EXPRESS

Daniel FREDON
Myriam MAUMY-BERTRAND
Frédéric BERTRAND

Mathématiques

Statistique et probabilités

en 30 fiches

Comprendre et s'entraîner facilement

DUNOD

Consultez nos parutions sur dunod.com

ISBN 978-2-10-054257-4

Avant-propos

L'organisation en crédits d'enseignement entraîne des variations entre les Universités.

Les deux premières années de licence (L1 et L2) ont cependant suffisamment de points communs pour proposer des livres utiles à tous.

Avec la collection Express, vous allez vite à l'essentiel.

Pour aller vite, il faut la taille mince et le prix léger.

Il faut aussi une organisation en fiches courtes et nombreuses pour vous permettre de ne retenir que les sujets du moment, semestre après semestre.

Il faut avoir fait des choix cohérents et organisés de ce qui est le plus couramment enseigné lors des deux premières années des licences de mathématiques, informatique, mais aussi de sciences physiques et dans les cycles préparatoires intégrés.

Il faut un index détaillé pour effacer rapidement un malencontreux trou de mémoire.

Dans la collection Express, il y a donc l'essentiel, sauf votre propre travail. Bon courage !

Toutes vos remarques, vos commentaires, vos critiques, et mêmes vos encouragements seront accueillis avec plaisir.

daniel.fredon@laposte.net
mmaumy@math.u-strasbg.fr
fbertran@math.u-strasbg.fr

Table des matières

Partie 1 : Analyse combinatoire

Partie 2 : Probabilités

Partie 3 : Statistique

FICHE 1

Le langage des ensembles

I Opérations sur les ensembles

– **Réunion**. La réunion des deux ensembles A et B est notée $A \cup B$ et est définie par :

$$x \in A \cup B \Leftrightarrow (x \in A \text{ ou } x \in B).$$

– **Intersection**. L'intersection des deux ensembles A et B est notée $A \cap B$ et est définie par :

$$x \in A \cap B \Leftrightarrow (x \in A \text{ et } x \in B).$$

Deux ensembles A et B sont **disjoints** si $A \cap B = \emptyset$.

– **Partition**. Une famille $(A_i)_{i \in I}$ de parties d'un ensemble Ω est une partition de Ω si

$$\begin{cases} \bigcup_{i \in I} A_i = \Omega \\ \forall (i,j) \in I^2, (i \neq j \Rightarrow A_i \cap A_j = \emptyset). \end{cases}$$

– **Complémentaire**. Soit A une partie d'un ensemble E, le complémentaire de A dans E est noté A^c et est défini par :

$$x \in A^c \Leftrightarrow (x \in E \text{ et } x \notin A).$$

– **Différence**. Soit A et B deux parties de E, nous notons $A \setminus B$ l'ensemble défini par :

$$x \in A \setminus B \Leftrightarrow (x \in A \text{ et } x \notin B).$$

Par conséquent, nous avons l'égalité $A \setminus B = A \cap B^c$.

– **Différence symétrique**. Soit A et B deux parties de E, nous notons $A \Delta B$ l'ensemble défini par :

$$x \in A \Delta B \Leftrightarrow [x \in (A \text{ ou } B)] \text{ et } [x \notin (A \text{ et } B)].$$

Par conséquent, nous avons l'égalité $A \Delta B = (A \cup B) \setminus (A \cap B)$.

– **Produit cartésien**. Le produit cartésien des deux ensembles E et F est noté $E \times F$. Il est défini par :

$$E \times F = \{(x,y)/x \in E \text{ et } y \in F\}.$$

– **Ensemble des parties**. L'ensemble des parties d'un ensemble E, noté $\mathcal{P}(E)$, est l'ensemble de tous les sous-ensembles de E.

$$\mathcal{P}(E) = \{F | F \subset E\}.$$

– **Règles de calcul**. Soit A, B et C trois parties de E.

1. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.
2. $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.
3. $(A^c)^c = A$.
4. $(A \cup B)^c = A^c \cap B^c$.
5. $(A \cap B)^c = A^c \cup B^c$.

II Ensembles et fonctions

- **Images et images réciproques**

L'ensemble des applications d'un ensemble E vers un ensemble F est noté $\mathcal{F}(E,F)$ ou F^E. Soit f de F^E, A une partie de E et B une partie de F.

– L'**image** de A par f est l'ensemble :

$$f(A) = \{y \in F/\exists x \in A \text{ tel que } y = f(x)\}.$$

– L'**image réciproque** de B par f est l'ensemble :

$$f^{-1}(B) = \{x \in E/f(x) \in B\}.$$

– **Règles de calcul**. Soit f une application de E dans F et A et B deux parties de F.

1. $f(\emptyset) = \emptyset, f^{-1}(\emptyset) = \emptyset$.
2. $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.
3. $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.
4. $f^{-1}(A^c) = \left(f^{-1}(A)\right)^c$ où A^c est le complémentaire de A dans F et $\left(f^{-1}(A)\right)^c$ est le complémentaire de $f^{-1}(A)$ dans E.

- **Fonction caractéristique d'une partie**

Soit A une partie d'un ensemble E. La fonction caractéristique de A, ou fonction indicatrice de A, notée 1_A, est une fonction définie sur E et à valeurs dans $\{0,1\}$ par :

$$1_A(x) = 1 \text{ si } x \in A \quad \text{et} \quad 1_A(x) = 0 \text{ si } x \notin A.$$

Soit A et B deux parties d'un ensemble E et A^c le complémentaire de A dans E.

– **Inclusion**.

$$A \subseteq B \quad \Leftrightarrow \quad \forall x \in E, \quad 1_A(x) \leqslant 1_B(x).$$

– **Complémentaire**.

$$\forall x \in E, \quad 1_{A^c}(x) = 1 - 1_A(x).$$

– **Différence** $A \setminus B$.

$$\forall x \in E, \quad 1_{A \setminus B}(x) = 1_A(x) - 1_A(x)1_B(x).$$

– **Intersection**.

$$\forall x \in E, \quad 1_{A \cap B}(x) = \min(1_A(x), 1_B(x)) = 1_A(x) \cdot 1_B(x).$$

En particulier

$$\forall x \in E, \quad 1_A(x) = 1_{A \cap A}(x) = 1_A(x) \cdot 1_A(x) = 1_A(x)^2.$$

– **Réunion**.

$$\begin{aligned} \forall x \in E, \quad 1_{A \cup B}(x) &= \max(1_A(x), 1_B(x)) \\ &= 1_A(x) + 1_B(x) - 1_A(x) \cdot 1_B(x). \end{aligned}$$

– **Différence symétrique**.

$$\forall x \in E, \quad 1_{A \Delta B}(x) = 1_A(x) + 1_B(x) - 2 \cdot 1_A(x) \cdot 1_B(x).$$

Proposition : L'ensemble des parties de E, $\mathcal{P}(E)$ est en bijection avec l'ensemble des fonctions de E dans $\{0,1\}$, $\mathcal{F}(E,\{0,1\})$.

Différence symétrique

Soit X et Y deux sous-ensembles d'un ensemble E. Montrez que l'égalité suivante est vraie :

$$A\Delta B = (A \cap B^c) \cup (A^c \cap B).$$

Solution

Soit $x \in A\Delta B$. x appartient à A ou B mais pas à A et B. Ainsi si $x \in A$, x n'appartient pas à B et de ce fait $x \in A \cap B^c$.
De même, si $x \in B$, x n'appartient pas à A et de ce fait $x \in A^c \cap B$.
Par conséquent x appartient à $(A \cap B^c) \cup (A^c \cap B)$ et $A\Delta B \subset (A \cap B^c) \cup (A^c \cap B)$.

Soit $x \in (A \cap B^c) \cup (A^c \cap B)$. Si $x \in A \cap B^c$ alors x appartient à A et pas à B et si $x \in B \cap A^c$ alors x appartient à B et pas à A. Dans ces deux cas, $x \in A \cup B$ et $x \notin A \cap B$ et par conséquent $x \in A\Delta B$ et $(A \cap B^c) \cup (A^c \cap B) \subset A\Delta B$.

Les deux ensembles étudiés sont donc égaux.

Utilisation des fonctions caractéristiques

Soit X, Y et Z trois sous-ensembles d'un ensemble E. À l'aide des fonctions caractéristiques de ces ensembles, montrez que :

$$(X \setminus Y) \setminus Z = X \setminus (Y \cup Z) \text{ et } X \cap (Y \setminus Z) = (X \cap Y) \setminus (X \cap Z).$$

Solution

$$\begin{aligned} 1_{(X\setminus Y)\setminus Z} &= 1_{X\setminus Y} - 1_{X\setminus Y}1_Z \\ &= 1_X - 1_X1_Y - 1_X1_Z + 1_X1_Y1_Z \\ &= 1_X - 1_X(1_Y + 1_Z - 1_Y1_Z) \\ &= 1_X - 1_X1_{Y\cup Z} = 1_{X\setminus(Y\cup Z)}. \end{aligned}$$

Les ensembles $(X \setminus Y) \setminus Z$ et $X \setminus (Y \cup Z)$ sont égaux puisque leurs fonctions caractéristiques coïncident.

$$\begin{aligned} 1_{(X\cap Y)\setminus(X\cap Z)} &= 1_{X\cap Y} - 1_{X\cap Y}1_{X\cap Z} = 1_X1_Y - 1_X1_Y1_X1_Z \\ &= 1_X1_Y - 1_X1_Y1_Z = 1_X(1_Y - 1_Y1_Z) \\ &= 1_{X\cap(Y\setminus Z)}. \end{aligned}$$

Les ensembles $X \cap (Y \setminus Z)$ et $(X \cap Y) \setminus (X \cap Z)$ sont égaux puisque leurs fonctions caractéristiques coïncident.

FICHE 2

Analyse combinatoire

I Ensembles au plus dénombrables

• *Définitions*

– Un ensemble E est **dénombrable** s'il existe une application bijective de E dans $\mathbb{N}$.

Un ensemble dénombrable est donc infini.

– Un ensemble E est **au plus dénombrable** s'il existe une application bijective de E dans une partie de $\mathbb{N}$.

Un ensemble au plus dénombrable est donc un ensemble fini ou un ensemble dénombrable au sens de la définition précédente.

• **Caractérisations :** Soit E un ensemble.

– S'il existe une injection de E dans $\mathbb{N}$, alors E est au plus dénombrable.
– S'il existe une surjection de $\mathbb{N}$ sur E, alors E est au plus dénombrable.

• *Propriétés*

– Toute partie infinie A de $\mathbb{N}$ est en bijection avec $\mathbb{N}$.
– $\mathbb{N}$ est en bijection avec $\mathbb{N}^2$.
– Le produit cartésien d'une famille finie d'ensembles dénombrables (resp. au plus dénombrables) est dénombrable (resp. au plus dénombrables).
– La réunion d'une famille finie d'ensembles au plus dénombrables est au plus dénombrable. Si l'un des ensembles de la famille finie est dénombrable, la réunion est dénombrable.
– La réunion d'une famille dénombrable d'ensembles dénombrables (resp. au plus dénombrables) est dénombrable (resp. au plus dénombrables).
– $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{N}^2$, $\mathbb{N}^n$, $n \geqslant 1$, $\mathbb{Q}[X]$ sont dénombrables.
– $\mathbb{R}$, $\mathbb{C}$, $\mathcal{P}(\mathbb{N})$, $\{0,1\}^{\mathbb{N}}$, $\mathbb{N}^{\mathbb{N}}$ ne sont pas au plus dénombrables.

- ***Théorème de Cantor-Bernstein :*** S'il existe une injection de E dans F et une injection de F dans E, alors il existe une bijection de E dans F.

II Ensembles finis

- ***Définition :*** Un ensemble E est **fini** s'il est vide ou s'il existe un entier naturel n et une bijection de $[\![1, n]\!]$ dans E. Si $E \neq \emptyset$, n est appelé cardinal de E, noté indifféremment $\text{Card}\,(E)$ ou $|E|$. Par convention $\text{Card}\,(\emptyset) = 0$.

- ***Propriétés des cardinaux***

  – Soit E un ensemble fini. Si F est un ensemble tel qu'il existe une bijection de E dans F, alors F est un ensemble fini et $\text{Card}\,(F) = \text{Card}\,(E)$.
  – Soit E un ensemble fini. Toute partie A de E est finie et $\text{Card}\,(A) \leqslant \text{Card}\,(E)$. Une partie A de E est égale à E si et seulement si le cardinal de A est égal à celui de E.
  – Soit A et B deux parties d'un ensemble fini E et A^c le complémentaire de A dans E.
  1. Si A et B sont disjointes, $\text{Card}\,(A \cup B) = \text{Card}\,(A) + \text{Card}\,(B)$.
  2. $\text{Card}\,(A \setminus B) = \text{Card}\,(A) - \text{Card}\,(A \cap B)$.
  3. $\text{Card}\,(A^c) = \text{Card}\,(E) - \text{Card}\,(A)$.
  4. $\text{Card}\,(A \cup B) = \text{Card}\,(A) + \text{Card}\,(B) - \text{Card}\,(A \cap B)$.
  5. $\text{Card}\,(A \Delta B) = \text{Card}\,(A) + \text{Card}\,(B) - 2\text{Card}\,(A \cap B)$.

  – **Formule du crible ou de Poincaré**. Si $(A_i)_{1 \leqslant i \leqslant n}$ est une famille de parties de l'ensemble fini E, alors

$$\text{Card}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \text{Card}\,\left(A_{i_1} \bigcap \ldots \bigcap A_{i_r}\right).$$

  – Soit E et F deux ensembles finis. Nous avons :

$$\begin{aligned} \text{Card}\,(E \times F) &= \text{Card}\,(E) \cdot \text{Card}\,(F) \\ \text{Card}\,(\mathcal{F}(E,F)) &= \text{Card}\,(F^E) = \text{Card}\,(F)^{\text{Card}\,(E)}. \end{aligned}$$

  – Soit E un ensemble fini. Nous avons :

$$\text{Card}\,(\mathcal{P}(E)) = \text{Card}\,(\mathcal{F}(E,\{0,1\})) = \text{Card}\,(\{0,1\}^E) = 2^{\text{Card}\,(E)}.$$

III Dénombrement

- **Nombre de permutations de n objets**

  – Nous appelons **factorielle** la fonction, notée !, de $\mathbb{N}$ dans $\mathbb{N}$ définie par $n! = 1 \times 2 \times \cdots \times n$. Par convention 0! est égal à 1.
  – Le nombre des bijections d'un ensemble E à n éléments sur lui-même, c'est-à-dire le **nombre de permutations de n objets**, vaut $n!$.

- **Nombre d'arrangements de n objets m à m**

  – Soit E et F deux ensembles finis tels que $\text{Card}\,(E) = m \leqslant n = \text{Card}\,(F)$. Alors le nombre des applications injectives de E dans F est $A_n^m = \frac{n!}{(n-m)!} = n(n-1)\cdots(n-m+1)$.
  Nous appelons A_n^m **nombre d'arrangements de n objets m à m**.
  – A_n^m est le nombre de suites distinctes à m éléments sans répétition qu'il est possible de former avec un ensemble de n éléments.

- **Nombre de combinaisons de n objets p à p**

  – Soit E un ensemble fini de cardinal n et $p \in \mathbb{N}$ tel que $p \leqslant n$. Le nombre des parties A de E comportant p éléments est égal à $\frac{n!}{p!(n-p)!} = \frac{n(n-1)\cdots(n-p+1)}{p(p-1)\cdots 1}$. Ce nombre, noté $\binom{n}{p}$, est le **nombre des combinaisons de n objets p à p**.
  – $\binom{n}{p}$ est le nombre de choix possibles de p éléments différents pris dans une collection de n objets, indépendamment de l'ordre dans lequel les éléments sont choisis.
  – Nous avons $\binom{n}{p} = \binom{n}{n-p}$, $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = \binom{n}{n-1} = n$.
  – Le triangle de Pascal, basé sur la relation $\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$, permet un calcul récursif des $\binom{n}{p}$.
  – **Formule du binôme de Newton.**
  Soit x et y deux éléments d'un anneau tels que $x\ y = y\ x$ et $n \in \mathbb{N}^*$, alors :
  $$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$
  – Nous avons les deux égalités $\displaystyle\sum_{p=0}^{n} \binom{n}{p} = 2^n$ et $\displaystyle\sum_{p=0}^{n} (-1)^p \binom{n}{p} = 0$.

Formule du crible

Soit $E_1,\dots,E_n$ n sous-ensembles d'un ensemble fini E. Pour toute partie H de $[\![1,\ n]\!]$, nous notons $E_H = \bigcap_{i\in H} E_i$. $\mathcal{P}([\![1,\ n]\!])$ est l'ensemble des parties de l'ensemble $[\![1,\ n]\!]$. Il s'agit de démontrer la formule du crible :

$$\text{Card}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{r=1}^{n}(-1)^{r-1} \sum_{1\leqslant i_1<\ldots<i_r\leqslant n} \text{Card}\left(E_{i_1}\bigcap\ldots\bigcap E_{i_r}\right).$$

1. Soit $a_1,\dots,a_n$ n réels. En développant le produit $(X-a_1)\cdots(X-a_n)$, démontrez que l'égalité suivante est vraie :

$$(1-a_1)\cdots(1-a_n) = \sum_{H\in\mathcal{P}([\![1,\ n]\!])} (-1)^{\text{Card }(H)} \prod_{i\in H} a_i.$$

2. Déduisez-en l'expression de la fonction caractéristique $1_{\bigcup_{i=1}^{n} E_i}$ à l'aide des différentes fonctions 1_{E_H}, pour $H \in \mathcal{P}([\![1,\ n]\!]) \setminus \emptyset$.

