} < p < \widehat{p}_n + \frac{a}{2\sqrt{n}}$$

3. Intervalles associés aux paramètres de la loi normale

On dispose d'un échantillon $(X_1, \ldots, X_n)$ d'une v.a. X qui suit une loi $N(m, \sigma)$.

3.1 Intervalle pour la moyenne d'une loi normale d'écart type connu

La moyenne empirique $\bar{X}_n$ est un estimateur de m dont on connaît la loi de probabilité, qui est la loi normale $N\left(m, \sigma/\sqrt{n}\right)$. L'intervalle de confiance pour le paramètre m est déterminé à partir de la valeur de sa probabilité : $1-\alpha = P\left\{-b < \bar{X}_n - m < a\right\}$. La loi de la v.a. $\bar{X}_n - m$ étant symétrique, on choisit $b = a$, qui doit vérifier la condition :

$$1 - \alpha = P\left\{-\frac{a}{\sigma/\sqrt{n}} < \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} < \frac{a}{\sigma/\sqrt{n}}\right\}$$

L'intervalle de confiance, de niveau $1 - \alpha$, pour le paramètre m est donc :

$$\bar{X}_n - u\frac{\sigma}{\sqrt{n}} < m < \bar{X}_n + u\frac{\sigma}{\sqrt{n}}$$

où u est le fractile d'ordre $1 - \alpha/2$ de la loi normale standard, soit $u = \Phi^{-1}(1 - \alpha/2)$, Φ étant la f.r. de la loi $N(0, 1)$.

3.2 Intervalle pour la moyenne d'une loi normale d'écart type inconnu

La statistique utilisée dans la situation précédente était la variable normale centrée-réduite :

$$\frac{\bar{X}_n - m}{\sigma/\sqrt{n}} = \sqrt{n}\,\frac{\bar{X}_n - m}{\sigma}$$

Le paramètre σ étant inconnu va être remplacé par un estimateur, basé sur l'estimateur sans biais de σ^2 :

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

On utilise donc comme nouvelle statistique :

$$\sqrt{n}\,\frac{\bar{X}_n - m}{S_n}$$

de loi de Student à $n - 1$ degrés de liberté. Par lecture de la table 6, page 200, on détermine la valeur de t telle que :

$$P\left[-t < \sqrt{n}\,\frac{\bar{X}_n - m}{S_n} < t\right] = 1 - \alpha$$

La valeur de t est le fractile d'ordre $1 - \alpha/2$ de la loi de Student T_{n-1}, ce qui fournit l'intervalle de confiance pour m de niveau $1 - \alpha$:

$$P\left(\bar{X}_n - t\frac{S_n}{\sqrt{n}} < m < \bar{X}_n + t\frac{S_n}{\sqrt{n}}\right) = 1 - \alpha$$

3.3 Intervalle pour la variance d'une loi normale d'espérance connue

L'estimateur sans biais de σ^2, basé sur l'échantillon $(X_1, \ldots, X_n)$ de cette loi $N(m, \sigma)$ où m est connu est :

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2$$

de loi connue : $n\widehat{\sigma}_n^2/\sigma^2 \rightsquigarrow \chi_n^2$. On peut donc déterminer les valeurs de a et b telles que :

$$P\left(a < n\frac{\widehat{\sigma}_n^2}{\sigma^2} < b\right) = 1 - \alpha$$

Ce qui conduit à l'intervalle de confiance défini par :

$$P\left(n\frac{\widehat{\sigma}_n^2}{b} < \sigma^2 < n\frac{\widehat{\sigma}_n^2}{a}\right) = 1 - \alpha$$

Cependant, il n'y a qu'une seule condition pour déterminer les deux valeurs a et b et il reste un degré d'incertitude puisque la loi utilisée n'est pas symétrique. Si on pose $\alpha_1 = P\left(\chi_n^2 < a\right)$ et $\alpha_2 = P\left(\chi_n^2 > b\right)$, la seule contrainte dans le choix de α_1 et α_2 est $\alpha_1 + \alpha_2 = \alpha$.

3.4 Intervalle pour la variance d'une loi normale d'espérance inconnue

Quand le second paramètre m de la loi normale est inconnu, l'estimateur sans biais de σ^2 qu'il faut retenir est :

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2$$

Sa loi est connue, $(n-1)\, S_n^2/\sigma^2 \rightsquigarrow \chi_{n-1}^2$, et on peut déterminer les valeurs de a et b telles que :

$$P\left\{a < (n-1)\,\frac{S_n^2}{\sigma^2} < b\right\} = 1 - \alpha$$

Ce qui permet d'en déduire l'intervalle de confiance défini par :

$$P\left\{(n-1)\,\frac{S_n^2}{b} < \sigma^2 < (n-1)\,\frac{S_n^2}{a}\right\} = 1 - \alpha$$

Là encore, il n'y a qu'une seule contrainte pour déterminer les valeurs de a et b. Si nous posons $\alpha_1 = P\left(\chi_{n-1}^2 < a\right)$ et $\alpha_2 = P\left(\chi_{n-1}^2 > b\right)$, la contrainte est $\alpha_1 + \alpha_2 = \alpha$.

POUVEZ-VOUS RÉPONDRE ?

	Vrai	Faux
1. On peut toujours trouver un intervalle de confiance qui contient le paramètre avec une probabilité égale à 1.	☐	☐
2. À un estimateur donné du paramètre, de loi connue, on peut associer une infinité d'intervalles de confiance de même niveau de confiance donné.	☐	☐
3. La fréquence empirique f_n permet de construire un intervalle de confiance pour le paramètre d'une loi de Bernoulli, de niveau exact $1 - \alpha$, avec $0 < \alpha < 1$.	☐	☐
4. C'est pour une proportion p voisine de 50 % que l'estimation de ce paramètre est la plus imprécise.	☐	☐
5. Pour un niveau de confiance donné, l'intervalle pour la moyenne d'une loi normale d'écart type connu est de longueur minimale si on le choisit symétrique.	☐	☐
6. Pour un niveau de confiance donné et une taille d'échantillon donnée, la longueur d'un intervalle bilatéral symétrique est fixée.	☐	☐
7. Un intervalle de confiance pour une moyenne a une longueur fixe pour un niveau donné et une taille d'échantillon n donnée, cette longueur tendant vers zéro avec $1/n$.	☐	☐

QUESTIONS DE RÉFLEXION

8. Sur un échantillon de 100 objets prélevés dans un lot de fabrication, on n'a observé aucun défectueux. La proportion p de défectueux admet donc comme estimation $\widehat{p}_{100} = 0$. Peut-on retenir, avec l'approximation normale, l'intervalle de confiance de niveau 0,95 qui se réduit à la valeur 0 ou l'intervalle approché le plus grand qui est $-0{,}10 < p < 0{,}10$?

9. Une société de vente par correspondance désire établir le profil de sa clientèle et extrait par tirage sans remise cent dossiers de son fichier clients. Malheureusement vingt d'entre eux s'avèrent inexploitables. Indiquer comment à partir d'un intervalle de confiance on peut déterminer le nombre minimum de dossiers à extraire pour qu'il y en ait au moins mille d'exploitables avec une probabilité de 0,95.

10. Le chiffre d'affaires moyen d'un commerçant, calculé sur les trente derniers jours, est de 2 000 francs, avec un écart type empirique de valeur $s_{30} = 300$ francs.

Si on admet que son chiffre d'affaires quotidien peut être représenté par une v.a. X de loi normale, d'espérance m et d'écart type σ inconnus, donner un intervalle de confiance de niveau 0,95 pour le paramètre m. Obtient-on le même intervalle si σ est connu, de valeur $\sigma = 300$?

11. Un sondage effectué auprès de 1 000 personnes, réparties en cinq groupes de 200 personnes, a donné comme valeur moyenne de dépenses mensuelles consacrées aux loisirs 301 F. N'ayant relevé que la dispersion empirique des moyennes empiriques de chacun des cinq groupes, soit l'écart type empirique $s_5 = 15{,}51$, construire un intervalle de confiance de niveau 0,95 pour cette dépense de loisirs. Comment expliquer l'amplitude de l'intervalle alors que l'enquête porte sur 1 000 individus ?

ENTRAÎNEMENT

Intervalles pour une proportion

12. Sur 100 personnes interrogées, 51 déclarent qu'elles voteront pour le candidat D. Magog aux prochaines élections présidentielles. Donner un intervalle de confiance de niveau 0,95 pour la proportion p d'intentions de vote pour ce candidat, dans la population. Même question si ce sondage a été effectué auprès d'un échantillon de 1 000 personnes. Combien aurait-il fallu interroger d'électeurs pour que le résultat soit fourni à 2 % près ?

Analyse de l'énoncé et conseils. L'intervalle pour p peut être déterminé par simple lecture de l'abaque 2, page 204, ou par l'approximation normale puisque la taille de l'échantillon le permet. La longueur de cet intervalle peut s'exprimer en fonction de la taille d'échantillon, dont la valeur minimale peut être déterminée pour que cette longueur soit inférieure à 4 %.

13. À la sortie d'une chaîne de montage, 20 véhicules automobiles tirés au sort sont testés de façon approfondie et 2 d'entre eux présentent des défauts importants. Donner un intervalle de confiance de niveau 0,95 pour la proportion p de véhicules fabriqués qui présentent des défauts importants.

Analyse de l'énoncé et conseils. L'interprétation du paramètre p conduit ici à construire un intervalle unilatéral à gauche, sans utiliser l'approximation normale puisque la taille d'échantillon est trop faible.

Moyenne d'une loi normale d'écart type connu

14. On a effectué 25 pesées d'un objet, avec une balance dont la précision est mesurée par l'écart type $\sigma = 1$. La totalisation de ces pesées étant de 30,25 g, donner un intervalle de confiance de niveau 0,90 pour le poids de cet objet.

Analyse de l'énoncé et conseils. Le résultat d'une pesée peut être considéré comme une v.a. qui suit une loi normale dont la moyenne est égale au vrai poids de l'objet et dont l'écart type est la caractéristique de précision fournie par le constructeur de la balance. On est alors ramené au problème classique de l'intervalle de confiance pour la moyenne d'une loi normale d'écart type connu.

Moyenne d'une loi normale d'écart type inconnu

15. Un fabricant d'ampoules électriques annonce une durée de vie moyenne de ses ampoules de 170 heures. Afin de vérifier cette affirmation, un organisme de défense des consommateurs prélève au hasard 100 ampoules dans un lot de fabrication et, à l'issue de l'expérimentation, constate une moyenne de durée de vie de 158 heures, avec un écart type empirique de 30 heures. Si on admet que cette durée de vie suit une loi normale, peut-on déduire de cette enquête que l'affirmation du fabricant est mensongère ?

Analyse de l'énoncé et conseils. L'affirmation du fabricant peut être validée à partir de la construction d'un intervalle de confiance qui devrait contenir la valeur 170. En choisissant un niveau de confiance de 99 % au lieu de la valeur habituelle 95 %, l'intervalle obtenu sera plus grand, ce qui est plus favorable au fabricant. Pour que le résultat ne puisse pas être contesté, on peut même choisir un intervalle unilatéral de la forme $m <$ constante.

Variance d'une loi normale de moyenne connue

16. Pour estimer la précision d'un thermomètre, on réalise $n = 100$ mesures indépendantes de la température d'un liquide maintenu à une température constante de 20 degrés Celsius. Les observations x_i, $1 \leq i \leq 100$, ayant conduit à la valeur $\sum_{i=1}^{100} x_i^2 = 40\,011$, construire un intervalle de confiance de niveau 0,95 pour la précision de ce thermomètre, mesurée par la dispersion σ^2 des mesures.

Analyse de l'énoncé et conseils. Si on fait l'hypothèse classique de normalité des erreurs de mesure, il s'agit de construire un intervalle de confiance pour la variance d'une loi normale, de moyenne connue $m = 20$. On peut donner un intervalle symétrique, mais compte tenu de l'interprétation du paramètre, il paraît plus logique de construire un intervalle donnant une borne supérieure de ce qui mesure en fait l'imprécision du thermomètre.

Variance d'une loi normale de moyenne inconnue

17. Afin de vérifier la régularité des horaires des bus, la RATP a effectué 71 relevés des écarts en secondes constatés entre l'arrivée des bus à un arrêt et l'horaire moyen constaté. L'estimation sans biais du paramètre σ^2 obtenue a été de 5 100. En déduire un intervalle de confiance de niveau 0,95 pour cette mesure de la régularité des horaires.

Analyse de l'énoncé et conseils. On fait toujours l'hypothèse classique de normalité des erreurs de mesure, donc il s'agit toujours de construire un intervalle de confiance pour la variance d'une loi normale, mais cette fois de moyenne inconnue. Seuls en effet sont ici relevés les écarts à la moyenne empirique.

Loi de Poisson

18. Le nombre d'interruptions du trafic de plus d'une minute, dans une journée, sur la ligne A du RER, est supposé suivre une loi de Poisson de paramètre inconnu, que l'on se propose d'estimer à partir du relevé de ces interruptions sur 200 journées. Les moments empiriques calculés sur cet échantillon ont pour valeurs $\bar{x}_{200} = 3$ et $s^2_{200} = 3{,}2$. En déduire un intervalle de confiance de niveau 0,95 pour le paramètre de cette loi de Poisson.

Analyse de l'énoncé et conseils. On sait que la moyenne empirique est un estimateur efficace de la moyenne de la loi de Poisson. La loi exacte de cet estimateur ne pouvant pas être utilisée, on construit un intervalle approché à partir de la loi asymptotique, déduite du théorème central limite. Pour que les inégalités obtenues ne soient pas trop difficiles à exploiter, on remplacera au dénominateur le paramètre par son estimateur.

Écart type de la loi normale

19. Soit X une v.a. de loi normale $N(0, \sigma)$ où σ est un paramètre réel strictement positif, que l'on se propose d'estimer à partir d'un échantillon $(X_1, \ldots, X_n)$ de cette loi. Construire, à partir de :

$$D_n = \frac{1}{n} \sum_{i=1}^{n} |X_i|$$

un estimateur sans biais de σ et montrer qu'il est convergent. À partir de l'observation $D_{100} = 1{,}084$, construire un intervalle de confiance de niveau 0,95 pour σ.

Analyse de l'énoncé et conseils. On détermine les moments de la v.a. $|X|$ pour construire l'estimateur demandé par la méthode des moments et étudier ensuite ses propriétés. L'intervalle de confiance sera obtenu à partir de sa loi asymptotique, déduite du théorème central limite.

Loi exponentielle

20. La durée de vie d'un certain matériel électronique est une v.a. X de loi exponentielle de paramètre $\theta > 0$ et $(X_1, \ldots, X_n)$ est un échantillon de cette loi. Déterminer, par la méthode des moments, un estimateur sans biais de θ et montrer qu'il est convergent. Construire un intervalle de confiance de niveau 0,95 pour θ à partir d'un échantillon où on a observé une moyenne empirique $\bar{x}_{100} = 0{,}79$.

Analyse de l'énoncé et conseils. On résout l'équation en θ obtenue par égalisation des moyennes théorique et empirique et on ajuste pour obtenir un estimateur sans biais. On construira l'intervalle de confiance à partir du théorème central limite donnant ensuite la loi asymptotique de cet estimateur à partir de la loi limite d'une suite image figurant dans les rappels de cours du chapitre 5.

Loi exponentielle tronquée

21. Soit X une v.a. de densité :

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

où θ est un paramètre réel inconnu. Si $(X_1, \ldots, X_n)$ est un échantillon de cette loi, déterminer à partir de $m_n = \min\{X_1, \ldots, X_n\}$ un estimateur sans biais de θ. Construire un intervalle de confiance de niveau 0,95 pour θ à partir d'un échantillon où on a observé $m_{100} = 1{,}26$.

Analyse de l'énoncé et conseils. Il faut au préalable déterminer la loi de probabilité de la valeur minimale de l'échantillon pour ensuite calculer ses moments et en déduire un estimateur sans biais. C'est en utilisant la fonction de répartition de m_n que l'on pourra obtenir un intervalle de confiance de niveau exact 0,95 pour le paramètre.

Loi de Pareto

22. Le revenu mensuel X, en francs constants, des ménages d'une population donnée est considéré comme une v.a. de densité :

$$f(x\,;\theta) = \begin{cases} \dfrac{3\theta^3}{x^4} & \text{si } x \geq \theta \\ 0 & \text{sinon} \end{cases}$$

où θ est le revenu minimum de cette population. Si $(X_1, \ldots, X_n)$ est un échantillon de cette loi, construire à partir de $m_n = \min\{X_1, \ldots, X_n\}$ un estimateur sans biais de θ. Construire un intervalle de confiance de niveau 0,98 pour θ à partir d'un échantillon où on a observé $m_{300} = 4{,}04$.

Analyse de l'énoncé et conseils. On détermine au préalable la loi de probabilité de la valeur minimale de l'échantillon pour ensuite calculer les moments et en déduire un estimateur sans biais. C'est en utilisant la fonction de répartition de m_n que l'on pourra obtenir un intervalle de confiance de niveau exact 0,98 pour le paramètre.

SOLUTIONS

1 ▸ Vrai. Il suffit de le choisir suffisamment grand. S'il s'agit d'une proportion, elle est contenue dans $[0, 1]$ avec une probabilité égale à 1. Si le paramètre est positif, l'intervalle $[0, +\infty[$ le contient avec une probabilité égale à 1.

2 ▸ Vrai. Comme il n'y a qu'une condition pour déterminer les deux bornes de l'intervalle de confiance, il y a une infinité de choix possibles. Pour fixer l'intervalle, il faut imposer une condition supplémentaire, comme par exemple l'intervalle unilatéral à droite ou l'intervalle bilatéral symétrique.

3 ▸ Faux. La loi binômiale de nf_n ne permet pas de construire un intervalle de confiance de niveau exact, car c'est une loi discrète. Les abaques permettent cependant d'obtenir un niveau très légérement supérieur à celui souhaité, d'autant plus proche de la valeur exacte que la taille d'échantillon n est élevée.

4 ▸ Vrai. La fréquence empirique f_n qui est l'estimateur de p a une variance proportionnelle à $p\left(1-p\right)$, produit de deux termes dont la somme est constante et qui est maximal quand ils sont égaux, soit $p = 0{,}5$. C'est donc au voisinage de cette valeur que l'intervalle de confiance sera de longueur maximum.

5 ▸ Vrai. La loi de l'estimateur moyenne empirique étant la loi normale symétrique, la longueur de l'intervalle symétrique par rapport à son centre sera minimale.

6 ▸ Faux. Si la loi de l'estimateur dépend d'un autre paramètre inconnu qu'il faut estimer, cette longueur peut être une variable aléatoire, comme dans le cas de l'intervalle de confiance pour la moyenne d'une loi normale d'écart type inconnu.

7 ▸ Faux. Si on prend l'exemple de l'intervalle de confiance pour la moyenne d'une loi normale, l'énoncé est vrai si l'écart type est connu, mais faux dans le cas contraire. En effet, dans ce dernier cas, l'intervalle est de longueur aléatoire $L_n = 2tS_n/\sqrt{n}$, qui tend cependant en probabilité vers zéro quand n tend vers l'infini.

8 ▸ Bien sûr, aucune de ces deux réponses n'est satisfaisante. Le paramètre p étant positif, on pourrait retenir l'intervalle $0 < p < 0{,}10$ qui est bien de niveau supérieur à 0,95 mais très approximatif puisqu'il correspond au cas, le plus défavorable, d'un paramètre voisin de 0,5 alors que manifestement ici il est proche de zéro ! Rappelons d'autre part que l'approximation normale n'est considérée comme justifiée que si $np(1-p) \geq 3$, donc il faudrait que p soit supérieur ici à 4 %. Il faut donc avoir recours à l'abaque 2, page 204, qui va fournir l'intervalle le plus exact, soit $0 < p < 0{,}04$. Cependant, si on ne dispose pas de cet abaque, on peut quand même partir de l'approximation normale donnant l'intervalle symétrique :

$$P\left(-a < \sqrt{n}\,\frac{\widehat{p}_n - p}{\sqrt{pq}} < a\right) = 1 - \alpha$$

avec ici $\widehat{p}_{100} = 0$. Mais on n'utilisera pas la seconde approximation de p par $\widehat{p}_{100}$ au dénominateur puisque cette valeur est nulle. Pour $1-\alpha = 0{,}95$ on lit $a = 1{,}96$ donc on est amené à résoudre les inégalités :

$$-1{,}96 < \sqrt{n}\,\frac{-p}{\sqrt{pq}} < 1{,}96$$

Comme $p > 0$, cela se réduit à :

$$1{,}96 > \sqrt{n}\sqrt{\frac{p}{1-p}}$$

soit l'intervalle de confiance :

$$0 < p < \frac{1{,}96^2}{n + 1{,}96^2} = 0{,}04$$

9 ▸ Si on associe à chaque dossier du fichier une variable de Bernoulli X qui prend la valeur 1 quand il n'est pas exploitable, on aura $P(X=1) = p$ qui représentera la proportion de dossiers inexploitables. Si on prélève n dossiers, il y en aura $n - \sum_{i=1}^{n} X_i$ qui pourront être utilisés. La condition imposée se traduit donc par $P\left(n - \sum_{i=1}^{n} X_i \geq 1\,000\right) = 0{,}95$ qui peut encore s'écrire $P\left(\bar{X}_n \leq 1 - 1\,000/n\right) = 0{,}95$. La quantité $1 - 1\,000/n$ peut donc s'interpréter comme la borne supérieure d'un intervalle de confiance unilatéral pour une proportion, de niveau 0,95. Ayant observé sur un échantillon de taille 100 la valeur $\widehat{p}_{100} = 0{,}2$, on obtient par lecture de l'abaque 1, page 203, l'intervalle $p < 0{,}28$, ce qui correspond à la valeur de n telle que $1\,000/n = 0{,}72$, soit $n = 1\,389$.

10 ▸ On applique la formule donnée dans les rappels de cours avec $\bar{x}_{30} = 2\,000$ et $s = 300$, la table 6 donnant comme fractile d'ordre 0,975 de la loi de Student à 29 degrés de liberté la valeur $t = 2{,}045$. L'intervalle de confiance obtenu est donc $1\,888 < m < 2\,112$. L'estimation de σ étant exactement égale à la valeur de ce paramètre, on aurait pu penser que l'intervalle ne serait pas modifié. Mais l'application des formules du cours dans le cas où σ est connu, avec comme fractile d'ordre 0,975 de la loi normale $u = 1{,}96$, conduit à l'intervalle $1\,893 < m < 2\,107$. Sa longueur est 214, légèrement inférieure à celle, 224, de l'intervalle précédent. Ainsi, même dans le cas d'une estimation sans erreur, le fait d'estimer un paramètre inconnu fait perdre de l'information et donc fournit un intervalle moins précis.

11 ▸ Si X est la v.a. qui représente la dépense consacrée aux loisirs, de loi supposée normale $N(m,\sigma)$, on ne dispose pas ici de 1 000 observations de cette variable, mais seulement de l'échantillon des moyennes empiriques M_j, $1 \leq j \leq 5$, c'est-à-dire d'une v.a. M qui suit une loi $N\left(m,\sigma/\sqrt{200}\right)$. L'intervalle de confiance pour m est donc basé sur la statistique $\sqrt{5}\left(\bar{M}_5 - m\right)/S_5$ qui suit une loi de Student à quatre degrés de liberté. Il est donc de la forme :

$$\bar{M}_5 - t\frac{S_5}{\sqrt{5}} < m < \bar{M}_5 + t\frac{S_5}{\sqrt{5}}$$

où t est le fractile d'ordre 0,975 de la loi T_4 lu dans la table 6, page 200, soit $t = 2{,}776$. L'intervalle obtenu est donc $282 < m < 320$. L'amplitude assez élevée de cet intervalle, malgré une taille d'échantillon importante, s'explique par une grande dispersion du caractère mesuré dans les cinq groupes, qui paraissent assez dissemblables quant à leur comportement vis-à-vis des loisirs.

12 ▸ La lecture de l'abaque pour $n = 100$ et $\widehat{p}_{100} = 0{,}51$ donne l'intervalle $0{,}42 < p < 0{,}62$. Le fractile d'ordre 0,975 de la loi normale a comme valeur $a = 1{,}96$, ce qui fournit comme nouvel intervalle approché $0{,}41 < p < 0{,}61$ de même longueur que le précédent mais décalé de 1 % vers la gauche. Enfin, en remplaçant $p\left(1-p\right)$ par sa borne supérieure $1/4$, on obtient le même intervalle, ce qui s'explique par la valeur exacte de p qui doit être effectivement proche de 0,5. Pour une taille d'échantillon $n = 1\,000$, on ne peut plus utiliser l'abaque, mais l'approximation normale fournit un intervalle précis, soit ici $0{,}48 < p < 0{,}54$ de longueur 6 %. Pour une taille d'échantillon n quelconque, la longueur de l'intervalle est :

$$l_n = 2a\sqrt{\frac{\widehat{p}_n\left(1-\widehat{p}_n\right)}{n}}$$

et la condition $l_n \leq 0{,}04$ impose $n \geq 2\,500a^2\widehat{p}_n\left(1-\widehat{p}_n\right)$. Pour une estimation voisine ici de 0,5, on aboutit à la contrainte $n \geq 2\,400$. Dans le cas le plus défavorable d'une proportion voisine de 50 %, il faut donc des tailles d'échantillon élevées pour l'estimer avec précision.

13 ▸ On obtient ici $\widehat{p}_{20} = 0{,}10$ et $n\widehat{p}_{20}\left(1-\widehat{p}_{20}\right) = 1{,}8$, ce qui n'autorise pas à admettre que $np\left(1-p\right) \geq 3$, condition de validité de l'approximation normale. L'intervalle unilatéral exact est fourni par l'abaque 1, page 203 : $p < 0{,}28$. L'utilisation de l'approximation normale avec le fractile d'ordre 0,95 $a = 1{,}644\,9$ aurait donné l'intervalle $p < 0{,}21$ assez éloigné de l'intervalle exact.

14 ▸ On peut considérer que la valeur indiquée par la balance est une v.a. X de loi normale, d'espérance m égale au poids réel de l'objet et d'écart type $\sigma = 1$. Le fractile d'ordre 0,95 de la loi normale étant $u = 1{,}644\,9$, l'observation $\sum_{i=1}^{25} x_i = 30{,}25$ conduit à l'estimation ponctuelle $\bar{x}_{25} = 1{,}21$ puis à l'intervalle de confiance $0{,}88 < m < 1{,}54$. Compte tenu du petit nombre de mesures et de la valeur élevée de l'écart type, cet intervalle est peu précis.

15 ▸ Pour construire un intervalle de confiance pour la moyenne m de cette loi normale d'écart type inconnu σ, on utilise la statistique $\sqrt{n}\left(\bar{X}_n - m\right)/S_n$ qui suit une loi de Student à $n-1$ degrés de liberté. Avec ici $n = 100$, on lit dans la table 6 le fractile d'ordre 0,995 : $t = 2{,}63$. On obtient comme intervalle bilatéral symétrique : $150 < m < 166$. La valeur indiquée par le fabricant est bien en dehors de cet intervalle. Si on choisit un intervalle de même niveau, mais unilatéral à gauche, il sera défini par :

$$P\left[\sqrt{n}\,\frac{\bar{X}_n - m}{S_n} > t\right] = 0{,}99$$

soit $t = 2{,}37$ et l'intervalle $m < 165$. Il y a donc 99 chances sur 100 que cette publicité soit mensongère.

16 ▸ On obtient comme estimation de la variance :

$$\widehat{\sigma}^2_{100} = \frac{1}{100}\sum_{i=1}^{100} x_i^2 - m^2 = 0{,}11$$

Les fractiles d'ordre 0,025 et 0,975 de la loi du khi-deux à 100 degrés de liberté, lus dans la table 5, page 199, sont respectivement $a = 74{,}22$ et $b = 129{,}56$, d'où l'intervalle bilatéral symétrique $0{,}08 < \sigma^2 < 0{,}15$. Si on retient un intervalle unilatéral à gauche, le fractile d'ordre 0,05 est $a = 77{,}93$, ce qui donne comme intervalle $\sigma^2 < 0{,}14$.

17 ▸ Pour un intervalle de confiance bilatéral symétrique, on relève dans la table 5 les fractiles d'ordres respectifs 0,025 et 0,975 à 70 degrés de liberté $a = 48{,}76$ et $b = 95{,}02$, d'où $3\,757 < \sigma^2 < 7\,322$. Pour un intervalle unilatéral à gauche, le fractile d'ordre 0,05 est $a = 51{,}74$, ce qui donne comme intervalle $\sigma^2 < 6\,900$.

18 ▸ D'après le théorème central limite, l'estimateur du paramètre θ, centré et réduit, converge en loi vers la loi normale centrée et réduite :

$$\sqrt{n}\,\frac{\bar{X}_n - \theta}{\sqrt{\theta}} \underset{\text{loi}}{\rightarrow} N(0,1)$$

Les inéquations que l'on peut en déduire s'exprimant en fonction de θ et $\sqrt{\theta}$, on remplace le paramètre au dénominateur par son estimateur et l'intervalle de confiance bilatéral symétrique se déduit de la condition :

$$P\left[-u < \sqrt{n}\,\frac{\bar{X}_n - \theta}{\sqrt{\bar{X}_n}} < u\right] = 0{,}95$$

où on retient comme valeur approchée de u le fractile d'ordre 0,975 de la loi $N(0,1)$, soit $u = 1{,}96$. L'intervalle obtenu est alors :

$$-u\sqrt{\frac{\bar{X}_n}{n}} < \bar{X}_n - \theta < u\sqrt{\frac{\bar{X}_n}{n}} \iff \bar{X}_n - u\sqrt{\frac{\bar{X}_n}{n}} < \theta < \bar{X}_n + u\sqrt{\frac{\bar{X}_n}{n}}$$

soit pour cet échantillon $2{,}76 < \theta < 3{,}24$.