3. Établissez alors la formule recherchée.

Solution

1. Considérons le polynôme $P(X) = (X-a_1)\cdots(X-a_n)$. Les relations coefficients-racines, voir la fiche 11 du tome d'algèbre et géométrie, permettent de relier les coefficients α_i de P, $P(X) = \sum_{i=0}^{n} \alpha_i X^i$, et les a_i. Notons σ_p la somme des produits p à p des a_i. Le coefficient α_p est égal à $(-1)^p\sigma_p$ et la somme σ_p des produits p à p des a_i peut s'écrire :

$$\sum_{H\in\mathcal{P}([\![1,\ n]\!])\text{ et Card }(H)=p}\ \prod_{i\in H} a_i.$$

Au final nous obtenons :

$$\begin{aligned}
\prod_{p=1}^{n}(1-a_p) &= \sum_{p=0}^{n}(-1)^p\sigma_p \\
&= \sum_{p=0}^{n}(-1)^p \sum_{H\in\mathcal{P}([\![1,\ n]\!])\text{ et Card }(H)=p}\ \prod_{i\in H} a_i \\
&= \sum_{p=0}^{n}\ \sum_{H\in\mathcal{P}([\![1,\ n]\!])\text{ et Card }(H)=p} (-1)^{\text{Card }(H)} \prod_{i\in H} a_i \\
&= \sum_{H\in\mathcal{P}([\![1,\ n]\!])} (-1)^{\text{Card }(H)} \prod_{i\in H} a_i.
\end{aligned}$$

2. Nous remarquons que $\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c$

$$\begin{aligned} 1_{\bigcup_{i=1}^{n} E_i} &= 1 - 1_{\left(\bigcup_{i=1}^{n} E_i\right)^c} = 1 - 1_{\bigcap_{i=1}^{n} E_i^c} \\ &= 1 - \prod_{i=1}^{n} 1_{E_i^c} = 1 - \prod_{i=1}^{n}(1 - 1_{E_i}). \end{aligned}$$

Utilisons la question **1.** pour développer le membre de droite.

$$\begin{aligned} 1_{\bigcup_{i=1}^{n} E_i} &= 1 - \sum_{H\in\mathcal{P}([\![1,n]\!])} (-1)^{\text{Card }(H)} \prod_{i\in H} 1_{E_i} \\ &= 1 - \sum_{H\in\mathcal{P}([\![1,n]\!])} (-1)^{\text{Card }(H)} 1_{\bigcap_{i\in H} E_i} \\ &= \sum_{H\in\mathcal{P}([\![1,n]\!])\setminus\emptyset} (-1)^{\text{Card }(H)-1} 1_{E_H}. \end{aligned}$$

3. Il suffit d'utiliser la relation entre cardinal et fonction indicatrice.

$$\begin{aligned} \text{Card}\left(\bigcup_{i=1}^{n} E_i\right) &= \sum_{x\in E} 1_{\bigcup_{i=1}^{n} E_i}(x) \\ &= \sum_{x\in E} \sum_{H\in\mathcal{P}([\![1,\ n]\!])\setminus\emptyset} (-1)^{\text{Card }(H)-1} 1_{E_H}(x) \\ &= \sum_{r=1}^{n}(-1)^{r-1} \sum_{1\leqslant i_1<\ldots<i_r\leqslant n} \text{Card}\left(E_{i_1}\bigcap\ldots\bigcap E_{i_r}\right). \end{aligned}$$

Nombre de surjections de $[\![1,\ n]\!]$ sur $[\![1,\ p]\!]$

1. Soit $B^i_{n,p}$ l'ensemble des applications f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ telles qu'un élément donné i n'ait pas d'antécédent par f. Quel est le cardinal de $B^i_{n,p}$?
2. Calculez le cardinal de $A^k_{n,p}$ l'ensemble des applications f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ telles qu'exactement k éléments donnés de $[\![1,\ p]\!]$ n'aient pas d'antécédent par f.
3. Soit $S_{n,p}$ l'ensemble des surjections de $[\![1,\ n]\!]$ sur $[\![1,\ p]\!]$.
Démontrez la relation :

$$\text{Card }(S_{n,p}) = p^n - \binom{p}{1}(p-1)^n + \cdots + (-1)^{p-1}\binom{p}{p-1}.$$

Solution

1. Fixons un élément i de $[\![1,\ p]\!]$. Une application f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ telle que l'élément i n'ait pas d'antécédent par f est également une application de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!] \setminus \{i\}$. Réciproquement, une application de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!] \setminus \{i\}$ définit également une seule application f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ telle que l'élément i n'ait pas d'antécédent par f. Le nombre d'applications recherchées est égal au nombre d'applications de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!] \setminus \{i\}$, c'est-à-dire à $(p-1)^n$.

2. Soit k un entier compris entre 0 et $p-1$ et B un ensemble de k éléments distincts de $[\![1,\ p]\!]$. Une application f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ telle que les éléments de B n'aient pas d'antécédent par f est également une application de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!] \setminus B$. Réciproquement, une application de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!] \setminus B$ définit également une seule application f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ telle que les éléments de B n'aient pas d'antécédent par f. Lorsque l'ensemble B est fixé à l'avance le nombre d'applications f telles que les éléments de B n'aient pas d'antécédent par f est égal au nombre d'applications de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!] \setminus B$, c'est-à-dire à $(p-k)^n$.

$$\text{Card}\ \left(A_{n,p}^k\right) = (p-k)^n.$$

3. Soit f une application de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ qui n'est pas surjective. Il existe au moins un élément de $[\![1,\ p]\!]$ qui n'a pas d'antécédent par f. Réciproquement, une application f de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$ qui appartient à l'un des $B_{n,p}^i$ n'est pas surjective. Posons $S_{n,p}^c$ le complémentaire de $S_{n,p}$ dans l'ensemble des applications de $[\![1,\ n]\!]$ dans $[\![1,\ p]\!]$. Nous avons $\text{Card}\ (S_{n,p}) = p^n - \text{Card}\ (S_{n,p}^c)$ et $S_{n,p}^c = \bigcup_{i=1}^p B_{n,p}^i$.

$$\begin{aligned}
\text{Card}\ (S_{n,p}) &= p^n - \text{Card}\ (S_{n,p}^c) = p^n - \text{Card}\left(\bigcup_{i=1}^p B_{n,p}^i\right) \\
&= p^n - \sum_{r=1}^p (-1)^{r-1} \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant p} \text{Card}\left(B_{n,p}^{i_1} \bigcap \ldots \bigcap B_{n,p}^{i_r}\right) \\
&= p^n - \sum_{r=1}^p (-1)^{r-1} \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant p} \text{Card}\ (A_{n,p}^r) \\
&= p^n - \sum_{r=1}^p (-1)^{r-1} \binom{p}{r} \text{Card}\ (A_{n,p}^r) \\
&= p^n - \sum_{r=1}^p (-1)^{r-1} \binom{p}{r} (p-r)^n \\
&= p^n - p(p-1)^n + \tbinom{p}{2}(p-2)^n - \cdots + (-1)^{p-1} p.
\end{aligned}$$

Équations en nombres entiers

Soit $n \in \mathbb{N}^*$ et $p \in \mathbb{N}$ fixés. Soit $E(n,p)$ (resp. $E'(n,p)$) l'ensemble des applications f de $[\![1,\ n]\!]$ dans $[\![0,\ p]\!]$ qui à k associe $f(k) = x_k$ et telles que

$$\sum_{1\leqslant k\leqslant n} x_k \leqslant p\left(\text{resp. } \sum_{1\leqslant k\leqslant n} x_k = p\right).$$

1. Soit A l'ensemble des applications strictement croissantes de $[\![1,\ n]\!]$ dans $[\![1, n+p]\!]$. Quel est le cardinal de A ?

2. À chaque fonction f de $E(n,p)$ nous associons la fonction g de $[\![1,\ n]\!]$ dans $[\![1, n+p]\!]$ qui à un entier k associe $y_k = k + x_1 + \cdots + x_k$.
Montrez que g est strictement croissante, puis que l'application Φ de $E(n,p)$ dans A qui à une fonction f associe $\Phi(f) = g$ est bijective.
Déduisez-en Card $(E(n,p))$.

3. Calculez Card $(E'(n,p))$.

4. Nous supposons $p \geqslant n$. Quel est le nombre de n-uplets $(a_1,\ldots,a_n)$ de nombres entiers strictement positifs tels que $a_1 + \cdots + a_n = p$?

Solution

1. Soit $X = \{x_1,\ldots,x_n\}$ une partie à n éléments de $[\![1,n+p]\!]$. Posons $Y = \{y_1,\ldots,y_n\}$ le n-uple dont les coordonnées sont les éléments de X ordonnés dans l'ordre croissant. Soit f l'application de $[\![1,\ n]\!]$ dans $[\![1,n+p]\!]$ définie par $f(i) = y_i$. Si $i < j$ alors $f(i) = y_i < y_j = f(j)$ car les y_i sont ordonnés dans l'ordre croissant et sont tous distincts sinon $n >$ Card $(Y) =$ Card $(X) = n$. Ainsi f est strictement croissante et $f \in A$. Réciproquement soit f de A. Comme f est strictement croissante, les $f(i)$, pour $1 \leqslant i \leqslant n$, sont distincts et $f([\![1,\ n]\!])$ est un sous-ensemble de $[\![1,n+p]\!]$ de cardinal n. En conclusion, A est en bijection avec les parties à n éléments de $[\![1,n+p]\!]$, un ensemble à $n+p$ éléments, et ainsi Card $(A) = \binom{n+p}{n}$.

2. Soit $f \in E(n,p)$. Pour tout $k \in [\![1,\ n-1]\!]$, $g(k+1) - g(k) = 1 + x_{k+1} > 0$ car tous les x_k sont positifs. Ainsi g est strictement croissante. De plus $g(1) = 1 + x_1 \geqslant 1$ et $g(n) = n + x_1 + \cdots + x_n \leqslant n + p$, donc $g([\![1,\ n]\!]) \subset [\![1,n+p]\!]$, par conséquent $g \in A$.
Soit f_1 et f_2 deux éléments de $E(n,p)$ tels que $\Phi(f_1) = \Phi(f_2)$. Nous avons $1 + f_1(1) = \Phi(f_1)(1) = \Phi(f_2)(1) = 1 + f_2(1)$, c'est-à-dire $f_1(1) = f_2(1)$. En évaluant $\Phi(f_1)$ et $\Phi(f_2)$ en 2, puis 3, …, et enfin n nous obtenons successivement $f_1(2) = f_2(2)$, puis $f_1(3) = f_2(3)$, …, et enfin $f_1(n) = f_2(n)$. Ainsi $f_1 = f_2$ et Φ est injective.

Soit $g \in A$. Définissons une application f de $[\![1, n]\!]$ dans $\mathbb{N}$ en posant, pour tout $k \in [\![0, n-1]\!]$, $f(k+1) = x_{k+1} = g(k+1) - g(k) - 1 \geqslant 0$ car g est à valeurs entières et strictement croissante. Enfin $g(n) \leqslant n+p$ car $g([\![1, n]\!]) \subset [\![1, n+p]\!]$ et de ce fait $n + x_1 + \cdots + x_n \leqslant n + p$ d'où $x_1 + \cdots + x_n \leqslant p$. Les x_i étant positifs, ils sont tous inférieurs à p. Ainsi $f \in E(n,p)$ et $\Phi(f) = g$: Φ est surjective de $E(n,p)$ dans A.

Φ est une bijection de $E(n,p)$ dans A. Ces deux ensembles finis ont donc le même cardinal. Card $(E(n,p)) =$ Card $(A) = \begin{pmatrix} n+p \\ n \end{pmatrix} = \begin{pmatrix} n+p \\ p \end{pmatrix}$.

3. Pour tous $x_1, \ldots, x_n$ de $[\![0, p]\!]$ tels que $x_1 + \cdots + x_n = p$, nous avons $x_1 + \cdots + x_{n-1} \leqslant p$. Réciproquement, pour tous $x_1, \ldots, x_{n-1}$ de $[\![0, p]\!]$ tels que $x_1 + \cdots + x_{n-1} \leqslant p$, posons $x_n = p - (x_1 + \cdots + x_{n-1}) \geqslant 0$, nous avons alors $x_1 + \cdots + x_n = p$. Ainsi, si $n \geqslant 2$, les solutions de $x_1 + \cdots + x_{n-1} \leqslant p$ sont en bijection avec celles de $x_1 + \cdots + x_n = p$: $E'(n,p) = E(n-1,p) = \begin{pmatrix} n+p-1 \\ p \end{pmatrix}$ $= \begin{pmatrix} n+p-1 \\ n-1 \end{pmatrix}$.

4. Commençons par associer aux entiers strictement positifs a_i, les entiers positifs ou nuls $x_i = a_i - 1$. L'équation $a_1 + \cdots + a_n = p$ devient $x_1 + 1 + \cdots + x_n + 1 = p$, c'est-à-dire $x_1 + \cdots + x_n = p - n \geqslant 0$. Cette équation admet $E'(n, p-n)$ $= \begin{pmatrix} p-1 \\ p-n \end{pmatrix} = \begin{pmatrix} p-1 \\ n-1 \end{pmatrix}$ solutions.

FICHE 3

Fonctions génératrices

- ***Définition :*** Nous appelons **fonction génératrice** d'une suite d'entiers naturels, $(a_k)_{k\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$, la série formelle notée $f(x)$ et définie par

$$f(x) = \sum_{k\in\mathbb{N}} a_k x^k = \sum a_k x^k.$$

- ***Propriétés :*** Soit $f(x)$ la fonction génératrice de $(a_0,a_1,a_2,a_3,\ldots,a_n,\ldots) \in \mathbb{N}^{\mathbb{N}}$ et $g(x)$ la fonction génératrice de $(b_0,b_1,b_2,b_3,\ldots,b_n,\ldots) \in \mathbb{N}^{\mathbb{N}}$.

– f et g sont égales si et seulement si pour entier naturel n, $a_n = b_n$.
– **Décalage vers la droite.** $(0,a_0,a_1,a_2,\ldots,a_{n-1},\ldots)$.

$$xf(x) = \sum_{k\in\mathbb{N}^*} a_{k-1}x^k.$$

– **Décalage vers la gauche**. $(a_1,a_2,a_3,a_4,\ldots,a_{n+1},\ldots)$.

$$\frac{f(x)-a_0}{x} = \sum a_{k+1}x^k.$$

– **Mise à l'échelle**. $\lambda \in \mathbb{R}$, $(a_0,\lambda a_1,\lambda^2 a_2,\lambda^3 a_3,\ldots,\lambda^n a_n,\ldots)$.

$$f(\lambda x) = \sum (\lambda^k a_k)x^k.$$

– **Addition**. $(a_0+b_0,a_1+b_1,a_2+b_2,a_3+b_3,\ldots,a_n+b_n,\ldots)$.

$$f(x)+g(x) = \sum (a_k+b_k)x^k.$$

– **Multiplication par l'indice**. $(a_1,2a_2,3a_3,4a_4,\ldots,(n+1)a_{n+1},\ldots)$.

$$f'(x) = \sum k a_k x^{k-1}.$$

– **Division par l'indice**. $(0,a_0/1,a_1/2,a_2/3,\ldots,a_{n-1}/n,\ldots)$.

$$\int_0^x f(u)\,du = \sum \frac{a_k}{k+1}x^{k+1}.$$

– **Différence.** $(a_0, a_1 - a_0, a_2 - a_1, a_3 - a_2, \ldots, a_n - a_{n-1}, \ldots)$.

$$(1-x)f(x) = a_0 + \sum (a_k - a_{k-1})x^k.$$

– **Convolution.** $(a_0b_0, a_1b_0 + a_0b_1, \ldots, \sum_{l=0}^{k} a_l b_{k-l}, \ldots)$.

$$f(x)g(x) = \sum_{k\in\mathbb{N}} \left(\sum_{l=0}^{k} a_l b_{k-l}\right) x^k.$$

– **Somme partielle.** $(a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots, \sum_{l=0}^{k} a_l, \ldots)$.

$$\frac{f(x)}{1-x} = \sum_{k\in\mathbb{N}} \left(\sum_{l=0}^{k} a_l\right) x^k.$$

- **Fonctions génératrices usuelles :** consultez la fiche sur les séries entières du tome d'analyse.

Jeu de dés

Nous lançons un dé rouge et un dé bleu à six faces.
1. Combien y a-t-il de combinaisons possibles ?
2. Développez $P(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6)^2$.
3. Considérons la somme des valeurs des dés, quel est le nombre obtenu le plus souvent ?

Solution

1. Les dés étant de deux couleurs différentes, il est possible de les distinguer. Nous obtenons ainsi un couple de résultats, chacune des coordonnées pouvant varier entre 1 et 6. Il y a donc $6 \times 6 = 36$ combinaisons possibles.

2. $P(x) = x^2 + 2x^3 + 3x^4 + 4x^5 + 5x^6 + 6x^7 + 5x^8 + 4x^9 + 3x^{10} + 2x^{11} + x^{12}$.

3. Chaque coefficient d'un monôme de degré n du polynôme précédent vaut le nombre de manières différentes qu'il est possible que la valeur de la somme des dés soit égale à n. Ainsi 7 est la somme la plus fréquente qu'il est possible d'obtenir avec deux dés à six faces.

Bonbons

1. Trouvez le coefficient de x^6 dans $(x + x^2)^4$.
2. Déterminez le nombre de façons de distribuer six bonbons à quatre enfants de telle manière que chacun en ait un ou deux.

Solution

1. $(x+x^2)^4 = x^4 + 4x^5 + 6x^6 + 4x^7 + x^8$.

2. Chaque enfant recevant un ou deux bonbons, nous lui associons le polynôme $x + x^2$. La fonction génératrice associée au problème, notée $Q(x) = \sum_{k=0}^{+\infty} q_k x^k$, est de ce fait égale à $(x+x^2)^4$. Un coefficient q_k de cette fonction est égal au nombre de façons de distribuer k bonbons à quatre enfants de telle manière que chacun en ait un ou deux. Ainsi le problème n'a de solution que si $4 \leqslant k \leqslant 8$ et alors :

Nombre de bonbons	$\leqslant 3$	4	5	6	7	8	$\geqslant 9$
Nombre de solutions	0	1	4	6	4	1	0

En particulier pour $k = 6$ bonbons il y 6 manières différentes de les répartir.

Beignets

Soit n un entier naturel inférieur ou égal à 12. Combien y a-t-il de façons de choisir n beignets de trois saveurs différentes (fraise, vanille et chocolat) s'il doit y avoir au moins deux beignets de chaque saveur et pas plus de quatre beignets au chocolat.

Solution

Utilisons une fonction génératrice pour résoudre ce problème. Cette fonction génératrice, notée $G(x)$ est le produit de trois facteurs, chacun d'entre eux correspondant aux contraintes sur le nombre de beignets d'un parfum donné.
Le facteur de la fonction génératrice correspondant au choix du nombre de beignets à la fraise est la série formelle $x^2 + x^3 + x^4 + x^5 + \cdots$.
Celui pour le choix du nombre de beignets à la vanille est la série formelle $x^2 + x^3 + x^4 + x^5 + \cdots$.
Enfin celui correspondant au choix du nombre de beignets au chocolat est le polynôme $x^2 + x^3 + x^4$ car le nombre de beignets au chocolat doit être compris entre 2 et 4.
Multiplions ces trois facteurs pour obtenir $G(x)$ qui est la fonction génératrice associée à notre problème :

$$\begin{aligned}
G(x) &= (x^2+x^3+x^4+\cdots)(x^2+x^3+x^4+\cdots)(x^2+x^3+x^4) \\
&= x^6\left(\sum_{k=0}^{+\infty} x^k\right)^2 (1+x+x^2) = x^6 \frac{1}{(1-x)^2}\frac{1-x^3}{1-x} = x^6 \frac{1-x^3}{(1-x)^3} \\
&= x^6(1-x^3)\sum_{k=0}^{+\infty}\binom{3+k-1}{k} x^k = x^6(1-x^3)\sum_{k=0}^{+\infty}\frac{(k+2)(k+1)}{2}x^k.
\end{aligned}$$

Pour un nombre total de beignets égal à n, le nombre de choix possibles respectant les contraintes de l'énoncé est le coefficient g_n de x^n dans $G(x)$. Calculons ces coefficients g_n pour $0 \leqslant n \leqslant 12$.

$$\begin{array}{rcl} G(x) & = & x^6(1-x^3)(1+3x+6x^2+10x^3+15x^4+21x^5+28x^6+\cdots) \\ & = & x^6+3x^7+6x^8+9x^9+12x^{10}+15x^{11}+18x^{12}+\cdots. \end{array}$$

Nous en déduisons le tableau suivant :

Nombre de beignets	$\leqslant 5$	6	7	8	9	10	11	12
Nombre de solutions	0	1	3	6	9	12	15	18

FICHE 4

Compléments sur les séries et les intégrales

I Quelques sommes finies classiques

– Sommes des n premiers entiers, de leurs carrés, de leurs cubes et de leurs puissances quatrièmes.

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6},$$

$$\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2 \text{ et } \sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}.$$

– Somme des premiers termes d'une suite géométrique. Soit $q \in \mathbb{C}$ et $n \in \mathbb{N}$:

$$1 + q + q^2 + \cdots + q^n = \begin{cases} \dfrac{1-q^{n+1}}{1-q} & \text{si} \quad q \neq 1 \\ n+1 & \text{si} \quad q = 1. \end{cases}$$

II Séries indexées par une partie de $\mathbb{Z}$

Soit I une partie de $\mathbb{N}$ ou de $\mathbb{Z}$ et $(x_i)_{i \in I}$ une famille de nombres réels. Pour définir la **convergence absolue de la somme** $\sum_{i \in I} x_i$ quatre cas sont à distinguer : la famille I peut être finie, minorée, majorée ou n'être ni minorée, ni majorée.

1. Si I est une partie finie de $\mathbb{Z}$ de cardinal k, $I = \{i_1, \ldots, i_k\}$, alors la somme $\sum_{i \in I} x_i = \sum_{k=1}^{n} x_{i_k}$ converge absolument.
2. Si I est une partie minorée de $\mathbb{Z}$, alors il existe n_0 tel que $I \subset [\![n_0, +\infty[\![$ et $I = \{i_k,\, k \in \mathbb{N}\}$, la somme $\sum_{i \in I} x_i$ converge absolument si la série $\sum_{k=0}^{+\infty} x_{i_k}$ converge absolument.

3. Si I est une partie majorée de $\mathbb{Z}$, alors $J = -I = \{-i,\, i \in I\}$ est une partie minorée de $\mathbb{Z}$. La somme $\sum_{i \in I} x_i$ converge absolument si $\sum_{j \in J} x_{-j}$ converge absolument.

4. Si I est une partie de $\mathbb{Z}$ ni majorée ni minorée, alors soit $I^- = \{i, i \in I \text{ et } I < 0\}$ et $I^+ = \{i,\, i \in I \text{ et } I \geqslant 0\}$. La somme $\sum_{i \in I} x_i$ converge absolument si $\sum_{i \in I^-} x_i$ et $\sum_{i \in I^+} x_i$ convergent absolument, ce qui a un sens puisque I^- est une partie de $\mathbb{Z}$ majorée et I^+ est une partie de $\mathbb{Z}$ minorée.

III Séries doubles

Définition : Soit $(u_{i,j})_{(i,j) \in I \times J}$ une famille de nombres réels positifs. Si pour tout $i \in I$, $\sum_{j \in J} u_{i,j}$ converge et est égale à a_i et si $\sum_{i \in I} a_i$ converge et a pour somme U, nous appelons $\sum_{(i,j) \in I \times J} u_{i,j}$ **série double convergente** de somme U.

Proposition : Si $\sum_{(i,j) \in I \times J} u_{i,j}$ est une série double convergente de somme U, alors pour tout $j \in J$, $\sum_{i \in I} u_{i,j}$ converge et est égale à b_j et $\sum_{j \in J} b_j$ converge et a pour somme U. Nous avons ainsi en cas de convergence $\sum_{(i,j) \in I \times J} u_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} u_{i,j} \right) = \sum_{j \in J} \left(\sum_{i \in I} u_{i,j} \right)$.

Définition : Soit $(u_{i,j})_{(i,j) \in I \times J}$ une famille de nombres réels. Nous appelons $\sum_{(i,j) \in I \times J} u_{i,j}$ **série double absolument convergente** lorsque la série double $\sum_{(i,j) \in I \times J} |u_{i,j}|$ est convergente.

Théorème de Fubini

Si $\sum_{(i,j) \in I \times J} u_{i,j}$ est une série double absolument convergente alors :

1. pour tout $i \in I$, $\sum_{j \in J} u_{i,j}$ converge et $\sum_{i \in I} \sum_{j \in J} u_{i,j}$ converge ;

2. pour tout $j \in J$, $\sum_{i \in I} u_{i,j}$ converge et $\sum_{j \in J} \sum_{i \in I} u_{i,j}$ converge ;

3. les égalités suivantes sont vérifiées :

$$\sum_{(i,j)\in I\times J} u_{i,j} = \sum_{i\in I}\left(\sum_{j\in J} u_{i,j}\right) = \sum_{j\in J}\left(\sum_{i\in I} u_{i,j}\right).$$

Théorème

Si $\sum\limits_{(i,j)\in I\times J} u_{i,j}$ est une série double absolument convergente alors pour toute partition $(A_k)_{k\in K}$ de $I\times J$:

1. pour tout $k\in K$, la série double $\sum\limits_{(i,j)\in A_k} u_{i,j}$ converge absolument ;

2. la série $\sum\limits_{k\in K}\left(\sum\limits_{(i,j)\in A_k} u_{i,j}\right)$ converge absolument ;

3. l'égalité suivante est vérifiée :

$$\sum_{(i,j)\in I\times J} u_{i,j} = \sum_{k\in K}\left(\sum_{(i,j)\in A_k} u_{i,j}\right).$$

IV Intégrales doubles et plus

- **Intégrale sur un pavé de $\mathbb{R}^n$**

Soit P un pavé compact de $\mathbb{R}^n$, $P=[a_1,b_1]\times\cdots\times[a_n,b_n]\subset\mathbb{R}^n$, avec $a_1\leqslant b_1$, $\ldots$, $a_n\leqslant b_n$. Le volume de P, noté $\mathrm{vol}(P)$, est égal à $\prod\limits_{j=1}^{n}|b_j-a_j|$.

Définitions : Soit A_i des pavés fermés et bornés contenus dans P, nous appelons **fonction en escalier sur** P, une fonction f de la forme

$$f(x)=\sum_{i}^{k}\lambda_i 1_{A_i}(x), \qquad \forall x\in P.$$

Nous appelons **intégrale sur** P **de la fonction** f **en escalier sur** P, notée $\int_P f(x)\,\mathrm{d}x$, le nombre

$$\int_P f(x)\,\mathrm{d}x = \sum_{i}^{k}\lambda_i \mathrm{vol}(A_i).$$

Définition : Nous appelons f **fonction intégrable sur** P une fonction bornée sur P et telle que, pour tout $\epsilon>0$, il existe des fonctions en escaliers ϕ et ψ définies sur

P telles que

$$\phi \leqslant f \leqslant \psi \text{ et } \int_P (\psi(x) - \phi(x))\ \mathrm{d}x < \epsilon.$$

Si f est intégrable sur P, alors $\sup\limits_{\phi \leqslant f} \int_P \phi(x)\,\mathrm{d}x = \inf\limits_{f \leqslant \psi} \int_P \psi(x)\,\mathrm{d}x$.
Cette valeur est appelée **intégrale de f sur P** et est notée

$$\int_P f(x)\,\mathrm{d}x \text{ ou } \int_P f(x_1,\ldots,x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n.$$

Théorème de Fubini

Soit f une fonction définie et continue sur $[a, b] \times [c, d] \subset \mathbb{R}^2$. Les fonctions

$$x \mapsto \int_c^d f(x,y)\,\mathrm{d}y \quad \text{et} \quad y \mapsto \int_a^b f(x,y)\,\mathrm{d}x$$

sont continues respectivement sur $[a, b]$ et $[c, d]$ et nous avons

$$\int_{[a,b]\times[c,d]} f(x,y)\,\mathrm{d}x\ \mathrm{d}y = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y.$$

- **Intégrale généralisée sur $\mathbb{R}^n$**

Définition : Soit f une fonction positive définie sur $\mathbb{R}^n$ et intégrable sur tout pavé compact de $\mathbb{R}^n$. f est appelée **fonction intégrable sur $\mathbb{R}^n$** si la borne supérieure des intégrales de f sur tout pavé compact de $\mathbb{R}^n$ est finie.

$$\int_{\mathbb{R}^n} f(x)\,\mathrm{d}x = \sup_{P \text{pavé compact} \subset \mathbb{R}^n} \int_P f(x)\,\mathrm{d}x.$$

Une fonction f définie sur $\mathbb{R}^n$ est **intégrable sur $\mathbb{R}^n$** si $|f|$ l'est.

Théorème de Fubini

Soit f une fonction définie et intégrable sur $\mathbb{R}^2$. Si pour tout x réel la fonction $y \mapsto f(x,y)$ est intégrable sur $\mathbb{R}$ et si la fonction

$$x \mapsto \int_{-\infty}^{+\infty} |f(x,y)|\,dy$$

est intégrable sur $\mathbb{R}$, alors nous avons

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Proposition : Interversion des intégrales

Soit f une fonction définie et intégrable sur $\mathbb{R}^2$. Si pour tout x réel la fonction $y \mapsto f(x,y)$ est intégrable sur $\mathbb{R}$ et pour tout y réel la fonction $x \mapsto f(x,y)$ est intégrable sur $\mathbb{R}$ et si les fonctions

$$x \mapsto \int_{-\infty}^{+\infty} |f(x,y)|\,\mathrm{d}y \text{ et } y \mapsto \int_{-\infty}^{+\infty} |f(x,y)|\,\mathrm{d}x$$

sont intégrables sur $\mathbb{R}$, alors nous avons

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\mathrm{d}y &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y.\end{aligned}$$

Sommabilité

Soit $\lambda \in \mathbb{R}$. La famille de nombre réels $\left(\dfrac{\lambda^{n+m}}{n!m!}\right)_{(n,m)\in\mathbb{N}^2}$ est-elle sommable ?

Solution

Pour tout réel λ, la série $\displaystyle\sum_{m=0}^{+\infty} \frac{\lambda^m}{m!}$ converge et sa somme vaut $\exp(\lambda)$.

Soit $a_n = \displaystyle\sum_{m=0}^{+\infty} \frac{\lambda^n}{n!}\frac{\lambda^m}{m!} = \frac{\lambda^n}{n!}\sum_{m=0}^{+\infty}\frac{\lambda^m}{m!} = \frac{\exp(\lambda)\lambda^n}{n!}$.

La série $\displaystyle\sum_{n=0}^{+\infty} a_n$ converge et sa somme est égale à $\exp(2\lambda)$.

Ainsi la famille $\left(\dfrac{\lambda^{n+m}}{n!m!}\right)_{(n,m)\in\mathbb{N}^2}$ est sommable. Nous en déduisons que

$$\sum_{(n,m)\in\mathbb{N}^2} \frac{\lambda^{n+m}}{n!m!} = \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} \frac{\lambda^{n+m}}{n!m!} = \sum_{m=0}^{+\infty}\sum_{n=0}^{+\infty} \frac{\lambda^{n+m}}{n!m!} = \exp(2\lambda).$$

Séries de nombres premiers

Pout tout $n \in \mathbb{N}$, $n \geqslant 2$, soit q_n le plus grand facteur premier de n. Si $k \in \mathbb{N}^*$, soit $\mathcal{E}_k = \{n \in \mathbb{N}^*,\ n \geqslant 2 \mid q_n = p_k\}$, où $(p_k)_{k\geqslant 1}$ désigne la suite strictement croissante des nombres premiers qui commence pour $p_1 = 2$. Soit enfin $M_k = \prod_{i=1}^{k}\left(\frac{1}{1-\frac{1}{p_i}}\right)$.

1. Montrez que, pour tout $k \in \mathbb{N}^*$, la famille $\left(\frac{1}{nq_n}\right)_{n\in\mathcal{E}_k}$ est sommable et que sa somme est $\frac{M_k}{p_k^2}$.

2. Montrez que la série $\sum_{n=2}^{+\infty}\frac{1}{nq_n}$ converge.

Solution

1. Si $n \in \mathcal{E}_k$, alors $q_n = p_k$. Comme l'application $(n_1,\ldots,n_k) \in \mathbb{N}^{k-1}\times\mathbb{N}^* \mapsto p_1^{n_1}\cdots p_k^{n_k} \in \mathcal{E}_k$ est une bijection, la famille $\left(\frac{1}{nq_n}\right)_{n\in\mathcal{E}_k}$ est sommable si et seulement si la famille $\left(\frac{1}{p_k p_1^{n_1}\cdots p_k^{n_k}}\right)_{(n_1,\ldots,n_k)\in\mathbb{N}^{k-1}\times\mathbb{N}^*}$ est sommable ce qui est le cas puisque

$$\frac{1}{p_k}\sum_{n_1=0}^{+\infty}\cdots\sum_{n_k=1}^{+\infty}\frac{1}{p_1^{n_1}}\cdots\frac{1}{p_k^{n_k}} = \frac{1}{p_k}\left(\sum_{n_1=0}^{+\infty}\frac{1}{p_1^{n_1}}\right)\cdots\left(\sum_{n_k=1}^{+\infty}\frac{1}{p_k^{n_k}}\right)$$
$$= \frac{1}{p_k}\prod_{i=1}^{k-1}\left(\frac{1}{1-\frac{1}{p_i}}\right)\frac{1}{p_k}\frac{1}{1-\frac{1}{p_k}} = \frac{M_k}{p_k^2}.$$

2. Posons $u_k = \frac{M_k}{p_k^2}$. Nous avons $\frac{p_{k+1}^3u_{k+1}^2}{p_k^3u_k^2} = \frac{p_{k+1}^3p_k^4p_{k+1}^2}{p_k^3(p_{k+1}-1)^2p_{k+1}^4} = \frac{p_kp_{k+1}}{(p_{k+1}-1)^2}$.

Si $k \geqslant 2$, alors $p_k \leqslant p_{k+1}-2$ et $p_kp_{k+1} \leqslant (p_{k+1}-2)p_{k+1} < p_{k+1}^2 - 2p_{k+1} + 1 = (p_{k+1}-1)^2$. La suite $(p_k^{3/2}u_k)_{k\geqslant 1}$ est donc bornée.