19 ▶ Pour calculer l'espérance de $|X|$, on fait le changement de variable $2\sigma^2 u = x^2$:

$$E(|X|) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} |x| \, \mathrm{e}^{-x^2/2\sigma^2} \, \mathrm{d}x = \frac{2\sigma}{\sqrt{2\pi}} \int_0^{+\infty} \mathrm{e}^{-u} \, \mathrm{d}u = \sigma\sqrt{\frac{2}{\pi}}$$

L'expression de la variance se déduit aisément de la formule développée :

$$V(|X|) = E\left(|X|^2\right) - E^2(|X|) = \sigma^2 - \frac{2}{\pi}\sigma^2 = \left(1 - \frac{2}{\pi}\right)\sigma^2$$

L'estimateur de σ obtenu par la méthode des moments est solution de $D_n = \sigma\sqrt{2/\pi}$, puisque D_n est la moyenne empirique de l'échantillon $(|X_1|, \ldots, |X_n|)$. La solution est l'estimateur, sans biais par construction :

$$T_n = \sqrt{\frac{\pi}{2}} D_n$$

Il est convergent, d'après la loi des grands nombres, puisque :

$$D_n \underset{p}{\rightarrow} E(|X|) = \sigma\sqrt{\frac{2}{\pi}}$$

Cet estimateur étant sans biais et de variance :

$$V(T_n) = \frac{\pi}{2} V(D_n) = \frac{\pi}{2n} V(|X|) = \left(\frac{\pi}{2} - 1\right)\frac{\sigma^2}{n}$$

qui tend vers zéro, est d'ailleurs aussi convergent en moyenne quadratique. Sa loi asymptotique va se déduire du théorème central limite :

$$\sqrt{n}\,\frac{D_n - E(|X|)}{\sigma(|X|)} \underset{\text{loi}}{\rightarrow} N(0,1)$$

En multipliant numérateur et dénominateur par $\sqrt{\pi/2}$, on obtient en effet :

$$\sqrt{n}\,\frac{T_n - \sigma}{\sigma\sqrt{\pi/2 - 1}} \underset{\text{loi}}{\rightarrow} N(0,1)$$

On peut alors obtenir un intervalle bilatéral symétrique de niveau voisin de 0,95 par la relation :

$$P\left[-u < \sqrt{n}\,\frac{T_n - \sigma}{\sigma\sqrt{\pi/2 - 1}} < u\right] = 0{,}95$$

en prenant pour valeur approchée de u le fractile d'ordre 0,975 de la loi $N(0,1)$, soit $u = 1{,}96$. La résolution de cette double inégalité conduit à :

$$-u\sigma\sqrt{\frac{\pi - 2}{2n}} < T_n - \sigma < u\sigma\sqrt{\frac{\pi - 2}{2n}} \Leftrightarrow$$

$$\sigma\left[1 - u\sqrt{\frac{\pi - 2}{2n}}\right] < T_n < \sigma\left[1 + u\sqrt{\frac{\pi - 2}{2n}}\right]$$

D'où l'intervalle pour l'écart type :

$$\frac{T_n}{1 + u\sigma\sqrt{(\pi - 2)/2n}} < \sigma < \frac{T_n}{1 - u\sigma\sqrt{(\pi - 2)/2n}}$$

Soit, pour cet échantillon, l'intervalle de confiance : $1{,}16 < \sigma < 1{,}56$.

20 ▸ On sait que la moyenne théorique de la loi exponentielle $\mathcal{E}(\theta)$ est $E(X) = 1/\theta$, donc la méthode des moments conduit à résoudre l'équation en θ : $1/\theta = \bar{X}_n$. La solution fournit l'estimateur $\widehat{\theta}_n = 1/\bar{X}_n$. La loi exponentielle est la loi $\gamma(1, \theta)$; donc $n\bar{X}_n = S_n = \sum_{i=1}^{n} X_i$ suit une loi $\gamma(n, \theta)$, dont la densité permet de calculer l'espérance de cet estimateur :

$$E\left(\widehat{\theta}_n\right) = nE\left(\frac{1}{S_n}\right) = \frac{n}{\Gamma(n)} \int_0^{+\infty} \theta^n e^{-\theta x} x^{n-2}\, dx$$

$$= \frac{n}{\Gamma(n)} \int_0^{+\infty} \theta u^{n-2} e^{-u}\, du = \theta \frac{n\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1}\theta$$

Cet estimateur est asymptotiquement sans biais, l'estimateur sans biais étant :

$$T_n = \frac{n-1}{n}\widehat{\theta}_n = \frac{n-1}{S_n}$$

La loi des grands nombres permet d'obtenir la convergence en probabilité de $\bar{X}_n$ vers $E(X) = 1/\theta$ et le théorème de Slutsky permet de conclure que $1/\bar{X}_n$ converge vers θ, donc aussi T_n puisque $(n-1)/n$ tend vers 1. On peut aussi calculer la variance de cet estimateur à partir de :

$$E\left(\frac{1}{S_n^2}\right) = \frac{1}{\Gamma(n)} \int_0^{+\infty} \theta^n e^{-\theta x} x^{n-3}\, dx = \frac{\theta^2}{\Gamma(n)} \int_0^{+\infty} u^{n-3} e^{-u}\, du$$

$$= \theta^2 \frac{\Gamma(n-2)}{\Gamma(n)} = \frac{\theta^2}{(n-1)(n-2)}$$

D'où on déduit $V(1/S_n) = \theta^2/(n-1)^2(n-2)$, puis :

$$V(T_n) = (n-1)^2\, V\left(\frac{1}{S_n}\right) = \frac{\theta^2}{n-2}$$

qui tend vers zéro quand n devient infini. Donc, T_n converge en moyenne quadratique vers θ, puisqu'il est sans biais. La loi de T_n n'est pas une loi tabulée qui permet de construire un intervalle de confiance, donc on utilise la loi asymptotique déduite du théorème central limite :

$$\sqrt{n}\,\frac{\bar{X}_n - 1/\theta}{1/\theta} \underset{\text{loi}}{\rightarrow} N(0, 1)$$

car, pour la loi exponentielle $\mathcal{E}(\theta)$, $V(X) = 1/\theta^2$. Pour obtenir la loi asymptotique de $\widehat{\theta}_n = 1/\bar{X}_n$ on applique le résultat relatif à la loi limite d'une suite image vue au chapitre 5, avec $\widehat{\theta}_n = g\left(\bar{X}_n\right)$ et $g(x) = 1/x$, soit $g'(x) = -1/x^2$ et $\left|g'(1/\theta)\right| = \theta^2$, d'où :

$$\sqrt{n}\,\frac{\widehat{\theta}_n - \theta}{\theta} \underset{\text{loi}}{\rightarrow} N(0,1)$$

Le résultat est le même pour $T_n = \dfrac{n-1}{n}\widehat{\theta}_n$ puisque $\dfrac{n-1}{n}$ tend vers 1 quand n devient infini. On obtient donc un intervalle de confiance bilatéral symétrique pour θ à partir de la condition :

$$P\left[-u < \sqrt{n}\,\frac{T_n - \theta}{\theta} < u\right] = 0{,}95$$

en remplaçant u par le fractile d'ordre 0,975 de la loi normale centrée-réduite, soit $u = 1{,}96$. La résolution de cette double inégalité pour θ conduit à l'intervalle :

$$-u\frac{\theta}{\sqrt{n}} < T_n - \theta < u\frac{\theta}{\sqrt{n}} \quad \Longleftrightarrow \quad \frac{T_n}{1 + u/\sqrt{n}} < \theta < \frac{T_n}{1 - u/\sqrt{n}}$$

Soit, pour cet échantillon $1{,}05 < \theta < 1{,}56$.

21 ▸ Pour déterminer la loi de $m_n = \min\{X_1, \ldots, X_n\}$, on considère l'événement $m_n > x$ qui est équivalent à l'événement tous les X_i sont supérieurs à x. Les v.a. X_i étant indépendantes et de même loi de f.r. F, on en déduit :

$$P(m_n > x) = P\left\{\bigcap_{i=1}^{n}(X_i > x)\right\} = \prod_{i=1}^{n} P(X_i > x) = [1 - F(x\,;\theta)]^n$$

La f.r. de X est nulle pour $x < \theta$ et d'expression :

$$F(x\,;\theta) = \int_\theta^x \mathrm{e}^{-(t-\theta)}\,\mathrm{d}t = 1 - \mathrm{e}^{-(x-\theta)}$$

pour $x \geq \theta$. La f.r. de m_n est donc aussi nulle pour $x < \theta$ et définie par :

$$G(x\,;\theta) = 1 - \mathrm{e}^{-n(x-\theta)}$$

pour $x \geq \theta$, sa densité étant $g(x\,;\theta) = n\mathrm{e}^{-n(x-\theta)}$. On peut alors calculer son espérance :

$$E(m_n) = n\int_\theta^{+\infty} x\mathrm{e}^{-n(x-\theta)}\,\mathrm{d}x = n\int_\theta^{+\infty}(x-\theta)\,\mathrm{e}^{-n(x-\theta)}\,\mathrm{d}x + \theta n\int_\theta^{+\infty}\mathrm{e}^{-n(x-\theta)}\,\mathrm{d}x$$
$$= \frac{1}{n}\int_0^{+\infty} u\mathrm{e}^{-u}\,\mathrm{d}u + \theta\int_\theta^{+\infty} g(x\,;\theta)\,\mathrm{d}x = \frac{1}{n} + \theta$$

la première intégrale étant l'espérance d'une loi exponentielle de paramètre 1 et la seconde l'intégrale de la densité de m_n. L'estimateur sans biais construit à partir de m_n est donc $T_n = m_n - 1/n$. Comme $m_n > \theta$, on aura $T_n > \theta - 1/n$; donc l'intervalle retenu sera de la forme $T_n - a < \theta < T_n + 1/n$. La f.r. de m_n va permettre de déterminer un intervalle de confiance de niveau $1 - \alpha$ par la condition :

$$\begin{aligned} 1 - \alpha &= P\left(-\frac{1}{n} < T_n - \theta < a\right) = P\left(\theta < m_n < \theta + a + \frac{1}{n}\right) \\ &= G\left(\theta + a + \frac{1}{n}\,;\theta\right) - G(\theta\,;\theta) = G\left(\theta + a + \frac{1}{n}\,;\theta\right) = 1 - \mathrm{e}^{-na-1} \end{aligned}$$

soit $\alpha = \mathrm{e}^{-(na+1)}$ ou $a = -\left(1 + \ln \alpha\right)/n$, ce qui correspond à l'intervalle :

$$m_n + \frac{\ln \alpha}{n} < \theta < m_n$$

Pour cet échantillon , on obtient l'intervalle : $1{,}23 < \theta < 1{,}26$.

22 ▸ Comme dans l'exercice précédent, on détermine la loi de $m_n = \min\{X_1, \ldots, X_n\}$ à partir de :

$$P(m_n > x) = P\left\{\bigcap_{i=1}^{n}(X_i > x)\right\} = \prod_{i=1}^{n} P(X_i > x) = [1 - F(x\,;\theta)]^n$$

La f.r. F de X est nulle pour $x < \theta$ et d'expression :

$$F(x\,;\theta) = 3\theta^3 \int_\theta^x \frac{\mathrm{d}t}{t^4} = \theta^3\left[-t^{-3}\right]_\theta^x = 1 - \frac{\theta^3}{x^3}$$

pour $x \geq \theta$. La f.r. de m_n est donc aussi nulle pour $x < \theta$ et définie par :

$$G(x\,;\theta) = 1 - \left(\frac{\theta}{x}\right)^{3n}$$

pour $x \geq \theta$, sa densité étant $g(x\,;\theta) = 3n\theta^{3n}/x^{3n+1}$. On peut alors calculer son espérance :

$$E(m_n) = 3n\int_\theta^{+\infty} \frac{\theta^{3n}}{x^{3n}}\,\mathrm{d}x = 3n\theta^{3n}\left[-\frac{1}{(3n-1)\,x^{3n-1}}\right]_\theta^{+\infty} = \frac{3n}{3n-1}\theta$$

L'estimateur sans biais est donc $T_n = \dfrac{3n-1}{3n}m_n$. En raison de la condition $m_n > \theta$ et la f.r. de m_n dépendant de x/θ, l'intervalle de confiance de niveau $1 - \alpha$ est déterminé par la condition :

$$\begin{aligned} 1 - \alpha &= P(0 < m_n - \theta < a\theta) = P(\theta < m_n < \theta + a\theta) \\ &= G(\theta + a\theta\,;\theta) - G(\theta\,;\theta) = G(\theta + a\theta\,;\theta) = 1 - \frac{1}{(1+a)^{3n}} \end{aligned}$$

soit $(1+a)^{3n} = 1/\alpha$ ou $\ln(1+a) = -\ln\alpha/3n$. L'intervalle pour le paramètre est :

$$\frac{m_n}{1+a} < \theta < m_n$$

Pour $\alpha = 0{,}02$ et $n = 300$, on obtient $a = 0{,}004$ et l'intervalle $4{,}02 < \theta < 4{,}04$, donc intervalle très étroit.

Théorie des tests 8

L'ESSENTIEL DU COURS

Bâtir un test nécessite, comme pour un problème d'estimation, de construire au préalable un modèle statistique où la v.a. X suit une loi de probabilité P_θ, qui dépend d'un paramètre θ inconnu. On fait l'hypothèse a priori *que la valeur de ce paramètre est égale à une valeur fixée θ_0 et on cherche à valider cette hypothèse, au vu d'un échantillon de la loi de X. Cette hypothèse qui est privilégiée, parce qu'elle paraît la plus vraisemblable* a priori, *est appelée hypothèse nulle et notée H_0. Construire un test va consister à découper l'ensemble $\mathbb{R}^n$ des réalisations possibles du n-échantillon en deux régions, celle où l'on décidera d'accepter H_0, et celle où on décidera de la rejeter, qui se nommera région critique du test. Pour délimiter ces deux régions, on fixera une valeur faible à la probabilité de l'erreur qui consiste à décider, au vu de l'échantillon, de rejeter l'hypothèse nulle alors que celle-ci est vérifiée. Cette probabilité se nomme risque de première espèce, sa valeur standard étant de 5 %. Lorsque le paramètre θ ne peut prendre que deux valeurs distinctes θ_0 et θ_1, c'est le théorème de Neyman et Pearson qui permet de déterminer la forme de la région critique, à partir du rapport des vraisemblances associées à chacune des deux valeurs possibles du paramètre.*

1. Définitions et principe général de construction

On considère un modèle statistique où la loi de probabilité P_θ de la v.a. X dépend d'un paramètre inconnu θ qui varie dans un sous-ensemble donné Θ de $\mathbb{R}$. On suppose que cet ensemble est partitionné en deux sous-ensembles donnés Θ_0 et Θ_1, auxquels vont être associées les deux *hypothèses* notées $H_0 : \theta \in \Theta_0$ et $H_1 : \theta \in \Theta_1$. Construire un *test* consiste à définir une *règle de décision* qui va associer une décision à un échantillon observé $(X_1, \ldots, X_n)$ de la loi de X, les deux décisions possibles étant D_0 : accepter H_0, et D_1 : accepter H_1. À chaque décision correspond une région de $\mathbb{R}^n$, qui va donc être partitionnée en deux sous-ensembles W et $\bar{W}$, c'est-à-dire que si la réalisation de l'échantillon est un point $(x_1, \ldots, x_n)$ de W, on décide D_1, donc on rejette H_0 ; dans le cas contraire, c'est-à-dire pour un point de $\bar{W}$, on décide D_0, donc on accepte H_0.

Définitions. La région W de rejet de l'hypothèse nulle H_0 se nomme *région critique* du test et la région $\bar{W}$ *région d'acceptation*.

La construction d'un test va donc consister à déterminer cette région critique et la méthode retenue dépendra des conséquences que l'on attribue à chacune des deux erreurs qui sont associées aux deux décisions possibles et qui sont les suivantes.

Définitions. L'*erreur de première espèce* consiste à décider D_1 alors que H_0 est vraie, soit rejeter à tort l'hypothèse nulle H_0. L'*erreur de seconde espèce* consiste à décider D_0 alors que H_1 est vraie, soit accepter à tort l'hypothèse nulle H_0.

2. Méthode de Bayes

On se place dans le cas où on a attribué des probabilités *a priori* p_0 et $p_1 = 1 - p_0$ à chacune des hypothèses respectives H_0 et H_1 et que l'on a également associé un coût à chaque décision, en fonction de l'hypothèse qui est effectivement réalisée. Le tableau suivant contient ces coûts, la décision prise figurant en colonne et l'hypothèse vraie en ligne :

	D_0	D_1
$H_0\,(p_0)$	C_{00}	C_{01}
$H_1\,(p_1)$	C_{10}	C_{11}

Après la réalisation $(x_1, \ldots, x_n)$ on peut calculer, à l'aide du théorème de Bayes, les probabilités *a posteriori* π_0 et π_1 des hypothèses H_0 et H_1 :

$$\pi_0 = \frac{p_0 L_0}{p_0 L_0 + p_1 L_1} \quad \text{et} \quad \pi_1 = \frac{p_1 L_1}{p_0 L_0 + p_1 L_1}$$

où on a noté L_0 la valeur de la vraisemblance $L(x_1, \ldots, x_n\,;\theta)$, quand $\theta \in \Theta_0$, et L_1, quand $\theta \in \Theta_1$. On peut alors calculer les espérances du coût de chaque décision pour cette distribution *a posteriori* :

$$E[C(D_0)] = C_{00}\pi_0 + C_{10}\pi_1 \quad \text{et} \quad E[C(D_1)] = C_{01}\pi_0 + C_{11}\pi_1$$

La *règle de décision de Bayes* consiste à associer à l'observation $(x_1, \ldots, x_n)$ la décision dont l'espérance de coût est la plus faible.

3. Méthode de Neyman et Pearson

3.1 Principe de la règle de Neyman et Pearson

On privilégie une hypothèse que l'on considère comme la plus vraisemblable et que l'on choisit comme *hypothèse nulle* H_0. Ce sera celle dont le rejet à tort sera le plus préjudiciable. L'autre hypothèse H_1 est l'*hypothèse alternative*.

Définitions. On appelle *risque de première espèce* la probabilité de rejeter à tort l'hypothèse nulle, soit :

$$\alpha = P_\theta(D_1|H_0) = P_\theta(H_1|H_0) = P_\theta(W|\theta \in \Theta_0)$$

On appelle *risque de seconde espèce* la probabilité d'accepter à tort l'hypothèse nulle, soit :

$$\beta = P_\theta(D_0|H_1) = P_\theta(H_0|H_1) = P_\theta(\bar{W}|\theta \in \Theta_1)$$

Privilégier l'hypothèse nulle revient à considérer que l'erreur la plus grave consiste à la rejeter à tort. La méthode de Neyman et Pearson fixe un seuil maximum α_0 au risque de première espèce et le test est alors déterminé par la recherche de la règle qui minimise l'autre risque de seconde espèce.

Définition. On appelle *puissance* d'un test la probabilité de refuser H_0 avec raison, c'est-à-dire lorsque H_1 est vérifiée, soit :

$$\eta = P_\theta(D_1|H_1) = P_\theta(H_1|H_1) = P_\theta(W|\theta \in \Theta_1) = 1 - \beta$$

La *règle de décision de Neyman et Pearson* consiste à déterminer la région critique W pour laquelle la puissance est maximum, sous la contrainte $\alpha \leq \alpha_0$.

3.2 Hypothèses simples

Une hypothèse est qualifiée de *simple* si la loi de la v.a. X est totalement spécifiée quand cette hypothèse est réalisée. Dans le cas contraire, elle est dite *multiple*. Nous allons examiner le cas où le paramètre θ ne peut prendre que deux valeurs θ_0 et θ_1, ce qui correspond au choix entre les deux hypothèses simples suivantes :

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases}$$

Théorème de Neyman et Pearson. Pour un risque de première espèce fixé à α_0, le test de puissance maximum entre les hypothèses simples ci-dessus est défini par la région critique :

$$W = \left\{ (x_1, \ldots, x_n) \,/ \frac{L_0\,(x_1, \ldots, x_n)}{L_1\,(x_1, \ldots, x_n)} \leq k \right\}$$

où la valeur de la constante k est déterminée par le risque fixé $\alpha_0 = P_\theta\,(W|\theta = \theta_0)$, ayant posé $L_0\,(x_1, \ldots, x_n) = L\,(x_1, \ldots, x_n\,;\theta_0)$ et $L_1\,(x_1, \ldots, x_n) = L\,(x_1, \ldots, x_n\,;\theta_1)$.

3.3 Hypothèses multiples

Nous allons seulement considérer le cas d'une hypothèse simple contre une hypothèse multiple de la forme suivante :

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$$

On détermine d'abord par la méthode de Neyman et Pearson la région critique W du test suivant :

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases}$$

où θ_1 est une valeur fixée différente de θ_0. Si W ne dépend pas de θ_1, alors on aura obtenu un *test uniformément le plus puissant (UPP)*, c'est-à-dire que pour toute autre région critique W', on aura $P_\theta\,(W|\theta \in \Theta_1) \geq P_\theta\left(W'|\theta \in \Theta_1\right)$ pour tout θ de Θ_1.

Théorème de Lehmann

$$\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$$

On suppose que la loi P_θ est à rapport des vraisemblances monotone, c'est-à-dire qu'il existe une statistique T_n telle que :

$$\frac{L\,(x_1, \ldots, x_n\ ;\theta)}{L\,(x_1, \ldots, x_n\ ;\theta')}$$

soit une fonction croissante de T_n pour toutes les valeurs $\theta > \theta'$.

Alors, il existe un test UPP de niveau α, de région critique définie par la condition :

$$T_n > k.$$

4. Tests du khi-deux

4.1 Test d'indépendance

Pour tester l'indépendance de deux caractères X et Y, qualitatifs ou quantitatifs, à respectivement r et s modalités, on relève le nombre n_{ij} d'individus d'une population de taille $n = \sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij}$ qui possèdent simultanément la modalité i, $1 \leq i \leq r$, du caractère X et la modalité j, $1 \leq j \leq s$, du caratère Y. Si p_{ij} est la probabilité théorique correspondante, les probabilités marginales sont $p_{i.} = \sum_{j=1}^{s} p_{ij}$ et $p_{.j} = \sum_{i=1}^{r} p_{ij}$ et l'indépendance de ces deux caractères se traduit par l'hypothèse nulle $H_0 : p_{ij} = p_{i.} \times p_{.j}$. Pour tester cette hypothèse contre l'alternative $H_1 : p_{ij} \neq p_{i.} \times p_{.j}$, on utilise la statistique :

$$D_n = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{\left(n_{ij} - n_{i.}n_{.j}/n\right)^2}{n_{i.}n_{.j}/n} = n\left(\sum_{i=1}^{r}\sum_{j=1}^{s} \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1\right)$$

dont la loi asymptotique sous H_0 est la loi du khi-deux à $(r-1)(s-1)$ degrés de liberté. On a posé $n_{i.} = \sum_{j=1}^{s} n_{ij}$ et $n_{.j} = \sum_{i=1}^{r} n_{ij}$. La région critique de ce test est de la forme $D_n \geq C$; pour un risque de première espèce $\alpha = P(D_n \geq C|H_0)$, la valeur de C est approximée par le fractile d'ordre $1-\alpha$ de la loi $\chi^2_{(r-1)(s-1)}$.

4.2 Test d'adéquation

Pour tester l'hypothèse nulle qu'une v.a. X admet pour fonction de répartition une fonction donnée F, on utilise un $n-$ échantillon de la loi de X. Si la loi est discrète, avec k valeurs possibles de probabilités $p_i = P(X = x_i)$, ou si les observations continues ont été réparties en k classes disjointes (a_i, a_{i+1}), avec cette fois $p_i = P\{X \in (a_i, a_{i+1})\} = F(a_{i+1}) - F(a_i)$, on compare cette distribution théorique avec la distribution empirique définie par les valeurs n_i/n, où n_i est le nombre d'observations x_i ou appartenant à la classe (a_i, a_{i+1}). On utilise comme distance entre ces deux distributions la statistique :

$$D_n = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{k} \frac{n_i^2}{np_i} - n$$

dont la loi asymptotique sous H_0 est la loi du khi-deux à $k-1$ degrés de liberté. La région critique de ce test est de la forme $D_n \geq C$. Pour un risque de première espèce $\alpha = P(D_n \geq C|H_0)$, la valeur de C est approximée par le fractile d'ordre $1-\alpha$ de la loi χ^2_{k-1}. On considère que cette approximation est possible si $np_i \geq 5$; en cas d'effectif de classe trop petit pour que cette condition soit remplie, on regroupe des classes contiguës. Si la loi de X sous H_0 dépend de r paramètres qu'il a fallu estimer au préalable, le nombre de degrés de liberté de la loi asymptotique devient $k-r-1$.

POUVEZ-VOUS RÉPONDRE ?

	Vrai	Faux
1. Pour un problème de test donné, choisir une valeur importante de la probabilité *a priori* p_0 de l'hypothèse nulle dans la méthode de Bayes est équivalent à choisir une valeur faible du risque de première espèce α dans la méthode de Neyman et Pearson.	☐	☐
2. Les deux hypothèses d'un test de Neyman et Pearson jouent des rôles symétriques.	☐	☐
3. Le risque de seconde espèce β est toujours plus grand que le risque de première espèce α.	☐	☐
4. Sans information *a priori*, chacune des deux hypothèses peut indifféremment être choisie comme hypothèse nulle, dans un test de risque de première espèce fixé α.	☐	☐

	Vrai	**Faux**
5. Dans le cas d'un test entre deux hypothèses simples, on peut toujours calculer sa puissance.	☐	☐
6. Dans le cas d'un test entre deux hypothèses simples, on peut toujours trouver un test de risque de première espèce fixé α.	☐	☐
7. Si on intervertit les deux hypothèses simples d'un test, la conclusion au vu du même échantillon peut être changée.	☐	☐

QUESTIONS DE RÉFLEXION

8. Le nombre X d'incidents graves par an survenus dans l'une des centrales nucléaires françaises, suit une loi de Poisson de paramètre λ. Disposant de statistiques sur les neuf dernières années, construire un test de risque de première espèce $\alpha = 0{,}05$ entre les deux hypothèses suivantes :

$$\begin{cases} H_0 : \lambda = 2 \\ H_1 : \lambda = 1 \end{cases}$$

9. Un générateur de nombres au hasard permet d'obtenir une simulation de la loi uniforme sur l'ensemble des dix chiffres $\{0, 1, \ldots, 9\}$. La distance entre la distribution empirique obtenue sur un échantillon de $n = 100$ chiffres et la distribution théorique a pour valeur $D_{100} = 2{,}1$. Le fractile d'ordre 0,95 de la loi χ^2_9 étant 16,9, on accepte l'adéquation à la loi uniforme. Si la même distribution empirique est obtenue sur un échantillon de taille $N = 1\,000$, expliquer pourquoi cette fois on refuse l'adéquation à la loi uniforme.

ENTRAÎNEMENT

Méthode de Bayes

10. Un magasin propose aux acheteurs d'un téléviseur de souscrire pour 300 F une garantie complémentaire de cinq ans pour le remplacement du tube-image, d'une valeur de 2 500 F. On sait par ailleurs que 80 % de ces téléviseurs sont fabriqués en Asie et que la probabilité de défaillance du tube-image est alors $p = 0{,}15$. Le reste des appareils est fabriqué en Europe, avec dans ce cas $p = 0{,}05$. En interrogeant vingt anciens acheteurs, on apprend que deux d'entre eux ont eu à changer leur tube dans les cinq ans qui ont suivi leur achat. Cette information conduit-elle à souscrire à cette garantie ? Une garantie à 350 F ferait-elle changer d'avis ?

Analyse de l'énoncé et conseils. Pour utiliser la méthode de Bayes, il faut construire le tableau des coûts. Si on ne prend pas la garantie, on détermine le coût moyen qui dépend de la valeur de p. La vraisemblance est associée ici à un échantillon d'une loi de Bernoulli qui prend la valeur un quand le tube est en panne dans les cinq années après l'achat.

11. Une entreprise possède cinq photocopieuses sur lesquelles neuf interventions ont été nécessaires l'année précédente, avec un coût moyen par intervention de 1 000 F. Une société lui propose un contrat d'entretien de 2 400 F par an et par machine. Le nombre annuel de pannes est supposé suivre une loi de Poisson de paramètre λ. Pour le quart des machines produites la valeur du paramètre est $\lambda = 3$. Une amélioration du procédé de fabrication a permis d'obtenir ensuite $\lambda = 1$ pour la production des trois quarts restants. Compte tenu de ces informations, quelle sera la décision de l'entreprise ?

Analyse de l'énoncé et conseils. Le coût moyen en l'absence de contrat dépend du paramètre λ qui représente l'espérance de la loi. La vraisemblance d'un échantillon d'une loi de Poisson permet ensuite de construire la règle de Bayes et de conclure au vu des observations faites.

Comparaison des méthodes de Bayes et de Neyman-Pearson

12. Le dirigeant d'une grande marque de vêtements pense qu'une campagne publicitaire choquante aurait une chance sur deux de réussir. Avant de lancer cette campagne à l'échelon national, une pré-campagne a été effectuée dans cinq grandes villes, ce qui a conduit à une augmentation de bénéfices pour cette marque de quatre millions de francs au total. On admet que l'accroissement de bénéfices dans une ville est une v.a. de loi normale d'écart type connu $\sigma = 2$ et de moyenne $m = 2$ si la campagne a été efficace et $m = 0$ dans le cas contraire. Sachant que le coût par ville de cette campagne est de 0,5 million de francs, quelle décision ces informations conduisent-elles à prendre ? Quelle serait la règle de décision dans l'optique de Neyman-Pearson où on souhaiterait avant tout limiter le risque de ne pas lancer une campagne qui serait efficace ? Même question dans le cas où la campagne est considérée comme efficace dès que $m > 0$?

Analyse de l'énoncé et conseils. On traduit d'abord les informations données pour construire le tableau des coûts. On écrit ensuite la vraisemblance d'un échantillon de taille cinq d'une loi normale d'écart type donné, ce qui permet d'établir la règle de Bayes et de conclure au vu des observations recueillies après la pré-campagne. On construit ensuite la règle de décision de Neyman-Pearson associée à l'hypothèse nulle retenue par ce dirigeant. Cette règle dépend bien sûr du risque de première espèce α retenu et on verra que la conclusion est différente selon les choix $\alpha = 0{,}05$ et $\alpha = 0{,}10$. Le changement d'hypothèse nulle proposé ensuite conduit à modifier très sensiblement cette règle.

Test d'une proportion

13. Entre les deux tours de l'élection présidentielle, le candidat D. Magog commande un sondage à une société spécialisée pour connaître ses chances d'être élu. Si p est la proportion d'électeurs qui ont l'intention de voter pour lui, cette société est conduite à effectuer le test entre les hypothèses simples suivantes :

$$\begin{cases} H_0 : p = 0{,}48 \\ H_1 : p = 0{,}52 \end{cases}$$

a) Indiquer le sens que l'on peut accorder à ce choix d'hypothèse nulle et déterminer la région critique du test de risque de première espèce $\alpha = 0{,}10$ dans le cas d'un sondage réalisé auprès de $n = 100$ électeurs. On calculera ensuite la puissance pour des valeurs de n égales à 100, 500 et 1 000.

b) On considère maintenant le test suivant :

$$\begin{cases} H_0 : p = 0{,}49 \\ H_1 : p = 0{,}51 \end{cases}$$

Déterminer la taille d'échantillon minimum pour que la puissance soit supérieure à 0,90 et indiquer dans ce cas la région critique.