Comme $p_k \geqslant k$, nous en déduisons que $k^{4/3}u_k$ tend vers 0 lorsque k tend vers $+\infty$. Le critère de Riemann permet d'affirmer que la série $\sum_{k=2}^{+\infty}u_k$ converge. Ainsi la série $\sum_{n=2}^{+\infty}\frac{1}{nq_n}$ converge et $\sum_{n=2}^{+\infty}\frac{1}{nq_n} = \sum_{k=1}^{+\infty}\left(\sum_{n\in\mathcal{E}_k}\frac{1}{nq_n}\right) = \sum_{k=1}^{+\infty}\frac{M_k}{p_k^2}$.

Des égalités remarquables

1. Montrez que pour tout $x \in [0,1]$, $\ln(1+x) = \int_0^1 \frac{x}{1+xy}\,\mathrm{d}y$.

2. Déduisez-en la valeur de l'intégrale $I = \int_0^1 \frac{\ln(1+x)}{1+x^2}\,\mathrm{d}x$.

Solution

1. Soit $x \in [0,\ 1]$ fixé, $\int_0^1 \frac{x}{1+xy}\,\mathrm{d}y = [\ln(1+xy)]_0^1 = \ln(1+x)$.

2. Nous remarquons que $I = \int_0^1 \frac{\ln(1+x)}{1+x^2}\,\mathrm{d}x = \int_0^1 \left(\int_0^1 \frac{x}{1+xy}\,\mathrm{d}y\right) \frac{1}{1+x^2}\,\mathrm{d}x$.

Ainsi $I = \int_0^1 \left(\int_0^1 \frac{x}{(1+xy)(1+x^2)}\,\mathrm{d}y\right) \mathrm{d}x$.

Posons $f(x,y) = \frac{x}{(1+xy)(1+x^2)}$ pour tout $(x,y) \in [0,1] \times [0,1]$. f est continue sur $[0,\ 1] \times [0,\ 1]$ et par conséquent le théorème de Fubini permet d'affirmer que

$$\int_0^1 \left(\int_0^1 \frac{x}{(1+xy)(1+x^2)\,\mathrm{d}x}\right) \mathrm{d}y = \int_0^1 \left(\int_0^1 \frac{x}{(1+xy)(1+x^2)}\,\mathrm{d}y\right) \mathrm{d}x$$
$$= \int_0^1 \left(\int_0^1 \frac{y}{(1+xy)(1+y^2)}\,\mathrm{d}x\right) \mathrm{d}y.$$

La dernière égalité ne résulte que de l'échange des variables muettes d'intégration x et y.

$$2I = \int_0^1 \left(\int_0^1 \frac{x}{(1+xy)(1+x^2)}\mathrm{d}x\right) \mathrm{d}y + \int_0^1 \left(\int_0^1 \frac{y}{(1+xy)(1+y^2)}\mathrm{d}x\right) \mathrm{d}y$$
$$= \int_0^1 \left(\int_0^1 \frac{(1+xy)(x+y)\mathrm{d}y}{(1+xy)(1+x^2)(1+y^2)}\right) \mathrm{d}x$$
$$= \int_0^1 \left(\int_0^1 \frac{x\,\mathrm{d}y}{(1+x^2)(1+y^2)}\right) \mathrm{d}x + \int_0^1 \left(\int_0^1 \frac{y\,\mathrm{d}y}{(1+x^2)(1+y^2)}\right) \mathrm{d}x$$
$$= \int_0^1 \frac{\ln(1+1) - \ln(1+0)}{2(1+y^2)}\,\mathrm{d}y + \int_0^1 \frac{y(\arctan 1 - \arctan 0)}{1+y^2}\,\mathrm{d}y.$$
$$I = \frac{1}{2}\frac{\pi \ln(2)}{8} + \frac{1}{2}\frac{\pi \ln(2)}{8} = \frac{\pi \ln(2)}{8}.$$

FICHE 5

Introduction aux probabilités

I Expérience aléatoire

Certaines expériences entraînent des résultats aléatoires, c'est-à-dire qui dépendent directement du hasard. Nous les appelons des **expériences aléatoires**. Dans la suite, $\mathcal{E}$ désignera une expérience aléatoire.

II Modéliser une expérience aléatoire

En théorie des probabilités, le terme **modéliser** désigne l'opération qui consiste à associer à $\mathcal{E}$ trois objets mathématiques, notés et appelés généralement Ω, l'univers, $\mathcal{F}$, l'ensemble des événements et $\mathbb{P}$, la probabilité.

- **Univers**

  Nous appelons **univers** associé à $\mathcal{E}$ l'ensemble de tous les résultats possibles de $\mathcal{E}$. Généralement, l'univers est noté Ω.

- **Événement**

  Pour représenter les événements, il est d'usage d'utiliser la théorie des ensembles. Un **événement** est donc associé à un sous-ensemble de l'univers, l'ensemble des résultats pour lesquels l'événement est réalisé.

Écriture logique	Écriture ensembliste
le contraire de A s'est réalisé	A^c
A et B se sont réalisés	$A \cap B$
A ou B s'est réalisé	$A \cup B$
B s'est réalisé mais pas A	$B \setminus A$

Terminologie des événements :

– $\emptyset$ est appelé **événement impossible**.

– Ω est appelé **événement certain**.

– Tout singleton $\{\omega\}$, où $\omega \in \Omega$, est appelé **événement élémentaire**.

- **Ensemble des événements**

Nous appelons **ensemble des événements** associés à $\mathcal{E}$ tout sous-ensemble $\mathcal{F}$ de $\mathcal{P}(\Omega)$ vérifiant les trois propriétés suivantes :

1. $\emptyset$ et $\Omega \in \mathcal{F}$;
2. si $A \in \mathcal{F}$, alors $A^c \in \mathcal{F}$;
3. soit I une partie finie ou infinie de $\mathbb{N}$ ou de $\mathbb{Z}$. Si $(A_i)_{i \in I}$ est une suite d'événements de $\mathcal{F}$, alors $\bigcup_{i \in I} A_i \in \mathcal{F}$ et $\bigcap_{i \in I} A_i \in \mathcal{F}$.

Un ensemble d'événements vérifiant les trois propriétés ci-dessous est appelé une σ-**algèbre** ou une **tribu**.

- **Espace probabilisable**

Nous appelons **espace probabilisable** lié à $\mathcal{E}$ le couple $(\Omega, \mathcal{F})$, où Ω est l'ensemble des résultats possibles de $\mathcal{E}$ et $\mathcal{F}$ un ensemble des événements liés à $\mathcal{E}$.

- **Événements incompatibles**

Nous appelons **événements incompatibles** ou **disjoints** deux événements A et B tels que $A \cap B = \emptyset$, c'est-à-dire qu'il est impossible que A et B se réalisent simultanément.

- **Probabilité**

Nous appelons **probabilité** sur $(\Omega, \mathcal{F})$ une application $\mathbb{P}$ de $\mathcal{F}$ dans $[0,1]$ vérifiant les deux propriétés suivantes (axiomes de Kolmogorov)

1. $\mathbb{P}(\Omega) = 1$;
2. pour tout suite $(A_i)_{i \in I}$ finie ou infinie dénombrable d'événements de $\mathcal{F}$ deux à deux incompatibles, nous avons

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbb{P}(A_i),$$

c'est-à-dire que la probabilité d'un événement qui est la réunion disjointe d'événements est égale à la somme des probabilités de ces événements (σ-additivité).

Les incontournables

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

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

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

Solution

1. a) Dans le cas du lancer d'une pièce de monnaie, les résultats possibles sont « face » **OU** « pile ».
b) Dans le cas du lancer d'une pièce de monnaie, l'univers Ω est égal à {« face », « pile »}.
c) « Nous avons obtenu face » et « nous avons obtenu pile » sont deux événements associés à cette expérience aléatoire.
2. a) Un résultat possible est l'un des chiffres inscrit sur la face supérieure du dé, ou encore 1 **OU** 2 **OU** 3 **OU** 4 **OU** 5 **OU** 6.
b) Dans le cas du lancer d'un dé cubique, l'univers Ω est égal à $\{1,2,3,4,5,6\}$.
c) « Le résultat du lancer est divisible par 2 », « le résultat du lancer est divisible par 3 », « le résultat du lancer est un nombre impair » sont trois événements associés à cette expérience.
3. L'univers Ω est égal à $\{1,2,3,4,5,6\}^3$.

Pratique du langage des événements

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

Solution

L'univers Ω associé à cette expérience expérience aléatoire est défini par :

$$\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} = \{1,2,3,4,5,6\}^2.$$

Pour obtenir un nombre impair en faisant la somme de deux nombres, il faut faire la somme d'un nombre pair et d'un nombre impair. Décrivons alors les éléments de E, F et G :

$$E = [\{1, 3, 5\} \times \{2, 4, 6\}] \cup [\{2,4,6\} \times \{1, 3, 5\}]\,,$$

$$F = [\{1\} \times \{1, 2, 3, 4, 5, 6\}] \cup [\{1, 2, 3, 4, 5, 6\} \times \{1\}]$$

et

$$G = \{(1, 4), (4, 1), (2, 3), (3, 2)\}.$$

Décrivons maintenant $E \cap F$, $E \cup F$, $F \cap G$, $E \cap F^c$ et $E \cap F \cap G$:

$$\begin{aligned}
E \cap F &= [\{1\} \times \{2, 4, 6\}] \cup [\{2, 4, 6\} \times \{1\}] \\
E \cup F &= [\{1\} \times \{1, 2, 3, 4, 5, 6\}] \cup [\{1, 2, 3, 4, 5, 6\} \times \{1\}] \\
&\quad \cup [\{3, 5\} \times \{2, 4, 6\}] \cup [\{2, 4, 6\} \times \{3, 5\}] \\
F \cap G &= \{(1, 4), (4, 1)\} \\
E \cap F^c &= [\{3, 5\} \times \{2, 4, 6\}] \cup [\{2, 4, 6\} \times \{3, 5\}] \\
E \cap F \cap G &= \{(1, 4), (4, 1)\}.
\end{aligned}$$

D'après une idée du livre de D. Foata et A. Fuchs, *Calcul des probabilités*, Dunod, 2ème édition, 1998

Soit Ω l'ensemble des couples mariés d'une ville donnée. Nous considérons les événements :

– A « la femme a plus de trente ans » ;
– B « le mari a plus de trente ans » ;
– C « le mari est plus jeune que la femme ».

1. Interprétez en fonction de A, B, C l'événement « la femme a plus de trente ans mais non son mari ».

2. Décrivez en langage ordinaire les événements $A \cap B^c \cap C$, $A \setminus (A \cap C)$, $A \cap B \cap C^c$, $A \cup C$.

3. Vérifiez que $A \cap B^c \subset C$.

Solution

1. L'événement recherché est la réalisation simultanée de l'événement A « la femme a plus de trente ans » et du complémentaire de l'événement B « l'homme a plus de trente ans ». Nous avons donc : « la femme a plus de trente ans mais non son mari » $= A \cap B^c$.

2. $A \cap B^c \cap C =$ « la femme a plus de trente ans mais non son mari (qui est donc plus jeune qu'elle) ».
$A \setminus (A \cap C) =$« la femme a plus de trente ans et son mari est plus agé qu'elle ».
$A \cap B \cap C^c =$ « la femme et son mari ont plus de trente ans et la femme est plus jeune que son mari ».
$A \cup C =$ « la femme a plus de trente ans ou son mari est plus jeune qu'elle ».

3. Nous avons remarqué que $A \cap B^c \cap C =$ « la femme a plus de trente ans mais non son mari » $= A \cap B^c$. Par conséquent, $A \cap B^c \subset C$.

Encore et toujours le dé cubique

L'expérience aléatoire consiste à lancer un dé cubique dont les faces sont numérotées de 1 à 6. Si le résultat obtenu est un 6, l'expérimentateur inscrit sur sa fiche 6 et l'expérience s'arrête. Si le résultat obtenu est un nombre entre 1 et 5, l'expérimentateur l'inscrit sur sa fiche et relance le dé.

1. Soit n un entier strictement positif. Décrivez l'événement E_n traduisant le fait que « l'expérience s'arrête au n-ème lancer ».

2. Décrivez l'univers Ω associé à cette expérience aléatoire.

Solution

1. Soit n un entier strictement positif. Décrivons E_n :

$$E_n = \{(J_1, J_2, \ldots, J_{n-1}, 6) \text{ où } J_i \in \{1, 2, 3, 4, 5\} \text{ pour } 1 \leqslant i \leqslant n-1\}.$$

2. L'univers Ω est égal à

$$\left(\bigcup_{n \geqslant 1} E_n\right) \bigcup F$$

où $F = \{(J_1, J_2, \ldots) \text{ où } J_i \in \{1, 2, 3, 4, 5\} \text{ pour } i \geqslant 1\}$.

Les trois joueurs

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

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

1. Donnez une interprétation des points de Ω.

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

Solution

1. {0000...} représente l'événement « personne ne gagne », les autres configurations correspondent à des événements où α, β ou γ gagne.

2. a) Les configurations où le numéro 1 se situe à la place $3n + 1$, pour un entier n, forment l'événement où « α gagne ».
b) Les configurations où le numéro 1 se situe à la place $3n + 2$, pour un entier n, forment l'événement où « β gagne ».
c) $(A \cup B)^c$ est l'événement: « γ gagne » ou « personne ne gagne ». Les configurations où le numéro 1 se situe à la place $3n$, pour un entier n non nul, forment l'événement où « γ gagne ».

FICHE 6

Espaces probabilisés

I Espace probabilisé fini

- **Espace probabilisé fini**

Nous appelons **espace probabilisé fini** associé à une expérience aléatoire $\mathcal{E}$ le triplet $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, lorsque l'univers Ω est fini.

- ***Propriétés***

Pour tous événements A et B de $\mathcal{P}(\Omega)$, nous avons

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, d'où $\mathbb{P}(\emptyset) = 0$;
2. $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
3. si $A \subset B$, alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$;
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
5. si A et B sont incompatibles, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. Plus généralement : si $A_1, \ldots, A_n$ sont n événements deux à deux incompatibles, $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.
6. **Formule du crible ou de Poincaré :** Pour toute suite d'événements $(A_i)_{1 \leqslant i \leqslant n}$, nous avons

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) + \cdots + (-1)^{k+1} \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} \mathbb{P}\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right) + \cdots + (-1)^{n+1} \mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n).$$

- **Équiprobabilité et probabilité uniforme**

Nous disons qu'il y a **équiprobabilité** lorsque les probabilités de tous les événements élémentaires sont égales.
Dans ce cas, $\mathbb{P}$ **est la probabilité uniforme** sur $(\Omega, \mathcal{P}(\Omega))$.

Conséquence : S'il y a équiprobabilité, pour tout événement A, nous avons alors

$$\mathbb{P}(A) = \frac{\operatorname{Card}(A)}{\operatorname{Card}(\Omega)} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

II Espace probabilisé quelconque

- **Espace probabilisé**

Nous appelons **espace probabilisé** associé à une expérience aléatoire $\mathcal{E}$ le triplet $(\Omega, \mathcal{F}, \mathbb{P})$, où Ω, $\mathcal{F}$ et $\mathbb{P}$ ont été définis dans la fiche 5.

- ***Propriétés***

Pour tous événements A et B de $\mathcal{F}$, nous avons encore les propriétés 1., 2., 3., 4. précédentes, et

5. si A et B sont incompatibles, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. Si I est une partie finie ou infinie de $\mathbb{N}$ ou de $\mathbb{Z}$ et si $(A_i)_{i \in I}$ est une suite d'événements deux à deux incompatibles, alors nous avons $\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbb{P}(A_i)$.
6. La **formule du crible** est encore valable.

- ***Propriété de la limite monotone de la probabilité $\mathbb{P}$***

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.
1. Si $(A_i)_{i \in \mathbb{N}}$ est une suite croissante (au sens de l'inclusion) d'événements de $\mathcal{F}$, la suite $(\mathbb{P}(A_i))_{i \in \mathbb{N}}$ converge et
$$\lim_{n \to +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{i=0}^{+\infty} A_i\right).$$
2. Si $(A_i)_{i \in \mathbb{N}}$ est une suite décroissante (au sens de l'inclusion) d'événements de $\mathcal{F}$, la suite $(\mathbb{P}(A_i))_{i \in \mathbb{N}}$ converge et
$$\lim_{n \to +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{i=0}^{+\infty} A_i\right).$$

Application directe du cours

Calculez la probabilité de l'union de trois événements A, B et C de deux manières différentes :

1. en utilisant la formule $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;

2. en appliquant directement la formule du crible.

Solution

1. Calculons la probabilité de l'union de trois événements A, B et C. Pour cela, commençons par écrire que

$$\mathbb{P}(A \cup (B \cup C)) = \mathbb{P}(A) + \mathbb{P}(B \cup C) - \mathbb{P}(A \cap (B \cup C)).$$

Comme l'événement $A \cap (B \cup C)$ est l'union des deux événements $A \cap B$ et $A \cap C$, nous avons alors

$$\begin{aligned}\mathbb{P}(A \cap (B \cup C)) &= \mathbb{P}(A \cap B) + \mathbb{P}(A \cap C) - \mathbb{P}((A \cap B) \cap (A \cap C))\\ &= \mathbb{P}(A \cap B) + \mathbb{P}(A \cap C) - \mathbb{P}(A \cap B \cap C).\end{aligned}$$

De même, nous avons

$$\mathbb{P}(B \cup C) = \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(B \cap C).$$

En regroupant toutes ces identités, nous obtenons finalement :

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(B \cap C) - \mathbb{P}(A \cap B)\\ - \mathbb{P}(A \cap C) + \mathbb{P}(A \cap B \cap C).\end{aligned}$$

2. Calculons la probabilité de l'union de trois événements A, B et C en appliquant la formule du crible, avec $n = 3$ et $A = A_1$, $B = A_2$, $C = A_3$.

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) &= \mathbb{P}\left(\bigcup_{i=1}^{3} A_i\right)\\ &= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3)\\ &\quad + (-1)^{2+1}\mathbb{P}(A_1 \cap A_2) + (-1)^{2+1}\mathbb{P}(A_1 \cap A_3)\\ &\quad + (-1)^{2+1}\mathbb{P}(A_2 \cap A_3)\\ &\quad + (-1)^{3+1}\mathbb{P}(A_1 \cap A_2 \cap A_3)\\ &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B)\\ &\quad - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).\end{aligned}$$

Le dé cubique non truqué

Soit n un entier naturel non nul. Nous effectuons n lancers d'un dé cubique bien équilibré dont les faces sont numérotées de 1 à 6.

1. Définissez l'espace probabilisé associé aux n lancers ainsi que la probabilité associée.

2. Calculez la probabilité des deux événements A_n et B_n définis par
– A_n : « Nous obtenons 1 pour la première fois au n-ième lancer. »
– B_n : « Nous n'obtenons aucun 1 lors des n lancers. »

Solution

1. L'espace probabilisé associé aux n lancers du dé est $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ avec :
– $\Omega = \{1, 2, 3, 4, 5, 6\}^n$;
– le dé est bien équilibré : chaque événement élémentaire a la même probabilité de se réaliser. La probabilité associée $\mathbb{P}$ est donc la probabilité uniforme sur $(\Omega, \mathcal{P}(\Omega))$.

2. Calculons la probabilité des deux événements A_n et B_n. Nous avons

$$\mathbb{P}(A_n) = \frac{\text{Card}(A_n)}{\text{Card}(\Omega)} \quad \text{et} \quad \mathbb{P}(B_n) = \frac{\text{Card}(B_n)}{\text{Card}(\Omega)},$$

où $\text{Card}(\Omega) = \text{Card}(\{1, 2, 3, 4, 5, 6\})^n = 6^n$. Il reste maintenant à déterminer $\text{Card}(A_n)$ et $\text{Card}(B_n)$. Nous avons

$$A_n = \{(x_1, x_2, \ldots, x_n) / \forall\, k \in [\![1, n-1]\!],\, x_k \in \{1, 2, 3, 4, 5\} \text{ et } x_n = 6\}.$$

Donc nous en déduisons que

$$\text{Card}(A_n) = 5^{n-1} \times 1 \quad \text{et} \quad \mathbb{P}(A_n) = \frac{5^{n-1}}{6^n}.$$

D'autre part, nous avons

$$B_n = \{1, 2, 3, 4, 5\}^n.$$

Donc nous en déduisons que

$$\text{Card}(B_n) = 5^n \quad \text{et} \quad \mathbb{P}(B_n) = \frac{5^n}{6^n}.$$

Probabilité conditionnelle et indépendance en probabilité

FICHE 7

I Probabilité conditionnelle

- ***Définition***

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé associé à $\mathcal{E}$. Soit A un événement de probabilité non nulle. Pour tout événement B, nous posons

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

L'application qui à un événement $B \in \mathcal{F}$ associe $\mathbb{P}(B|A)$ est une probabilité sur $(\Omega, \mathcal{F})$, appelée **probabilité conditionnelle relative à** A ou **probabilité sachant** A.

- ***Propriétés***

1. $\mathbb{P}(B^c|A) = 1 - \mathbb{P}(B|A)$;
2. $\mathbb{P}(B \setminus C|A) = \mathbb{P}(B|A) - \mathbb{P}(B \cap C|A)$;
3. $\mathbb{P}\left(B_1 \cup B_2|A\right) = \mathbb{P}(B_1|A) + \mathbb{P}(B_2|A) - \mathbb{P}\left(B_1 \cap B_2|A\right)$;
4. si B et C sont incompatibles, alors $\mathbb{P}\left(B \cup C|A\right) = \mathbb{P}(B|A) + \mathbb{P}(C|A)$.
   Plus généralement, si I est une partie finie ou infinie de $\mathbb{N}$ ou de $\mathbb{Z}$ et si $(B_i)_{i\in I}$ est une suite d'événements deux à deux incompatibles, $\mathbb{P}\left(\bigcup_{i\in I} B_i \middle| A\right) = \sum_{i\in I} \mathbb{P}(B_i|A)$;
5. la **formule du crible ou de Poincaré** se généralise en remplaçant $\mathbb{P}$ par $\mathbb{P}(\cdot|A)$;

6. si $(B_i)_{i\in\mathbb{N}}$ est une suite d'événements, alors nous avons

$$\lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{i=0}^{n} B_i \middle| A\right) = \mathbb{P}\left(\bigcup_{i=0}^{+\infty} B_i \middle| A\right),$$

$$\lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{i=0}^{n} B_i \middle| A\right) = \mathbb{P}\left(\bigcap_{i=0}^{+\infty} B_i \middle| A\right).$$

- **Formule des probabilités composées**

Soit $(A_i)_{1\leqslant i\leqslant n}$ une suite finie d'événements de $\mathcal{F}$ telle que

$$\mathbb{P}\left(A_1 \cap A_2 \cap \ldots \cap A_{n-1}\right) \neq 0.$$

Nous avons l'égalité suivante :

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\ldots\mathbb{P}\left(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}\right).$$

- **Système complet d'événements**

Soit $(\Omega, \mathcal{F})$ un espace probabilisable. Soit $(A_i)_{i\in I}$ une suite d'événements de $\mathcal{F}$. Nous appelons la famille $(A_i)_{i\in I}$ un système complet d'événements lorsque :

1. $\forall (i,j) \in I^2$ tel que $i \neq j, A_i \cap A_j = \emptyset$;
2. $\displaystyle\bigcup_{i\in I} A_i = \Omega$.

- **Formule des probabilités totales**

Soit $(A_i)_{i\in I}$ un système complet d'événements de probabilités toutes non nulles. Pour tout événement B, nous avons :

$$\mathbb{P}(B) = \sum_{i\in I} \mathbb{P}\left(B \cap A_i\right) = \sum_{i\in I} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

- **Formule de Bayes**

Pour tout système complet d'événements $(A_i)_{i\in I}$ de probabilités non nulles et pour tout événement B de probabilité non nulle, nous avons :

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\displaystyle\sum_{j\in I} \mathbb{P}(A_j)\mathbb{P}(B|A_j)}.$$

II Indépendance en probabilité

- **Événements indépendants**

  Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Nous appelons **événements indépendants pour la probabilité** $\mathbb{P}$ des événements A et B qui vérifient :

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

Pour généraliser la notion d'indépendance à plus de deux événements, nous utilisons la notion de σ-algèbre introduite à la fiche 5.

- **σ-algèbre engendrée par une famille d'événements**

  Soit $\mathcal{S}$ une famille de parties de Ω. L'intersection de toutes les σ-algèbres sur Ω qui contiennent $\mathcal{S}$ est une σ-algèbre sur Ω qui contient $\mathcal{S}$. Nous la notons $\mathcal{F}(\mathcal{S})$ et nous l'appelons la σ-**algèbre engendrée par** $\mathcal{S}$.

- **σ-algèbres indépendantes**

  Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Soit $(\mathcal{F}_i)_{1 \leqslant i \leqslant n}$ une famille de n σ-algèbres contenues dans $\mathcal{F}$. Les σ-algèbres $\mathcal{F}_i$ sont **indépendantes pour la probabilité** $\mathbb{P}$ lorsque pour toute famille $(A_i)_{1 \leqslant i \leqslant n}$ d'événements telle $A_i \in \mathcal{F}_i$ pour tout $i \in [\![1,n]\!]$, nous avons

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

- **Indépendance mutuelle de n événements**

  Soit n événements $A_1, \ldots, A_n$ d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Les événements $A_1, \ldots, A_n$ sont **mutuellement indépendants** pour la probabilité $\mathbb{P}$ lorsque les σ-algèbres $\mathcal{F}_1, \ldots, \mathcal{F}_n$ engendrées par les familles $(A_1), \ldots, (A_n)$ sont indépendantes.

  Théorème

  $A_1, \ldots, A_n$ sont mutuellement indépendants si et seulement si pour tout partie $I \subset [\![1,n]\!]$ l'égalité suivante est vérifiée

$$\mathbb{P}\left(\bigcap_{i \in I}^{n} A_i\right) = \prod_{i \in I}^{n} \mathbb{P}(A_i).$$

  Propriété : Soit n événements $A_1, \ldots, A_n$ mutuellement indépendants. Alors les événements $B_1, \ldots, B_n$ où pour tout $i \in \{1, \ldots, n\}$, $B_i = A_i$ ou $B_i = A_i^c$, sont mutuellement indépendants.

- **Indépendance mutuelle d'une suite infinie d'événements**

Soit $(A_i)_{i\in\mathbb{N}}$ une suite infinie d'événements. Les événements A_i sont **mutuellement indépendants** pour la probabilité $\mathbb{P}$ lorsque pour tout $n \in \mathbb{N}$, $A_0,\ldots,A_n$ sont des événements mutuellement indépendants pour la probabilité $\mathbb{P}$.

Théorème

$(A_i)_{i\in\mathbb{N}}$ est une famille d'événements mutuellement indépendants si et seulement si pour tout partie finie $I \subset \mathbb{N}$ l'égalité suivante est vérifiée

$$\mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathbb{P}(A_i).$$

- **Indépendance deux à deux de n événements**

Soit n événements $A_1,\ldots,A_n$ d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Les événements $A_1,\ldots,A_n$ sont **deux à deux indépendants** pour la probabilité $\mathbb{P}$ lorsque pour tout $(i,j) \in [\![1,n]\!]^2$, les événements A_i et A_j sont indépendants pour la probabilité $\mathbb{P}$.

Remarque : Trois événements A, B et C peuvent être deux à deux indépendants sans que A, B et C soient mutuellement indépendants.

Le paradoxe des prisonniers, proposé par J. Pearl

Trois prisonniers sont dans une cellule. Ils savent que deux vont être condamnés à mort et un gracié, mais ils ne savent pas qui. L'un d'entre eux va voir le gardien et lui demande :

– « Je sais bien que tu ne peux rien me dire, mais tu peux au moins me montrer un de mes compagnons qui sera exécuté. »

Le gardien réfléchit, se dit que de toutes manières au moins l'un des deux autres prisonniers sera condamné, et s'exécute.

Le prisonnier lui déclare alors :

– « Merci, avant, j'avais une chance sur trois d'être gracié, et maintenant, j'ai une chance sur deux. »

A-t-il raison de croire que sa probabilité d'être exécuté a varié ?

Solution

Nous supposerons équiprobables les chances des prisonniers. Nous excluons également le mensonge ou une forme de préférence dans la réponse du gardien. Désignons par A le prisonnier qui a conversé avec le gardien et par B et C les deux autres. Le pri-

sonnier A a demandé un nom de condamné, mais nous ne savons pas quel nom le gardien va choisir si les deux autres prisonniers, B et C sont condamnés. Si nous voulons décrire complètement la situation en termes probabilistes, il faut dire comment s'effectue le choix du nom. En l'absence de toute autre information le prisonnier doit faire une hypothèse, la plus plausible étant celle qui tient compte de son ignorance totale : si B et C sont condamnés, le gardien choisira de nommer B avec la probabilité de 1/2 et par conséquent il nommera C avec la probabilité de 1/2. La formalisation du problème est maintenant possible. L'ensemble fondamental est $\Omega = \{(X, Y)\}$ où X prend ses valeurs dans l'ensemble $\{AB, AC, BC\}$ (la paire de prisonniers condamnés) et Y prend ses valeurs dans $\{B, C\}$ (le prisonnier nommé par le gardien).
L'événement « A est condamné » s'écrit

$$\{(X,Y) : X = AB \text{ ou } X = AC\}.$$

Nous voulons obtenir la probabilité pour que A ait été condamné sachant que B a été nommé par le gardien

$$\mathbb{P}(\{X = AB\} \cup \{X = AC\} \big| Y = B).$$

Ne sachant rien *a priori* sur la décision du tribunal, l'hypothèse suivante est raisonnable

$$\mathbb{P}(X = AB) = \mathbb{P}(X = AC) = \mathbb{P}(X = BC) = \frac{1}{3}\cdot$$

Si $X = AB$, le nom du prisonnier nommé par le gardien sera forcément B, donc

$$\mathbb{P}(Y = B \big| X = AB) = 1.$$

De même, nous avons

$$\mathbb{P}(Y = C \big| X = AC) = 1.$$

Par contre, si $X = BC$, le gardien tire au sort B ou C, avec la même probabilité 1/2 pour B et C, d'où

$$\mathbb{P}(Y = B \big| X = BC) = \mathbb{P}(Y = C \big| X = BC) = \frac{1}{2}\cdot$$

Ensuite, nous pouvons donc écrire que

$$\mathbb{P}\left(\{X = AB\} \cup \{X = AC\} \big| Y = B\right)$$
$$= \mathbb{P}(X = AB | Y = B) + \mathbb{P}(X = AC | Y = B).$$

Le deuxième terme du deuxième membre de cette égalité disparaît étant donné que

$$\mathbb{P}\left(X = AC \big| Y = B\right) = \frac{\mathbb{P}\left(\{X = AC\} \cap \{Y = B\}\right)}{\mathbb{P}\left(Y = B\right)} = 0$$

puisque $\{X = AC\} \cap \{Y = B\} = \emptyset$ (le geôlier ne ment pas).

D'autre part, nous avons

$$\mathbb{P}(X = AB | Y = B) = \frac{\mathbb{P}(\{X = AB\} \cap \{Y = B\})}{\mathbb{P}(Y = B)}.$$

Mais si $X = AB$ alors nécessairement $Y = B$. L'égalité précédente se réduit donc à

$$\frac{\mathbb{P}(X = AB)}{\mathbb{P}(Y = B)}.$$

Nous savons que $\mathbb{P}(X = AB) = 1/3$. Il reste donc à calculer $\mathbb{P}(Y = B)$. Nous appliquons la formule des probabilités totales puisque les événements $\{X = AB\}$, $\{X = AC\}$ et $\{X = BC\}$ forment une partition de Ω,

$$\begin{aligned}\mathbb{P}(Y = B) = \mathbb{P}(Y = B|X = AB) \times \mathbb{P}(X = AB) + \mathbb{P}(Y = B|X = AC)\\ \times \mathbb{P}(X = AC) + \mathbb{P}(Y = B|X = BC) \times \mathbb{P}(X = BC)\\ = 1 \times 1/3 + 0 \times 1/3 + 1/2 \times 1/3 = 1/2.\end{aligned}$$

La probabilité recherchée vaut donc

$$\mathbb{P}(\{X = AB\} \cup \{X = AC\}|Y = B) = \frac{1/3}{1/2} = \frac{2}{3}.$$

Ainsi, la probabilité pour que A soit condamné sachant que $Y = B$ est la même que la probabilité *a priori* pour que A soit condamné. Le prisonnier a donc tort de croire que sa probabilité d'être exécuté a varié.

Monty Hall, inspiré du jeu télévisé américain Let's Make a Deal

Une voiture est mise en jeu pendant un divertissement à la télévision. Elle a été cachée, suite à un tirage au sort aux issues équiprobables, par l'animateur derrière l'une des trois portes qui sont présentées au candidat. Pour gagner la voiture, le candidat doit retrouver l'unique porte qui permet d'atteindre la voiture.
Le jeu se déroule de la manière suivante. Le présentateur invite un candidat à se présenter et à choisir l'une des trois portes, sans l'ouvrir. Le présentateur indique, parmi les deux portes qui n'ont pas été choisies par le candidat, une porte ne permettant pas d'accéder à la voiture puis l'ouvre. Il offre alors au concurrent la possibilité de conserver la porte qu'il a choisie au début du jeu ou de la remplacer par celle qui reste encore fermée.
Quel est le meilleur choix pour le candidat ?

Solution

Nous supposerons équiprobables les chances qu'une porte masque la voiture. Nous excluons également le mensonge ou une forme de préférence dans le choix du présentateur. Désignons par A la porte choisie initialement par le candidat et par B et C les deux autres.
Nous ne savons pas laquelle des portes le présentateur va choisir si les deux portes B et C ne masquent pas la voiture. Si nous voulons décrire complètement la situation en termes probabilistes, il faut dire comment s'effectue le choix de la porte.
En l'absence de toute autre information le candidat doit faire une hypothèse, la plus plausible étant celle qui tient compte de son ignorance totale : si B et C ne masquent pas la voiture, le présentateur choisira de nommer B avec la probabilité de 1/2 et par conséquent il nommera C avec la probabilité de 1/2.
Maintenant nous pouvons formaliser le problème. L'ensemble fondamental est $\Omega = \{(X, Y)\}$ où X prend ses valeurs dans l'ensemble $\{AB, AC, BC\}$ (la paire de portes ne masquant pas la voiture) et Y prend ses valeurs dans $\{B, C\}$ (la porte ouverte par le présentateur).
Un calcul identique à celui de l'exercice précédent montre que la probabilité pour que A soit la porte qui ne masque pas la voiture sachant que $Y = B$ est égale à :

$$\mathbb{P}(\{X = AB\} \cup \{X = AC\} | Y = B) = \frac{1/3}{1/2} = \frac{2}{3}.$$

Ainsi, la probabilité pour que A soit la porte derrière laquelle la voiture est dissimulée sachant que $Y = B$ est égale à :

$$p_1 = 1 - \mathbb{P}(\{X = AB\} \cup \{X = AC\} | Y = B) = \frac{1}{3}.$$

D'autre part, comme la porte B a été ouverte, si la porte A ne masque pas la voiture alors elle est dissimulée derrière la porte C. Par conséquent, la probabilité p_2 pour que C soit la porte derrière laquelle la voiture est dissimulée sachant que $Y = B$ est égale à :

$$p_2 = \mathbb{P}(\{X = AB\} \cup \{X = AC\} | Y = B) = \frac{2}{3}.$$

Donc la probabilité de gagner en changeant de porte vaut $p_2 = 2/3$ tandis que la probabilité de gagner sans changer de porte vaut $p_1 = 1/3$. Le meilleur choix pour le candidat est donc de changer de porte !