Analyse de l'énoncé et conseils. On utilise le théorème de Neyman et Pearson pour déterminer la forme de la région critique de ce test entre deux hypothèses simples. On introduit pour cela une variable de Bernoulli indicatrice de l'intention de vote pour le candidat D. Magog, de paramètre p. La taille de l'échantillon permet d'utiliser la loi asymptotique de la statistique qui sert à délimiter cette région critique, pour déterminer une valeur approchée du seuil en fonction du risque de première espèce fixé. Pour calculer la puissance du second test en fonction de différentes tailles d'échantillon, on exprimera ce seuil en fonction de la taille n.

Test de la moyenne d'une loi normale d'écart type connu

14. On dispose d'un échantillon $(X_1, \ldots, X_n)$ d'une v.a. X qui suit une loi normale d'espérance inconnue m et d'écart type connu $\sigma = 2$ pour choisir entre les deux hypothèses :

$$\begin{cases} H_0 : m = 2 \\ H_1 : m = 3 \end{cases}$$

Déterminer la région critique du test de Neyman et Pearson et calculer sa puissance dans le cas où $n = 100$ et $\alpha = 0{,}05$. Quelle devrait être la taille d'échantillon minimum pour que cette puissance soit supérieure à 0,95 ? 0,99 ?

Analyse de l'énoncé et conseils. On détermine la forme de la région critique de ce test par le théorème de Neyman et Pearson, en faisant le rapport des vraisemblances associées aux deux hypothèses simples. On transforme les inégalités successives obtenues en ne conservant que les termes dépendant de $x_1, \ldots, x_n$ et en remplaçant les autres par une nouvelle constante. On aboutit à la définition d'une région critique qui fait intervenir une statistique dont on connaît la loi de probabilité sous l'hypothèse nulle, ce qui permet de déterminer le seuil de la région critique en fonction du risque de première espèce fixé. On peut ensuite calculer la puissance du test comme probabilité de cette région critique, cette fois en se plaçant dans l'hypothèse alternative.

Test de la moyenne d'une loi normale d'écart type inconnu

15. On dispose d'un échantillon $(X_1, \ldots, X_n)$ d'une v.a. X qui suit une loi normale d'espérance inconnue m et d'écart type inconnu σ, pour choisir entre les deux hypothèses :

$$\begin{cases} H_0 : m = 1 \\ H_1 : m = 2 \end{cases}$$

Déterminer la région critique du test de Neyman et Pearson et calculer sa puissance dans le cas où $n = 25$ et $\alpha = 0{,}05$.

Analyse de l'énoncé et conseils. La démarche est identique à celle de l'exercice précédent, mais la région critique est déterminée à partir d'une statistique dont la loi n'est pas connue sous l'hypothèse nulle. Il faudra donc modifier cette statistique par l'utilisation d'un estimateur de l'écart type, qui est inconnu ici, et définir la région critique à l'aide d'une statistique de loi connue sous H_0.

Test de l'écart type d'une loi normale de moyenne connue

16. Un fabricant de thermomètres prélève un échantillon de taille $n = 76$ dans un lot qu'il vient de produire. La valeur indiquée par ces thermomètres plongés dans un liquide maintenu à la température constante de 16 degrés est considérée comme une v.a. X de loi normale dont l'écart type σ est inconnu. Il désire effectuer le test entres les deux hypothèses suivantes :

$$\begin{cases} H_0 : \sigma = 1 \\ H_1 : \sigma = 2 \end{cases}$$

Pour un risque de première espèce $\alpha = 0{,}05$, indiquer ce que sera sa conclusion pour des observations de températures telles que $\sum_{i=1}^{76} x_i = 1\,140$ et $\sum_{i=1}^{76} x_i^2 = 17\,195$. Que se produit-il si on intervertit les deux hypothèses ?

Analyse de l'énoncé et conseils. Le rapport des vraisemblances fournit la forme de la région critique de ce test. Cette région est définie à l'aide d'une statistique dont la loi est connue sous l'hypothèse nulle, ce qui permet de déterminer le seuil critique en fonction du risque de première espèce α.

Paramètre d'une loi uniforme

17. Un générateur de nombres au hasard est censé simuler une loi uniforme sur $[0,1]$. S'il est déréglé, il fournira une simulation d'une loi uniforme sur $[0,\theta]$ avec $\theta = 1{,}1$. S'il fournit les valeurs 0,95 ; 0,24 ; 0,83 ; 0,52 et 0,68 doit-on considérer qu'il est bien réglé ?

Analyse de l'énoncé et conseils. Le problème posé se ramène à un test sur la valeur du paramètre d'une loi uniforme sur $[0,\theta]$, où les hypothèses confrontées sont $\theta = 1$ et $\theta = 1{,}1$. Le rapport des vraisemblances n'étant pas défini partout, on construira la région critique sans référence au théorème de Neyman et Pearson en utilisant un estimateur du paramètre θ.

18. La v.a. X suivant une loi uniforme sur $[\theta, 2\theta]$, donner la région critique et la puissance du test de risque de première espèce α entre les hypothèses suivantes :

$$\begin{cases} H_0 : \theta = 1 \\ H_1 : \theta = 0{,}9 \end{cases}$$

Peut-on en déduire la région critique du test suivant :

$$\begin{cases} H_0 : \theta = 1 \\ H_1 : \theta < 1 \end{cases}$$

Analyse de l'énoncé et conseils. Il n'est pas question de déterminer ici un estimateur efficace puisque l'ensemble des valeurs possibles pour la variable dépend du paramètre à estimer. Cependant, entre les différents estimateurs sans biais possibles pour θ, obtenus par la méthode des moments ou à partir des valeurs extrêmes $m_n = \min\{X_1, \ldots, X_n\}$ et $M_n = \max\{X_1, \ldots, X_n\}$ de l'échantillon $(X_1, \ldots, X_n)$ de X, on fait le choix de celui qui est le plus efficace. On admettra qu'il s'agit de l'estimateur $T_n = (n+1)\,M_n/\,(2n+1)$. On construira donc la région critique à partir de cet estimateur dont la loi est connue sous H_0.

Hypothèse multiple pour une proportion

19. Sur un échantillon de 900 naissances, on constate qu'il y a 470 garçons. Un généticien décide d'utiliser ces données pour effectuer le test suivant relatif aux proportions p et $1-p$ de naissances respectivement masculines et féminines :

$$\begin{cases} H_0 : p = 0{,}50 \\ H_1 : p = 0{,}48 \end{cases}$$

Quelle est la conclusion, au vu de cet échantillon, et pourquoi est-ce peu satisfaisant ?

Quelle est la conclusion retenue si on effectue le test suivant :

$$\begin{cases} H_0 : p = 0{,}5 \\ H_1 : p \neq 0{,}5 \end{cases}$$

Analyse de l'énoncé et conseils. On procède comme dans l'exercice 4, page 127, en utilisant la loi limite pour déterminer la région critique, ce qui se justifie par la taille élevée de l'échantillon. Le calcul de la puissance permet d'apprécier la qualité de ce test. On retient ensuite comme alternative $p = p_1$ en étudiant successivement les cas $p_1 > 0{,}5$ et $p_1 < 0{,}5$ pour en déduire la région critique du test d'alternative $p \neq 0{,}5$.

Hypothèses multiples pour la moyenne d'une loi normale

20. Un fabricant de conserves de petis pois produit des boîtes où l'étiquette annonce un poids net égoutté de 560 g. Il prélève un lot de 200 boîtes pour s'assurer qu'il n'aura pas d'ennuis à l'issue d'un contrôle éventuel. Formaliser le problème de test et indiquer s'il existe un test UPP. Pour un risque de première espèce retenu $\alpha = 0{,}01$ quelle sera sa conclusion pour des observations $(x_1, \dots, x_n)$ telles que $\sum_{i=1}^{200} x_i = 111\,140$ et $\sum_{i=1}^{200} (x_i - \bar{x})^2 = 17\,959{,}75$? Peut-on déterminer la fonction puissance de ce test si le procédé de fabrication est tel que sa précision est caractérisée par un écart type connu $\sigma = 10$?

Analyse de l'énoncé et conseils. On pourra considérer le poids net égoutté comme une v.a. de loi normale dont les paramètres m et σ sont inconnus, l'étiquette annonçant que $m = 560$ qui sera donc l'hypothèse privilégiée par le fabricant. Il aura des ennuis seulement si le poids est inférieur à cette valeur annoncée, donc l'hypothèse alternative est $m < 560$. On résout d'abord le test entre deux hypothèses simples $m = m_0$ contre $m = m_1$ avec $m_0 = 560$ et m_1 qui est une valeur quelconque, inférieure à 560. Si la région critique obtenue par la méthode de Neyman et Pearson est indépendante de la valeur m_1, ce sera celle du test le plus puissant contre l'alternative $m < 560$. Sans hypothèse supplémentaire, on ne peut pas déterminer la fonction puissance puisque le seuil critique dépend d'une statistique. Pour cela nous allons tester l'autre hypothèse $\sigma = 10$ et, si elle est acceptée, on calculera la fonction puissance en faisant comme si l'écart type était connu.

21. Un industriel produit des pièces dont la caractéristique principale peut être considérée comme une v.a. X de loi normale, de moyenne $m = 1$ et d'écart type $\sigma = 1$. Il prélève au hasard 25 pièces qui ont été produites, pour tester si le réglage de la machine est toujours à $m = 1$. Sachant que la précision est toujours caractérisée par $\sigma = 1$, indiquer les trois tests que l'on peut envisager pour contrôler un dérèglement éventuel et représenter leur fonction puissance pour un risque de première espèce $\alpha = 0{,}05$. Préciser à chaque fois la décision qui sera prise pour un échantillon où on a observé $\sum_{i=1}^{25} x_i = 30{,}25$.

Analyse de l'énoncé et conseils. Pour savoir si la machine est déréglée, on effectue un test dont l'hypothèse nulle est $m = 1$, car c'est bien entendu l'hypothèse qui est privilégiée. L'hypothèse alternative sera une hypothèse multiple, différente selon que l'on a ou non des informations supplémentaires sur la nature du déréglement. Si l'on pense *a priori* que la valeur de réglage a été augmentée, l'alternative sera $m > 1$; elle sera $m < 1$ si l'on pense qu'elle a diminué et on retiendra $m \neq 1$ sans information *a priori*. Pour déterminer la région critique des deux premiers tests, on se place d'abord dans le cas d'une hypothèse alternative simple, conforme à l'hypothèse multiple (voir exercice précédent) ; pour le troisième test on retiendra l'union des régions critiques des deux précédents.

Comparaison de proportions

22. Un sondage effectué auprès de 2 000 personnes indique que 19 % d'entre elles connaissent la marque de lessive Omopaic. Après une campagne publicitaire, un sondage analogue auprès de 1 000 personnes montre que 230 d'entre elles connaissent cette marque. Peut-on considérer que cette campagne a été efficace ?

Analyse de l'énoncé et conseils. Il faut associer deux variables de Bernoulli, indicatrices de notoriété de cette marque de lessive, avant et après la campagne et faire un test de comparaison entre les paramètres de ces deux lois. L'hypothèse nulle est l'égalité de ces paramètres, l'alternative correspond à une valeur plus grande après la campagne. On utilisera la loi approchée de l'estimateur de la différence de ces deux paramètres pour construire la région critique.

Comparaison d'échantillons

23. Pour comparer deux marques d'ampoules électriques, on a effectué des essais sur des échantillons prélevés au hasard. La durée de vie moyenne observée sur un échantillon de 25 ampoules de la marque 1 est de 1 012 heures, avec un écart type empirique de 241 h. Sur un échantillon de 13 ampoules de la marque 2, on a obtenu respectivement 1 104 h et 319 h. Peut-on considérer que ces marques sont de qualités équivalentes ?

Analyse de l'énoncé et conseils. Nous ferons l'hypothèse que les durées de vie des ampoules des marques 1 et 2 sont des v.a. X et Y de lois normales et l'hypothèse nulle correspond à l'égalité des moyennes de ces deux lois. Les tailles d'échantillon sont trop faibles pour pouvoir faire des approximations de loi ; donc, pour effectuer ce test, il faut au préalable tester l'égalité des variances de ces deux lois. Si cette hypothèse est acceptée, on pourra réaliser le test de l'égalité des moyennes de deux lois normales de même écart type, car la loi de la statistique de test dans ce cas sera connue.

Test d'indépendance

24. Le chef d'établissement d'un grand lycée parisien désire comparer les taux de succès au baccalauréat des trois sections générales. Les statistiques de l'année précédente figurent dans le tableau suivant :

	L	ES	S	Total
Réussite	41	59	54	154
Échec	21	36	75	132
Total	62	95	129	286

Au vu de ce tableau, peut-on considérer que les taux de réusssite sont indépendants de la section du baccalauréat ?

Analyse de l'énoncé et conseils. Pour répondre à cette question, il faut utiliser le test d'indépendance du khi-deux. On remplit donc le tableau qui serait obtenu dans l'hypothèse nulle à tester d'indépendance des deux caractéristiques, réussite au baccalauréat et section choisie, puis on calcule la distance du khi-deux entre ces deux tableaux.

Tests d'adéquation

25. Pendant 200 minutes, on note toutes les minutes le nombre de voitures qui ont franchi un poste de péage sur l'autoroute, ce qui donne la distribution observée ci-après :

Nombre de voitures	0	1	2	3	4	5	6	7	8	9	10	11
Effectif observé	1	15	30	46	38	30	16	13	5	3	2	1

Déterminer la loi théorique que l'on pourrait retenir pour modéliser ce phénomène d'arrivées par minute à un péage.

Analyse de l'énoncé et conseils. Parmi les lois usuelles, c'est la loi de Poisson qui paraît convenir pour modéliser ces arrivées au poste de péage. Il faut estimer le paramètre de cette loi et ensuite tester son adéquation à la loi observée, en étant attentif au nombre de degrés de liberté de la loi utilisée et aux effectifs de classes.

26. À la sortie d'une chaîne de fabrication, on contrôle toutes les trente minutes le nombre de pièces défectueuses contenues dans un lot de vingt pièces produites.

Les résultats observés sur 200 lots indépendants sont résumés dans le tableau suivant :

Nombre de défectueux	0	1	2	3	4	5	6	7
Nombre de lots	26	52	56	40	20	2	0	4

Quelle loi théorique du nombre de défectueux par lot de vingt peut-on retenir au vu de ces observations ?

Analyse de l'énoncé et conseils. Il s'agit de tester l'adéquation de cette distribution empirique à une distribution théorique qui n'est pas donnée ici. Cependant, le nombre de défectueux dans un lot de 20 pièces produites suit une loi binômiale de paramètres $n = 20$ et p qui représente la probabilité qu'une pièce soit défectueuse (c'est-à-dire la proportion de pièces défectueuses fabriquées par cette chaîne). Comme l'un des paramètres de cette loi n'est pas fixé, il faudra l'estimer et en tenir compte pour le nombre de degrés de liberté de la loi asymptotique de la statistique utilisée dans ce test.

SOLUTIONS

1 ▸ Vrai. Si l'hypothèse nulle est privilégiée dans la méthode de Bayes par une valeur forte de p_0, elle doit donc être rejetée avec un risque très faible d'erreur qui est justement le risque de première espèce α dans la méthode de Neyman et Pearson.

2 ▸ Faux. Le principe de la méthode de Neyman et Pearson consiste au contraire à faire jouer des rôles dissymétriques aux deux hypothèses, l'hypothèse nulle étant *a priori* celle qui paraît la plus vraisemblable. Il faut donc des contre-preuves fortes pour la rejeter.

3 ▸ Vrai. L'erreur la plus grave consistant à rejeter à tort l'hypothèse nulle avec une probabilité α, il serait contraire à l'esprit de la méthode de Neyman et Pearson que l'autre erreur ait une probabilité β plus faible.

4 ▸ Faux. Dans le cas de deux hypothèses simples et sans information *a priori* l'affirmation est vraie. Mais si une seule des hypothèses est simple, c'est celle que l'on doit retenir comme hypothèse nulle pour que l'on puisse construire la région critique de risque de première espèce fixé à une valeur donnée α.

5 ▸ Vrai. Si on retient comme définition d'une hypothèse simple celle qui est fournie dans les rappels de cours, la loi de X est totalement spécifiée sous l'alternative H_1 et le calcul de la puissance est donc possible. Par contre, si on retient la définition plus généralement admise où une hypothèse simple correspond au cas où le paramètre θ ne peut prendre qu'une seule valeur, cette affirmation peut devenir erronée. Si nous prenons l'exemple d'une v.a. X qui suit une loi normale d'écart type inconnu et que le test porte sur la moyenne, soit $H_0 : m = m_0$ contre $m = m_1$, la loi sous H_1 n'est pas totalement déterminée car elle dépend de deux paramètres, et il n'est donc pas possible de calculer la puissance.

6 ▸ Faux. Dans le cas d'une loi discrète, les valeurs du risque de première espèce ne varient pas de façon continue et ne comporteront pas en général la valeur exacte α.

7 ▸ Vrai. La région critique étant construite à partir de l'hypothèse nulle, l'inversion des hypothèses peut conduire à des conclusions inverses, pour le même échantillon, si on intervertit les deux hypothèses.

8 ▸ L'expression de la vraisemblance est :

$$L\left(x_1, \ldots, x_n ; \lambda\right) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{\lambda^{\sum_1^n x_i}}{\prod_{i=1}^{n} x_i!}$$

Le rapport des vraisemblances, L_0 sous H_0, et L_1 sous H_1, s'écrit donc :

$$\frac{L_0}{L_1} = e^{-2n} 2^{\sum_1^n x_i}$$

Le théorème de Neyman et Pearson, qui permet d'obtenir le test de puissance maximum, définit la région critique par la condition $\frac{L_0}{L_1} \leq k$, qui est donc ici équivalente à :

$$\sum_{i=1}^{n} x_i \leq C$$

La valeur du seuil C qui définit cette région critique W est déterminée par la condition :

$$\alpha = P\left(W|H_0\right) = P\left[\sum_{i=1}^{n} X_i \leq C|\lambda = 2\right]$$

avec $\sum_{i=1}^{n} X_i$ qui suit une loi de Poisson de paramètre $2n$. Pour $n = 9$, on lit dans les tables de la loi de Poisson $\mathcal{P}(18)$ les probabilités $P\left(\sum_{i=1}^{9} X_i \leq 10\right) = 0{,}030\,4$ et $P\left(\sum_{i=1}^{9} X_i \leq 11\right) = 0{,}054\,9$. Comme la loi est discrète, on constate donc que l'on ne peut pas trouver un test de niveau exact $\alpha = 0{,}05$. Si on retient comme région critique l'ensemble des points $(x_1, \ldots, x_9)$ tels que :

$$\sum_{i=1}^{9} x_i \leq 10$$

on obtiendra un test de risque de première espèce inférieur à 5 %. Si on tient absolument à avoir un test de risque exact $\alpha = 0{,}05$, il faut construire ce qu'on appelle un *test mixte*. Pour cela on procède de la façon suivante. Pour des observations telles que $\sum_{i=1}^{9} x_i \leq 10$, on rejette l'hypothèse nulle. Dans le cas où $\sum_{i=1}^{9} x_i = 11$, on effectue un tirage au sort où on rejette l'hypothèse H_0 avec une probabilité γ (on l'accepte donc avec une probabilité $1-\gamma$). Le test sera bien de risque exact $\alpha = 0{,}05$ si on choisit γ tel que :

$$\alpha = 0{,}05 = P\left[\sum_{i=1}^{9} X_i \leq 10 | \lambda = 2\right] + \gamma P\left[\sum_{i=1}^{9} X_i = 11 | \lambda = 2\right] = 0{,}030\,4 + 0{,}024\,5\gamma$$

soit $\gamma = \dfrac{196}{245} = 0{,}8$.

9 ▸ L'expression de la statistique de test est :

$$D_n = n\left[\sum_{i=1}^{k} \frac{(n_i/n)^2}{p_i} - 1\right]$$

Pour un échantillon de taille N où la distribution empirique est caractérisée par les mêmes valeurs n_i/n, la statistique associée aura pour valeur :

$$D_N = N\left[\sum_{i=1}^{k} \frac{(n_i/n)^2}{p_i} - 1\right]$$

Dans le cas où $N = 10n$, on aura donc $D_N = 10D_n$, soit pour les valeurs données $D_{1\,000} = 21$ qui est supérieur au seuil 16,9, ce qui conduit à rejeter l'adéquation à la loi uniforme. Pour une taille d'échantillon plus élevée, la valeur de la statistique augmente et la distribution empirique obtenue doit donc être plus proche de la distribution théorique pour qu'on accepte l'adéquation.

10 ▸ Le coût de la décision D_0 de prendre la garantie est de 300 F et le coût moyen associé à la décision D_1 de ne pas la prendre est de $2\,500p$, car p qui représente la probabilité de panne est aussi la valeur moyenne de la variable indicatrice de panne définie par :

$$X = \begin{cases} 1 & \text{si panne} \quad (p) \\ 0 & \text{sinon} \quad (1-p) \end{cases}$$

Les renseignements sur la provenance du téléviseur se traduisent par les probabilités *a priori* 0,8 et 0,2 affectées aux deux valeurs possibles de p qui sont $p = 0{,}15$ et $p = 0{,}05$. Tout ceci est résumé dans le tableau de coûts suivant :

	D_0	D_1	prob. *a priori*
$H_0 : p = 0{,}15$	300	375	0,8
$H_1 : p = 0{,}05$	300	125	0,2

La vraisemblance de l'échantillon $(X_1, \ldots, X_n)$ de cette loi de Bernoulli s'écrit :

$$L\left(x_1, \ldots, x_n\ ; p\right) = p^{\sum_{i=1}^N x_i} \left(1 - p\right)^{N - \sum_{i=1}^N x_i}$$

avec $x_i \in \{0, 1\}$, $1 \leq i \leq n$. On note L_0 la vraisemblance pour $p = 0{,}15$ et L_1 pour $p = 0{,}05$. Les espérances *a posteriori* des coûts sont :

$$E\left[C\left(D_0\right)\right] = 300\pi_0 + 300\pi_1 \quad \text{et} \quad E\left[C\left(D_1\right)\right] = 375\pi_0 + 125\pi_1$$

La règle de Bayes conduit à prendre la garantie si $E\left[C\left(D_0\right)\right] < E\left[C\left(D_1\right)\right]$ soit :

$$175\pi_1 < 75\pi_0 \Longleftrightarrow 7 \times 0{,}2 \times L_1 < 3 \times 0{,}8 \times L_0 \Longleftrightarrow 7L_1 < 12L_0$$

$$\Longleftrightarrow \left(\ln 3 + \ln \frac{19}{17}\right) \sum_{i=1}^{20} x_i > 20 \ln \frac{19}{17} - \ln \frac{12}{7} \Longleftrightarrow \sum_{i=1}^{20} x_i > 1{,}39$$

Les x_i ne prenant que des valeurs entières, on décide de prendre la garantie pour des observations telles que $\sum_{i=1}^{20} x_i \geq 2$, ce qui est le cas pour cet échantillon. Pour un coût de garantie de 350 F, la relation $E\left[C\left(D_0\right)\right] < E\left[C\left(D_1\right)\right]$ est vraie pour $225\pi_1 < 25\pi_0$, soit $9 \times 0{,}2 \times L_1 < 0{,}8 \times L_0$, ce qui est équivalent à :

$$\left(\ln 3 + \ln \frac{19}{17}\right) \sum_{i=1}^{20} x_i > 20 \ln \frac{19}{17} + \ln \frac{9}{4} \Longleftrightarrow \sum_{i=1}^{20} x_i > 2{,}51 \Longleftrightarrow \sum_{i=1}^{20} x_i \geq 3$$

et cette fois on ne prend donc pas la garantie.

11 ▸ Les données fournies conduisent à affecter les probabilités *a priori* 0,75 et 0,25 aux deux valeurs possibles $\lambda = 1$ et $\lambda = 3$ du paramètre de cette loi de Poisson. Le nombre annuel moyen de pannes pour cette loi est $E(X) = \lambda$. Si on prend la décision D_0 de ne pas souscrire de contrat d'entretien, cela correspond aux coûts moyens pour chaque hypothèse de 1 000 F et $3 \times 1\,000 = 3\,000$ F. Tout cela est résumé dans le tableau de coûts suivant :

	D_0	D_1	prob. *a priori*
$H_0 : \lambda = 1$	1 000	2 400	0,75
$H_1 : \lambda = 3$	3 000	2 400	0,25

Les espérances *a posteriori* des coûts sont :

$$E\left[C\left(D_0\right)\right] = 1\,000\pi_0 + 3\,000\pi_1 \quad \text{et} \quad E\left[C\left(D_1\right)\right] = 2\,400\pi_0 + 2\,400\pi_1$$

La règle de Bayes conduit à ne pas souscrire le contrat d'entretien si $E\left[C\left(D_0\right)\right] < E\left[C\left(D_1\right)\right]$ soit :

$$600\pi_1 < 1\,400\pi_0 \Longleftrightarrow 3 \times 0{,}25 \times L_1 < 7 \times 0{,}75 \times L_0 \Longleftrightarrow L_1 < 7L_0$$

La vraisemblance de l'échantillon de la loi de Poisson s'écrit :

$$L\left(x_1, \ldots, x_5\ ; \lambda\right) = \mathrm{e}^{-5\lambda} \lambda^{\sum_{i=1}^5 x_i} \Big/ \prod_{i=1}^5 x_i!$$

L'inégalité précédente est donc équivalente, en prenant les logarithmes, à :

$$\left[\sum_{i=1}^{5} x_i\right] \ln 3 < 10 + \ln 7 \Longleftrightarrow \sum_{i=1}^{5} x_i < 10{,}87 \Longleftrightarrow \sum_{i=1}^{5} x_i \leq 10$$

puisque les x_i ne prennent que des valeurs entières. L'observation $\sum_{i=1}^{5} x_i = 9$ conduit donc à ne pas souscrire le contrat d'entretien.

12 ▸ Dans l'optique de Bayes, les hypothèses $m = 0$ et $m = 2$ ont la même probabilité *a priori* 1/2. La décision D_0 de ne pas lancer la campagne a un coût nul. La décision D_1 d'entreprendre cette campagne a un coût de 0,5 diminué d'un bénéfice moyen de 2, dans le cas où la campagne a été efficace. Ce qui conduit au tableau des coûts suivant :

	D_0	D_1	prob. *a priori*
$H_0 : m = 0$	0	0,5	0,5
$H_1 : m = 2$	0	−1,5	0,5

Les espérances *a posteriori* des coûts sont $E\left[C\left(D_0\right)\right] = 0$ et $E\left[C\left(D_1\right)\right] = 0{,}5\pi_0 - 1{,}5\pi_1$; donc on aura $E\left[C\left(D_0\right)\right] < E\left[C\left(D_1\right)\right]$ si $\pi_0 > 3\pi_1$, soit $L_0 > 3L_1$. L'accroissement de bénéfices par ville est une v.a. X de loi $N(m, 2)$ soit une vraisemblance :

$$L\left(x_1, \ldots, x_n\,;\, m\right) = \frac{1}{2\sqrt{2\pi}} \exp\left\{-\frac{1}{8}\sum_{i=1}^{5}\left(x_i - m\right)^2\right\}$$

En prenant les logarithmes, la dernière inégalité est équivalente à :

$$-\frac{1}{8}\sum_{i=1}^{5}\left[x_i^2 - \left(x_i - 2\right)^2\right] > \ln 3 \Longleftrightarrow \sum_{i=1}^{5} x_i < 5 - 2\ln 3 = 2{,}80$$

L'observation $\sum_{i=1}^{5} x_i = 4$ conduit donc à lancer la campagne publicitaire si on adopte cette règle de décision de Bayes.

Dans l'optique de Neyman-Pearson, l'hypothèse nulle à ne pas rejeter à tort est celle d'une campagne qui serait efficace, ce qui correspond à l'hypothèse simple $H_0 : m = 2$. L'hypothèse alternative est alors $H_1 : m = 0$. Le théorème de Neyman et Pearson définit la forme de la région critique par $L_0/L_1 \leq k$, soit ici en prenant les logarithmes :

$$-\frac{1}{8}\sum_{i=1}^{5}\left[\left(x_i - 2\right)^2 - x_i^2\right] < k_1 \Longleftrightarrow \sum_{i=1}^{5} x_i < C$$

la valeur de la constante C étant déterminée par le risque de première espèce $\alpha = P\left(\sum_{i=1}^{5} X_i < C | m = 2\right)$. Sous l'hypothèse nulle, la moyenne empirique $\bar{X}_5$ suit une loi normale $N\left(2, 2/\sqrt{5}\right)$; donc C est défini par :

$$\alpha = P\left[\frac{\bar{X}_5 - 2}{2/\sqrt{5}} < \frac{C/5 - 2}{2/\sqrt{5}}\right]$$

soit $C = 10 + 2u\sqrt{5}$ où u est le fractile d'ordre α de la loi normale centrée-réduite. Pour un risque $\alpha = 0{,}05$, on a $u = -1{,}6449$ et $C = 2{,}64$; donc la règle de Neyman conduit à accepter l'hypothèse nulle, c'est-à-dire à lancer la campagne. Par contre, pour un risque de première espèce plus élevé $\alpha = 0{,}10$, on lit dans la table 2, page 195, le fractile $u = -1{,}2816$, d'où $C = 4{,}27$. L'observation $\sum_{i=1}^{5} x_i = 4$ conduit cette fois à ne pas lancer la campagne.

Le changement de $m = 2$ en $m > 0$ oblige à retenir l'hypothèse simple $m = 0$ comme hypothèse nulle du test de Neyman-Pearson. La région critique $L_0/L_1 \leq k$ est cette fois définie par :

$$-\frac{1}{8}\sum_{i=1}^{5}\left[x_i^2 - (x_i - m)^2\right] < k_1 \Longleftrightarrow \sum_{i=1}^{5} x_i > C$$

où la constante C est déterminée par la condition $\alpha = P\left(\sum_{i=1}^{5} X_i > C | m = 0\right)$. On obtient cette fois $C = 2u\sqrt{5}$, où u est le fractile d'ordre $1 - \alpha$ de la loi normale centrée-réduite. Pour un risque $\alpha = 0{,}10$, le fractile est $u = 1{,}2816$ et $C = 5{,}73$; donc même avec ce risque de première espèce élevé on décide de ne pas lancer la campagne. Le changement d'hypothèse nulle, dû à des raisons techniques, amène à ne la rejeter que pour des très grandes valeurs observées d'augmentation de bénéfices, même si on choisit un risque α élevé.

13 ▸ a) Choisir $p = 0{,}48$ comme hypothèse nulle, au lieu de $p = 0{,}52$, c'est vouloir se prémunir avant tout contre le risque de conclure à son élection, en acceptant l'hypothèse $p = 52\,\%$ alors qu'en fait on aurait $p = 48\,\%$. Ce candidat préfère être agréablement surpris, plutôt que d'entretenir un faux espoir. On introduit ici des variables de Bernoulli associées à chaque électeur interrogé i, $1 \leq i \leq n$, qui prennent la valeur 1 quand celui-ci déclare son intention de vote pour le candidat D. Magog :

$$X_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

La vraisemblance de l'échantillon $(X_1, \ldots, X_n)$ de cette loi de Bernoulli s'écrit :

$$L\left(x_1, \ldots, x_n\,;p\right) = \prod_{i=1}^{n} p^{x_i}\left(1-p\right)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}\left(1-p\right)^{n-\sum_{i=1}^{n} x_i}$$

avec $x_i \in \{0, 1\}$, $1 \leq i \leq n$. On note L_0 la vraisemblance pour $p = 0{,}48 = p_0$ et L_1 pour $p = 0{,}52 = p_1$. Le rapport des vraisemblances s'écrit :

$$\frac{L_0}{L_1} = \left(\frac{p_0}{p_1}\right)^{\sum_{i=1}^{n} x_i}\left(\frac{1-p_0}{1-p_1}\right)^{n-\sum_{i=1}^{n} x_i}$$

La région critique est définie par le théorème de Neyman et Pearson, après avoir pris les logarithmes, par l'inégalité :

$$\left[\sum_{i=1}^{n} x_i\right]\ln\frac{p_0}{p_1} + \left[n - \sum_{i=1}^{n} x_i\right]\ln\frac{1-p_0}{1-p_1} \leq \ln k$$

qui est équivalente à :

$$\left(\ln\frac{p_0}{p_1}+\ln\frac{1-p_1}{1-p_0}\right)\sum_{i=1}^{n}x_i\leq k_1$$

Comme $p_0 < p_1$ et $1-p_1 < 1-p_0$, on obtient la condition $\sum_{i=1}^{n} x_i \geq k_2$. On peut définir également la région critique à l'aide de la moyenne empirique, qui est un estimateur sans biais du paramètre à tester p et dont on pourra déterminer la loi asymptotique. C'est l'ensemble des points $(x_1,\ldots,x_n)$ tels que :

$$\bar{x}_n \geq C$$

La valeur de la constante C, qui est le seuil critique de ce test, est définie à partir du risque de première espèce, soit :

$$\alpha = P\left(\bar{X}_n \geq C|p=0{,}48\right)$$

Si on utilise la loi asymptotique, déduite du théorème central limite, la v.a. $\bar{X}_n$ suit approximativement une loi normale d'espérance p_0 et de variance p_0q_0/n avec $q_0 = 1-p_0$. On obtient donc une valeur approchée de C par la condition :

$$\alpha = P\left[U \geq \frac{C-p_0}{\sqrt{p_0q_0/n}}\right]$$

où U est une v.a. de loi $N(0,1)$. Ainsi, u désignant le fractile d'ordre $1-\alpha$ de cette loi, la région critique est définie par la condition :

$$\bar{x}_n \geq p_0 + u\sqrt{\frac{p_0q_0}{n}}$$

Pour $\alpha = 0{,}10$ on lit dans la table 2, page 195, le fractile $u = 1{,}281\,6$ et le seuil critique pour $n = 100$ est alors $C = 0{,}48 + 1{,}281\,6/20 = 0{,}544$. En dépit d'un risque de rejet de l'hypothèse nulle assez élevé puisque $\alpha = 10\,\%$, on ne rejette l'hypothèse $p = 48\,\%$ que si plus de 54,4 % des personnes interrogées déclarent vouloir voter pour ce candidat, valeur bien supérieure à $p_1 = 52\,\%$. Ceci s'explique par une taille d'échantillon trop faible pour séparer deux hypothèses voisines de 50 %. La puissance de ce test se calcule par :

$$\eta = P\left(\bar{X}_n \geq C|p=p_1\right) \simeq P\left[U \geq \frac{C-p_1}{\sqrt{p_1q_1/n}}\right]$$

Les valeurs de p_0 et p_1 étant proches de $1/2$, on peut remplacer C par $p_0 + u/2\sqrt{n}$ et on calcule alors une valeur approchée de la puissance par la probabilité :

$$\eta = P\left\{U \geq u + \left(p_0 - p_1\right)2\sqrt{n}\right\}$$

ou $\eta = 1 - \Phi\left(u - 0{,}08\sqrt{n}\right)$, Φ désignant la f.r. de la loi $N(0,1)$. Pour $n = 100$, on obtient $\eta = 0{,}544$, soit un risque de seconde espèce élevé $\beta = 0{,}456$. Le risque de première espèce étant fixé, la puissance augmentera si la région critique s'agrandit, c'est-à-dire si le seuil C diminue, ce qui se produira en augmentant la taille de l'échantillon. Le tableau suivant fournit les différents seuils et puissances pour les tailles d'échantillon proposées :

n	100	500	1 000
C	0,544	0,509	0,500 3
η	0,544	0,695	0,894

b) La région critique, de ce test est définie par :

$$\bar{x}_n \geq 0{,}49 + u/2\sqrt{n}$$

Sa puissance est calculée approximativement par $\eta = 1 - \Phi\left(u - 0{,}04\sqrt{n}\right)$. Pour que cette puissance soit supérieure à 0,90, il faut que la taille d'échantillon vérifie $1{,}281\,6 - 0{,}04\sqrt{n} < -1{,}281\,6$, soit $\sqrt{n} > 64{,}08$ ou $n \geq 4\,107$. Pour cette valeur élevée de n, on obtient l'égalité des risques $\alpha = \beta = 0{,}10$ et un seuil critique $C = 0{,}50$ à équidistance des deux hypothèses. Il faut donc environ quatre fois plus d'observations que pour le test précédent afin de séparer ces deux hypothèses, encore plus proches, avec les mêmes risques.

14 ▸ La vraisemblance s'écrit :

$$L\left(x_1, \ldots, x_n\ ;\ m\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - m\right)^2\right\}$$

La forme de la région critique, donnée par le théorème de Neyman et Pearson, est $L_0/L_1 \leq k$, ce qui en passant aux logarithmes conduit à l'inégalité :

$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[\left(x_i - 2\right)^2 - \left(x_i - 3\right)^2\right] \leq \ln k$$

Elle est équivalente, après développement du crochet, à :

$$-2\sum_{i=1}^{n} x_i + 5n \leq 2\sigma^2 \ln k$$

Pour l'instant, on cherche seulement à déterminer la forme de la région critique et l'expression des constantes qui apparaissent dans cette inégalité est sans intérêt. On remplace donc cette inégalité par $\sum_{i=1}^{n} x_i \geq k_1$, qui lui est équivalente. La région critique est donc définie à partir de la valeur de la somme des observations. Cependant, comme c'est la loi de la moyenne empirique $\bar{X}_n$ qui est connue, il est préférable de faire intervenir cette moyenne dans la définition de la région critique W, qui sera donc l'ensemble des points $\left(x_1, \ldots, x_n\right)$ tels que :

$$\bar{x}_n \geq C$$

Il reste à préciser la valeur de cette constante C, qui sera le seuil de cette région, à partir du risque de première espèce α qui est fixé. Il représente la probabilité de cette région, lorsque l'hypothèse nulle est réalisée, et par conséquent, la condition qui définit C s'écrit :

$$\alpha = P\left(W|m=2\right) = P\left(\bar{X}_n \geq C|m=2\right)$$

Sous H_0, la moyenne empirique $\bar{X}_n$ suit la loi $N\left(2, 2/\sqrt{n}\right)$, donc en centrant et réduisant on obtient la condition :

$$\alpha = P\left(U \geq \frac{C-2}{2/\sqrt{n}}\right)$$

où U est une v.a. de loi $N\left(0,1\right)$. Ainsi la constante C est définie par :

$$C = 2 + 2\frac{u}{\sqrt{n}}$$

où u est le fractile d'ordre $1-\alpha$ de la loi $N\left(0,1\right)$. Pour un risque $\alpha = 0{,}05$, on lit dans la table 2, page 195, le fractile $u = 1{,}644\,9$, d'où une région critique définie pour $n = 100$ par :

$$W = \left\{\left(x_1, \ldots, x_{100}\right)/\bar{x}_{100} \geq 2{,}33\right\}$$

Dans ce cas, la puissance du test est la probabilité de cette région dans l'hypothèse alternative, soit :

$$\eta = P\left(\bar{X}_{100} \geq 2{,}33|m=3\right) = P\left(U \geq \frac{2{,}33-3}{2/\sqrt{100}}\right) = 0{,}996$$

On obtient ainsi un risque de seconde espèce de valeur $\beta = 0{,}004$ très inférieure à celle du risque de première espèce, ce qui est contraire à l'optique de Neyman et Pearson qui a permis de bâtir ce test. Si on veut que, plus logiquement, $\beta \geq \alpha$, il faut choisir une valeur plus faible pour α. Si on retient par exemple $\alpha = 0{,}001$, le fractile d'ordre 0,999 est $u = 3{,}090\,2$, soit $C = 2{,}62$. La puissance est alors $\eta = 0{,}971\,9$, donc $\beta = 0{,}028\,1$ est bien cette fois supérieur à α, ce qui est plus cohérent.

Pour un risque $\alpha = 0{,}05$, la condition $\eta \geq 0{,}95$ conduit à choisir une taille d'échantillon n telle que :

$$\eta = P\left(\bar{X}_n \geq 2 + 2\frac{u}{\sqrt{n}}|m=3\right) = P\left(U \geq u - \frac{\sqrt{n}}{2}\right) \geq 0{,}95$$

ce qui est équivalent $1{,}644\,9 - \sqrt{n}/2 \leq -1{,}644\,9$, soit $\sqrt{n} \geq 6{,}6$, donc une taille d'échantillon supérieure à $n_0 = 44$ pour laquelle on aura $\beta = 0{,}05 = \alpha$ avec $C = 2{,}5$ seuil équidistant des deux hypothèses. Si on choisit une taille d'échantillon supérieure à n_0, ce qui était le cas ici, le risque β diminue et devient inférieur à $\alpha = 0{,}05$, ce qui est contraire à l'optique dans laquelle a été construit ce test. La condition $\eta \geq 0{,}99$ impose de choisir n tel que cette fois $1{,}644\,9 - \sqrt{n}/2 \leq -2{,}326\,3$, soit $n \geq 64$. On avait d'ailleurs obtenu $\eta = 0{,}996$ pour $n = 100$.

15 ▸ La méthode de Neyman et Pearson conduit à la même région critique que dans l'exercice précédent. Mais cette fois il n'est pas possible de déterminer le seuil C car la loi de $\bar{X}_n$ dépend du paramètre inconnu σ. On fait donc intervenir l'estimateur sans biais $S_n^2 = \dfrac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2$ de la variance σ^2 et la variable centrée et réduite $\sqrt{n}\left(\bar{X}_n - m\right)/\sigma$ est remplacée par la statistique $\sqrt{n}\left(\bar{X}_n - m\right)/S_n$ qui suit une loi de Student à $n-1$ degrés de liberté. La région critique W est donc l'ensemble des points $(x_1, \ldots, x_n)$ tels que :

$$\sqrt{n}\,\frac{\bar{x}_n - 1}{s_n} \geq C$$

où la valeur du seuil C est déterminée par la condition :

$$\alpha = P\left\{\sqrt{n}\,\frac{\bar{X}_n - m}{S_n} \geq C | m = 1\right\}$$

c'est-à-dire que C est le fractile d'ordre $1-\alpha$ de la loi de Student à $n-1$ degrés de liberté. Pour un risque $\alpha = 0{,}05$ et $n = 25$, on lit dans la table 6 le fractile $C = 1{,}711$, d'où une région critique définie par :

$$W = \{(x_1, \ldots, x_{25}) \,/\, \bar{x}_{25} \geq 1 + 0{,}34 s_{25}\}$$

La puissance représente la probabilité de cette région dans l'hypothèse alternative, soit $\eta = P\left\{\sqrt{n}\left(\bar{X}_n - 1\right)/S_n \geq C | m = 2\right\} = P\left\{\sqrt{n}\left(\bar{X}_n - 2\right)/S_n \geq C - \sqrt{n}/S_n\right\}$. La v.a. $\sqrt{n}\left(\bar{X}_n - 2\right)/S_n$ suit une loi de Student à $n-1$ degrés de liberté, mais on ne peut pas calculer cette probabilité car la borne de l'intervalle dépend de S_n qui est une variable aléatoire.

16 ▸ Pour effectuer ce test, nous disposons d'un échantillon d'une loi normale d'espérance 16 et d'écart type $\sigma = 1$ sous H_0 et $\sigma = 2$ sous H_1. Le rapport des vraisemblances a donc pour expression :

$$\frac{L_0}{L_1} = 2^n \exp\left[\frac{1}{8}\sum_{i=1}^n (x_i - 16)^2 - \frac{1}{2}\sum_{i=1}^n (x_i - 16)^2\right] = 2^n \exp -\frac{3}{8}\sum_{i=1}^n (x_i - 16)^2$$

La région critique est donc définie par le théorème de Neyman et Pearson comme l'ensemble des points $(x_1, \ldots, x_n)$ tels que :

$$\sum_{i=1}^n (x_i - 16)^2 \geq C$$

où C est le fractile d'ordre $1-\alpha$ de la loi χ_n^2 suivie par $\sum_{i=1}^n (X_i - 16)^2$ sous l'hypothèse nulle. Pour $\alpha = 0{,}05$, on lit dans la table 5, page 199, par interpolation pour $n = 76$, la valeur $C = 97$. Les observations donnent $\sum_{i=1}^{76} (x_i - 16)^2 = 17\,195 - 32 \times 1\,140 + 76 \times 16^2 = 171$, ce qui conduit à refuser l'hypothèse nulle $\sigma = 1$. L'interversion des hypothèses change le sens de l'inégalité, la région critique étant définie par :

$$\sum_{i=1}^n (x_i - 16)^2 \leq C$$

Mais cette fois c'est la variable $\sum_{i=1}^{n}(X_i - 16)^2/4$ qui suit une loi χ_n^2 sous H_0 et c'est donc $C/4$ qui est le fractile d'ordre α de cette loi. Pour $\alpha = 0{,}05$, on déduit de la table 5, page 199, par interpolation pour $n = 76$, la valeur $C = 227$. On refuse donc à nouveau l'hypothèse nulle, qui est cette fois $\sigma = 2$. L'expression des deux régions critiques nous montre que pour toutes les observations telles que $97 < \sum_{i=1}^{76}(x_i - 16)^2 < 227$ l'interversion des hypothèses entraînera l'interversion des conclusions. Ceci s'explique par la remarque qu'il s'agit d'une région de probabilité très faible, quelle que soit l'hypothèse retenue, puisque :

$$P\left\{\sum_{i=1}^{76}(X_i - 16)^2 > 150|\sigma = 1\right\} \simeq 0$$

et :

$$P\left\{\sum_{i=1}^{76}(X_i - 16)^2 < 180|\sigma = 2\right\} \simeq 0{,}002\,6$$

17 ▸ Le paramètre θ représentant la plus grande valeur possible de la variable X de loi uniforme sur $[0, \theta]$, l'estimateur de θ basé sur un n-échantillon de cette loi sera la plus grande valeur observée, soit $M_n = \max\{X_i/1 \leq i \leq n\}$. Bien entendu, on rejettera l'hypothèse nulle $H_0 : \theta = 1$ si on observe $M_n > 1$. La forme de la région critique est donc $M_n > C$, la valeur de $C \leq 1$ étant déterminée par $\alpha = P(M_n > C|\theta = 1)$ ou $1 - \alpha = P(M_n \leq C|\theta = 1) = C^n$, soit $C = (1-\alpha)^{1/n}$. Pour $\alpha = 0{,}05$ et $n = 5$, on obtient $C = 0{,}99$; donc on accepte bien ici l'hypothèse $\theta = 1$.

18 ▸ La valeur du paramètre étant plus petite dans l'hypothèse alternative, on acceptera donc celle-ci pour les plus petites valeurs de l'estimateur, d'où une région critique de la forme $T_n < C$, la valeur de C étant déterminée par :

$$\alpha = P(T_n < C|\theta = 1) = P\left(M_n < \frac{2n+1}{n+1}C|\theta = 1\right)$$

La loi de M_n est obtenue par $P(M_n < x) = F^n(x)$ où F est la f.r. de la loi uniforme sur $[\theta, 2\theta]$, définie par :

$$F(x) = \begin{cases} 0 & \text{si } x \leq \theta \\ \dfrac{x}{\theta} - 1 & \text{si } \theta \leq x \leq 2\theta \\ 1 & \text{si } 2\theta \leq x \end{cases}$$

On obtient donc :

$$\alpha = \left(\frac{2n+1}{n+1}C - 1\right)^n$$

Soit la région critique définie par :

$$M_n < 1 + \alpha^{1/n}$$

La puissance de ce test se calcule par :

$$\eta = P\left(M_n < 1 + \alpha^{1/n} | \theta = 0,9\right) = \left(\frac{1 + \alpha^{1/n}}{0,9} - 1\right)^n$$

La région critique précédente étant indépendante de la valeur $\theta = \theta_1$ dans l'hypothèse alternative, avec seulement la condition $\theta_1 < 1$, est aussi celle du test d'alternative $H_1 : \theta < 1$. Cependant, cette alternative étant une hypothèse multiple, on ne peut pas calculer la puissance de ce test.

19 ▸ On introduit ici une variable de Bernoulli, indicatrice d'une naissance masculine, de paramètre p. Nous avons vu dans l'exercice 4, page 127, que le rapport des vraisemblances conduit à une région critique définie par :

$$\left(\ln \frac{p_0}{p_1} + \ln \frac{1 - p_1}{1 - p_0}\right) \sum_{i=1}^{n} x_i \leq k$$

avec ici $p_0 = 0,50$ et $p_1 = 0,48$, donc de façon équivalente $\sum_{i=1}^{n} x_i \leq C$. La constante C est définie par :

$$\alpha = P\left[\sum_{i=1}^{n} X_i \leq C | H_0\right] = P\left[\bar{X}_n \leq \frac{C}{n} | p = 0,5\right]$$

On utilise la loi approchée de $\bar{X}_n$, qui est la loi normale de paramètres $p = 0,5$ et d'écart type $\sqrt{pq/n} = 1/2\sqrt{n}$. Donc, en centrant sur 1/2 et en réduisant, on obtient comme valeur approchée du seuil critique :

$$C = \frac{n}{2} + u\frac{\sqrt{n}}{2}$$

où u est le fractile d'ordre α de la loi $N(0, 1)$ et α le risque de première espèce retenu. Pour $\alpha = 0,05$, le fractile est $u = -1,644\,9$, donc pour $n = 900$, on obtient $C = 425$, ce qui conduit ici à retenir l'hypothèse nulle $p = 50\,\%$, ce qui paraît logique puisqu'on a observé 52 % de naissances masculines dans cet échantillon. Cependant, le calcul de la puissance va permettre de juger de la qualité de ce test :

$$\eta = P\left[\sum_{i=1}^{n} X_i \leq C | H_1\right] = P\left[\bar{X}_n \leq \frac{C}{n} | p = 0,48\right] \simeq P\left[U \leq \frac{C/n - 0,48}{\sqrt{0,48 \times 0,52/n}}\right]$$

soit ici comme valeur approchée $\Phi(-0,467) = 0,32$ où Φ est la f.r. de la loi $N(0, 1)$. Il y a donc 68 chances sur 100 d'accepter l'hypothèse $p = 0,50$ dans le cas où $p = 0,48$. Le risque de seconde espèce est très élevé.

Suivant que $p_1 < 0,5$ ou $p_1 > 0,5$, la région critique sera de la forme $\bar{x}_n < k_1$ ou $\bar{x}_n > k_2$. On retient donc pour l'alternative $p \neq 0,5$ une région critique symétrique par rapport à l'hypothèse nulle, donc définie par :

$$|\bar{x}_n - 0,5| > C$$

où C est définie par $\alpha = P\left(\left|\bar{X}_n - 0,5\right| > C | H_0\right)$ donc de valeur approchée $u/2\sqrt{n}$ où u est le fractile d'ordre $1 - \alpha/2$ de la loi $N(0, 1)$. Pour $\alpha = 0,05$, on lit $u = 1,96$ et

$C = 0{,}033$. Pour l'observation $\bar{x}_{900} = 0{,}52$, on retient à nouveau l'hypothèse $p = 0{,}5$. Cependant, l'autre risque d'erreur a pour expression :

$$\begin{aligned}\beta &= P\left(0{,}467 < \bar{X}_n < 0{,}533|p = 0{,}48\right)\\ &= P\left(\sqrt{n}\,\frac{0{,}467 - 0{,}48}{\sqrt{0{,}48 \times 0{,}52}} < U < \sqrt{n}\,\frac{0{,}533 - 0{,}48}{\sqrt{0{,}48 \times 0{,}52}}\right)\end{aligned}$$

soit $\beta = 0{,}78$ valeur encore plus élevée que dans le test précédent. Il serait plus logique ici de retenir comme alternative $p > 0{,}5$, la région critique étant alors définie par $\bar{x}_{900} > 0{,}527\,5$. On accepterait encore l'hypothèse $p = 0{,}5$.

20 ▸ D'après ce qui a été obtenu dans l'exercice 6, page 128, la région critique du test entre les hypothèses :

$$\begin{cases} H_0 : m = 560 \\ H_1 : m = m_1 \end{cases}$$

avec $m_1 < 560$ est l'ensemble des points $(x_1, \ldots, x_n)$ tels que :

$$\sqrt{n}\,\frac{\bar{x}_n - 560}{s_n} \leq C$$

La constante C est définie par la condition :

$$\alpha = P\left[\sqrt{n}\,\frac{\bar{X}_n - 560}{S_n} \leq C|m = 560\right]$$

c'est-à-dire que c'est le fractile d'ordre α de la loi de Student à 199 degrés de liberté. Pour $\alpha = 0{,}01$, on lit dans la table 6, page 200, le fractile $C = -2{,}345$. Pour cet échantillon, on obtient $\sqrt{200}\,(\bar{x}_{200} - 560)\,/s_{200} = -6{,}40$, valeur très inférieure au seuil critique ; donc on rejette l'hypothèse nulle, bien que le risque α ait été choisi très faible. Comme la région critique est indépendante de la valeur de m_1, pourvu que $m_1 < 560$, ce test qui est de puissance maximum d'après le théorème de Neyman et Pearson est un test UPP pour l'alternative $m < 560$. On ne peut pas calculer sa puissance, qui est définie par :

$$\eta = P\left[\sqrt{n}\,\frac{\bar{X}_n - 560}{S_n} \leq C|m < 560\right] = P\left[\sqrt{n}\,\frac{\bar{X}_n - m}{S_n} \leq C + \sqrt{n}\,\frac{560 - m}{S_n}\right]$$

En effet, si la loi de $\sqrt{n}\left(\bar{X}_n - m\right)/S_n$ est connue, loi de Student à $n - 1$ degrés de liberté, la borne de l'intervalle dépend de S_n, qui est une v.a. On va donc tester l'hypothèse $\sigma = 10$, équivalente à $\sigma^2 = 100$, contre l'alternative $\sigma^2 \neq 100$. On utilise pour cela la variance empirique et la région d'acceptation de l'hypothèse nulle de ce test est de la forme :

$$a < \sum_{i=1}^{n} \frac{\left(X_i - \bar{X}_n\right)^2}{100} < b$$

région de probabilité $1 - \alpha$. Donc a et b sont respectivement les fractiles d'ordre $\alpha/2$ et $1 - \alpha/2$ de la loi du khi-deux à $n - 1$ degrés de liberté. Pour $n = 200$, on utilise l'approximation par la loi $N(0,1)$ de la loi de $\sqrt{2Y} - \sqrt{2\upsilon - 1}$ avec Y de loi χ^2_υ et ici $\upsilon = n - 1$. On obtient $a = 150{,}49$ et $b = 253{,}14$. Comme $\sum_{i=1}^{200}(x_i - \bar{x}_{200})^2/100 = 179{,}60$, on accepte l'hypothèse $\sigma = 10$ et on est ramené à un problème de test de la moyenne d'une loi normale d'écart type connu. La région critique est alors définie par :

$$\sqrt{n}\,\frac{\bar{x}_n - 560}{10} \leq C$$

et C est le fractile d'ordre α de la loi $N(0,1)$. Pour $\alpha = 0{,}01$, on lit $C = -2{,}326\,3$ et la région critique est définie par $\bar{x}_n < 558{,}36$. La fonction puissance est alors définie en fonction de m par :

$$\eta = P\left[\sqrt{n}\,\frac{\bar{X}_n - m}{10} \leq C + \sqrt{n}\,\frac{560 - m}{10}\right] = \Phi\left[\sqrt{2}\,(558{,}36 - m)\right]$$

où Φ est la f.r. de la loi $N(0,1)$.

21 ▸ On effectue le test entre les hypothèses simples :

$$\begin{cases} H_0 : m = m_0 \\ H_1 : m = m_1 \end{cases}$$

pour une v.a. X de loi $N(m,\sigma)$ avec l'écart type connu $\sigma = 1$. Le théorème de Neyman et Pearson conduit à une région critique définie par :

$$(m_0 - m_1)\sum_{i=1}^{25} x_i \leq k$$

La forme de la région critique de ce test, de puissance maximum, ne dépend pas de la valeur précise de m_1, mais de sa position par rapport à $m_0 = 1$. Ainsi, pour le test de :

$$\begin{cases} H_0 : m = 1 \\ H_1 : m > 1 \end{cases}$$

toutes les valeurs de m correspondant à l'alternative sont supérieures à m_0 ; donc la région critique du test précédent est toujours la même et on a obtenu un test UPP :

$$W_1 = \{(x_1, \ldots, x_{25}) \,/\, \bar{x}_{25} \geq C\}$$

où C est défini par :

$$\alpha = P(W_1|H_0) = P\left(\bar{X}_{25} \geq C|m = 1\right) = P\{U \geq 5(C-1)\} = 1 - \Phi\{5(C-1)\}$$

avec U v.a. de loi $N(0,1)$, de f.r. Φ. Le seuil de ce test est donc $C = 1 + u/5$, où u est le fractile d'ordre $1 - \alpha$ de la loi $N(0,1)$. Pour $\alpha = 0{,}05$, on lit dans la table 2, page 195, le fractile $u = 1{,}644\,9$ et on obtient $C = 1{,}33$. Sur cet échantillon on a observé $\bar{x}_{25} = 1{,}21$; donc on accepte l'hypothèse nulle d'un bon réglage. La fonction puissance de ce test est définie par :

$$\begin{aligned}\eta_1(m) &= P(W_1|H_1) = P\left(\bar{X}_{25} \geq C|m > 1\right)\\ &= P\{U \geq 5(C-m)\} = 1 - \Phi\{u - 5(m-1)\}\end{aligned}$$

Pour le test de :

$$\begin{cases} H_0 : m = 1 \\ H_1 : m < 1 \end{cases}$$

on obtient aussi un test UPP de région critique :

$$W_2 = \{(x_1, \ldots, x_{25})\,/\bar{x}_{25} \leq C\}$$

avec toujours $C = 1 + u/5$, mais cette fois u étant le fractile d'ordre α de la loi $N(0,1)$. Pour $\alpha = 0{,}05$, on a donc $u = -1{,}644\,9$ et $C = 0{,}67$, l'hypothèse nulle étant bien sûr toujours acceptée. La fonction puissance de ce test est définie pour $m < 1$ par :

$$\eta_2(m) = \Phi\{u + 5(1-m)\}$$

Pour le test de :

$$\begin{cases} H_0 : m = 1 \\ H_1 : m \neq 1 \end{cases}$$

on ne peut pas cette fois déduire la forme de la région critique de celle du test entre les hypothèses simples, car elle dépend de la position inconnue de m par rapport à 1. On va retenir une région critique qui sera l'union des précédentes, c'est-à-dire définie par $\bar{x}_{25} \geq C_1$ ou $\bar{x}_{25} \leq C_2$, mais qui ne sera pas celle d'un test UPP. En raison de la symétrie de la statistique de test et de l'alternative, on retient comme région critique :

$$W = \{(x_1, \ldots, x_{25})\,/\,|\bar{x}_{25} - 1| \geq C\}$$

où C est définie par :

$$\alpha = P(W|H_0) = P\left(\left|\bar{X}_{25} - 1\right| \geq C|m = 1\right) = P(|U| \geq 5C) = 2\,[1 - \Phi(5C)]$$

Donc $C = u/5$ où u est le fractile d'ordre $1 - \alpha/2$ de la loi $N(0,1)$. Pour $\alpha = 0{,}05$, on lit $u = 1{,}96$ et $C = 0{,}392$, soit une région critique définie par $\bar{x}_{25} \leq 0{,}61$ ou $\bar{x}_{25} \geq 1{,}39$ et une nouvelle fois on accepte l'hypothèse nulle. La fonction puissance de ce test est définie pour m quelconque par :

$$\eta(m) = 1 - \Phi\{u + 5(1-m)\} + \Phi\{5(1-m) - u\}$$

Le graphe des trois fonctions puissances sur la figure 8.1, page ci-contre, montre bien que ce dernier n'est pas UPP puisque $\eta_1(m) \geq \eta(m)$ pour $m > 1$ et $\eta_2(m) \geq \eta(m)$ pour $m < 1$.