FICHE 8

Variables aléatoires réelles. Variables aléatoires discrètes

I Variables aléatoires réelles

- **Variable aléatoire réelle**

  Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Nous appelons **variable aléatoire réelle** (v.a.r.) toute application de Ω dans $\mathbb{R}$ telle que : pour tout intervalle I de $\mathbb{R}$, l'ensemble $X^{-1}(I) = \{\omega \in \Omega | X(\omega) \in I\}$ est un événement de $\mathcal{F}$.

- ***Notations***

  – Les variables aléatoires sont notées avec des lettres **majuscules** et les quantités déterministes avec des lettres **minuscules**.

  – Nous notons $X(\Omega)$ l'ensemble des valeurs prises par la variable aléatoire X définie sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

  – $X^{-1}(\{x\}) = \{\omega \in \Omega | X(\omega) = x\}$ se note $\{X = x\}$.

  – $X^{-1}(]-\infty, a]) = \{\omega \in \Omega | X(\omega) \leqslant a\}$ se note $\{X \leqslant a\}$.

  – $X^{-1}(]a, b]) = \{\omega \in \Omega | a < X(\omega) \leqslant b\}$ se note $\{a < X \leqslant b\}$.

- **Fonction de répartition**

  Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Nous appelons **fonction de répartition de la variable aléatoire** X la fonction numérique réelle F_X définie par

$$\forall x \in \mathbb{R}, \quad F_X(x) = \mathbb{P}(X \leqslant x).$$

Propriétés

1. $\forall x \in \mathbb{R}, \quad F_X(x) \in [0, 1]$;
2. F_X est une fonction croissante sur $\mathbb{R}$;

3. $\lim_{x\to-\infty} F_X(x) = 0$ et $\lim_{x\to+\infty} F_X(x) = 1$ et F_X est continue à droite et admet une limite à gauche en tout point de $\mathbb{R}$;
4. pour tous réels a et b, $\mathbb{P}(a < X \leqslant b) = F_X(b) - F_X(a)$.

II Variables aléatoires discrètes

- **Variable aléatoire discrète**

Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Nous appelons **variable aléatoire discrète** (v.a.d.) X si l'ensemble de ses valeurs $X(\Omega)$, est au plus dénombrable (voir la fiche 2 pour cette notion).

- **Loi d'une variable aléatoire discrète**

Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Nous appelons **loi de la variable aléatoire** X la donnée d'une suite numérique $(\mathbb{P}(X = k) = p_X(k))_{k\in X(\Omega)}$ telle que

1. $\forall k \in X(\Omega), \quad p_X(k) \geqslant 0$;
2. $\displaystyle\sum_{k\in X(\Omega)} p_X(k) = 1$;
3. pour tout réel x, $\mathbb{P}(X \leqslant x) = \displaystyle\sum_{k\leqslant x} p_X(k)$ où $\displaystyle\sum_{k\leqslant x}$ désigne la sommation sur l'ensemble des $k \in X(\Omega)$ inférieurs ou égaux à x.

Pour commencer

Soit X une variable aléatoire discrète prenant les valeurs 2, 4, 6 et 8. Déterminez la loi de probabilité de X sachant que :

$$\mathbb{P}(X < 6) = \frac{1}{3}, \quad \mathbb{P}(X > 6) = \frac{1}{2}, \quad \mathbb{P}(X = 2) = \mathbb{P}(X = 4).$$

Solution

Pour déterminer la loi de X, nous allons procéder de la manière suivante :

k	1	2	3	4
x_k	2	4	6	8
$\mathbb{P}(X = x_k)$	a	b	c	d

Nous savons que :

– $\mathbb{P}(X < 6) = \dfrac{1}{3}$, donc nous en déduisons que $a + b = \dfrac{1}{3}\cdot$

– $\mathbb{P}(X > 6) = \dfrac{1}{2}$, donc nous en déduisons que $d = \dfrac{1}{2}\cdot$

– $\mathbb{P}(X = 2) = \mathbb{P}(X = 4)$, donc nous en déduisons que $a = b$.

De plus, nous avons la relation suivante :

$$a + b + c + d = 1 \quad \text{car} \quad \sum_{k=1}^{4} \mathbb{P}(X = x_k) = 1.$$

Nous en déduisons $a = b = \dfrac{1}{6}$, $d = \dfrac{1}{2}$ et $c = \dfrac{1}{6}\cdot$

Le dé cubique truqué

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

1. Déterminez la loi de X.

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

Solution

1. X prend les valeurs $k \in [\![1, 6]\!]$. Si $p = \mathbb{P}(X = 1)$, alors nous avons $\mathbb{P}(X = k) = kp$ pour tout $k \in [\![1, 6]\!]$. Nous devons avoir :

$$\sum_{k=1}^{6} \mathbb{P}(X = k) = 1, \quad \text{donc} \quad \sum_{k=1}^{6} kp = 1,$$

c'est-à-dire $p \times \dfrac{6 \times 7}{2} = 1$. D'où $p = \dfrac{1}{21}$. La loi de X est donnée par le tableau suivant :

k	1	2	3	4	5	6
$\mathbb{P}(X = k)$	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$

2. Y prend les valeurs $y_k = \dfrac{1}{k}$, avec $k \in [\![1, 6]\!]$ et pour tout $k \in [\![1, 6]\!]$, $\mathbb{P}\left(Y = \dfrac{1}{k}\right) = \mathbb{P}(X = k)$. Donc la loi de Y est donnée par le tableau suivant :

y_k	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$
$\mathbb{P}\ (X = y_k)$	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$

La constante

Soit $a \in \mathbb{R}$ et X une v.a.r. à valeurs dans $\mathbb{N}$, telle que :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = \frac{a}{2^k \times k!}.$$

Déterminez a.

Solution

Posons $p_X(k) = \mathbb{P}(X = k) = \dfrac{a}{2^k \times k!}$. $(p_X(k))_{k\in\mathbb{N}}$ est la loi d'une v.a.r. X si et seulement si :

$$\forall k \in \mathbb{N}, \quad p_X(k) \geqslant 0 \quad \text{et} \quad \sum_{k=0}^{+\infty} p_X(k) = 1.$$

$$\forall k \in \mathbb{N}, \quad p_X(k) = \frac{a}{2^k \times k!} \geqslant 0 \quad \text{équivaut à} \quad a \geqslant 0.$$

$$\sum_{k=0}^{+\infty} p_X(k) = a \sum_{k=0}^{+\infty} \frac{1}{2^k \times k!} = a \sum_{k=0}^{+\infty} \frac{\left(\frac{1}{2}\right)^k}{k!} = a \exp(1/2).$$

Donc :

$$\sum_{k=0}^{+\infty} p_X(k) = 1 \quad \text{équivaut à} \quad a \exp(1/2) = 1 \quad \text{équivaut à} \quad a = \frac{1}{\sqrt{e}}.$$

Par suite, l'énoncé définit la loi d'une v.a.r. X à valeurs dans $\mathbb{N}$ si et seulement si $a = \dfrac{1}{\sqrt{e}}$.

FICHE 9

Moments et fonctions génératrices d'une v.a. discrète

Ici, $(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace probabilisé et X une variable aléatoire discrète (v.a.d.) définie sur Ω. $X(\Omega) = \{x_i / i \in I\}$ avec I une partie de $\mathbb{N}$ ou de $\mathbb{Z}$. Nous posons $\mathbb{P}(X = x_i) = p_X(x_i)$.

I Moments d'une variable aléatoire discrète

- **Espérance mathématique**

Nous disons que **la variable aléatoire** X **admet une espérance mathématique** lorsque I est fini ou lorsque la série $\sum_i x_i p_X(x_i)$ est absolument convergente.

Nous appelons alors **espérance mathématique de** X, la moyenne pondérée notée $\mathbb{E}(X)$ et définie par

$$\mathbb{E}(X) = \sum_i x_i p_X(x_i).$$

$\mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ désigne l'ensemble des variables aléatoires discrètes définies sur Ω et admettant une espérance. C'est un $\mathbb{R}$-espace vectoriel.

Remarque : Si $X \notin \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$, c'est-à-dire si la série $\sum_i x_i p_X(x_i)$ n'est pas absolument convergente, nous disons que X ne possède pas d'espérance mathématique.

- **Espérance d'une fonction d'une variable aléatoire**

Théorème du transfert : Soit g une fonction définie sur $X(\Omega)$ à valeurs dans $\mathbb{R}$. Lorsque $g(X)$ admet une espérance mathématique, c'est-à-dire $g(X) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$, nous avons

$$\mathbb{E}(g(X)) = \sum_i g(x_i) p_X(x_i).$$

Remarque : Pour calculer $\mathbb{E}(g(X))$, il suffit de changer x_i en $g(x_i)$.

- **Linéarité de l'opérateur $\mathbb{E}(\cdot)$**

L'espérance est une forme linéaire sur $\mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$: pour tous réels a et b et pour toutes variables aléatoires discrètes X et Y admettant une espérance mathématique, nous avons

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

Remarques

1. Pour tous réels a et b et pour toute v.a.d. X admettant une espérance mathématique, la formule ci-dessus se particularise en $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$. Car b est constant et donc $\mathbb{E}(b) = b$.
2. Soit g et h deux fonctions définies sur $X(\Omega)$ à valeurs dans $\mathbb{R}$ telles que $g(X), h(X) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$. Nous avons

$$\mathbb{E}(ag(X) + bh(X)) = a\mathbb{E}(g(X)) + b\mathbb{E}(h(X)), \quad \forall\, (a,b) \in \mathbb{R}.$$

Il faut noter que $\mathbb{E}(g(X))$ n'est généralement pas égal à $g(\mathbb{E}(X))$. En particulier nous avons le plus souvent $\mathbb{E}(X^2) \neq (\mathbb{E}(X))^2$.

- **Moment simple et moment centré d'ordre r**

Soit $r \in \mathbb{N}^*$. Sous réserve d'existence, nous appelons **moment simple d'ordre r de la variable aléatoire** X la valeur $\mu_r(X) = \mathbb{E}(X^r)$.

Remarques

1. Le moment d'ordre 1 de la variable aléatoire X, noté plus simplement μ ou encore μ_X s'il y a plusieurs variables aléatoires à distinguer, est l'espérance mathématique de X.
2. Si une variable aléatoire possède un moment simple d'ordre r, elle possède un moment simple pour tout ordre $s \in [\![1, r]\!]$.
3. Si $X(\Omega)$ est un ensemble fini, alors la variable aléatoire X admet des moments simples à tous les ordres $r \in \mathbb{N}^*$.

Soit $r \in \mathbb{N}^*$. Sous réserve d'existence, nous appelons **moment centré d'ordre r de** X le réel $_{\mu}m_r(X) = \mathbb{E}\left((X - \mathbb{E}(X))^r\right) = \mathbb{E}\left((X - \mu)^r\right)$.

Remarques

1. Pour $r = 1$, nous avons ${}_{\mu}m_1(X) = \mathbb{E}(X - \mu) = 0$.
Pour $r = 2$, ${}_{\mu}m_2(X)$ est appelée la variance de X, voir ci-dessous. Les caractéristiques de forme des distributions sont associées aux moments centrés d'ordres $r = 3$ et $r = 4$.
2. Si une variable aléatoire possède un moment centré d'ordre r, elle possède un moment centré pour tout ordre $s \in [\![1, r]\!]$.
3. Si $X(\Omega)$ est un ensemble fini, alors la variable aléatoire X admet des moments centrés à tous les ordres $r \in \mathbb{N}^*$.
4. Si la variable aléatoire X ne possède pas d'espérance mathématique, alors la variable X ne peut admettre de moment centré.

- **Variance et écart-type**

Nous disons que **la variable aléatoire discrète X admet une variance** lorsque X admet un moment centré d'ordre 2. Nous appelons alors **variance de X** la valeur

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left((X - \mu)^2\right) = {}_{\mu}m_2(X).$$

Nous la notons également σ^2 ou σ_X^2 s'il y a plusieurs variables aléatoires à distinguer.
Nous appelons **écart-type de X** la valeur $\sqrt{\mathbb{V}\mathrm{ar}(X)}$, que nous noterons σ ou σ_X selon les cas.

Théorème

X admet une variance si et seulement si X^2 admet une espérance mathématique. La **formule de Huygens** est alors valable :

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X).$$

Propriété

Pour tous réels a et b et pour toute variable aléatoire discrète X admettant une variance, nous avons

$$\mathbb{V}\mathrm{ar}(aX + b) = a^2\mathbb{V}\mathrm{ar}(X).$$

- ***Vocabulaire***

– Si $\mathbb{E}(X) = 0$, alors X **est une variable aléatoire centrée.**
– Si $\mathbb{V}\mathrm{ar}(X) = 1$, alors X **est une variable aléatoire réduite.**
– Si X admet une variance non nulle, la variable $X^* = \dfrac{X - \mu_X}{\sigma_X}$ est appelée **variable centrée et réduite associée à X.**

- **Moments d'ordres 3 et 4**

Nous mentionnons les moments d'ordres 3 et 4 car ils sont utilisés en statistique.

1. Le moment centré d'ordre 3, lorsqu'il existe, de la v.a.d. X dont $\sigma > 0$ fournit le **coefficient d'asymétrie**

$$\frac{\mathbb{E}\left((X-\mu)^3\right)}{\sigma^3}.$$

2. Le moment centré d'ordre 4, lorsqu'il existe, de la v.a.d. X dont $\sigma > 0$ fournit le **coefficient d'aplatissement**

$$\frac{\mathbb{E}\left((X-\mu)^4\right)}{\sigma^4} - 3.$$

II Fonctions génératrices de la variable aléatoire discrète X

- **Fonction génératrice des moments**

Considérons la fonction $g_X(t) = \mathbb{E}(\exp(tX))$ définie pour l'ensemble des valeurs de t pour lesquelles $\mathbb{E}(\exp(tX)) < +\infty$. $g_X(t)$ est la **fonction génératrice des moments de la variable aléatoire discrète** X lorsqu'elle est définie dans un voisinage de l'origine.

Propriété

Le moment d'ordre r de X est donné par :

$$\mu_r = g_X^{(r)}(0) \text{ où } g_X^{(r)}(0) \text{ est la dérivée d'ordre } r \text{ de } g_X.$$

- **Fonction génératrice**

Nous appelons **fonction génératrice de la variable aléatoire discrète** X à valeurs dans $\mathbb{N}$, la série entière

$$G_X(t) = \sum_{k=0}^{+\infty} t^k \mathbb{P}(X=k) = \mathbb{E}\left(t^X\right), \quad t \in [-1,+1].$$

La fonction génératrice ne dépend que de la loi de X. Elle caractérise la loi de X

La constante (reprise de l'exercice fiche 8)

Soit X une v.a.r. à valeurs dans $\mathbb{N}$, telle que :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = \frac{1}{\sqrt{e} \times 2^k \times k!}.$$

1. X admet-elle une espérance ? Si oui, la déterminez.

2. X admet-elle une variance ? Si oui, la déterminez.

Solution

1. X admet une espérance si et seulement si $\sum_k k\mathbb{P}(X = k)$ est absolument convergente.

$$\forall k \in \mathbb{N}^*, \quad k\mathbb{P}(X = k) = \frac{k}{\sqrt{e} \times 2^k \times k!} = \frac{1}{2\sqrt{e}} \times \frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!} \geqslant 0$$

et $\sum_k \frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!}$ converge, donc $\mathbb{E}(X)$ existe. Nous avons donc :

$$\mathbb{E}(X) = \frac{1}{2\sqrt{e}} \sum_{k=1}^{+\infty} \frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!} = \frac{1}{2\sqrt{e}} \sum_{k=0}^{+\infty} \frac{\left(\frac{1}{2}\right)^{k}}{k!} = \frac{1}{2\sqrt{e}}\sqrt{e} = \frac{1}{2}.$$

2. X admet une variance si et seulement si X^2 admet une espérance, c'est-à-dire si la série $\sum_k k^2\mathbb{P}(X = k)$ est absolument convergente.

$$\forall k \in \mathbb{N}^*, \quad k^2\mathbb{P}(X = k) = \frac{k^2}{\sqrt{e} \times 2^k \times k!} = \frac{1}{2\sqrt{e}} \times \frac{k\left(\frac{1}{2}\right)^{k-1}}{(k-1)!} \geqslant 0\,.$$

Nous pouvons écrire $k = (k-1) + 1$, ainsi nous avons :

$$\forall k \geqslant 2, \quad k^2\mathbb{P}(X = k) = \frac{1}{2\sqrt{e}} \times \left(\frac{\frac{1}{2}\left(\frac{1}{2}\right)^{k-2}}{(k-2)!} + \frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!} \right)$$

et $\sum_k \frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!}$ et $\sum_k \frac{\left(\frac{1}{2}\right)^{k-2}}{(k-2)!}$ convergent, donc $\mathbb{E}\left(X^2\right)$ existe.