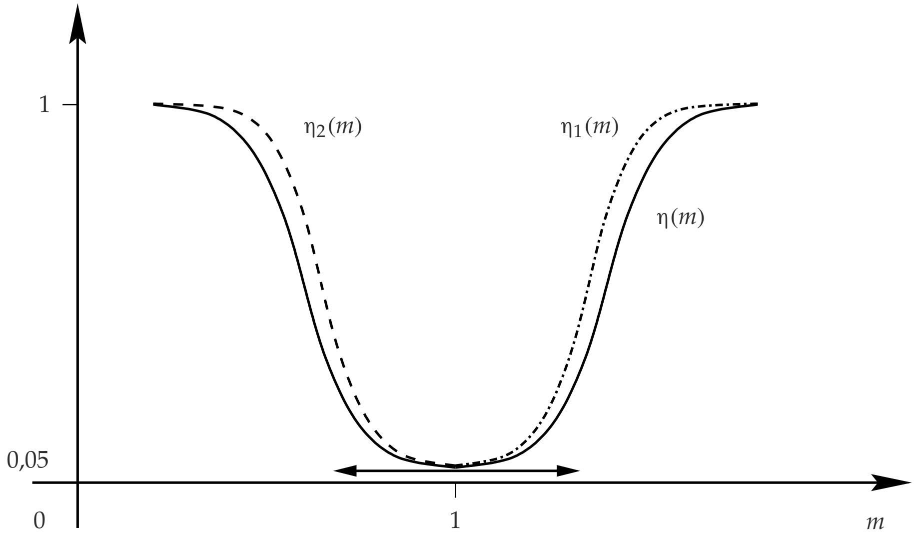


Figure 8.1

22 ▸ Les variables indicatrices de notoriété de la marque Omopaic sont X et Y respectivement avant et après la campagne, de paramètres p_1 et p_2. On dispose d'échantillons de ces deux lois, d'effectifs respectifs $n_1 = 2\,000$ et $n_2 = 1\,000$, pour effectuer le test :

$$\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 < p_2 \end{cases} \iff \begin{cases} H_0 : p_1 - p_2 = 0 \\ H_1 : p_1 - p_2 < 0 \end{cases}$$

Les moyennes empiriques $\bar{X}$ et $\bar{Y}$ de ces deux échantillons permettent de définir l'estimateur sans biais $\bar{X} - \bar{Y}$ du paramètre à tester $\theta = p_1 - p_2$. Sa loi approchée est une loi normale d'espérance θ et d'écart type inconnu $\sigma = \sqrt{p_1q_1/n_1 + p_2q_2/n_2}$ où $q_1 = 1 - p_1$ et $q_2 = 1 - p_2$. Sous l'hypothèse nulle, on estime la valeur commune $p = p_1 = p_2$ par la réunion des deux échantillons :

$$\widehat{p} = \frac{n_1\bar{X} + n_2\bar{Y}}{n_1 + n_2} = \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n_1} X_i + \sum_{i=1}^{n_2} Y_i\right]$$

Cet estimateur permet aussi d'estimer dans ce cas l'écart type :

$$\sigma = \sqrt{p\left(1 - p\right)\left(1/n_1 + 1/n_2\right)}$$

On effectue alors le test à l'aide de la v.a. normalisée :

$$\widehat{\theta} = \frac{\bar{X} - \bar{Y}}{\sqrt{\widehat{p}\left(1 - \widehat{p}\right)\left(1/n_1 + 1/n_2\right)}}$$

dont on peut admettre, compte tenu des tailles d'échantillon élevées, qu'elle suit approximativement une loi normale standard dans l'hypothèse nulle. La région critique est définie par $\widehat{\theta} < C$, où on retient comme valeur approchée de C le fractile d'ordre α de la loi $N(0,1)$ pour un risque de première espèce α fixé. Pour $\alpha = 0{,}05$, la région critique est définie par $\widehat{\theta} < -1{,}644\,9$ et on obtient pour cet échantillon $\bar{x} = 0{,}19$, $\bar{y} = 0{,}23$, $\widehat{p} = 0{,}203$, $\sqrt{\widehat{p}\left(1-\widehat{p}\right)(1/n_1 + 1/n_2)} = 0{,}015\,6$ et $\widehat{\theta} = -2{,}57$, donc on accepte l'hypothèse de l'efficacité de la campagne. Pour un risque plus faible $\alpha = 0{,}01$, le seuil critique est $C = -2{,}33$ et l'hypothèse nulle est encore rejetée.

23 ▸ Les durées de vie des ampoules des marques 1 et 2 sont respectivement les v.a. X et Y de lois respectives $N(m_1, \sigma_1)$ et $N(m_2, \sigma_2)$, ces quatre paramètres étant inconnus. Le test à effectuer est :

$$\begin{cases} H_0 : m_1 = m_2 \\ H_1 : m_1 \neq m_2 \end{cases} \iff \begin{cases} H_0 : m_1 - m_2 = 0 \\ H_1 : m_1 - m_2 \neq 0 \end{cases}$$

et utilise pour cela l'estimateur sans biais $\bar{X} - \bar{Y}$ de $m_1 - m_2$. Cet estimateur suit une loi normale centrée sous H_0, mais de variance inconnue $\sigma_1^2/n_1 + \sigma_2^2/n_2$, où n_1 et n_2 sont les effectifs des échantillons respectifs des lois de X et Y. On utilise donc les estimateurs sans biais de σ_1^2 et σ_2^2 qui sont respectivement :

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 \quad \text{et} \quad S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2$$

Si on remplace l'écart type inconnu par son estimateur $\sqrt{S_1^2/n_1 + S_2^2/n_2}$ pour réduire $\bar{X} - \bar{Y}$, on n'obtiendra pas une loi de Student. Il faut pour cela que l'écart type des deux échantillons soit le même, donc on doit faire un test préalable d'égalité des variances :

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

On accepte l'hypothèse nulle si le rapport S_1^2/S_2^2 est voisin de 1, soit une région d'acceptation de la forme :

$$a < \frac{S_1^2}{S_2^2} < b$$

Les valeurs des constantes a et b sont définies par :

$$\frac{\alpha}{2} = P\left[\frac{S_1^2}{S_2^2} < a | H_0\right] = P\left[\frac{S_1^2}{S_2^2} > b | H_0\right]$$

avec S_1^2/S_2^2 qui suit une loi $F(n_1 - 1, n_2 - 1)$ sous H_0. Pour $\alpha = 0{,}10$, on lit dans la table 7, page 201, le fractile $a = \dfrac{1}{2{,}18} = 0{,}46$ et on obtient par interpolation $b = 2{,}50$. La valeur calculée de S_1^2/S_2^2 pour cet échantillon est 0,55, donc on accepte l'égalité des variances. On retient alors comme estimateur sans biais de la variance commune $\sigma^2 = \sigma_1^2 = \sigma_2^2$:

$$S^2 = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2}$$

La région critique du test initial est alors définie par :

$$\frac{\left|\bar{X} - \bar{Y}\right|}{S\sqrt{1/n_1 + 1/n_2}} > t$$

où t est déterminé par :

$$\alpha = P\left(\frac{\left|\bar{X} - \bar{Y}\right|}{S\sqrt{1/n_1 + 1/n_2}} > t|H_0\right)$$

donc est le fractile d'ordre $1 - \alpha/2$ de la loi de Student à $n_1 + n_2 - 2$ degrés de liberté. Pour $\alpha = 0{,}05$, on lit dans la table 6, page 200, le fractile $t = 2{,}028$. Pour cet échantillon, on observe $s^2 = 77\,081{,}06$ et :

$$\frac{\bar{x} - \bar{y}}{s\sqrt{1/n_1 + 1/n_2}} = -\frac{92}{94{,}93} = -0{,}97$$

Donc on accepte l'égalité des moyennes de ces deux lois.

24 ▶ Dans l'hypothèse nulle d'indépendance des deux caractéristiques de ces candidats au baccalauréat, le tableau est obtenu à partir du produit des effectifs marginaux. Par exemple, le nombre d'étudiants ayant réussi en section L serait dans ce cas $\frac{154 \times 62}{286}$, arrondi à 33 puisqu'il s'agit d'effectifs entiers. On aboutit ainsi au tableau suivant :

	L	ES	S	Total
Réussite	33	51	70	154
Échec	29	44	59	132
Total	62	95	129	286

La région critique est de la forme $D_n \geq C$ où C a comme valeur approchée le fractile d'ordre $1 - \alpha$ de la loi du khi-deux à deux degrés de liberté, pour un test de risque de première espèce α. Pour $\alpha = 0{,}05$, on lit dans la table 5, page 199, la valeur $C = 5{,}991$. La valeur de la statistique utilisée pour ce test est ici :

$$D_n = \frac{(41 - 33)^2}{33} + \frac{(59 - 51)^2}{51} + \frac{(54 - 70)^2}{70} + \frac{(21 - 29)^2}{29} + \frac{(36 - 44)^2}{44} + \frac{(75 - 59)^2}{59}$$

soit $D_n = 14{,}85$; donc on rejette l'hypothèse d'indépendance.

25 ▸ La moyenne empirique de cette distribution vaut $\frac{800}{200} = 4$, donc on teste l'adéquation à une loi de Poisson de paramètre $\lambda = 4$. La table 4, page 198, nous permet de calculer les effectifs théoriques en multipliant les probabilités individuelles par 200 et en arrondissant. Par exemple pour $x = 1$, la probabilité vaut $p_1 = 0{,}073\,3$, soit un effectif théorique $np_1 = 14{,}66$ arrondi à 15. On obtient ainsi le tableau des deux distributions, empirique et théorique :

Nombre de voitures	0	1	2	3	4	5	6	7	8	9	10	11
Effectif observé	1	15	30	46	38	30	16	13	5	3	2	1
Effectif théorique	4	15	29	39	39	31	21	12	6	3	1	0

Les effectifs des valeurs extrêmes étant inférieurs à 5, on regroupe les deux premières et les quatre dernières valeurs. Il reste donc seulement huit classes de valeurs et comme le paramètre de la loi théorique retenue a été estimé, la statistique qui définit la région critique $D_n \geq C$ suit approximativement une loi du khi-deux à six $(8 - 1 - 1)$ degrés de liberté. Pour $\alpha = 0{,}05$, le seuil critique est le fractile d'ordre 0,95 de cette loi, soit $C = 12{,}6$. La valeur observée ici est :

$$D_n = \frac{(16-19)^2}{19} + \frac{(30-29)^2}{29} + \frac{49}{39} + \frac{1}{39} + \frac{1}{31} + \frac{25}{21} + \frac{1}{12} + \frac{1}{10} = 3{,}196$$

donc on retient comme loi théorique la loi $\mathcal{P}(4)$ pour la v.a. X qui représente le nombre de voitures qui arrivent par minute à ce poste de péage.

26 ▸ Si on considère que le nombre de défectueux est une v.a. X de loi binômiale $\mathcal{B}(20, p)$, on estime le paramètre p à partir de la moyenne empirique $\bar{x}_{200} = 2{,}01$ et on retient $p = 0{,}1$. Les probabilités individuelles p_i lues dans la table 3, page 197, de la loi binômiale $\mathcal{B}(20\,;0{,}1)$ permettent de calculer les effectifs théoriques np_i, arrondis en valeurs entières. Par exemple, pour la valeur $x = 0$, on lit $p_0 = 0{,}121\,6$, soit $np_0 = 24{,}32$ arrondi à 24. On obtient ainsi le tableau des deux distributions, empirique et théorique :

Nombre de défectueux	0	1	2	3	4	5	6	7
Nombre de lots	26	52	56	40	20	2	0	4
Effectif théorique	24	54	57	38	18	6	2	1

On regroupe les trois dernières valeurs pour obtenir un effectif théorique supérieur à cinq. La région critique est de la forme $D_n \geq C$. La loi asymptotique de la statistique D_n est une loi du khi-deux à quatre degrés de liberté car il y six classes de valeurs distinctes retenues et un paramètre estimé. Pour $\alpha = 0{,}05$, la valeur de C est le fractile d'ordre 0,95 de la loi χ^2_4, soit $C = 9{,}5$. La valeur observée ici est :

$$D_n = \frac{(26-24)^2}{24} + \frac{(52-54)^2}{54} + \frac{1}{57} + \frac{4}{38} + \frac{4}{18} + \frac{9}{9} = 1{,}59$$

Donc on accepte comme loi théorique la loi binômiale $\mathcal{B}$ (20 ; 0,1). Pour $\alpha = 0{,}10$, la conclusion serait identique puisque le seuil serait $C = 7{,}8$.

La valeur retenue pour p étant ici très faible, on pourrait aussi considérer que la présence d'une pièce défectueuse est un événement rare susceptible d'être modélisé par la loi de Poisson. La variance empirique a pour valeur 1,98 très proche de celle 2,01 de la moyenne empirique, ce qui conforte cette possibilité. On va donc calculer les effectifs théoriques associés à une loi de Poisson $\mathcal{P}$ (2) :

Nombre de défectueux	0	1	2	3	4	5	6	7
Nombre de lots	26	52	56	40	20	2	0	4
Effectif théorique	27	54	54	36	18	7	2	2

On regroupe encore les trois dernières valeurs et on obtient $D_n = 3{,}13$, valeur un peu plus élevée que dans le cas de la loi binômiale, mais on peut également retenir cette loi théorique.

Annales corrigées Université Panthéon-Assas, Paris II

9

SUJETS D'EXAMEN

LICENCE ÉCONOMIE ET GESTION, 2e ANNÉE

Septembre 2007

1. Deux événements A et B ont comme probabilités respectives $1/4$ et $1/3$ de se réaliser seuls (l'un et pas l'autre). Sachant que la probabilité que l'un ou l'autre se réalise est de $3/4$, préciser s'ils sont indépendants ou non.

2. Quatre joueurs A, B, C et D tirent successivement et sans remise une carte d'un lot de 6 cartes numérotées de 1 à 6. Si un joueur tire une carte paire, il a gagné et le jeu s'arrête.

a) Calculer la probabilité de gain de chaque joueur.

b) Soit N la v.a. qui représente le nombre de tirages effectués à l'issue de ce jeu. Calculer $E(N)$.

3. La loi d'une v.a. X est définie par $P(X = n) = \frac{1}{n} - \frac{1}{n+1}$ pour tout entier $n \geqslant 1$. Calculer $P(X < n)$ pour tout entier n, puis $E(X)$.

4. a) Soit X une v.a. de loi normale, telle que $P(X < -2) = 0{,}3085$ et $P(X < 2) = 0{,}9332$. Trouver le nombre réel a tel que $P\left(X^2 + 2X < a\right) = 0{,}975$.

b) Soit U une v.a. de loi normale centrée et réduite et V une v.a. qui suit une loi du khi-deux à 16 degrés de liberté. Ces deux v.a. étant indépendantes, trouver le nombre réel b tel que $P\left(\frac{U}{\sqrt{V}} < b\right) = 0{,}975$.

5. On effectue deux tirages successifs sans remise dans une urne qui contient 3 boules noires et 2 boules rouges. On note X la v.a. qui représente le nombre de boules rouges tirées et Y la v.a. qui vaut 1 si on tire une boule noire au premier tirage et 0 sinon. Déterminer la loi de probabilité du couple (X, Y) puis calculer $cov(X, Y)$. Les v.a. X et Y sont-elles indépendantes ?

6. Soit X une v.a. de loi normale d'espérance θ et de variance $\theta(1 - \theta)$, où θ est un paramètre réel inconnu, avec $0 < \theta < 1$.

a) À partir d'un échantillon $(X_1, ..., X_n)$ de X déterminer deux estimateurs T_n et T'_n de θ par la méthode des moments, en utilisant les moments non centrés d'ordres 1 et 2. Étudier leurs propriétés.

b) Comparer ces 2 estimateurs et indiquer celui qu'il faut choisir. L'un d'eux est-il efficace ? On donne $V\left(X^2\right) = 2\theta^2\left(1 - \theta^2\right)$.

7. Soit X une v.a. de densité :

$$f(x; \theta) = \begin{cases} 0 & \text{si } x < \theta \\ e^{-(x-\theta)} & \text{si } \theta \leqslant x \end{cases}$$

où θ est un paramètre réel inconnu.

a) Calculer $E(X)$.

b) À partir d'un échantillon $(X_1, ..., X_n)$ de X déterminer un estimateur T_n de θ par la méthode des moments et étudier ses propriétés.

c) Déterminer la loi limite de T_n quand n devient infini et en déduire un intervalle de confiance bilatéral symétrique de niveau voisin de $1 - \alpha$ pour θ, dans le cas où on a observé un échantillon de taille supérieure à 100. On donne $V(X) = 1$.

8. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. X de densité :

$$f(x; \theta) = \begin{cases} \dfrac{3x^2}{\theta^3} & \text{si } 0 \leqslant x \leqslant \theta \\ 0 & \text{sinon} \end{cases}$$

où θ est un paramètre réel inconnu strictement positif.

a) Déterminer la fonction de répartition de X, puis celle de $M_n = \max\{X_1, ..., X_n\}$.

b) Représenter l'allure générale du graphe de la vraisemblance $L(x_1, ..., x_n; \theta)$ pour $(x_1, ..., x_n)$ fixé et θ variant. En déduire l'estimateur du maximum de vraisemblance, puis un estimateur sans biais T_n du paramètre θ. Est-il efficace ?

c) Déterminer en fonction de α_1 et α_2 les réels t_1 et t_2 compris entre 0 et 1 qui vérifient les deux conditions :

$$\alpha_1 = P(M_n - \theta < -\theta t_1) \text{ et } \alpha_2 = P(M_n - \theta > -\theta t_2)$$

En déduire un intervalle de confiance bilatéral symétrique de niveau $1 - \alpha$ pour θ.

Mai 2007

1. Un magasin souhaite s'équiper d'un détecteur de faux billets de 500 euros. On note p la probabilité qu'un détecteur repère effectivement un billet qui est faux.

a) On présente 100 faux billets à un détecteur bon marché et seulement 80 d'entre eux sont détectés. Construire un intervalle de confiance pour p, de niveau $1 - \alpha = 0{,}95$, à l'aide d'un abaque puis en utilisant une approximation que l'on justifiera, si l'on ne disposait pas de cet abaque.

b) Ce caissier présente à nouveau 50 faux billets à un détecteur plus sophistiqué qui les a tous détectés. Construire à l'aide d'un abaque un intervalle de confiance de niveau $1 - \alpha = 0{,}95$ pour le paramètre p associé à ce détecteur. Quelle approximation pourrait-on utiliser ici si l'on ne disposait pas de cet abaque ?

2. Soit X une v.a. de loi normale d'espérance θ et de variance 2θ, où θ est un paramètre réel inconnu strictement positif.

a) Étudier les propriétés de l'estimateur $\overline{X}_n$ construit à partir d'un échantillon $(X_1, ..., X_n)$ de X.

b) Étudier les propriétés de l'estimateur sans biais T_n de θ construit à partir de la variance empirique modifiée :

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$$

c) Comparer les estimateurs T_n et $\overline{X}_n$ et indiquer celui qu'il faut choisir. L'un d'eux est-il efficace ?

3. Une étude statistique effectuée pour la chaîne de magasins *Farrecour* a établi que le bénéfice mensuel moyen du magasin $i, 1 \leqslant i \leqslant n$, pouvait être représenté par une v.a. de loi normale d'espérance m_i et de variance σ^2, ces valeurs étant déterminées par l'étude. Une campagne publicitaire nationale a conduit à une augmentation θ du bénéfice. Pour estimer ce paramètre réel inconnu θ, on observe sur une certaine période les nouveaux bénéfices mensuels moyens de chaque magasin $i, 1 \leqslant i \leqslant n$, représentés cette fois par une v.a. X_i de loi normale d'espérance $m_i + \theta$ et de variance σ^2.

a) Déterminer un estimateur T_n de θ par la méthode du maximum de vraisemblance. Étudier ses propriétés. Est-il efficace ?

b) Déterminer un intervalle de confiance bilatéral symétrique de niveau $1-\alpha$ pour θ. Application : $\sigma^2 = 1$, $\sum_{i=1}^{36}(x_i - m_i) = 270$ et $1-\alpha = 0{,}95$.

Janvier 2007

1. Dans une certaine population, la probabilité de naissance d'un garçon est de $4/7$. Déterminer la proportion de familles de trois enfants où les enfants ne sont pas tous du même sexe, en précisant l'ensemble fondamental Ω retenu pour modéliser ce problème.

2. Soit p la proportion de bovins d'un département français atteints de la maladie ESB (« vache folle »). On utilise un test de dépistage qui se révèle positif avec une probabilité $1-\alpha$ si l'animal est effectivement malade et avec une probabilité β si l'animal est sain.

a) Déterminer la probabilité $\pi\left(p\right)$ qu'un animal pris au hasard dans ce département soit effectivement malade, sachant que le test s'est révélé positif.

b) On suppose à partir de maintenant que $\alpha = \beta$. Déterminer $\pi\left(p\right)$ puis calculer $\pi(\alpha)$. Commenter le résultat.

c) Calculer la valeur de $\pi(0{,}005)$ pour $\alpha = 0{,}02$, puis pour $\alpha = 0{,}01$. Commenter le résultat.

3. La fonction de répartition d'une v.a. X est nulle pour $x \leqslant 3$ et pour tous les entiers $n \geqslant 3$ a pour valeur :

$$F(x) = 1 - \frac{1}{2^{n-2}} \quad \text{si } n < x \leqslant n+1$$

Déterminer la loi de probabilité de X.

4. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \dfrac{1}{4} + \dfrac{3}{4}x^2 & \text{si } -1 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

a) Déterminer la densité de la v.a. $Y = X^2$.

b) Calculer $E(Y)$.

5. Soit X une v.a. de loi normale centrée et réduite.

a) Déterminer la loi de probabilité des v.a. $Y = 2X, X+Y, X+Y+Z$ avec $Z = -3X$.

b) Calculer $cov(X, Y)$ et $V(X+Y+Z)$.

Septembre 2006

1. On lance un dé rouge et un dé vert et on considère les événements A « obtenir un chiffre pair sur le dé rouge », B « obtenir un chiffre pair sur le dé vert » et C « la somme des chiffres obtenus est paire ». Les événements A, B et C sont-ils indépendants deux à deux ? dans leur ensemble ? On précisera l'ensemble fondamental Ω retenu pour modéliser ce problème.

2. On effectue l'expérience suivante sur la mémoire des rats. Un rat peut appuyer sur l'une des trois manettes placées devant lui et dans un seul cas il recevra de la nourriture. On note X la variable aléatoire qui représente le nombre de tentatives effectuées par le rat jusqu'à ce qu'il obtienne de la nourriture.

a) Déterminer la loi de probabilité de X dans les 3 situations suivantes :
– il choisit au hasard une manette à chaque tentative (aucune mémoire) ;
– il exclut la « mauvaise » manette choisie à la tentative précédente (mémoire courte) ;
– il exclut toutes les « mauvaises » manettes choisies aux tentatives précédentes (mémoire parfaite).

b) Calculer dans chacune de ces situations la probabilité qu'il fasse plus de 3 tentatives pour obtenir de la nourriture.

3. Soit X une v.a. de densité :

$$f(x) = \begin{cases} x & \text{si } 0 \leqslant x \leqslant 1 \\ 2 - x & \text{si } 1 \leqslant x \leqslant 2 \\ 0 & \text{sinon} \end{cases}$$

a) Déterminer la fonction de répartition de X et en déduire la médiane de cette loi.

b) Calculer $P\{|X-1| < x\}$ pour tout réel x.

4. Soit (X, Y) un couple de v.a. de densité :

$$f(x, y) = \begin{cases} 2 & \text{si } 0 \leqslant x \leqslant y \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

a) Déterminer les densités marginales de X et Y. Les v.a. X et Y sont-elles indépendantes ?

b) Déterminer les densités conditionnelles de X sachant que $Y = y$ et de Y sachant que $X = x$.

5. Pour connaître la proportion p d'utilisateurs d'internet dans une région française, on effectue un sondage auprès de N personnes qui ont été tirées au hasard et avec remise dans un fichier des habitants de cette région qui est découpée en deux zones rurale et urbaine. On tire donc deux sous-échantillons indépendants de tailles respectives n et m, avec $n + m = N$. On note T_n et T'_m les estimateurs associés de p. On retient comme estimateur construit sur l'échantillon total $\widehat{p}_N = aT_n + bT'_m$. Déterminer le réel b en fonction du réel a pour que cet estimateur soit sans biais. Déterminer ensuite a pour que cet estimateur soit optimal.

6. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. X de densité :

$$f(x;\theta) = \begin{cases} \dfrac{2x}{\theta^2} & \text{si } 0 \leqslant x \leqslant \theta \\ 0 & \text{sinon} \end{cases}$$

où θ est un paramètre réel inconnu strictement positif.

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

b) Déterminer un estimateur T_n de θ par la méthode des moments. Étudier ses propriétés. Est-il efficace ?

c) Déterminer un nouvel estimateur sans biais T'_n de θ construit à partir de $M_n = \max\{X_1, ..., X_n\}$. Comparer les estimateurs T_n et T'_n et indiquer celui qu'il faut choisir.

7. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. X de loi normale centrée et de variance θ, où θ est un paramètre réel strictement positif.

a) Déterminer un estimateur T_n de θ par la méthode des moments. Étudier ses propriétés. Est-il efficace ?

b) Déterminer un intervalle de confiance de niveau $1 - \alpha = 0,95$ pour θ. Application : $\sum_{i=1}^{50} x_i^2 = 100$.

Mai 2006

1. Soit p la probabilité qu'une pièce produite par une machine présente un défaut grave.

a) On prélève au hasard 150 pièces produites pour effectuer un contrôle de qualité. Sachant que 6 d'entre elles présentent un défaut grave, construire un intervalle de confiance pour p de niveau 0,95.

b) Une nouvelle machine est installée et sur 30 pièces prélevées au hasard aucune ne présente un défaut grave. Donner un intervalle de confiance de niveau 0,95 pour ce nouveau paramètre p. Peut-on considérer que cette nouvelle machine est de meilleure qualité ?

2. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. X de densité :

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

où θ est un paramètre réel inconnu.

a) Déterminer l'estimateur $\widehat{\theta}_n$ du maximum de vraisemblance du paramètre θ.

b) Calculer la quantité d'information de Fisher $I_n(\theta)$. L'estimateur $\widehat{\theta}_n$ est-il efficace ?

3. Pour connaître le budget mensuel consacré aux loisirs, on a effectué un sondage dans neuf villes françaises auprès de 100 ménages. On note X la variable aléatoire qui représente le budget loisir d'un ménage tiré au sort et Y celle qui représente la moyenne d'un échantillon de taille 100 de ces budgets dans une ville donnée. On dispose donc d'un échantilllon $(Y_1, ..., Y_9)$ de cette v.a. Y.

a) Indiquer la loi qu'il est logique de retenir pour cette v.a. Y et en déduire alors la loi de la moyenne empirique $\overline{Y}_9$ de l'échantillon.

b) Déterminer un estimateur de chacun des paramètres $m = E(X)$ et $\sigma^2 = V(X)$.

c) Déterminer un intervalle de confiance bilatéral symétrique de niveau $1 - \alpha$ pour m. Application : $\sum_{i=1}^{9} y_i = 2709$, $\sum_{i=1}^{9} \left(y_i - \overline{y}_9\right)^2 = 1800$ et $1 - \alpha = 0{,}95$.

Janvier 2006

1. Parmi les patients venus consulter un ophtalmologue, 10 % étaient myopes, 20 % presbytes, les autres ayant une bonne vue, les proportions de ceux déclarant souffrir de maux de tête étant respectivement de 90 %, 50 % et 5 %. Quelle est la probabilité qu'un patient souffrant de maux de tête n'ait pas besoin de porter de lunettes ?

2. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \dfrac{x}{2} & \text{si } 0 \leqslant x \leqslant 1 \\ \dfrac{1}{2} & \text{si } 1 \leqslant x \leqslant 2 \\ \dfrac{1}{2}(3 - x) & \text{si } 2 \leqslant x \leqslant 3 \\ 0 & \text{sinon} \end{cases}$$

a) Déterminer la fonction de répartition de X.

b) Calculer $E(X)$.

3. Soit X une v.a. de loi normale, telle que $P(X < 3) = 0{,}1587$ et $P(X > 12) = 0{,}0228$. Calculer $P(X > 0)$ puis trouver le nombre réel a tel que $P\left([X - E(X)]^2 < a\right) = 0{,}975$.

4. On lance 2 boules qui tombent chacune au hasard dans l'un des 2 tiroirs A ou B. On note X la v.a. qui représente le nombre de boules tombées dans le tiroir A et Y la v.a. qui représente le nombre de tiroirs non vides. Déterminer la loi de probabilité du couple (X, Y) puis calculer $cov(X, Y)$. Les v.a. X et Y sont-elles indépendantes ?

Septembre 2005

1. Une urne contient trois boules rouges et deux boules noires. On effectue des tirages successifs sans remise et on arrête dès qu'on a obtenu 2 boules rouges. Préciser l'ensemble fondamental Ω retenu pour modéliser ce problème et calculer la probabilité de chaque événement élémentaire.

2. Soit $(X_1, ..., X_n)$ des v.a. indépendantes et de même loi que X de densité :

$$f(x) = \begin{cases} 2(1-x) & \text{si } 0 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

Déterminer la densité de la v.a. $M_n = \max\{X_1, ..., X_n\}$.

3. Soit X une v.a. de loi normale, telle que $P(X < -1) = 0{,}500$ et $P(-1{,}6 < X < -0{,}4) = 0{,}770$. Calculer $P(X > 0)$ puis trouver le nombre réel a tel que $P\left(X^2 + 2X < a\right) = 0{,}995$.

4. Soit (X, Y) un couple de v.a. de densité :

$$f(x, y) = \begin{cases} e^{y-x} & \text{si } y \leqslant x \text{ et } 0 \leqslant y \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

Déterminer les densités marginales de X et Y. Les v.a. X et Y sont-elles indépendantes ?

5. Un joueur participe à une loterie où sa probabilité de gain est $1/\theta$, avec $\theta > 1$. Il joue jusqu'à ce qu'il ait gagné et on note X la v.a. qui représente le nombre de tirages indépendants auxquels il a participé.

a) Déterminer la loi de probabilité de X et calculer $E(X)$.

b) À partir d'un échantillon $(X_1, ..., X_n)$ de X déterminer un estimateur T_n de θ par la méthode des moments et étudier ses propriétés.

6. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. de loi normale d'espérance θ et de variance 2θ, où θ est un paramètre réel strictement positif. On pose :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 \quad T_{n-1} = \sqrt{n}\frac{\overline{X}_n - \theta}{S_n}$$

Déterminer à partir de T_{n-1} un intervalle de confiance bilatéral symétrique de niveau $1 - \alpha = 0{,}95$ pour θ dans le cas où $n = 30$.

7. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. X de densité :

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} x^{-1/\theta - 1} & \text{si } 1 \leqslant x \\ 0 & \text{sinon} \end{cases}$$

où θ est un paramètre réel inconnu strictement positif. Déterminer un estimateur T_n de θ par la méthode du maximum de vraisemblance. Étudier ses propriétés. Est-il efficace ?

Mai 2005

1. Soit p la probabilité qu'un système d'alarme se déclenche à bon escient.

a) Sur 250 essais indépendants effectués dans une bijouterie il y a eu 240 déclenchements justifiés de l'alarme. Construire un intervalle de confiance pour p de niveau 0,95 à l'aide d'un abaque, puis en utilisant une approximation que l'on justifiera, si l'on ne disposait pas de cet abaque.

b) Un nouvel équipement est installé dans cette bijouterie et sur 30 essais indépendants effectués, tous ont provoqué un déclenchement justifié. Donner un intervalle de confiance pour ce nouveau paramètre p de niveau 0,95. Peut-on considérer que ce nouvel équipement est de meilleure qualité ?

2. Pour comparer deux modèles de pèse-lettre, notés A et B, on effectue des pesées successives indépendantes d'une lettre de poids exact connu 100 g. On note X la v.a. qui représente le poids affiché par un pèse-lettre.

a) Indiquer la loi qu'il est logique de retenir pour cette v.a. X et préciser alors le paramètre θ qui peut servir d'indicateur de qualité d'un modèle de pèse-lettre. Déterminer un estimateur T_n de θ construit à partir d'un échantillon $(X_1, ..., X_n)$ de X et étudier ses propriétés.

b) Pour le modèle A on a effectué 30 pesées indépendantes, les valeurs observées $x_1, ..., x_{30}$ conduisant à $\sum_{i=1}^{30} x_i = 3020$ et $\sum_{i=1}^{30} x_i^2 = 315400$. Déterminer un intervalle de confiance de niveau $1 - \alpha = 0{,}95$ pour θ.

c) Pour le modèle B on a effectué 50 pesées indépendantes, les valeurs observées $y_1, ..., y_{50}$ conduisant à $\sum_{i=1}^{50} y_i = 4950$ et $\sum_{i=1}^{50} y_i^2 = 510000$. Déterminer un intervalle de confiance de niveau $1 - \alpha = 0{,}95$ pour le paramètre θ associé à ce modèle. Quel modèle vous paraît de meilleure qualité ?

3. Soit $(X_1, ..., X_n)$ un échantillon d'une v.a. X de densité :

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} x^{1/\theta - 1} & \text{si } 0 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

où θ est un paramètre réel inconnu strictement positif.

a) Déterminer la densité de la v.a. $Y = -\ln X$.

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

c) Déterminer un estimateur T_n de θ par la méthode du maximum de vraisemblance. Étudier ses propriétés. Est-il efficace ?

Janvier 2005

1. Le jeu de l'ambassadeur consiste à transmettre oralement à son voisin, sans que les autres entendent, un message compliqué pour s'amuser des déformations qu'il peut subir. Le message initial contient une quantité d'information notée I_0 (par exemple le nombre de phonèmes) et à chaque relais il y a une probabilité p qu'il soit transmis correctement. Avec une probabilité $q = 1 - p$ la quantité d'information est diminuée et devient aI_0, avec $0 < a < 1$. On note I_n la v.a. qui représente la quantité d'information obtenue après n relais. Déterminer la loi de probabilité de I_1, I_2 puis I_n pour $n > 2$.

2. Une information est transmise sous forme de codage logique en 0 ou 1. Il y a une probabilité p d'erreur au cours de chaque relais de transmission, c'est-à-dire de transformer le 0 (resp. 1) en 1 (resp. 0). On note V_n l'événement « l'information transmise après n relais est exacte, c'est-à-dire identique à l'information d'origine » et on pose $p_n = P(V_n)$.

a) Calculer p_1 et p_2.

b) Calculer p_n en fonction de p_{n-1} pour $n \geqslant 2$.

c) Calculer $u_n = p_n - \frac{1}{2}$ en fonction de u_{n-1} et en déduire la valeur de u_n, puis celle de p_n, en fonction de n et p.

Calculer la limite de p_n quand n devient infini.

3. Soit X une v.a. de densité :

$$f(x) = \begin{cases} \frac{1}{4} + \frac{3}{4}x^2 & \text{si } -1 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

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

b) Déterminer la fonction de répartition de X puis calculer la probabilité que X appartienne à l'intervalle $[x_1, x_2]$, en fonction des réels x_1 et x_2 tels que $x_1 < x_2$.

4. Soit (X, Y) un couple de v.a. discrètes dont la loi de probabilité est donnée par le tableau ci-après.

Calculer $E(X)$, $E(Y)$ puis $cov(X, Y)$. Les v.a. X et Y sont-elles indépendantes ?

Y \ X	–1	0	1
–2	1/8	1/8	1/8
2	1/4	1/8	1/4

ÉLÉMENTS DE CORRECTION

Septembre 2007

1 ▸ Les hypothèses se traduisent par $P\left(A \cap \overline{B}\right) = \frac{1}{4}, P\left(\overline{A} \cap B\right) = \frac{1}{3}, P(A \cup B) = \frac{3}{4}$.

On en déduit :

$$P(A \cap B) = P(A \cup B) - P\left(A \cap \overline{B}\right) - P\left(\overline{A} \cap B\right) = \frac{1}{6}$$

$$P(A) = P(A \cap B) + P\left(A \cap \overline{B}\right) = \frac{5}{12}$$

$$P(B) = P(A \cap B) + P\left(\overline{A} \cap B\right) = \frac{1}{2}$$

Comme $P(A \cap B) \neq P(A)P(B)$, les événements A et B sont dépendants.

2 ▸ a) Un joueur gagne s'il tire une carte paire et si le(s) précédent(s) a (ont) tiré une carte impaire. On obtient ainsi :

$$P(A) = \frac{3}{6}, \quad P(B) = \frac{3}{6} \times \frac{3}{5} = \frac{6}{20}, \quad P(C) = \frac{3}{6} \times \frac{2}{5} \times \frac{3}{4} = \frac{3}{20},$$
$$P(D) = \frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} \times \frac{3}{3} = \frac{1}{20}$$

b) On en déduit :

$$E(N) = \frac{3}{6} + 2 \times \frac{6}{20} + 3 \times \frac{3}{20} + 4 \times \frac{1}{20} = \frac{35}{20}$$

3 ▸ On a $P(X < 0) = P(X < 1) = 0$ et, pour $n \geqslant 2$:

$$P(X < n) = \sum_{k=1}^{n-1} P(X = k) = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots$$
$$+ \left(\frac{1}{n-2} - \frac{1}{n-1}\right) + \left(\frac{1}{n-1} - \frac{1}{n}\right)$$
$$= 1 - \frac{1}{n}$$

On obtient :

$$E(X) = \sum_{n=1}^{\infty} nP(X = n) = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty$$

c'est-à-dire que cette loi n'admet pas d'espérance.

4 ▸ a) On pose $m = E(X)$ et $\sigma^2 = V(X)$ et on considère la variable centrée et réduite ; les conditions s'écrivent :

$$P\left(\frac{X-m}{\sigma} < \frac{-2-m}{\sigma}\right) = 1 - 0{,}6915 \quad P\left(\frac{X-m}{\sigma} < \frac{2-m}{\sigma}\right) = 0{,}9332$$

La table 1 permet d'obtenir $\frac{2+m}{\sigma} = \frac{1}{2}$ et $\frac{2-m}{\sigma} = \frac{3}{2}$ d'où $m = -1$ et $\sigma = 2$. La v.a. $\left(\frac{X+1}{2}\right)^2$ suit une loi du khi-deux à 1 degré de liberté ; la condition :

$$P\left\{\left(\frac{X+1}{2}\right)^2 < \frac{a+1}{4}\right\} = 0{,}975$$

permet d'obtenir, à partir de la table 5, $a = 19{,}08$.

b) La v.a. $U\big/\left(\sqrt{V}/4\right)$ suit une loi de Student à 16 degrés de liberté, donc la table 6 permet d'obtenir $b = 0{,}53$.

5 ▸ La loi du couple (X, Y) figure dans le tableau ci-après, avec les lois de X et Y dans les marges de ce tableau :

Y \ X	0	1	2	
0	0	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{4}{10}$
1	$\frac{3}{10}$	$\frac{3}{10}$	0	$\frac{6}{10}$
	$\frac{3}{10}$	$\frac{6}{10}$	$\frac{1}{10}$	1

Avec par exemple $P(X = 1, Y = 0) = P(RN) = \frac{2}{5} \times \frac{3}{4} = \frac{3}{10}$. On calcule à partir du tableau $E(X) = \frac{4}{5}$, $E(Y) = \frac{3}{5}$ et $E(XY) = P(X = 1, Y = 1) = \frac{3}{10}$.

Ainsi $cov(X, Y) = -\frac{9}{50}$ et les variables X et Y sont dépendantes.

6 ▸ a) Comme $E(X) = \theta$, l'estimateur T_n obtenu par la méthode des moments est la moyenne empirique $\overline{X}_n$, estimateur sans biais et convergent d'après la loi des grands nombres. De même, $E\left(X^2\right) = V(X) + E^2(X) = \theta$, donc $T'_n = \dfrac{1}{n}\sum_{i=1}^n X_i^2$ est aussi un estimateur sans biais et convergent de θ d'après la loi des grands nombres.

b) Pour comparer ces deux estimateurs sans biais, nous formons le rapport des variances :

$$\frac{V(T'_n)}{V(T_n)} = 2\theta(\theta + 1)$$

Comme cette quantité dépend de la valeur du paramètre à estimer, on ne peut pas conclure. Ces deux estimateurs ne sont pas comparables. Aucun d'eux ne peut être efficace, car il serait alors toujours le meilleur.

7 ▸ a) On calcule $E(X) = \theta + 1$.

b) L'estimateur $T_n = \overline{X}_n - 1$ obtenu par la méthode des moments est sans biais et convergent d'après la loi des grands nombres.

c) D'après le théorème central-limite :

$$\sqrt{n}\frac{\overline{X}_n - E(X)}{\sigma(X)} \underset{\text{loi}}{\rightarrow} N(0,1)$$

donc on obtient :

$$\sqrt{n}(T_n - \theta) \underset{\text{loi}}{\rightarrow} N(0,1)$$

L'estimateur T_n suit asymptotiquement la loi normale $N\left(\theta, \dfrac{1}{\sqrt{n}}\right)$. On construit donc un intervalle de confiance bilatéral symétrique, puisque cette loi est symétrique. On peut trouver la valeur aprochée de la constante u telle que :

$$1 - \alpha = P\left(-u < \sqrt{n}(T_n - \theta) < u\right)$$

L'intervalle de confiance est alors défini par :

$$T_n - \frac{u}{\sqrt{n}} < \theta < T_n + \frac{u}{\sqrt{n}}$$

où u est approximé par le fractile d'ordre $1 - \dfrac{\alpha}{2}$ de la loi normale standard $N(0,1)$.

8 ▸ a) La f.r. F de X est nulle pour $x \leqslant 0$, a pour valeur 1 pour $x \geqslant \theta$ et a pour expression, quand $0 \leqslant x \leqslant \theta$:

$$F(x) = \frac{3}{\theta^3}\int_0^x t^2 dt = \frac{x^3}{\theta^3}$$

On calcule la f.r. de M_n :

$$G(x) = P(M_n < x) = P\left\{\bigcap_{i=1}^{n}(X_i < x)\right\} = \prod_{i=1}^{n} P(X_i < x) = F^n(x)$$

On obtient la densité par dérivation :

$$g(x) = G'(x) = nF^{n-1}(x) f(x)$$

Donc pour $0 \leqslant x \leqslant \theta$:

$$g(x) = 3n\frac{x^{3n-1}}{\theta^{3n}}$$

b) L'expression de la vraisemblance est :

$$L(x_1, ..., x_n; \theta) = \left(\frac{3}{\theta^3}\right)^n \prod_{i=1}^{n} x_i^2$$

si $0 \leqslant x_i \leqslant \theta$ pour $1 \leqslant i \leqslant n$. La vraisemblance, considérée comme une fonction de θ, est nulle pour $0 \leqslant \theta < \max\{x_1, ..., x_n\}$ puis décroissante pour $\theta \geqslant \max\{x_1, ..., x_n\}$ (voir figure 9.1) :

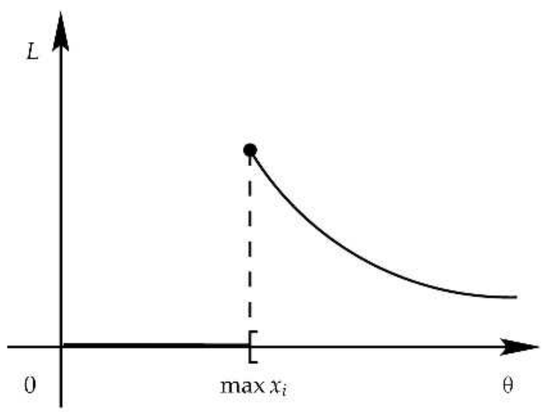


Figure 9.1

Il y a un maximum au point de discontinuité de la vraisemblance et l'estimateur du maximum de vraisemblance est donc M_n. Comme $E(M_n) = \frac{3n}{3n+1}\theta$, l'estimateur sans biais est $T_n = \frac{3n+1}{3n}M_n$. La question de l'efficacité ne se pose pas car l'ensemble des valeurs possibles pour X dépend de θ.

c) Des deux conditions :

$$\alpha_1 = G[\theta(1 - t_1)] = (1 - t_1)^{3n}$$
$$1 - \alpha_2 = G[\theta(1 - t_2)] = (1 - t_2)^{3n}$$

on déduit :

$$t_1 = 1 - \alpha_1^{1/3n} \quad t_2 = 1 - (1 - \alpha_2)^{1/3n}$$

L'intervalle de confiance est défini par la condition :

$$1 - \alpha = P\{-\theta t_1 < M_n - \theta < -\theta t_2\}$$

S'il est bilatéral symétrique $\alpha_1 = \alpha_2 = \alpha/2$, d'où l'intervalle de confiance pour θ :

$$\frac{M_n}{(1 - \alpha/2)^{1/3n}} < \theta < \frac{M_n}{(\alpha/2)^{1/3n}}$$

Mai 2007

1 ▸ **a)** L'intervalle de confiance unilatéral obtenu avec l'abaque 1 est $p > 72\ \%$. On peut ici utiliser le théorème central limite :

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \underset{\text{loi}}{\rightarrow} N(0,1)$$

On utilise cette loi asymptotique pour trouver la valeur approchée de la constante u telle que :

$$0{,}95 = P\left(\widehat{p}_n - u\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} < p\right)$$

On a ici $\widehat{p}_{100} = 0{,}8$ et on lit dans la table 2 $u = 1{,}645$ d'où $p > 73\ \%$.

b) On lit sur l'abaque $p > 93\ \%$. Pour cet échantillon, on obtient $\widehat{p}_{50} = 1$, donc on ne peut pas remplacer $p(1-p)$ par $\widehat{p}_n(1-\widehat{p}_n)$. L'intervalle de confiance serait obtenu par résolution de l'inégalité :

$$\widehat{p}_{50} - u\sqrt{\frac{p(1-p)}{n}} < p$$

2 ▸ **a)** L'estimateur $\overline{X}_n$ est sans biais et convergent d'après la loi des grands nombres.

b) L'estimateur $T_n = \dfrac{1}{2}S_n^2$ obtenu par la méthode des moments est sans biais. La v.a. $\dfrac{n-1}{2\theta}S_n^2$ suit une loi du khi-deux à $n-1$ degrés de liberté, donc $V(T_n) = \dfrac{2\theta^2}{n-1}$. Cet estimateur est sans biais et sa variance tend vers zéro donc il est convergent.

c) Pour comparer ces deux estimateurs sans biais, nous formons le rapport des variances :

$$\frac{V(T_n)}{V(\overline{X}_n)} = \frac{n}{n-1}\theta$$

Comme cette quantité dépend de la valeur du paramètre à estimer, on ne peut pas conclure. Ces deux estimateurs ne sont pas comparables. Aucun d'eux ne peut être efficace, car il serait alors toujours le meilleur.

3 ▸ a) L'expression de la vraisemblance est :

$$L(x_1, ..., x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp - \frac{(x_i - m_i - \theta)^2}{2\sigma^2}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m_i - \theta)^2$$

La log-vraisemblance est donc :

$$\ln L(x_1, ..., x_n; \theta) = -n \ln\left(\sigma\sqrt{2\pi}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m_i - \theta)^2$$

On obtient comme dérivée :

$$\frac{\partial \ln L}{\partial \theta} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - m_i - \theta)$$

Elle s'annule pour $\theta = \frac{1}{n} \sum_{i=1}^{n} (x_i - m_i)$. On dérive une nouvelle fois :

$$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{n}{\sigma^2} < 0$$

Le point qui annule la dérivée première correspond donc à un maximum et l'estimateur du maximum de vraisemblance est :

$$T_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - m_i)$$

Les v.a. $Y_i = X_i - m_i$ suivent des lois normales d'espérance θ et de variance σ^2, donc $T_n = \overline{Y}_n$ est un estimateur sans biais et convergent d'après la loi des grands nombres. On calcule alors la quantité d'information de Fisher :

$$I_n(\theta) = E\left(-\frac{\partial^2 \ln L}{\partial \theta^2}\right) = \frac{n}{\sigma^2}$$

L'estimateur est donc efficace puisque :

$$V(T_n) = \frac{\sigma^2}{n} = \frac{1}{I_n(\theta)}$$

b) L'estimateur $T_n = \overline{Y}_n$ suit une loi normale d'espérance θ et de variance σ^2/n. On construit donc un intervalle de confiance bilatéral symétrique puisque cette loi est symétrique. On doit trouver la valeur de la constante u telle que :

$$1 - \alpha = P\left(-u < \sqrt{n}\frac{T_n - \theta}{\sigma} < u\right)$$

L'intervalle de confiance est alors défini par :

$$-\sigma\frac{u}{\sqrt{n}} < T_n - \theta < \sigma\frac{u}{\sqrt{n}}$$

Ce qui est équivalent à :

$$T_n - \sigma\frac{u}{\sqrt{n}} < \theta < T_n + \sigma\frac{u}{\sqrt{n}}$$

où u est le fractile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale standard $N(0,1)$.

Application : $7{,}17 < \theta < 7{,}83$.

Janvier 2007

1 ▸ L'ensemble fondamental est constitué de tous les événements élémentaires possibles, écrits sous la forme de triplets dont la première lettre désigne le sexe du premier né. On a donc :

$$\Omega = \{F, G\}^3$$

Il y a équiprobabilité donc :

$$P(GGG) = \left(\frac{4}{7}\right)^3 \qquad P(FFF) = \left(\frac{3}{7}\right)^3$$

La probabilité que les enfants ne soient pas tous du même sexe est donc :

$$1 - \left(\frac{4}{7}\right)^3 - \left(\frac{3}{7}\right)^3 = \frac{36}{49}$$

2 ▸ a) Avec des notations évidentes :

$$\pi(p) = P\left(M \middle| T^+\right) = \frac{P\left(M \cap T^+\right)}{P(T^+)} = \frac{P(M)\,P\left(T^+ | M\right)}{P(M)\,P(T^+ | M) + P\left(\overline{M}\right) P\left(T^+ \middle| \overline{M}\right)}$$

$$= \frac{p(1-\alpha)}{p(1-\alpha) + (1-p)\,\beta}$$

b) Dans le cas où $\alpha = \beta$:

$$\pi(p) = \frac{p(1-\alpha)}{\alpha + p(1-2\alpha)}$$

On obtient $\pi(\alpha) = 1/2$, c'est-à-dire que dans le cas où le risque d'erreur du test est égal à la proportion d'animaux malades, il y a seulement une chance sur deux qu'un animal dont le test est positif soit réellement malade.

c) On a :

$$\pi(0{,}005) = \frac{1-\alpha}{1+198\alpha}$$

Soit $\pi(0{,}005) = 49/248(\simeq 1/5)$ pour $\alpha = 0{,}02$ et $\pi(0{,}005) = 99/298(\simeq 1/3)$ pour $\alpha = 0{,}01$. Même avec un risque d'erreur de 1 % le taux de détection exact est de l'ordre de 30 %, c'est-à-dire très faible. Cela est dû au fait que la maladie est rare.

3 ▸ La fonction de répartition est constante par morceaux, donc la loi de probabilité est discrète, avec ici $X(\Omega) = \{3, 4, ..., n, ...\}$ et pour tous les entiers $n \geqslant 3$:

$$\begin{aligned} p_n &= P(X = n) = P(X < n+1) - P(X < n) \\ &= F(n+1) - F(n) = \left(1 - \frac{1}{2^{n-2}}\right) - \left(1 - \frac{1}{2^{n-3}}\right) = \frac{1}{2^{n-2}} \end{aligned}$$

4 ▸ a) On calcule la f.r. de Y :

$$G(y) = P(Y < y) = P\left(X^2 < y\right)$$

On a donc $G(y) = 0$ pour $y \leqslant 0, G(y) = 1$ pour $y \geqslant 1$ et, pour $0 \leqslant y \leqslant 1$:

$$G(y) = P\left(-\sqrt{y} < X < \sqrt{y}\right) = F\left(\sqrt{y}\right) - F\left(-\sqrt{y}\right)$$

La densité est obtenue par dérivation :

$$g(y) = \frac{1}{2\sqrt{y}}\left[f\left(\sqrt{y}\right) + f\left(-\sqrt{y}\right)\right] = \frac{1}{4\sqrt{y}}(1+3y)$$

b) On obtient comme espérance :

$$E(Y) = \frac{1}{4}\int_0^1 \left(\sqrt{y} + 3y\sqrt{y}\right) dy = \frac{7}{15}$$

5 ▸ a) Toute fonction linéaire d'une v.a. normale suit une loi normale, donc Y suit une loi normale centrée de variance 4 , $X + Y = 3X$ suit une loi normale centrée de variance 9 et $X + Y - 3X = 0$ suit une loi normale dégénérée (de variance nulle), c'est-à-dire est une variable certaine, $P(X + Y + Z = 0) = 1$.

b) On a $cov(X, Y) = E(XY) = E\left(2X^2\right) = 2V(X) = 2$ et $V(X + Y + Z) = 0$ car c'est une variable certaine.

Septembre 2006

1 ▸ L'ensemble fondamental est constitué de tous les événements élémentaires possibles, écrits sous la forme de couples dont la première lettre désigne le résultat du dé rouge :

$$\Omega = \{P, I\}^2$$

Les événements considérés sont $A = \{(PP), (PI)\}$, $B = \{(PP), (IP)\}$ et $C = \{(PP), (II)\}$. Il y a équiprobabilité donc :

$$P(A) = P(B) = P(C) = \frac{2}{4}$$

D'autre part, $A \cap B = A \cap C = B \cap C = \{(PP)\}$ donc :

$$P(A \cap B) = \frac{1}{4} = P(A)P(B) \quad P(A \cap C) = \frac{1}{4} = P(A)P(C)$$

$$P(B \cap C) = \frac{1}{4} = P(B)P(C)$$

donc les événements A, B et C sont indépendants deux à deux. Enfin, $A \cap B \cap C = \{(PP)\}$ donc :

$$P(A \cap B \cap C) = \frac{1}{4} \neq P(A)P(B)P(C) = \frac{1}{8}$$

donc les événements A, B et C ne sont pas mutuellement indépendants.

2 ▸ a) Avec des notations évidentes, on obtient dans le cas aucune mémoire (am) :

$$P(X = n) = P\left(\overline{N}...\overline{N}N\right) = \left(\frac{2}{3}\right)^{n-1}\frac{1}{3}, \quad n \geqslant 1$$

Dans le cas mémoire courte (mc) :

$$P(X = 1) = \frac{1}{3}, \quad P(X = n) = P\left(\overline{N}...\overline{N}N\right) = \frac{2}{3}\left(\frac{1}{2}\right)^{n-1}, \quad n \geqslant 2$$

Dans le cas mémoire parfaite (mp) :

$$P(X = 1) = \frac{1}{3}, \quad P(X = 2) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}, \quad P(X = 3) = \frac{2}{3} \times \frac{1}{2} \times 1 = \frac{1}{3}$$
$$P(X = n) = 0 \quad n \geqslant 4$$

b) On calcule ensuite aisément, dans le cas am :

$$P(X \leqslant 3) = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} = \frac{19}{27} \quad p_{am} = \frac{8}{27}$$

Dans le cas mc :

$$P(X \leqslant 3) = \frac{1}{3} + \frac{1}{3} + \frac{1}{6} = \frac{5}{6} \quad p_{mc} = \frac{1}{6}$$

Dans le cas mp :

$$P(X \leqslant 3) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1 \quad p_{mp} = 0$$

3 ▸ a) La fonction de répartition de X est nulle pour $x \leqslant 0$. Son expression pour $0 < x \leqslant 1$ est :

$$F(x) = \int_0^x t dt = \frac{x^2}{2}$$

Pour $1 \leqslant x \leqslant 2$:

$$F(x) = F(1) + \int_1^x (2-t)\, dt = -\frac{x^2}{2} + 2x - 1$$

On a ensuite $F(x) = 1$ pour $x > 2$. Puisque $F(1) = \frac{1}{2}$, la médiane vaut 1 (c'est le centre de symétrie de cette loi).

b) On a $P\{|X-1| < x\} = 0$ pour tout réel $x \leqslant 0$. Pour $x > 0$:

$$P\{|X-1| < x\} = P\{1 - x < X < 1 + x\} = F(1+x) - F(1-x)$$

Donc $P\{|X-1| < x\} = 1$ pour tout réel $x \geqslant 1$. Pour $0 < x < 1$:

$$P\{|X-1| < x\} = 1 - (1-x)^2 = x(2-x)$$

4 ▸ a) Les densités marginales de X et Y s'obtiennent par intégration de la densité du couple (voir figure 9.2) :

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\, dy = 2\int_x^1 dy = 2(1-x) \quad \text{si } 0 \leqslant x \leqslant 1$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\, dx = 2\int_0^y dx = 2y \quad \text{si } 0 \leqslant y \leqslant 1$$

On constate que $f(x,y) \neq f_X(x) f_Y(y)$ donc les v.a. X et Y ne sont pas indépendantes.

b) Les densités conditionnelles s'obtiennent en divisant la densité du couple par la densité marginale de la v.a. qui conditionne :

$$f_X(x | Y = y) = \frac{f(x,y)}{f_Y(y)} = \frac{1}{y} \qquad \text{pour } 0 < x \leqslant y$$

Il s'agit de la loi uniforme sur l'intervalle $[0, y]$.

La loi de Y sachant que $X = x$ a pour densité :

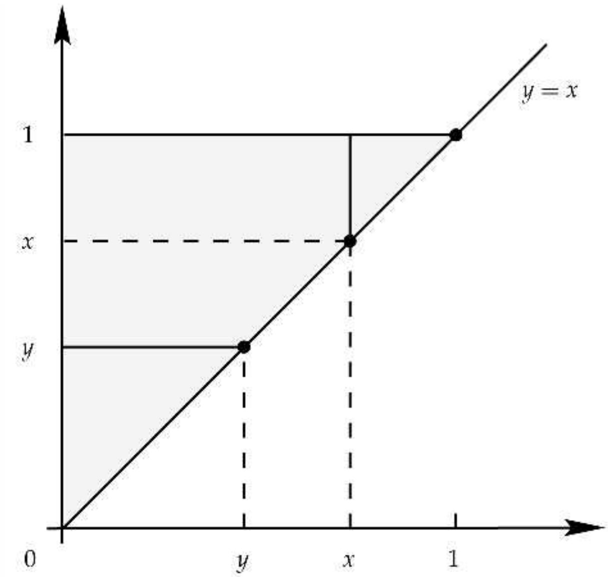


Figure 9.2

$$f_Y\left(y\,|X=x\right)=\frac{f\left(x,y\right)}{f_X\left(x\right)}=\frac{1}{1-x}\qquad\text{pour } x\leqslant y<1$$

Il s'agit de la loi uniforme sur l'intervalle $[x,1]$.

5 ▸ Les estimateurs T_n et T'_m sont sans biais, donc $E\left(\widehat{p}_N\right)=aE\left(T_n\right)+bE\left(T'_m\right)$ $=(a+b)\,p$. L'estimateur $\widehat{p}_N$ sera aussi sans biais si $b=1-a$. Les estimateurs T_n et T'_m sont indépendants, donc :

$$V\left(\widehat{p}_N\right)=a^2V\left(T_n\right)+b^2V\left(T'_m\right)=p\left(1-p\right)\left[\frac{a^2}{n}+\frac{(1-a)^2}{m}\right]$$

L'estimateur optimal est celui de variance minimale ; on doit donc minimiser la fonction de a :

$$f\left(a\right)=\frac{a^2}{n}+\frac{(1-a)^2}{m}$$

C'est une fonction de dérivée :

$$f'\left(a\right)=\frac{2a}{n}-\frac{2\left(1-a\right)}{m}$$

La dérivée s'annule pour $a=\dfrac{n}{n+m}$ avec $f''\left(a\right)=\dfrac{2}{n}+\dfrac{2}{m}>0$ donc l'estimateur sans biais optimal est :

$$\widehat{p}_N=\frac{nT_n+mT'_m}{N}$$

6 ▸ a) Nous calculons l'espérance de cette loi :

$$E(X) = \frac{2}{\theta^2}\int_0^\theta x^2 dx = \frac{2\theta}{3}$$

Pour calculer la variance, nous avons besoin de :

$$E\left(X^2\right) = \frac{2}{\theta^2}\int_0^\theta x^3 dx = \frac{\theta^2}{2}$$

D'où :

$$V(X) = E\left(X^2\right) - E^2(X) = \frac{\theta^2}{18}$$

b) En égalant $E(X)$ avec la moyenne empirique $\overline{X}_n$, on obtient comme solution $\theta = \frac{3}{2}\overline{X}_n$. L'estimateur obtenu par la méthode des moments est donc :

$$T_n = \frac{3}{2}\overline{X}_n$$

Cet estimateur est sans biais, $E(T_n) = \frac{3}{2}E\left(\overline{X}_n\right) = \frac{3}{2}E(X) = \theta$, et convergent, car d'après la loi des grands nombres :

$$\overline{X}_n \underset{p}{\rightarrow} E(X)$$

d'où on déduit en multipliant par $\frac{3}{2}$:

$$T_n = \frac{3}{2}\overline{X}_n \underset{p}{\rightarrow} \frac{3}{2}E(X) = \theta$$

On peut également établir la convergence de cet estimateur sans biais en montrant que sa variance tend vers 0 :

$$V(T_n) = V\left(\frac{3}{2}\overline{X}_n\right) = \frac{9}{4}V\left(\overline{X}_n\right) = \frac{9V(X)}{4n} \rightarrow 0$$

L'ensemble des valeurs possibles pour la variable est ici $X(\Omega) = [0, \theta]$ qui dépend du paramètre à estimer. L'inégalité FDCR n'est donc pas vérifiée et la question de l'efficacité ne se pose pas.

c) La f.r. F de X est nulle pour $x \leqslant 0$, a pour valeur 1 pour $x \geqslant \theta$ et a pour expression, quand $0 \leqslant x \leqslant \theta$:

$$F(x) = \frac{2}{\theta^2}\int_0^x t dt = \frac{x^2}{\theta^2}$$

On calcule ensuite la f.r. de M_n :

$$G(x) = P(M_n < x) = P\left\{\bigcap_{i=1}^{n}(X_i < x)\right\} = \prod_{i=1}^{n} P(X_i < x) = F^n(x)$$

On obtient la densité de M_n par dérivation de sa f.r. :

$$g(x) = G'(x) = nF^{n-1}(x) f(x) = 2n\frac{x^{2n-1}}{\theta^{2n}} \qquad 0 \leqslant x \leqslant \theta$$

Ainsi :

$$E(M_n) = \frac{2n}{\theta^{2n}}\int_0^\theta x^{2n}dx = \frac{2n}{2n+1}\theta$$

La valeur de $E(M_n)$ montre que l'estimateur sans biais est :

$$T'_n = \frac{2n+1}{2n}M_n$$

Pour les mêmes raisons que dans la question précédente, la question de l'efficacité ne se pose pas. On calcule ensuite :

$$E\left(M_n^2\right) = \frac{2n}{\theta^{2n}}\int_0^\theta x^{2n+1}dx = \frac{2n}{2n+2}\theta^2$$

D'où la variance :

$$V(M_n) = E\left(M_n^2\right) - E^2(M_n) = \frac{n}{(n+1)(2n+1)^2}\theta^2$$

La variance de cet estimateur est :

$$V\left(T'_n\right) = V\left(\frac{2n+1}{2n}M_n\right) = \frac{(2n+1)^2}{4n^2}V(M_n) = \frac{\theta^2}{4n(n+1)}$$

Nous allons comparer ces deux estimateurs sans biais en formant le rapport de leurs variances :

$$\frac{V(T'_n)}{V(T_n)} = \frac{2}{n+1} \to 0$$

L'estimateur T'_n est donc infiniment plus efficace que T_n.