Nous avons donc :

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= \frac{1}{2\sqrt{e}}\left(\frac{1}{2}\sum_{k=2}^{+\infty}\frac{\left(\frac{1}{2}\right)^{k-2}}{(k-2)!}+\sum_{k=1}^{+\infty}\frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!}\right)\\ &= \frac{1}{2\sqrt{e}}\left(\frac{1}{2}\sqrt{e}+\sqrt{e}\right)=\frac{3}{4}.\end{aligned}$$

X possède donc une variance égale à :

$$\mathbb{V}\mathrm{ar}(X)=\mathbb{E}\left(X^2\right)-\mathbb{E}^2(X)=\frac{3}{4}-\frac{1}{2^2}=\frac{1}{2}.$$

Le trousseau de clés

Dans un hôtel, une coupure de courant vient de se produire. Le gardien de nuit doit ouvrir dans le noir une chambre qu'un client vient de louer.
Il possède un trousseau de clés sur lequel se trouvent toutes les clés des n chambres de l'hôtel. Chaque clé ouvre une et une seule chambre.

1. Le gardien de nuit essaye les clés une à une sans utiliser deux fois la même. Donnez la loi de probabilité du nombre X d'essais nécessaires. Calculez l'espérance et la variance de X.

2. Lorsque le gardien de nuit n'est pas bien réveillé et que la coupure de courant persiste, il mélange toutes les clés à chaque tentative. Donnez la loi de probabilité de X. Calculez l'espérance et la variance de X.

Solution

1. La position de la bonne clé est au hasard sur les n positions possibles du trousseau de clé.

– Loi de X

Introduisons l'événement R_k = « la porte s'ouvre au k-ème essai », $k \in \{1,\ldots,n\}$.
Avec cette notation, l'événement $\{X=k\}$ s'écrit $R_k \cap R_{k-1}^c \cap \ldots \cap R_1^c$.
Donc,

$$\begin{aligned}\mathbb{P}(X=k) &= \mathbb{P}(R_k \cap R_{k-1}^c \cap \ldots \cap R_1^c)\\ &= \mathbb{P}(R_k|R_{k-1}^c \cap \ldots \cap R_1^c)\mathbb{P}(R_{k-1}^c|R_{k-2}^c \cap \ldots \cap R_1^c)\ldots\\ &\quad \ldots\mathbb{P}(R_2^c|R_1^c)\mathbb{P}(R_1^c).\end{aligned}$$

Les clés étant essayées au hasard, nous avons

$$\mathbb{P}(R_1^c) = \frac{n-1}{n},$$
$$\mathbb{P}(R_2^c|R_1^c) = \frac{n-2}{n-1},$$
$$\ldots,$$
$$\mathbb{P}(R_{k-1}^c|R_{k-2}^c \cap \ldots \cap R_1^c) = \frac{n+1-k}{n+2-k}$$
$$\mathbb{P}(R_k|R_{k-1}^c \cap \ldots \cap R_1^c) = \frac{1}{n-k+1}.$$

Donc la loi du nombre X d'essais nécessaires est donc égale à

$$\forall k \in [\![1,n]\!], \quad \mathbb{P}(X=k) = \frac{1}{n-k+1}\prod_{i=0}^{k-2}\frac{n-1-i}{n-i} = \frac{1}{n}.$$

– **Espérance de** X

$$\mathbb{E}(X) = \sum_{k=1}^{n} k\mathbb{P}(X=k) = \frac{1}{n}\sum_{k=1}^{n} k = \frac{1}{n} \times \frac{n\times(n+1)}{2} = \frac{n+1}{2}.$$

– **Variance de** X

$$\mathbb{E}\left(X^2\right) = \sum_{k=1}^{n} k^2\mathbb{P}(X=k) = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{(n+1)(2n+1)}{6}.$$

D'où

$$\begin{aligned}\mathbb{V}\mathrm{ar}(X) &= \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X)\\ &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}.\end{aligned}$$

2. Maintenant le gardien de nuit essaye les clés éventuellement plusieurs fois chacune. Soit A_k l'événement : « le gardien de nuit choisit la bonne clé lors de la k-ème tentative ». Les événements A_k, pour $k \geqslant 1$, sont indépendants et de probabilité égale à $1/n$.

– **Loi de** X

La loi de probabilité de X est donc égale à

$$\forall k \geqslant 1, \quad \mathbb{P}\,(X=k) = \left(1-\frac{1}{n}\right)^{k-1}\frac{1}{n}.$$

– Espérance de X

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^{+\infty} k\mathbb{P}(X=k) = \frac{1}{n}\sum_{k=1}^{+\infty} k\left(1-\frac{1}{n}\right)^{k-1} \\ &= \frac{1}{n}\frac{1}{\left(1-\left(1-\frac{1}{n}\right)\right)^2} = n,\end{aligned}$$

car pour tout réel q tel que $|q|<1$, nous avons

$$\sum_{k=1}^{+\infty} kq^{k-1} = \frac{1}{(1-q)^2}.$$

– Variance de X

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= \sum_{k=1}^{+\infty} k^2\mathbb{P}(X=k) = \frac{1}{n}\sum_{k=1}^{+\infty} k^2\left(1-\frac{1}{n}\right)^{k-1} \\ &= \frac{1}{n}\left(2\left(1-\frac{1}{n}\right)n^3+n^2\right) = 2n^2-n,\end{aligned}$$

car

$$k^2\left(1-\frac{1}{n}\right)^{k-1} = \left(1-\frac{1}{n}\right)k(k-1)\left(1-\frac{1}{n}\right)^{k-2} + k\left(1-\frac{1}{n}\right)^{k-1}$$

et pour tout réel q tel que $|q|<1$, nous avons

$$\sum_{k=2}^{+\infty} k(k-1)q^{k-2} = \frac{2}{(1-q)^3}.$$

D'où

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X) = 2n^2-n-n^2 = n(n-1).$$

Remarque :

Nous avons utilisé l'astuce classique :

$$k^2 = k(k-1)+k.$$

FICHE 10

Couples de variables aléatoires discrètes. Indépendance

I Lois associées à un couple de v.a.d.

- **Loi d'un couple** (X,Y)

Nous appelons **loi d'un couple** (X,Y) ou encore **loi conjointe de** X **et** Y la famille $(p_{(X,Y)}(x_i,y_j))_{(x_i,y_j)\in(X,Y)(\Omega)}$
où $p_{(X,Y)}(x_i,y_j) = \mathbb{P}([X=x_i]\cap[Y=y_j])$ et $(X,Y)(\Omega) = X(\Omega)\times Y(\Omega)$.

Remarque

Lorsque $X(\Omega)$ et $Y(\Omega)$ sont deux ensembles de petit cardinal, alors les probabilités $p_{(X,Y)}(x_i,y_j)$ peuvent être représentées dans un **tableau à double entrée**.

Théorème

$(p_{(X,Y)}(x_i,y_j))_{(x_i,y_j)\in(X,Y)(\Omega)}$ est la loi d'un couple de v.a.d. (X,Y) si et seulement si

1. $p_{(X,Y)}(x_i,y_j) \geqslant 0$ pour tout$(i,j)\in I\times J$.
2. $\displaystyle\sum_{i\in I}\sum_{j\in J} p_{(X,Y)}(x_i,y_j) = 1$.

- **Lois marginales**

Les v.a. X et Y sont appelées **variables marginales** du couple (X,Y). La loi de la v.a. X, respectivement la loi de la v.a. Y du couple (X,Y) est appelée **loi marginale** de X, respectivement **loi marginale** de Y.

Notations : $\mathbb{P}(X=x_i) = p_{i,\bullet}$ et $\mathbb{P}(Y=y_j) = p_{\bullet,j}$.

Propriétés

1. Pour tout $i\in I$, $p_{i,\bullet} = \displaystyle\sum_{j\in J} p_{(X,Y)}(x_i,y_j) = p_{(X,Y)}(x_i,\bullet)$.

2. Pour tout $j \in J$, $p_{\bullet,j} = \sum_{i \in I} p_{(X,Y)}(x_i, y_j) = p_{(X,Y)}(\bullet, y_j)$.

- **Lois conditionnelles**

Soit X une v.a.d. sur $(\Omega, \mathcal{F}, \mathbb{P})$ et $A \in \mathcal{F}$. Nous appelons **loi de X conditionnée par A**, ou **loi de X sachant A**, la suite numérique $(\mathbb{P}_A([X = x_i]))_{x_i \in X(\Omega)}$. Nous appelons **loi de X conditionnée par Y**, ou **sachant Y**, la famille $(\mathbb{P}_{[Y=y_j]}([X = x_i]))_{(x_i, y_j) \in (X,Y)(\Omega)}$.

II Indépendance de v.a.d.

- **Indépendance de deux v.a.d. X et Y**

Deux v.a.d. X et Y sont dites **indépendantes** si :

$$\forall (i, j) \in I \times J \qquad \mathbb{P}([X = x_i] \cap [Y = y_j]) = \mathbb{P}([X = x_i])\mathbb{P}([Y = y_j]).$$

Propriétés

1. X et Y sont indépendantes si et seulement si pour tout couple (x_i, y_j) de $(X,Y)(\Omega)$, $p_{(X,Y)}(x_i, y_j) = p_{i,\bullet}p_{\bullet,j}$.
2. Si X et Y sont indépendantes alors la **loi de X sachant** Y est égale à la loi de X.
3. Si X et Y sont indépendantes et f et g sont deux fonctions à valeurs réelles définies respectivement sur $X(\Omega)$ et sur $Y(\Omega)$, alors $f(X)$ et $g(Y)$ sont deux v.a.d. indépendantes.

- **Indépendance mutuelle d'une famille de v.a.d.**

1. Si l'ensemble K est fini, les **v.a.d.** $(X_k)_{k \in K}$ sont dites **mutuellement indépendantes**, lorsque, pour tout $(x_k)_{k \in K} \in \prod_{k \in K} X_k(\Omega)$,

$$\mathbb{P}\left(\bigcap_{k \in K} [X_k = x_k]\right) = \prod_{k \in K} \mathbb{P}([X_k = x_k]).$$

2. La famille $(X_k)_{k \in K}$ est une **famille infinie de v.a.d. mutuellement indépendantes** si, pour tout ensemble fini $L \subset K$, les v.a. $(X_l)_{l \in L}$, sont mutuellement indépendantes.

Théorème

Si les v.a.d. $X_1, \ldots, X_n$ sont mutuellement indépendantes, alors pour toutes les fonctions ϕ et ψ définies respectivement sur $(X_1, \ldots, X_p)(\Omega)$ et $(X_{p+1}, \ldots, X_n)(\Omega)$, les v.a.d. $\phi(X_1, \ldots, X_p)$ et $\psi(X_{p+1}, \ldots, X_n)$ sont indépendantes.

- **Indépendance deux à deux d'une famille de v.a.d.**

  Les **v.a.d.** $(X_k)_{k\in K}$ sont dites **deux à deux indépendantes** si, pour tout $(i,j) \in K$, X_i et X_j sont indépendantes.

  > **Remarque :** L'indépendance mutuelle implique l'indépendance deux à deux. La réciproque n'est pas vraie (voir le premier exercice de cette fiche).

III Somme et covariance de deux v.a.d.

Soit (X,Y) un couple de v.a.d. et g une fonction de $(X,Y)(\Omega)$ dans $\mathbb{R}$.

$$\mathbb{P}(g(X,Y)=z) = \sum_{\substack{(x_i,y_j)\in(X,Y)(\Omega) \\ \text{tels que } g(x_i,y_j)=z}} \mathbb{P}([X=x_i]\cap[Y=y_j]).$$

Nous en déduisons en particulier **la loi de la somme de X et de Y**

$$\mathbb{P}(X+Y=z) = \sum_{\substack{(x_i,y_j)\in(X,Y)(\Omega) \\ \text{tels que } x_i+y_j=z}} \mathbb{P}([X=x_i]\cap[Y=y_j]).$$

Théorème

Si la série double $\displaystyle\sum_{(x_i,y_j)\in(X,Y)(\Omega)} g(x_i,y_j)\mathbb{P}([X=x_i]\cap[Y=y_j])$ est absolument convergente, alors $g(X,Y)$ admet une espérance et

$$\mathbb{E}\big(g(X,Y)\big) = \sum_{(x_i,y_j)\in(X,Y)(\Omega)} g(x_i,y_j)\mathbb{P}([X=x_i]\cap[Y=y_j]).$$

Propriétés

1. Si deux v.a.d. X et Y ont une variance alors XY admet une espérance et

$$\mathbb{E}(XY) = \sum_{(x_i,y_j)\in(X,Y)(\Omega)} x_i y_j \mathbb{P}([X=x_i]\cap[Y=y_j]) = \mathbb{E}(YX).$$

2. Si deux v.a.d. X et Y ont une espérance et sont indépendantes alors XY admet une espérance et $\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$.

- **Covariance de v.a.d.**

Soit X et Y deux v.a.d. ayant une variance. La **covariance de X et de Y** est le nombre réel défini par

$$\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right].$$

Proposition : Soit X et Y deux v.a.d. ayant une variance.

$$\mathbb{C}\mathrm{ov}(X,X) = \mathbb{V}\mathrm{ar}(X) \quad \text{et} \quad \mathbb{C}\mathrm{ov}(X,Y) = \mathbb{C}\mathrm{ov}(Y,X).$$

Définition : Deux v.a.d. X et Y ayant une variance ne sont **pas corrélées** si $\mathbb{C}\mathrm{ov}(X,Y) = 0$.

Propriétés : Soit X et Y deux v.a.d. ayant une variance.

1. **Formule de Huygens :** $\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
   Si X et Y sont indépendantes, elles ne sont pas corrélées. La réciproque n'est pas vraie (voir le second exercice de cette fiche).
2. $X + Y$ admet une variance et
$$\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}(X) + 2\mathbb{C}\mathrm{ov}(X,Y) + \mathbb{V}\mathrm{ar}(Y).$$
3. De plus, si X et Y sont **indépendantes**,
$$\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y).$$
4. **Généralisation :** Si $(X_k)_{k\in K}$ est une famille finie de variables aléatoires deux à deux indépendantes et qui admettent une variance, alors $\sum_{k\in K} X_k$ admet une variance et
$$\mathbb{V}\mathrm{ar}\left(\sum_{k\in K} X_k\right) = \sum_{k\in K} \mathbb{V}\mathrm{ar}(X_k).$$

Propriétés : Soit X, Y et Z trois v.a.d. ayant une variance et a, b, c et d quatre nombres réels.

1. $\mathbb{C}\mathrm{ov}(aX + bY, Z) = a\,\mathbb{C}\mathrm{ov}(X,Z) + b\,\mathbb{C}\mathrm{ov}(Y,Z)$.
2. $\mathbb{C}\mathrm{ov}(aX + b, cY + d) = ac\,\mathbb{C}\mathrm{ov}(X,Y)$.

2 à 2 indépendantes $\not\Rightarrow$ mutuellement indépendantes

Soit X et Y deux v.a.d. indépendantes telles que $\mathbb{P}([X=0]) = \mathbb{P}([X=1]) = \mathbb{P}([Y=0]) = \mathbb{P}([Y=1]) = 1/2$ et $\mathbb{P}([X=x]) = \mathbb{P}([Y=x]) = 0$ pour tout $x \in \mathbb{R}\setminus\{0,1\}$. Posons Z la variable aléatoire qui prend pour valeur 1 si $X = Y$ et pour valeur 0 si $X \neq Y$.

Montrez que les variables aléatoires X, Y et Z sont deux à deux indépendantes mais pas mutuellement indépendantes.

Solution

Z est une v.a.d. qui peut prendre les valeurs 0 et 1.

$$\begin{aligned}
\mathbb{P}([Z=1]) &= \mathbb{P}([(X,Y)=(0,0)]) + \mathbb{P}([(X,Y)=(1,1)]) \\
&= \mathbb{P}([X=0])\mathbb{P}([Y=0]) + \mathbb{P}([X=1])\mathbb{P}([Y=1]) = \frac{1}{2} \\
\mathbb{P}([Z=0]) &= \mathbb{P}([(X,Y)=(1,0)]) + \mathbb{P}([(X,Y)=(0,1)]) \\
&= \mathbb{P}([X=1])\mathbb{P}([Y=0]) + \mathbb{P}([X=0])\mathbb{P}([Y=1]) = \frac{1}{2}.
\end{aligned}$$

Nous avons les égalités suivantes
$[(X,Z)=(0,0)] = [(X,Y)=(0,1)]$, $[(X,Z)=(1,0)] = [(X,Y)=(1,0)]$,
$[(X,Z)=(0,1)] = [(X,Y)=(0,0)]$, $[(X,Z)=(1,1)] = [(X,Y)=(1,1)]$.
X et Z sont indépendantes puisque pour tout $(a,b) \in \{0,1\}^2$,

$$\mathbb{P}([(X,Z)=(a,b)]) = \frac{1}{4} = \mathbb{P}([X=a])\mathbb{P}([Z=b]).$$

Nous montrons de la manière que Y et Z sont indépendantes.
Ainsi X, Y et Z sont deux à deux indépendantes.
Enfin si X, Y et Z étaient mutuellement indépendantes nous aurions

$$\mathbb{P}([(X,Y,Z)=(1,0,1)]) = \mathbb{P}([X=1])\mathbb{P}([Y=0])\mathbb{P}([Z=1]),$$

ce qui n'est pas le cas.
En effet $[(X,Y,Z)=(1,0,1)] = \emptyset$ et donc $\mathbb{P}([(X,Y,Z)=(1,0,1)]) = 0$ alors que $\mathbb{P}([X=1])\mathbb{P}([Y=0])\mathbb{P}([Z=1]) = 1/8$.