7 ▸ a) On a $E(X^2) = \theta$, donc l'estimateur T_n de θ par la méthode des moments est la moyenne empirique :

$$T_n = \overline{X_n^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Cet estimateur est sans biais :

$$E(T_n) = E\left(X^2\right) = \theta$$

Il est aussi convergent d'après la loi des grands nombres :

$$\overline{X_n^2} \underset{p}{\rightarrow} E(X^2) = \theta$$

Pour savoir s'il est efficace, on détermine d'abord l'expression de la vraisemblance :

$$L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta}} \exp -\frac{x_i^2}{2\theta}$$

$$= \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp -\frac{1}{2\theta}\sum_{i=1}^{n} x_i^2$$

La log-vraisemblance est donc :

$$\ln L(x_1, \ldots, x_n; \theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^{n} x_i^2$$

Soit en dérivant :

$$\frac{\partial \ln L}{\partial \theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2$$

On dérive une nouvelle fois :

$$\frac{\partial^2 \ln L}{\partial \theta^2} = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n} x_i^2$$

On calcule alors la quantité d'information de Fisher :

$$I_n(\theta) = E\left(-\frac{\partial^2 \ln L}{\partial \theta^2}\right) = -\frac{n}{2\theta^2} + \frac{1}{\theta^3}\sum_{i=1}^{n} E\left(X_i^2\right) = \frac{n}{2\theta^2}$$

La v.a. X^2/θ suit une loi du khi-deux à 1 degré de liberté, donc $V(X^2) = 2\theta^2$ et :

$$V(T_n) = \frac{2\theta^2}{n} = \frac{1}{I_n(\theta)}$$

L'estimateur T_n est donc efficace.

b) La v.a. nT_n/θ suit une loi du khi-deux à n degrés de liberté. Le paramètre θ mesurant un degré d'imprécision, on choisit un intervalle unilatéral défini par :

$$1 - \alpha = P\left\{\frac{nT_n}{\theta} > a\right\}$$

Application : $T_{50} = 2$ et $a = 34{,}76$ d'où l'intervalle $\theta < \dfrac{100}{a} = 2{,}88$.

Mai 2006

1 ▸ a) L'intervalle de confiance est unilatéral à gauche en raison de l'interprétation du paramètre. La taille d'échantillon ($n = 150$) permet d'utiliser le théorème central limite :

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \underset{\text{loi}}{\rightarrow} N(0,1)$$

On utilise cette loi asymptotique pour trouver la valeur approchée de la constante u telle que :

$$0{,}95 = P\left(p < \widehat{p}_n + u\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}\right)$$

On a ici $\widehat{p}_{150} = 4\ \%$ et on lit dans la table 1 $u = 1{,}645$ d'où $p < 7\ \%$.

b) On doit utiliser l'abaque car $n = 30$ est trop petit ; on lit $p < 8\ \%$. L'estimation de p est ici de 0 % contre 4 % dans le cas précédent. Cependant, cela conduit malgré tout à un intervalle de confiance plus grand car la taille d'échantillon est très petite. On ne peut pas conclure à une meilleure qualité de la nouvelle machine.

2 ▸ a) L'expression de la vraisemblance est :

$$L(x_1, ..., x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \exp - \sum_{i=1}^{n}(x_i - \theta) = e^{n\theta}\exp\left(-\sum_{i=1}^{n} x_i\right)$$

si $\theta \leqslant x_i$ pour $1 \leqslant i \leqslant n$. Cette fonction de θ est croissante pour $\theta \leqslant \min\{x_1, ..., x_n\}$ puis nulle pour $\theta > \min\{x_1, ..., x_n\}$ (voir figure 9.3) :

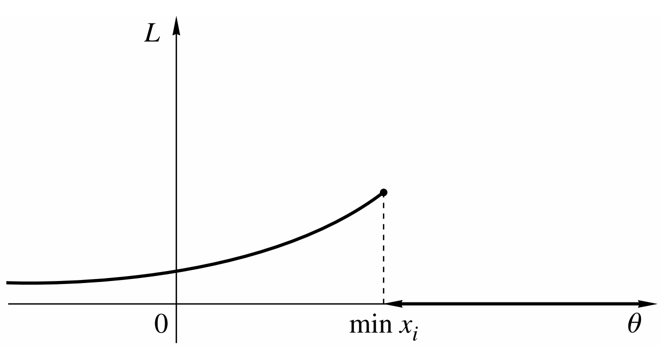


Figure 9.3

Il y a un maximum au point de discontinuité de la vraisemblance et l'estimateur du maximum de vraisemblance est donc $\widehat{\theta}_n = \min\{X_1, ..., X_n\}$.

b) La question de l'efficacité ne se pose pas car l'ensemble des valeurs possibles pour X dépend de θ. On peut cependant calculer la quantité d'information de Fisher :

$$\frac{\partial \ln L}{\partial \theta} = n$$

$$I_n(\theta) = E\left(\frac{\partial \ln L}{\partial \theta}\right)^2 = n^2$$

3 ▸ a) La v.a. Y étant la moyenne de 100 variables indépendantes d'espérance m et d'écart type σ, on peut considérer qu'elle suit la loi asymptotique déduite du théorème central limite, c'est-à-dire qu'elle suit une loi normale d'espérance m et d'écart type $\sigma/10$. La moyenne empirique $\overline{Y}_9$ suit donc une loi normale d'espérance m et d'écart type divisé par $\sqrt{9}$, soit $\sigma/30$.

b) La moyenne empirique $\overline{Y}_9$ est un estimateur sans biais de m. La question de la convergence ne se pose pas puisque la taille d'échantillon est ici fixée à 9 ; cependant il s'agit quand même d'un bon estimateur car cela correspond à 900 observations individuelles. La variance empirique modifiée :

$$S_9^2 = \frac{1}{8}\sum_{i=1}^{9}\left(Y_i - \overline{Y}_9\right)^2$$

est un estimateur sans biais de la variance de Y qui vaut $\sigma^2/100$.

c) Pour construire un intervalle de confiance bilatéral pour m, on utilise la statistique :

$$\sqrt{9}\frac{\overline{Y}_9 - m}{S_9}$$

qui suit une loi de Student à 8 degrés de liberté. On peut donc trouver la valeur t telle que :

$$1 - \alpha = P\left(-t < \sqrt{9}\frac{\overline{Y}_9 - m}{S_9} < t\right)$$

L'intervalle de confiance est alors :

$$\overline{Y}_9 - \frac{t}{3}S_9 < m < \overline{Y}_9 + \frac{t}{3}S_9$$

Pour $1 - \alpha = 0{,}95$ le fractile est $t = 2{,}306$ et on obtient l'intervalle :

$$289{,}5 < m < 312{,}5$$

Janvier 2006

1 ▸ Avec des notations évidentes, les hypothèses se traduisent par :

$$P(MY) = \frac{1}{10} \quad P(PR) = \frac{2}{10} \quad P(BV) = \frac{7}{10}$$

$$P(MT|MY) = \frac{9}{10} \quad P(MT|PR) = \frac{5}{10} \quad P(MT|BV) = \frac{1}{20}$$

On en déduit :

$$P(BV|MT) = \frac{P(BV \cap MT)}{P(MT)} = \frac{\frac{7}{10} \times \frac{1}{20}}{\frac{1}{10} \times \frac{9}{10} + \frac{2}{10} \times \frac{5}{10} + \frac{7}{10} \times \frac{1}{20}} = \frac{7}{45}$$

2 ▸ a) La fonction de répartition de X est nulle pour $x \leqslant 0$. Son expression pour $0 < x \leqslant 1$ est :

$$F(x) = \int_0^x \frac{t}{2} dt = \frac{x^2}{4}$$

Pour $1 \leqslant x \leqslant 2$:

$$F(x) = F(1) + \int_1^x \frac{1}{2} dt = \frac{2x-1}{4}$$

Pour $2 \leqslant x \leqslant 3$:

$$F(x) = F(2) + \int_2^x \frac{3-t}{2} dt = \frac{-x^2 + 6x - 5}{4}$$

On a ensuite $F(x) = 1$ pour $x > 3$.

b) On obtient :

$$E(X) = \int_0^1 \frac{x^2}{2} dx + \int_1^2 \frac{x}{2} dx + \int_2^3 \frac{3x - x^2}{2} dx = \frac{3}{2}$$

3 ▸ On pose $m = E(X)$ et $\sigma^2 = V(X)$ et on considère la variable centrée et réduite ; les conditions s'écrivent :

$$P\left(\frac{X-m}{\sigma} < \frac{3-m}{\sigma}\right) = 1 - 0{,}8413 \quad P\left(\frac{X-m}{\sigma} < \frac{12-m}{\sigma}\right) = 0{,}9772$$

La table 1 permet d'obtenir $\frac{3-m}{\sigma} = -1$ et $\frac{12-m}{\sigma} = 2$ d'où $m = 6$ et $\sigma = 3$. On en déduit :

$$P(X > 0) = P\left(\frac{X-6}{3} > -2\right) = P\left(\frac{X-6}{3} < 2\right) = 0{,}9772$$

La v.a. $\left(\frac{X-6}{3}\right)^2$ suit une loi du khi-deux à 1 degré de liberté ; la condition :

$$P\left\{\left(\frac{X-6}{3}\right)^2 < \frac{a}{9}\right\} = 0{,}975$$

permet d'obtenir, à partir de la table 5, $a = 45{,}18$.

4 ▸ La loi du couple (X, Y) est donnée dans le tableau suivant, avec les lois de X et Y dans les marges :

$Y \backslash X$	0	1	2	
1	1/4	0	1/4	1/2
2	0	1/2	0	1/2
	1/4	1/2	1/4	1

Avec par exemple $P(X = 1, Y = 2) = 2 \times \left(\frac{1}{2}\right)^2$. On calcule à partir du tableau $E(X) = 1, E(Y) = \frac{3}{2}$ et $E(XY) = \frac{3}{2}$. Ainsi $cov(X, Y) = 0$ et les variables X et Y sont cependant dépendantes, car par exemple $P(X = 1, Y = 1) \neq P(X = 1)\,P(Y = 1)$.

Septembre 2005

1 ▸ La construction d'un arbre permet d'obtenir aisément $\Omega = \{RR,\ RNR,\ RNNR,\ NNRR,\ NRNR,\ NRR\}$, avec :

$$P(RR) = \frac{3}{10} \quad P(RNR) = \frac{1}{5} \quad P(RNNR) = \frac{1}{10}$$
$$P(NNRR) = \frac{1}{10} \quad P(NRNR) = \frac{1}{10} \quad P(NRR) = \frac{1}{5}$$

On peut vérifier que la somme de ces probabilités est bien égale à 1.

2 ▸ La f.r. F de X est nulle pour $x \leqslant 0$, a pour valeur 1 pour $x \geqslant 1$ et a pour expression, quand $0 \leqslant x \leqslant 1$:

$$F(x) = 2\int_0^x (1-t)\,dt = 2\left[t - \frac{t^2}{2}\right]_0^x = 2x - x^2$$

On calcule ensuite la f.r. de M_n :

$$G(x) = P(M_n < x) = P\left\{\bigcap_{i=1}^{n}(X_i < x)\right\} = \prod_{i=1}^{n} P(X_i < x) = F^n(x)$$

On obtient la densité de M_n par dérivation de sa f.r. :

$$g(x) = G'(x) = nF^{n-1}(x)f(x) = 2n(1-x)\left(2x - x^2\right)^{n-1} \qquad 0 \leqslant x \leqslant 1$$

3 ▸ La moyenne $m = E(X)$ étant le centre de symètrie de la loi normale, on en déduit $m = -1$. On pose $\sigma^2 = V(X)$ et on considère la variable centrée et réduite ; l'autre condition s'écrit :

$$P\left(-\frac{0{,}6}{\sigma} < \frac{X+1}{\sigma} < \frac{0{,}6}{\sigma}\right) = 2\Phi\left(\frac{0{,}6}{\sigma}\right) - 1 = 0{,}770$$

La table 1 permet d'obtenir $\sigma = 1/2$. On en déduit :

$$P(X > 0) = P(2X + 2 > 2) = 0{,}0228$$

La v.a. $4(X+1)^2$ suit une loi du khi-deux à 1 degré de liberté ; la condition :

$$P\left\{4X^2 + 8X + 4 < 4a + 4\right\} = 0{,}995$$

permet d'obtenir, à partir de la table 5, $a = 0{,}97$.

4 ▸ Les densités marginales de X et Y s'obtiennent par intégration de la densité du couple (voir figure 9.4) :

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_0^x e^{y-x}dy = 1 - e^{-x} \quad \text{si } 0 \leqslant x \leqslant 1$$

$$f_X(x) = \int_0^1 e^{y-x}dy = (e-1)\,e^{-x} \quad \text{si } 1 \leqslant x$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \int_y^{+\infty} e^{y-x}dx = 1 \quad \text{si } 0 \leqslant y \leqslant 1$$

On constate que $f(x,y) \neq f_X(x)f_Y(y)$ donc les v.a. X et Y ne sont pas indépendantes.

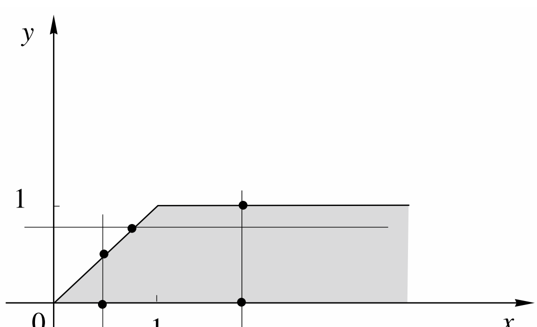


Figure 9.4

5 ▸ a) La v.a. X suit une loi géométrique de paramètre $1/\theta$:

$$P(X=k)=\frac{1}{\theta}\left(1-\frac{1}{\theta}\right)^{k-1} \qquad k\in\mathbb{N}^*$$

avec $E(X)=\theta$.

b) En égalant $E(X)$ avec la moyenne empirique $\overline{X}_n$, on obtient comme solution $\theta=\overline{X}_n$. L'estimateur obtenu par la méthode des moments est donc $T_n=\overline{X}_n$. Cet estimateur est sans biais, $E(T_n)=E\left(\overline{X}_n\right)=E(X)=\theta$, et convergent d'après la loi des grands nombres.

6 ▸ Pour construire un intervalle de confiance bilatéral pour θ, on utilise la statistique T_{n-1} qui suit une loi de Student à $n-1$ degrés de liberté. On peut donc trouver la valeur t telle que :

$$1-\alpha=P\left(-t<\sqrt{n}\frac{\overline{X}_n-\theta}{S_n}<t\right)$$

où t est le fractile d'ordre $1-\dfrac{\alpha}{2}$. L'intervalle de confiance est alors :

$$\overline{X}_n-t\frac{S_n}{\sqrt{n}}<\theta<\overline{X}_n+t\frac{S_n}{\sqrt{n}}$$

Pour $1-\alpha=0{,}95$ et $n=30$ le fractile est $t=2{,}045$.

7 ▸ L'expression de la vraisemblance est :

$$L(x_1,\ldots,x_n;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}x_i^{-1/\theta-1}=\left(\frac{1}{\theta}\right)^n\left(\prod_{i=1}^{n}x_i\right)^{-1/\theta-1}$$

La log-vraisemblance est :

$$\ln L(x_1,\ldots,x_n;\theta)=-n\ln\theta-\left(\frac{1}{\theta}+1\right)\sum_{i=1}^{n}\ln x_i$$

On obtient comme dérivée :

$$\frac{\partial\ln L}{\partial\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}\ln x_i$$

Elle s'annule pour $\theta=\dfrac{1}{n}\sum_{i=1}^{n}\ln x_i$. La dérivée seconde est :

$$\frac{\partial^2\ln L}{\partial\theta^2}=\frac{n}{\theta^2}-\frac{2}{\theta^3}\sum_{i=1}^{n}\ln x_i$$

Au point qui annule la dérivée première, la dérivée seconde est négative, donc cela correspond à un maximum et l'estimateur du maximum de vraisemblance est :

$$T_n = \frac{1}{n}\sum_{i=1}^{n} \ln X_i$$

Nous avons $E(T_n) = E(\ln X)$, donc nous calculons :

$$E(\ln X) = \frac{1}{\theta}\int_1^{+\infty} x^{-1/\theta-1}\ln x dx$$

Soit en intégrant par parties :

$$E(\ln X) = \left[-x^{-1/\theta}\ln x\right]_1^{+\infty} + \int_1^{+\infty} x^{-1/\theta-1}dx = \theta$$

Cet estimateur est donc sans biais et convergent d'après la loi des grands nombres :

$$T_n \underset{p}{\rightarrow} E(\ln X) = \theta$$

Pour savoir s'il est efficace on calcule alors la quantité d'information de Fisher :

$$I_n(\theta) = E\left(-\frac{\partial^2 \ln L}{\partial\theta^2}\right) = -\frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^{n} E(\ln X_i) = \frac{n}{\theta^2}$$

Il faut maintenant calculer :

$$E(\ln X)^2 = \frac{1}{\theta}\int_1^{+\infty} x^{-1/\theta-1}(\ln x)^2 dx$$

Soit en intégrant par parties :

$$E(\ln X)^2 = \left[-x^{-1/\theta}(\ln x)^2\right]_1^{+\infty} + 2\int_1^{+\infty} x^{-1/\theta-1}\ln x dx = 2\theta E(\ln X) = 2\theta^2$$

Ainsi $V(\ln X) = \theta^2$ et :

$$V(T_n) = \frac{V(\ln X)}{n} = \frac{\theta^2}{n} = \frac{1}{I_n(\theta)}$$

donc l'estimateur est efficace.

Mai 2005

1 ▸ **a)** L'intervalle de confiance unilatéral obtenu avec l'abaque 1 est $p > 93\ \%$.

On peut ici utiliser le théorème central limite :

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \underset{\text{loi}}{\rightarrow} N(0,1)$$

On utilise cette loi asymptotique pour trouver la valeur approchée de la constante u telle que :

$$0{,}95 = P\left(\widehat{p}_n - u\sqrt{\frac{\widehat{p}_n\left(1-\widehat{p}_n\right)}{n}} < p\right)$$

On a ici $\widehat{p}_{250} = 0{,}96$ et on lit dans la table 1 $u = 1{,}645$ d'où $p > 94\ \%$, résultat très proche du résultat exact précédent.

b) On doit utiliser l'abaque car $n = 30$ est trop petit ; on lit sur l'abaque 1 $p > 90\ \%$ avec pour cet échantillon $\widehat{p}_{30} = 1$. Bien que l'estimation soit plus élevée que dans le cas précédent, cela conduit malgré tout à un intervalle de confiance plus grand car la taille d'échantillon est très petite. On ne peut pas conclure à une meilleure qualité du nouvel équipement.

2 ▸ a) On considère que le poids affiché peut s'écrire $X = 100 + \varepsilon$, où on fera l'hypothèse que l'erreur de mesure ε suit une loi normale centrée. La variance notée θ de cette loi sera un bon indicateur de qualité ; moins les mesures affichées seront dispersées autour de la vraie valeur, meilleur sera le pèse-lettre. On retient comme estimateur sans biais de cette variance théorique la variance empirique :

$$T_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - 100)^2$$

La v.a. nT_n/θ suit une loi du khi-deux à n degrés de liberté, donc $E(T_n) = \theta$ et $V(T_n) = 2\theta^2/n$.

b) Le paramètre mesurant un degré d'imprécision, on choisit un intervalle unilatéral défini par :

$$1 - \alpha = P\left\{\frac{nT_n}{\theta} > a\right\}$$

Pour $n = 30$ on lit $a = 18{,}49$ et on obtient comme estimation :

$$T_{30}^{A} = \frac{1}{30}(315400 - 200 \times 3020) + 100^2 = 380$$

d'où l'intervalle $\theta < 617$.

c) Pour $n = 50$ on lit $a = 34{,}76$ et on obtient comme estimation :

$$T_{50}^{B} = \frac{1}{50}(510000 - 200 \times 4950) + 100^2 = 400$$

d'où l'intervalle $\theta < 575$. Bien que l'estimation soit ici plus grande, l'intervalle de confiance est plus petit car la taille d'échantillon est plus grande. Le modèle B paraît donc de meilleure qualité.

3 ▸ a) La f.r. de la v.a. $Y = -\ln X$ est nulle pour $y \leqslant 0$ et a pour expression pour $y > 0$:

$$\begin{aligned} G(y) &= P(-\ln X < y) = P(\ln X > -y) = P\left(X > \mathrm{e}^{-y}\right) \\ &= 1 - F\left(\mathrm{e}^{-y}\right) \end{aligned}$$

où F est la f.r. de la v.a. X. On en déduit par dérivation la densité de Y, qui est nulle pour $y \leqslant 0$ et qui a pour expression pour $y > 0$:

$$g(y) = \mathrm{e}^{-y} f\left(\mathrm{e}^{-y}\right) = \frac{1}{\theta}\mathrm{e}^{-y/\theta}$$

On reconnaît la densité de la loi exponentielle de paramètre $1/\theta$.

b) Nous calculons l'espérance :

$$E(X) = \frac{1}{\theta}\int_0^1 x^{1/\theta} dx = \frac{1}{\theta + 1}$$

Pour calculer la variance, nous avons besoin de :

$$E\left(X^2\right) = \frac{1}{\theta}\int_0^1 x^{1/\theta+1} dx = \frac{1}{2\theta + 1}$$

D'où :

$$V(X) = E\left(X^2\right) - E^2(X) = \frac{\theta^2}{(\theta+1)^2(2\theta+1)}$$

c) L'expression de la vraisemblance est :

$$L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n \frac{1}{\theta} x_i^{1/\theta - 1} = \left(\frac{1}{\theta}\right)^n \left(\prod_{i=1}^n x_i\right)^{1/\theta - 1}$$

La log-vraisemblance est :

$$\ln L(x_1, \ldots, x_n; \theta) = -n \ln\theta + \left(\frac{1}{\theta} - 1\right)\sum_{i=1}^n \ln x_i$$

On obtient comme dérivée :

$$\frac{\partial \ln L}{\partial \theta} = -\frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^n \ln x_i$$

Elle s'annule pour $\theta = -\frac{1}{n}\sum_{i=1}^n \ln x_i$. La dérivée seconde est :

$$\frac{\partial^2 \ln L}{\partial \theta^2} = \frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^n \ln x_i$$

Au point qui annule la dérivée première, la dérivée seconde est négative, donc cela correspond à un maximum et l'estimateur du maximum de vraisemblance est :

$$T_n = -\frac{1}{n}\sum_{i=1}^{n} \ln X_i$$

Nous avons $E(T_n) = E(-\ln X) = E(Y) = \theta$ donc cet estimateur est sans biais ; il a pour variance :

$$V(T_n) = \frac{V(-\ln X)}{n} = \frac{\theta^2}{n}$$

Il est donc convergent car cette variance tend vers 0 avec $\frac{1}{n}$. Pour savoir s'il est efficace, on calcule la quantité d'information de Fisher :

$$I_n(\theta) = E\left(-\frac{\partial^2 \ln L}{\partial\theta^2}\right) = -\frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^{n} E(-\ln X_i) = \frac{n}{\theta^2}$$

Ainsi :

$$V(T_n) = \frac{\theta^2}{n} = \frac{1}{I_n(\theta)}$$

donc l'estimateur est efficace.

Janvier 2005

1 ▸ On a :

$$P(I_1 = I_0) = p \quad P(I_1 = aI_0) = 1 - p$$

$$P(I_2 = I_0) = p^2 \quad P(I_2 = aI_0) = 2p(1-p) \quad P\left(I_2 = a^2 I_0\right) = (1-p)^2$$

Plus généralement, pour tout entier k tel que $0 \leqslant k \leqslant n$:

$$P\left(I_n = a^k I_0\right) = \binom{n}{k} p^{n-k}(1-p)^k$$

2 ▸ a) On obtient $p_1 = 1 - p$ et :

$$p_2 = P(V_2 V_1) + P\left(V_2 \overline{V_1}\right) = (1-p)^2 + p^2$$

b) Pour $n \geqslant 2$:

$$\begin{aligned} p_n &= P\left(V_n V_{n-1}\right) + P\left(V_n \overline{V}_{n-1}\right) \\ &= P\left(V_{n-1}\right) P\left(V_n \,|V_{n-1}\right) + P\left(\overline{V}_{n-1}\right) P\left(V_n \,|\overline{V}_{n-1}\right) \\ &= p_{n-1}(1-p) + (1 - p_{n-1})\,p = p + (1-2p)\,p_{n-1} \end{aligned}$$

c) On obtient :

$$u_n = p_n - \frac{1}{2} = (1-2p)\,u_{n-1} = (1-2p)^{n-1} u_1 = \frac{1}{2}(1-2p)^n$$

Ainsi, pour $0 < p < 1$:

$$p_n = \frac{1}{2} + \frac{1}{2}(1-2p)^n \to \frac{1}{2} \text{ car } |1-2p| < 1$$

avec $p_n = 1$ si $p = 0$ et $p_{2k} = 1$, $p_{2k+1} = 0$ si $p = 1$.

3 ▸ **a)** On obtient $E(X) = 0$ car la loi est symétrique et :

$$V(X) = E\left(X^2\right) = \frac{2}{4}\int_0^1 \left(x^2 + 3x^4\right) dx = \frac{7}{15}$$

b) La f.r. F de X est nulle pour $x \leqslant -1$ a pour valeur 1 pour $x \geqslant 1$ et a pour expression, quand $-1 \leqslant x \leqslant 1$:

$$F(x) = \int_{-1}^{x} \left(\frac{1}{4} + \frac{3}{4}t^2\right) dt = \frac{x^3}{4} + \frac{x}{4} + \frac{1}{2}$$

La probabilité demandée se calcule par :

$$p = P\{X \in [x_1, x_2]\} = \int_{x_1}^{x_2} f(t)\, dt = F(x_2) - F(x_1)$$

On a $p = 0$ pour $x_2 < -1$ et $1 < x_1$, les autres valeurs étant indiquées dans le tableau suivant, en fonction de la position de x_1 et x_2 :

$x_1 < -1 < x_2 < 1$	$x_1 < -1 < 1 < x_2$	$-1 < x_1 < x_2 < 1$	$-1 < x_1 < 1 < x_2$
$\frac{1}{4}\left(x_2^3 + x_2 + 2\right)$	1	$\frac{1}{4}(x_2 - x_1) + \frac{1}{4}\left(x_2^3 - x_1^3\right)$	$-\frac{1}{4}\left(x_1^3 + x_1 - 2\right)$

4 ▸ On obtient $E(X) = 0$ car la loi est symétrique, $E(Y) = \frac{1}{2}$, $cov(X, Y) = E(XY) = 0$. Les variables X et Y sont cependant dépendantes, car par exemple $P(X = -1, Y = -2) = \frac{1}{8} \neq P(X = -1)\,P(Y = -2) = \frac{3}{8} \times \frac{3}{8}$.

Tables statistiques

Table 1 : Fonction de répartition de la loi normale centrée réduite

Table 2 : Fractiles de la loi normale centrée réduite

Table 3 : Loi binômiale

Table 4 : Loi de Poisson

Table 5 : Fractiles de la loi χ^2_ν

Table 6 : Fractiles de la loi de Student T_ν

Table 7 : Fractiles d'ordre 0,95 de la loi de Fisher-Snedecor

Table 7 (suite) : Fractiles d'ordre 0,975 de la loi de Fisher-Snedecor

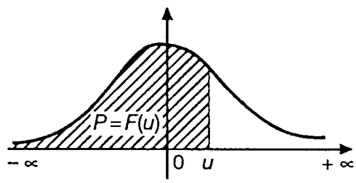
Abaque 1 : Intervalles de confiance pour une proportion (bilatéral de niveau 0,90 ou unilatéral de niveau 0,95)

Abaque 2 : Intervalles de confiance pour une proportion (bilatéral de niveau 0,95 ou unilatéral de niveau 0,975)

Table 1

Fonction de répartition de la loi normale centrée réduite

Probabilité $F(u)$ d'une valeur inférieure à u



u	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
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Tables pour les grandes valeurs de u

u	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,8	4,0	4,5
$F(u)$	0,99865	0,99904	0,99931	0,99952	0,99966	0,99976	0,999841	0,999928	0,999968	0,999997

Table 2
Fractiles d'ordre P de la loi normale centrée réduite

P	0,000	0,001	0,002	0,003	0,004	0,005	0,006	0,007	0,008	0,009	0,010	
0,00	∞	3,0902	2,8782	2,7478	2,6521	2,5758	2,5121	2,4573	2,4089	2,3656	2,3263	0,99
0,01	2,3263	2,2904	2,2571	2,2262	2,1973	2,1701	2,1444	2,1201	2,0969	2,0749	2,0537	0,98
0,02	2,0537	2,0335	,2,0141	1,9954	1,9774	1,9600	1,9431	1,9268	1,9110	1,8957	1,8808	0,97
0,03	1,8808	1,8663	1,8522	1,8384	1,8250	1,8119	1,7991	1,7866	1,7744	1,7624	1,7507	0,96
0,04	1,7507	1,7392	1,7279	1,7169	1,7060	1,6954	1,6849	1,6747	1,6646	1,6546	1,6449	0,95
0,05	1,6449	1,6352	1,6258	1,6164	1,6072	1,5982	1,5893	1,5805	1,5718	1,5632	1,5548	0,94
0,06	1,5548	1,5464	1,5382	1,5301	1,5220	1,5141	1,5063	1,4985	1,4909	1,4833	1,4758	0,93
0,07	1,4758	1,4684	1,4611	1,4538	1,4466	1,4395	1,4325	1,4255	1,4187	1,4118	1,4051	0,92
0,08	1,4051	1,3984	1,3917	1,3852	1,3787	1,3722	1,3658	1,3595	1,3532	1,3469	1,3408	0,91
0,09	1,3408	1,3346	1,3285	1,3225	1,3165	1,3106	1,3047	1,2988	1,2930	1,2873	1,2816	0,90
0,10	1,2816	1,2759	1,2702	1,2646	1,2591	1,2536	1,2481	1,2426	1,2372	1,2319	1,2265	0,89
0,11	1,2265	1,2212	1,2160	1,2107	1,2055	1,2004	1,1952	1,1901	1,1850	1,1800	1,1750	0,88
0,12	1,1750	1,1700	1,1650	1,1601	1,1552	1,1503	1,1455	1,1407	1,1359	1,1311	1,1264	0,87
0,13	1,1264	1,1217	1,1170	1,1123	1,1077	1,1031	1,0985	1,0939	1,0893	1,0848	1,0803	0,86
0,14	1,0803	1,0758	1,0714	1,0669	1,0625	1,0581	1,0537	1,0494	1,0450	1,0407	1,0364	0,85
0,15	1,0364	1,0322	1,0279	1,0237	1,0194	1,0152	1,0110	1,0069	1,0027	0,9986	0,9945	0,84
0,16	0,9945	0,9904	0,9863	0,9822	0,9782	0,9741	0,9701	0,9661	0,9621	0,9581	0,9542	0,83
0,17	0,9542	0,9502	0,9463	0,9424	0,9385	0,9346	0,9307	0,9269	0,9230	0,9192	0,9154	0,82
0,18	0,9154	0,9116	0,9078	0,9040	0,9002	0,8965	0,8927	0,8890	0,8853	0,8816	0,8779	0,81
0,19	0,8779	0,8742	0,8705	0,8669	0,8633	0,8596	0,8560	0,8524	0,8488	0,8452	0,8416	0,80
0,20	0,8416	0,8381	0,8345	0,8310	0,8274	0,8239	0,8204	0,8169	0,8134	0,8099	0,8064	0,79
0,21	0,8064	0,8030	0,7995	0,7961	0,7926	0,7892	0,7858	0,7824	0,7790	0,7756	0,7722	0,78
0,22	0,7722	0,7688	0,7655	0,7621	0,7588	0,7554	0,7521	0,7488	0,7454	0,7421	0,7388	0,77
0,23	0,7388	0,7356	0,7323	0,7290	0,7257	0,7225	0,7192	0,7160	0,7128	0,7095	0,7063	0,76
0,24	0,7063	0,7031	0,6999	0,6967	0,6935	0,6903	0,6871	0,6840	0,6808	0,6776	0,6745	0,75
0,25	0,6745	0,6713	0,6682	0,6651	0,6620	0,6588	0,6557	0,6526	0,6495	0,6464	0,6433	0,74
0,26	0,6433	0,6403	0,6372	0,6341	0,6311	0,6280	0,6250	0,6219	0,6189	0,6158	0,6128	0,73
0,27	0,6128	0,6098	0,6068	0,6038	0,6008	0,5978	0,5948	0,5918	0,5888	0,5858	0,5828	0,72
0,28	0,5828	0,5799	0,5769	0,5740	0,5710	0,5681	0,5651	0,5622	0,5592	0,5563	0,5534	0,71
0,29	0,5534	0,5505	0,5476	0,5446	0,5417	0,5388	0,5359	0,5330	0,5302	0,5273	0,5244	0,70
	0,010	0,009	0,008	0,007	0,006	0,005	0,004	0,003	0,002	0,001	0,000	*P*

Table 2 (suite)
Fractiles d'ordre P de la loi normale centrée réduite

P	0,000	0,001	0,002	0,003	0,004	0,005	0,006	0,007	0,008	0,009	0,010	
0,30	0,5244	0,5215	0,5187	0,5158	0,5129	0,5101	0,5072	0,5044	0,5015	0,4987	0,4959	0,69
0,31	0,4959	0,4930	0,4902	0,4874	0,4845	0,4817	0,4789	0,4761	0,4733	0,4705	0,4677	0,68
0,32	0,4677	0,4649	0,4621	0,4593	0,4565	0,4538	0,4510	0,4482	0,4454	0,4427	0,4399	0,67
0,33	0,4399	0,4372	0,4344	0,4316	0,4289	0,4261	0,4234	0,4207	0,4179	0,4152	0,4125	0,66
0,34	0,4125	0,4097	0,4070	0,4043	0,4016	0,3989	0,3961	0,3934	0,3907	0,3880	0,3853	0,65
0,35	0,3853	0,3826	0,3799	0,3772	0,3745	0,3719	0,3692	0,3665	0,3638	0,3611	0,3585	0,64
0,36	0,3585	0,3558	0,3531	0,3505	0,3478	0,3451	0,3425	0,3398	0,3372	0,3345	0,3319	0,63
0,37	0,3319	0,3292	0,3266	0,3239	0,3213	0,3186	0,3160	0,3134	0,3107	0,3081	0,3055	0,62
0,38	0,3055	0,3029	0,3002	0,2976	0,2950	0,2924	0,2898	0,2871	0,2845	0,2819	0,2793	0,61
0,39	0,2793	0,2767	0,2741	0,2715	0,2689	0,2663	0,2637	0,2611	0,2585	0,2559	0,2533	0,60
0,40	0,2533	0,2508	0,2482	0,2456	0,2430	0,2404	0,2378	0,2353	0,2327	0,2301	0,2275	0,59
0,41	0,2275	0,2250	0,2224	0,2198	0,2173	0,2147	0,2121	0,2096	0,2070	0,2045	0,2019	0,58
0,42	0,2019	0,1993	0,1968	0,1942	0,1917	0,1891	0,1866	0,1840	0,1815	0,1789	0,1764	0,57
0,43	0,1764	0,1738	0,1713	0,1687	0,1662	0,1637	0,1611	0,1586	0,1560	0,1535	0,1510	0,56
0,44	0,1510	0,1484	0,1459	0,1434	0,1408	0,1383	0,1358	0,1332	0,1307	0,1282	0,1257	0,55
0,45	0,1257	0,1231	0,1206	0,1181	0,1156	0,1130	0,1105	0,1080	0,1055	0,1030	0,1004	0,54
0,46	0,1004	0,0979	0,0954	0,0929	0,0904	0,0878	0,0853	0,0828	0,0803	0,0778	0,0753	0,53
0,47	0,0753	0,0728	0,0702	0,0677	0,0652	0,0627	0,0602	0,0577	0,0552	0,0527	0,0502	0,52
0,48	0,0502	0,0476	0,0451	0,0426	0,0401	0,0376	0,0351	0,0326	0,0301	0,0276	0,0251	0,51
0,49	0,0251	0,0226	0,0201	0,0175	0,0150	0,0125	0,0100	0,0075	0,0050	0,0025	0,0000	0,50
	0,010	0,009	0,008	0,007	0,006	0,005	0,004	0,003	0,002	0,001	0,000	P

Grandes valeurs de u

P	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
u_p	3,7190	4,2649	4,7534	5,1993	5,6120	5,9978

Table 3
Loi binômiale

Probabilités cumulées

n	c	Probabilités cumulées $\Pr(X \leq c) = \sum_{k=0}^{k=c} C_n^k p^k (1-p)^{n-k}$									
		p=1%	p=2%	p=3%	p=4%	p=5%	p=6%	p=7%	p=8%	p=9%	p=10%
10	0	0,9044	0,8171	0,7374	0,6648	0,5987	0,5386	0,4840	0,4344	0,3894	0,3486
	1	0,9957	0,9838	0,9655	0,9418	0,9139	0,8824	0,8483	0,8121	0,7746	0,7361
	2	0,9999	0,9991	0,9972	0,9938	0,9885	0,9812	0,9717	0,9599	0,9460	0,9298
	3	1	1	0,9999	0,9996	0,9990	0,9980	0,9964	0,9942	0,9912	0,9872
	4			1	1	0,9999	0,9998	0,9997	0,9994	0,9990	0,9984
	5					1	1	1	1	0,9999	0,9999
	6									1	1
20	0	0,8179	0,6676	0,5438	0,4420	0,3585	0,2901	0,2342	0,1887	0,1516	0,1216
	1	0,9831	0,9401	0,8802	0,8103	0,7358	0,6605	0,5869	0,5169	0,4516	0,3917
	2	0,9990	0,9929	0,9790	0,9561	0,9245	0,8850	0,8390	0,7879	0,7334	0,6769
	3		0,9994	0,9973	0,9926	0,9841	0,9710	0,9529	0,9294	0,9007	0,8670
	4		1	0,9997	0,9990	0,9974	0,9944	0,9893	0,9817	0,9710	0,9568
	5			1	0,9999	0,9997	0,9991	0,9981	0,9962	0,9932	0,9887
	6				1	1	0,9999	0,9997	0,9994	0,9987	0,9976
	7						1	1	0,9999	0,9998	0,9996
	8								1	1	0,9999
	9										1
30	0	0,7397	0,5455	0,4010	0,2939	0,2146	0,1563	0,1134	0,0820	0,0591	0,0424
	1	0,9639	0,8794	0,7731	0,6612	0,5535	0,4555	0,3694	0,2958	0,2343	0,1837
	2	0,9967	0,9783	0,9399	0,8831	0,8122	0,7324	0,6488	0,5654	0,4855	0,4114
	3	0,9998	0,9971	0,9881	0,9694	0,9392	0,8974	0,8450	0,7842	0,7175	0,6474
	4	0,9999	0,9996	0,9982	0,9937	0,9844	0,9685	0,9447	0,9126	0,8723	0,8245
	5	1	1	0,9997	0,9989	0,9967	0,9921	0,9838	0,9707	0,9519	0,9268
	6			1	0,9999	0,9994	0,9983	0,9960	0,9918	0,9848	0,9742
	7				1	0,9999	0,9997	0,9992	0,9980	0,9959	0,9922
	8					1	0,9999	0,9999	0,9996	0,9910	0,9980
	9						1	1	0,9999	0,9998	0,9995
	10								1	1	0,9999
	11										1
50	0	0,6050	0,3642	0,2181	0,1299	0,0769	0,0453	0,0266	0,0155	0,0090	0,0052
	1	0,9106	0,7358	0,5553	0,4005	0,2794	0,1900	0,1265	0,0827	0,0532	0,0338
	2	0,9862	0,9216	0,8106	0,6767	0,5405	0,4162	0,3108	0,2260	0,1605	0,1117
	3	0,9984	0,9822	0,9372	0,8609	0,7604	0,6473	0,5327	0,4253	0,3303	0,2503
	4	0,9999	0,9968	0,9832	0,9510	0,8964	0,8206	0,7290	0,6290	0,5277	0,4312
	5	1	0,9995	0,9963	0,9856	0,9622	0,9224	0,8650	0,7919	0,7072	0,6161
	6		0,9999	0,9993	0,9964	0,9882	0,9711	0,9417	0,8981	0,8404	0,7702
	7		1	0,9999	0,9992	0,9968	0,9906	0,9780	0,9562	0,9232	0,8779
	8			1	0,9999	0,9992	0,9973	0,9927	0,9834	0,9672	0,9421
	9				1	0,9998	0,9993	0,9978	0,9944	0,9875	0,9755
	10					1	0,9998	0,9994	0,9983	0,9957	0,9906
	11						1	0,9999	0,9995	0,9987	0,9968
	12							1	0,9999	0,9996	0,9990
	13								1	0,9999	0,9997
	14									1	0,9999
	15										1

Table 4
Loi de Poisson

k	Probabilités individuelles $\Pr(k) = e^{-m}\dfrac{m^k}{k!}$								
	m=0,1	m=0,2	m=0,3	m=0,4	m=0,5	m=0,6	m=0,7	m=0,8	m=0,9
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494
4		0,0003	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111
5				0,0001	0,0002	0,0004	0,0007	0,0012	0,0020
6							0,0001	0,0002	0,0003

k	Probabilités individuelles $\Pr(k) = e^{-m}\dfrac{m^k}{k!}$								
	m=1,0	m=1,5	m=2,0	m=2,5	m=3,0	m=3,5	m=4,0	m=4,5	m=5,0
0	0,3679	0,2231	0,1353	0,0821	0,0498	0,0302	0,0183	0,0111	0,0067
1	0,3679	0,3347	0,2707	0,2052	0,1494	0,1057	0,0733	0,0500	0,0337
2	0,1839	0,2510	0,2707	0,2565	0,2240	0,1850	0,1465	0,1125	0,0842
3	0,0613	0,1255	0,1804	0,2138	0,2240	0,2158	0,1954	0,1687	0,1404
4	0,0153	0,0471	0,0902	0,1336	0,1680	0,1888	0,1954	0,1898	0,1755
5	0,0031	0,0141	0,0361	0,0668	0,1008	0,1322	0,1563	0,1708	0,1755
6	0,0005	0,0035	0,0120	0,0278	0,0504	0,0771	0,1042	0,1281	0,1462
7	0,0001	0,0008	0,0034	0,0099	0,0216	0,0385	0,0595	0,0824	0,1044
8		0,0001	0,0009	0,0031	0,0081	0,0169	0,0298	0,0463	0,0653
9			0,0002	0,0009	0,0027	0,0066	0,0132	0,0232	0,0363
10				0,0002	0,0008	0,0023	0,0053	0,0104	0,0181
11					0,0002	0,0007	0,0019	0,0043	0,0082
12					0,0001	0,0002	0,0006	0,0016	0,0034
13						0,0001	0,0002	0,0006	0,0013
14							0,0001	0,0002	0,0005
15								0,0001	0,0002
16									0,0001

Table 5

Fractiles d'ordre P de la loi χ^2_ν

ν \ P	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,7	0,75	0,9	0,95	0,975	0,99	0,995	0,999
1	0,000	0,0002	0,001	0,004	0,016	0,102	0,455	1,074	1,32	2,706	3,841	5,02	6,635	7,88	10,827
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386	2,408	2,77	4,605	5,991	7,38	9,210	10,60	13,815
3	0,072	0,115	0,216	0,352	0,584	1,21	2,366	3,665	4,11	6,251	7,815	9,35	11,345	13,00	16,266
4	0,207	0,297	0,484	0,711	1,064	1,92	3,357	4,878	5,39	7,779	9,488	11,14	13,277	15,00	18,467
5	0,412	0,554	0,831	1,15	1,610	2,67	4,351	6,064	6,63	9,236	11,070	12,83	15,086	16,86	20,515
6	0,676	0,872	1,24	1,64	2,204	3,45	5,348	7,231	7,84	10,645	12,592	14,45	16,812	18,65	22,457
7	0,989	1,24	1,69	2,17	2,833	4,25	6,346	8,383	9,04	12,017	14,067	16,01	18,475	20,37	24,322
8	1,34	1,64	2,18	2,73	3,490	5,07	7,344	9,524	10,22	13,362	15,507	17,53	20,090	22,03	26,125
9	1,73	2,09	2,70	3,33	4,168	5,90	8,343	10,656	11,39	14,684	16,919	19,02	21,666	23,66	27,877
10	2,16	2,56	3,25	3,94	4,865	6,73	9,342	11,781	12,55	15,987	18,307	20,48	23,209	25,25	29,588
11	2,60	3,05	3,82	4,57	5,578	7,58	10,341	12,899	13,70	17,275	19,675	21,92	24,725	26,82	31,264
12	3,07	3,57	4,40	5,23	6,304	8,43	11,340	14,011	14,85	18,549	21,026	23,34	26,217	28,35	32,909
13	3,57	4,11	5,01	5,89	7,042	9,30	12,340	15,119	15,98	19,812	22,362	24,74	27,688	29,87	34,528
14	4,07	4,66	5,63	6,57	7,790	10,16	13,339	16,222	17,12	21,064	23,685	26,12	29,141	31,37	36,123
15	4,60	5,23	6,26	7,26	8,547	11,03	14,339	17,322	18,25	22,307	24,996	27,49	30,578	32,85	37,697
16	5,14	5,81	6,91	7,96	9,312	11,91	15,338	18,418	19,37	23,542	26,296	28,85	32,000	34,31	39,252
17	5,70	6,41	7,56	8,67	10,085	12,79	16,338	19,511	20,49	24,769	27,587	30,19	33,409	35,76	40,790
18	6,26	7,01	8,23	9,39	10,865	13,67	17,338	20,601	21,60	25,989	28,869	31,53	34,805	37,19	42,312
19	6,84	7,63	8,91	10,12	11,651	14,56	18,338	21,689	22,72	27,204	30,144	32,85	36,191	38,62	43,820
20	7,43	8,26	9,59	10,85	12,443	15,45	19,337	22,775	23,83	28,412	31,410	34,17	37,566	40,03	45,315
21	8,03	8,90	10,28	11,56	13,240	16,34	20,337	23,858	24,93	29,615	32,671	35,48	38,932	41,43	46,797
22	8,64	9,54	10,98	12,34	14,041	17,24	21,337	24,939	26,04	30,813	33,924	36,78	40,289	42,83	48,268
23	9,26	10,20	11,69	13,09	14,848	18,13	22,337	26,018	27,14	32,007	35,172	38,08	41,638	44,21	49,728
24	9,89	10,86	12,40	13,85	15,659	19,03	23,337	27,096	28,24	33,196	36,415	39,36	42,980	45,59	51,179
25	10,56	11,52	13,12	14,61	16,47	19,940	24,34	28,172	29,340	34,38	37,652	40,65	44,314	46,96	52,620
26	11,20	12,20	13,84	15,38	17,29	20,840	25,34	29,246	30,430	35,56	38,885	41,92	45,642	48,32	54,052
27	11,84	12,88	14,57	16,15	18,114	21,75	26,336	30,319	31,53	36,741	40,113	43,19	46,963	49,67	55,476
28	12,49	13,56	15,31	16,93	18,939	22,66	27,336	31,391	32,62	37,916	41,337	44,46	48,278	51,02	56,893
29	13,15	14,26	16,05	17,71	19,768	23,56	28,236	32,461	33,71	39,087	42,557	45,72	49,588	52,36	58,302
30	13,82	14,95	16,79	18,49	20,599	24,48	29,336	33,530	34,80	40,256	43,773	46,98	50,892	53,70	59,703
40	20,73	22,16	24,43	26,51	29,05	33,66	39,34	44,16	45,62	51,80	55,76	59,34	63,69	66,78	73,44
50	28,01	29,71	32,36	34,76	37,69	42,94	49,33	54,72	56,33	63,17	67,50	71,42	76,15	79,50	86,69
60	35,55	37,48	40,48	43,19	46,46	52,29	59,33	65,23	66,98	74,40	79,08	83,30	88,38	91,96	99,63
70	43,29	45,44	48,76	51,74	55,33	61,70	69,33	75,69	77,58	85,53	90,53	95,02	100,42	104,22	112,34
80	51,18	53,54	57,15	60,39	64,28	71,14	79,33	86,12	88,13	96,58	101,88	106,63	112,33	116,33	124,86
90	59,21	61,75	65,65	69,13	73,29	80,62	89,33	96,52	98,64	107,56	113,14	118,14	124,12	128,31	137,22
100	67,34	70,06	74,22	77,93	82,36	90,13	99,33	106,91	109,14	118,50	124,34	129,56	135,81	140,18	149,46

Table 6
Fractiles d'ordre P de la loi de Student T_ν

ν \ P	0,60	0,70	0,80	0,90	0,95	0,975	0,990	0,995	0,999	0,9995
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	4,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
32	0,256	0,530	0,853	1,309	1,694	2,037	2,449	2,738	3,365	3,622
34	0,255	0,529	0,852	1,307	1,691	2,032	2,441	2,728	3,348	3,601
36	0,255	0,529	0,852	1,306	1,688	2,028	2,434	2,719	3,333	3,582
38	0,255	0,529	0,851	1,304	1,686	2,024	2,429	2,712	3,319	3,566
40	0,255	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,261	3,496
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
70	0,254	0,527	0,847	1,294	1,667	1,994	2,381	2,648	3,211	3,435
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
90	0,254	0,526	0,846	1,291	1,662	1,987	2,368	2,632	3,183	3,402
100	0,254	0,526	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Table 7

Fractiles d'ordre 0,95 de la loi de Fisher-Snedecor $F(\nu_1, \nu_2)$

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	10	12	20	40	60	100	∞
1	161	200	216	225	230	234	237	239	242	244	248	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,79	8,74	8,66	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	5,96	5,91	5,80	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,74	4,68	4,56	4,46	4,43	4,41	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,06	4,00	3,87	3,77	3,74	3,71	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,64	3,57	3,44	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,35	3,28	3,15	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,14	3,07	2,94	2,83	2,79	2,76	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	2,98	2,91	2,77	2,66	2,62	2,59	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,85	2,79	2,65	2,53	2,49	2,46	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,75	2,69	2,54	2,43	2,38	2,35	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,67	2,60	2,46	2,34	2,30	2,26	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,60	2,53	2,39	2,27	2,22	2,19	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,54	2,48	2,33	2,20	2,16	2,12	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,49	2,42	2,28	2,15	2,11	2,07	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,45	2,38	2,23	2,10	2,06	2,02	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,41	2,34	2,19	2,06	2,02	1,98	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,38	2,31	2,16	2,03	1,98	1,94	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,35	2,28	2,12	1,99	1,95	1,91	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,32	2,25	2,10	1,96	1,92	1,88	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,30	2,23	2,07	1,94	1,89	1,85	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,27	2,20	2,05	1,91	1,86	1,82	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,25	2,18	2,03	1,89	1,84	1,80	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,24	2,16	2,01	1,87	1,82	1,78	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,22	2,15	1,99	1,85	1,80	1,76	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,20	2,13	1,97	1,84	1,79	1,74	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,19	2,12	1,96	1,82	1,77	1,73	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,18	2,10	1,94	1,81	1,75	1,71	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,16	2,09	1,93	1,79	1,74	1,70	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,08	2,00	1,84	1,69	1,64	1,59	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	1,99	1,92	1,75	1,59	1,53	1,48	1,39
120	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,91	1,83	1,65	1,49	1,42	1,36	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,83	1,75	1,57	1,39	1,32	1,24	1,00

Table 7 (suite)

Fractiles d'ordre 0,975 de la loi de Fisher-Snedecor $F(\nu_1, \nu_2)$

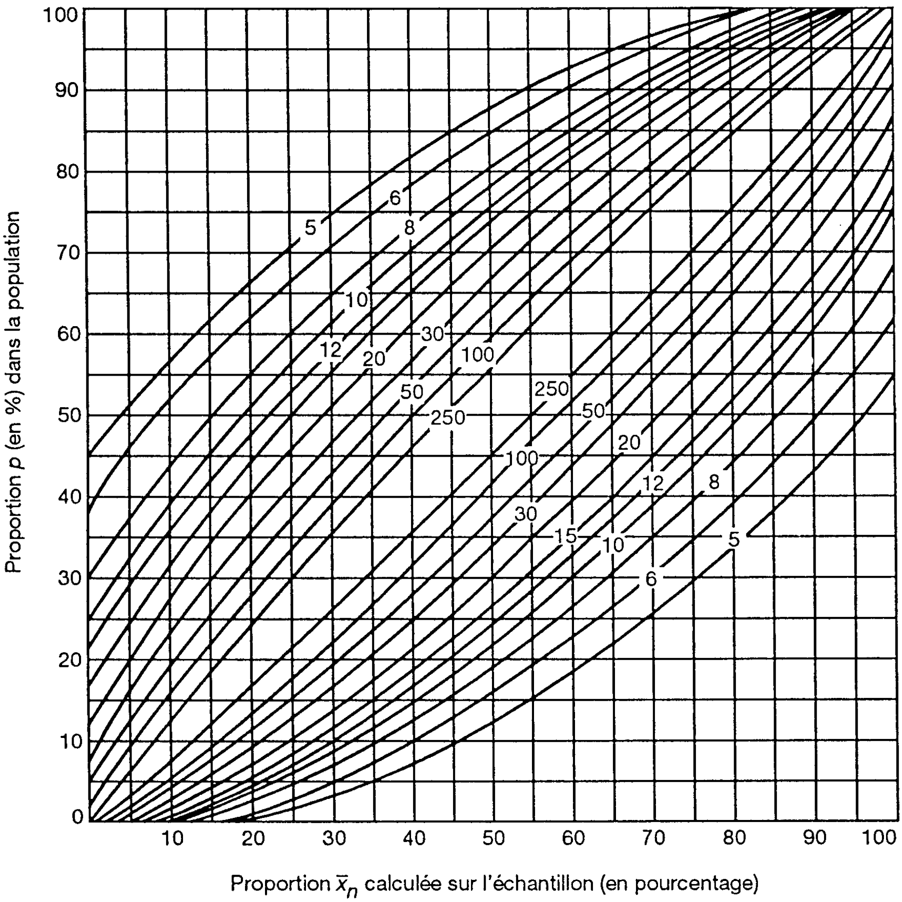
$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	10	12	20	40	60	100	∞
1	648	800	864	900	922	937	948	957	969	977	993	1006	1010	1013	1018
2	38,5	39,0	39,2	39,2	39,3	39,3	39,4	39,4	39,4	39,4	39,4	39,5	39,5	39,5	39,5
3	17,4	16,0	15,4	15,1	14,9	14,7	14,6	14,5	14,4	14,3	14,2	14,0	14,0	14,0	13,9
4	12,2	10,6	9,98	9,60	9,36	9,20	9,07	8,98	8,84	8,75	8,56	8,41	8,36	8,32	8,26
5	10,0	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,62	6,52	6,33	6,18	6,12	6,08	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,46	5,37	5,17	5,01	4,96	4,92	4,85
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,76	4,67	4,47	4,31	4,25	4,21	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,30	4,20	4,00	3,84	3,78	3,74	3,67
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	3,96	3,87	3,67	3,51	3,45	3,40	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,72	3,62	3,42	3,26	3,20	3,15	3,08
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,53	3,43	3,23	3,06	3,00	2,96	2,88
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,37	3,28	3,07	2,91	2,85	2,80	2,72
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,25	3,15	2,95	2,78	2,72	2,67	2,60
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,15	3,05	2,84	2,67	2,61	2,56	2,49
15	6,20	4,76	4,15	3,80	3,58	3,41	3,29	3,20	3,06	2,96	2,76	2,58	2,52	2,47	2,40
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	2,99	2,89	2,68	2,51	2,45	2,40	2,32
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,92	2,82	2,62	2,44	2,38	2,33	2,25
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,87	2,77	2,56	2,38	2,32	2,27	2,19
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,82	2,72	2,51	2,33	2,27	2,22	2,13
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,77	2,68	2,46	2,29	2,22	2,17	2,09
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,73	2,64	2,42	2,25	2,18	2,13	2,04
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,70	2,60	2,39	2,21	2,14	2,09	2,00
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,67	2,57	2,36	2,18	2,11	2,06	1,97
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,64	2,54	2,33	2,15	2,08	2,02	1,94
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,61	2,51	2,30	2,12	2,05	2,00	1,91
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,59	2,49	2,28	2,09	2,03	1,97	1,88
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,57	2,47	2,25	2,07	2,00	1,94	1,85
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,55	2,45	2,23	2,05	1,98	1,92	1,83
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,53	2,43	2,21	2,03	1,96	1,90	1,81
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,51	2,41	2,20	2,01	1,94	1,88	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,39	2,29	2,07	1,88	1,80	1,74	1,64
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,27	2,17	1,94	1,74	1,67	1,60	1,48
120	5,15	3,80	3,22	2,89	2,67	2,51	2,39	2,30	2,15	2,05	1,82	1,61	1,52	1,45	1,31
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,05	1,94	1,71	1,48	1,39	1,30	1,00

Abaque 1

Intervalles de confiance pour une proportion p

Intervalle bilatéral de niveau de confiance 0,90

Intervalles unilatéraux de niveau de confiance 0,95

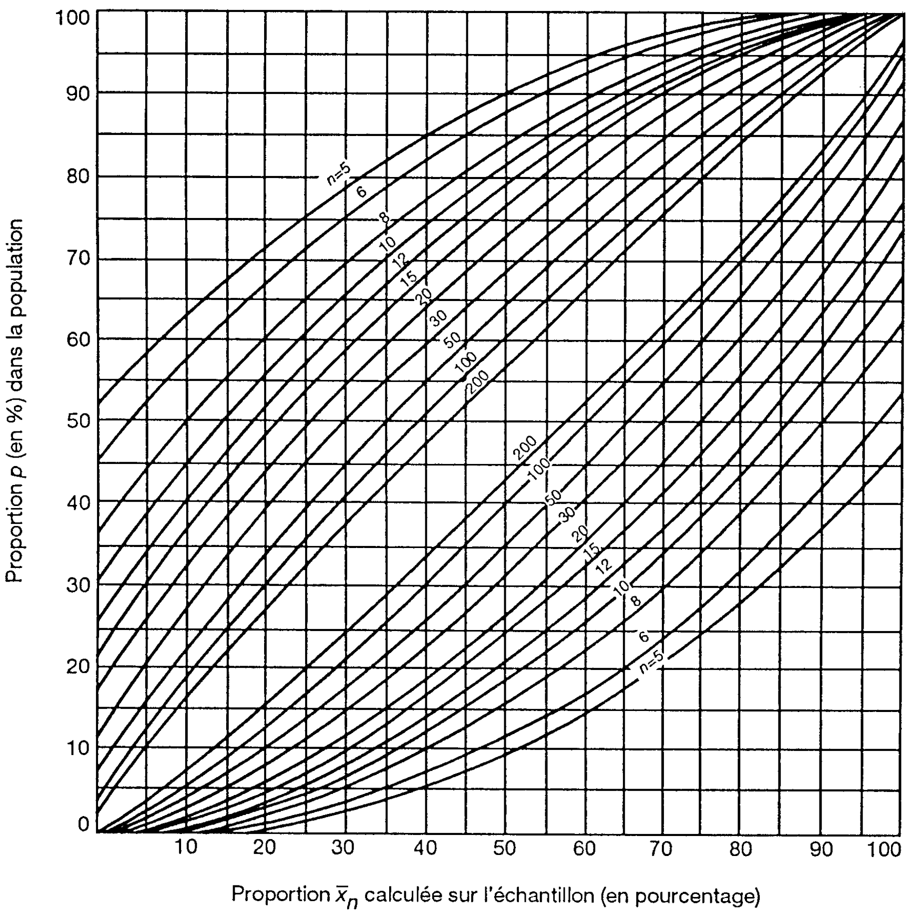


Abaque 2

Intervalles de confiance pour une proportion p

Intervalle bilatéral de niveau de confiance 0,95

Intervalles unilatéraux de niveau de confiance 0,975



Index

I

L

M

P

Q

R

TD Éco Sup

4e édition

Jean-Pierre Lecoutre

STATISTIQUE ET PROBABILITÉS

PUBLIC
- Licence Économie-Gestion, AES, MASS
- MIAGE

Les ouvrages de la collection **TD** proposent des **résumés de cours**, de **nombreux QCM**, des questions de réflexion, ainsi que des **exercices d'entraînement** et des **sujets d'annales corrigés**.

Cette **4e édition**, complétée par des sujets d'annales récents, couvre en **180 questions et exercices** les bases de la statistique et des probabilités :

- notion de probabilité ;
- variable aléatoire discrète ;
- variable aléatoire continue ;
- couple et vecteur aléatoires ;
- notions de convergence ;
- estimation ponctuelle ;
- estimation par intervalle de confiance ;
- théorie des tests.

JEAN-PIERRE LECOUTRE
Maître de conférences à l'université Panthéon-Assas (Paris II).

30% Cours

70% Applications
- QCM
- questions de réflexion
- exercices d'entraînement
- annales
+ corrigés

ISBN 978-2-10-054431-8

www.dunod.com

DUNOD