Corrélation et indépendance

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

Solution

La loi de variable aléatoire X est :

x_k	-1	0	1
$\mathbb{P}([X = x_k])$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

La loi de la variable aléatoire $Y = X^2$ est :

y_k	0	1
$\mathbb{P}([Y = y_k])$	$\frac{1}{3}$	$\frac{2}{3}$

La loi du couple de variables aléatoires (X,Y) est donnée par le tableau suivant :

x_k	-1	0	1
$\mathbb{P}([(X, Y) = (x_k, 0)])$	0	$\frac{1}{3}$	0
$\mathbb{P}([(X, Y) = (x_k, 1)])$	$\frac{1}{3}$	0	$\frac{1}{3}$

Calculons $\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.

$$\begin{aligned}
\mathbb{E}(X) &= -1 \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = 0.\\
\mathbb{E}(Y) &= 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}.\\
\mathbb{E}(XY) &= (-1 \times 1) \times \frac{1}{3} + (0 \times 0) \times \frac{1}{3} + (1 \times 1) \times \frac{1}{3} = 0.
\end{aligned}$$

Par conséquent, nous obtenons $\mathbb{C}\mathrm{ov}(X,Y) = 0 - 0 \times \dfrac{2}{3} = 0$.

Les variables aléatoires X et Y ne sont donc pas corrélées. Si les variables aléatoires X et Y étaient indépendantes, nous aurions

$$\mathbb{P}[(X,Y) = (1,0)] = \mathbb{P}(X = 1)\mathbb{P}(Y = 0).$$

Or $\mathbb{P}[(X,Y) = (1,0)] = 0 \neq 1/9 = \mathbb{P}(X = 1)\mathbb{P}(Y = 0)$.
Les variables aléatoires X et $Y = X^2$ ne sont donc pas indépendantes.

FICHE 11

Lois discrètes usuelles finies

I Loi de Bernouilli de paramètre p

- **Loi de Bernoulli**

  Soit $p \in [0,1]$.Une variable aléatoire X suit une **loi de Bernoulli de paramètre** p, notée $\mathcal{B}(1,p)$, si la variable aléatoire X prend la valeur 1 avec la probabilité p et la valeur 0 avec la probabilité $1-p=q$.

- ***Propriétés***

  L'espérance et la variance d'une variable aléatoire X suivant une loi de Bernoulli de paramètre p sont égales respectivement à

$$\mathbb{E}(X)=p \quad \text{et} \quad \mathbb{V}\text{ar}(X)=p(1-p)=pq \quad \text{où} \quad q=1-p.$$

II Loi uniforme discrète

- **Loi uniforme discrète**

  Soit $n \in \mathbb{N}^*$. Une variable aléatoire X suit une **loi uniforme discrète** si la variable aléatoire X prend n valeurs possibles $k_1,k_2,\ldots,k_n$ avec la probabilité égale à $1/n$ pour n'importe quelle valeur k_i.

- ***Propriétés***

  Si X suit une loi uniforme discrète sur $[\![a,b]\!]$, alors nous avons

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

- **Remarque**

  La loi uniforme discrète sur $[\![1,n]\!]$ peut s'appliquer à toute situation aléatoire à n issues équiprobables dès que celles-ci peuvent être associées aux nombres $1,2,\ldots,n$. Nous avons alors

$$\mathbb{E}(X)=\frac{n+1}{2} \quad \text{et} \quad \mathbb{V}\text{ar}(X)=\frac{n^2-1}{12}.$$

III Loi binomiale de paramètres n et p

- **Loi binomiale**
  Soit n un entier naturel et $p \in [0,1]$. Une variable aléatoire X suit une **loi binomiale de paramètres n et p**, notée $\mathcal{B}(n,p)$, si la variable aléatoire X prend la valeur k avec la probabilité égale à $\binom{n}{k} p^k (1-p)^{n-k}$. Nous notons q le nombre $1-p$.
  Cette loi caractérise le nombre de boules blanches obtenues dans un tirage **avec remise** de n boules dans une urne.

- **Stabilité pour la somme**
  Si $X_1,\dots,X_m$ sont m variables aléatoires mutuellement indépendantes et si, pour tout entier $k \in [\![1,m]\!]$, X_k suit la loi $\mathcal{B}(n_k,p)$ alors la variable aléatoire $\displaystyle\sum_{k=1}^{m} X_k$ suit la loi $\mathcal{B}(n,p)$ où $n = \displaystyle\sum_{k=1}^{m} n_k$.

- ***Propriétés***

  L'espérance et la variance d'une variable aléatoire X suivant une loi binomiale de paramètres n et p sont égales respectivement à :

$$\mathbb{E}(X) = np \quad \text{et} \quad \mathbb{V}\mathrm{ar}(X) = np(1-p) = npq.$$

IV Loi hypergéométrique de paramètres N, n et p

- **Loi hypergéométrique**

  Soit N et n deux entiers naturels tels que $n \leqslant N$ et $p \in [0,1]$ tel que Np soit entier. Une variable aléatoire X suit une **loi hypergéométrique de paramètres** N, n et p, notée $\mathcal{H}(N,n,p)$, si la variable aléatoire X prend la valeur k avec la probabilité égale à $\dfrac{\binom{Np}{k}\binom{N-Np}{n-k}}{\binom{N}{n}}$. Nous notons q le nombre $1-p$.

- **Remarques**
  1. La loi hypergéométrique modélise toutes les situations qui s'apparentent à un tirage sans remise.

2. La loi hypergéométrique prend les valeurs entières comprises entre $\max(0, n - N + Np)$ et $\min(Np, n)$.

- ***Propriétés***

  L'espérance et la variance d'une variable aléatoire X suivant une loi hypergéométrique de paramètres N, n, p sont égales respectivement à

$$\mathbb{E}[X] = np \quad \text{et} \quad \mathbb{V}\text{ar}[X] = np(1-p)\frac{N-n}{N-1} = npq\frac{N-n}{N-1}.$$

- **Remarques**

  1. L'espérance d'une loi hypergéométrique $\mathcal{H}(N, n, p)$ est égale à l'espérance d'une loi binomiale $\mathcal{B}(n, p)$.

  2. La variance d'une loi hypergéométrique $\mathcal{H}(N, n, p)$ est égale à la variance d'une loi binomiale $\mathcal{B}(n, p)$ au facteur multiplicatif $\dfrac{N-n}{N-1}$ près. Dans un contexte statistique, ce facteur s'appelle le facteur d'exhaustivité. Il est toujours inférieur ou égal à 1.

L'oral d'un concours

Le candidat se présente à un oral de concours. Le jury qui l'a convoqué lui pose 20 questions. Pour chaque question, le même nombre $k \geqslant 2$ de réponses lui sont proposées dont une et une seule est la bonne. Le candidat qui n'a pas travaillé son oral, choisit au hasard une des réponses proposées.

1. Le jury attribue un point par bonne réponse. Soit X le nombre de points obtenus à l'issue de l'oral. Quelle est la loi de la variable aléatoire X ?

2. Lorsque le candidat donne une mauvaise réponse, il peut choisir à nouveau une des autres réponses proposées. Le jury lui attribue alors $\frac{1}{2}$ point par bonne réponse. Soit Y le nombre de $\frac{1}{2}$ points obtenus lors de ces secondes tentatives. Quelle est la loi de la variable aléatoire Y ?

3. Soit T le nombre total de points obtenus. Déterminez k pour que le candidat obtienne en moyenne une note supérieure ou égale à 10 sur 20 afin qu'il soit admis.

Solution

1. Notons X_i le résultat du candidat à la question i. $X_1, \ldots, X_{20}$ sont des variables aléatoires discrètes de Bernoulli **mutuellement indépendantes** et **de même paramètre** $p = 1/k$.

Comme $X = \sum_{i=1}^{20} X_i$, la loi de la variable aléatoire X est la loi binomiale $\mathcal{B}(20, 1/k)$.

2. Si $X_i = 0$, alors introduisons Y_i le résultat du candidat à la question i lors de la seconde tentative. Si $X_i = 1$, alors $Y_i = 0$. Ainsi Y qui représente le nombre de $\frac{1}{2}$ points obtenus lors de ces secondes tentatives est égal à $\sum_{i=1}^{20} Y_i$, nombre de bonnes réponses aux secondes tentatives. $Y_1, \ldots, Y_{20}$ sont des variables aléatoires discrètes **mutuellement indépendantes** et de même loi. Comme les variables aléatoires Y_i sont à valeurs dans $[\![0,1]\!]$, il s'agit de variables aléatoires de Bernoulli. Il reste à déterminer le paramètre $p = \mathbb{P}(Y_i = 1)$:

$$\begin{aligned} \mathbb{P}(Y_i = 1) &= \mathbb{P}(Y_i = 1 \cap X_i = 0) = \mathbb{P}(Y_i = 1 | X_i = 0)\, \mathbb{P}(X_i = 0) \\ &= \frac{1}{k-1} \frac{k-1}{k} = \frac{1}{k}. \end{aligned}$$

Donc la loi de la variable aléatoire Y est la loi binomiale $\mathcal{B}(20, 1/k)$.

3. Le nombre total de points obtenus par le candidat est égal à

$$T = X + \frac{1}{2} Y.$$

Il vient alors que

$$\mathbb{E}(T) = \mathbb{E}\left(X + \frac{1}{2} Y\right) = \mathbb{E}(X) + \frac{1}{2}\mathbb{E}(Y) = \frac{20}{k} + \frac{1}{2} \times \frac{20}{k} = \frac{30}{k}.$$

Pour que le candidat obtienne en moyenne une note supérieure ou égale à 10 sur 20, le paramètre k doit être inférieur ou égal à 3.

Capture-recapture

Un lac contient un nombre N inconnu de poissons. Nous souhaitons faire une estimation de N sans compter tous les poissons. Une méthode qui permet de le faire est appelée « capture-recapture ». Cette méthode a été introduite par Laplace en 1768 pour estimer la population de la France.
Elle se déroule en deux étapes :

- **Première étape :** Nous capturons M d'entre eux, nous les marquons et nous les relâchons.
- **Deuxième étape :** Nous capturons de nouveau n poissons et nous en trouvons X qui portent la marque, parmi ces n poissons.

Nous supposons que la population du lac est fermée, c'est-à-dire que le nombre N de poissons n'a pas varié.

1. Quelle est la loi de la variable aléatoire X ?

2. Nous constatons que $X = k$. Faire une estimation de N par la **méthode du maximum de vraisemblance** (méthode que nous développerons dans la fiche 26), c'est-à-dire prendre comme estimation de la taille de la population, la valeur $\widehat{N}$, si elle existe, qui rende maximale la probabilité $\mathbb{P}(X = k)$, où k est la valeur observée de X.

3. Montrez que l'espérance de $\widehat{N}$ est infinie.

4. Application numérique : $M = n = 10\,000$, $k = 100$.

Solution

1. Il faut avoir pris k poissons marqués parmi M et $n-k$ non marqués parmi $N-M$. Ainsi la loi de la variable aléatoire X est égale à

$$\mathbb{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad 0 \leqslant k \leqslant n.$$

Nous reconnaissons une loi hypergéométrique de paramètres N, n, $\dfrac{M}{N}$.

2. Nous fixons k et nous cherchons le maximum sur N de $\mathbb{P}(X = k)$ quantité que nous notons F_N. F_N est strictement croissante tant que

$$F_{N+1} > F_N,$$

ou encore

$$\frac{(N+1-M)(N+1-n)}{(N+1-M-n+k)(N+1)} > 1.$$

Ainsi, nous avons

$$F_{N+1} > F_N \quad \text{si} \quad k(N+1) < Mn.$$

De la même manière, nous montrons que F_N est strictement décroissante

$$F_{N+1} < F_N \quad \text{si} \quad k(N+1) > Mn.$$

Supposons $k \neq 0$. Il faut alors prendre comme estimation de N,

$$\widehat{N} = \frac{Mn}{k} \quad \text{si } \frac{Mn}{k} \text{ est entier}$$

et choisir entre la partie entière de $\frac{Mn}{k}$ et la partie entière de $\frac{Mn}{k} + 1$, suivant la valeur de $F_{\widehat{N}}$ en ces points.
Si $k = 0$, alors pour tout $N \in \mathbb{N}$ $F_{N+1} > F_N$ puisque pour tout $N \in \mathbb{N}$:

$$\frac{F_{N+1}}{F_N} = \frac{(N+1-M)(N+1-n)}{(N-n-M+1)(N+1)} > 1.$$

Pour $k = 0$, nous trouvons une estimation de N infinie.

3. De plus, nous avons

$$\mathbb{P}(X = k = 0) = \frac{\binom{N-M}{n}}{\binom{N}{n}} > 0$$

et les autres termes de l'espérance sont finis. Nous avons donc que l'espérance de $\widehat{N}$ est infinie.

4. Application numérique : $\widehat{N} = Mn/k = (10\,000)^2/100 = 10^6$.

FICHE 12

Lois discrètes usuelles infinies

I Loi géométrique de paramètre p

- **Loi géométrique**

Soit $p \in]0, 1[$. Une variable aléatoire X suit une **loi géométrique de paramètre** p, notée $\mathcal{G}(p)$, si la variable aléatoire X prend la valeur $k \geqslant 1$ avec la probabilité égale à $(1-p)^{k-1}p$.

- **Remarques**

1. Cette loi sert généralement lorsque nous nous intéressons au temps d'attente du premier succès, c'est-à-dire au nombre d'essais nécessaires pour obtenir un succès, lors d'une succession d'expériences aléatoires indépendantes n'ayant que deux issues possibles : le succès avec une probabilité p et l'échec avec une probabilité $1-p$.
2. La loi géométrique est parfois utilisée pour modéliser des durées de vie. Par exemple, la loi géométrique est le modèle discret de la mort d'une particule radioactive. La loi géométrique est la version discrète d'une loi absolument continue : la loi exponentielle (voir fiche 16).

- ***Propriétés***

L'espérance et la variance d'une variable aléatoire X suivant une loi géométrique de paramètre p sont égales respectivement à

$$\mathbb{E}(X) = \frac{1}{p}$$

$$\mathbb{V}\text{ar}(X) = \frac{1-p}{p^2}.$$

II Loi de Poisson de paramètre λ

- **Loi de Poisson**

  Soit $\lambda > 0$. Une variable aléatoire X suit une **loi de Poisson de paramètre** λ, notée $\mathcal{P}(\lambda)$, si la variable aléatoire X prend la valeur k, $k \in \mathbb{N}$ avec la probabilité égale à $\exp(-\lambda)\dfrac{\lambda^k}{k!}$.

- **Remarques**

  1. La loi de Poisson est utilisée pour décrire plusieurs types de phénomènes comme le nombre d'appels reçus par un standard téléphonique pendant une période donnée, etc.
  2. La loi de Poisson est encore utilisée lorsque nous étudions le nombre d'apparitions de certains phénomènes rares.

- ***Propriétés***

  L'espérance et la variance d'une variable aléatoire suivant une loi de Poisson de paramètre λ sont égales respectivement à

$$\mathbb{E}(X) = \lambda \quad \text{et} \quad \mathbb{V}\text{ar}(X) = \lambda.$$

- **Stabilité pour l'addition**

  Si X suit une loi $\mathcal{P}(\lambda_1)$ et Y une loi $\mathcal{P}(\lambda_2)$, et si X et Y sont indépendantes alors $X + Y$ suit la loi $\mathcal{P}(\lambda_1 + \lambda_2)$.

  Plus généralement, si $X_1, \ldots, X_n$ sont n variables aléatoires mutuellement indépendantes et si, pour tout entier $k \in [\![1, n]\!]$, X_k suit une loi $\mathcal{P}(\lambda_k)$ alors la variable aléatoire $\displaystyle\sum_{k=1}^{n} X_k$ suit une loi $\mathcal{P}(\lambda)$ avec $\lambda = \displaystyle\sum_{k=1}^{n} \lambda_k$.

Mistral gagnant
d'après une idée du livre de G. Frugier

En sortant du lycée, Guy achète chaque jour un « mistral » dans une confiserie proche du lycée. La probabilité pour un « mistral » d'être un « mistral gagnant » est égale à p. Soit X le nombre de « mistral » que Guy doit acheter pour tomber sur un « mistral gagnant ».

1. Établissez la loi de probabilité de la variable aléatoire X.
2. Calculez l'espérance mathématique et la variance de la variable aléatoire X.

Solution

1. Les « mistral » peuvent être considérés comme indépendants. Si le « mistral gagnant » arrive au bout de k achats, alors la variable aléatoire X prend la valeur k. La probabilité de l'événement correspondant est donc :

$$\mathbb{P}(X = k) = (1 - p)^{k-1} p, \quad \text{avec } k \in \mathbb{N}^*.$$

Donc la loi de probabilité de la variable aléatoire X est la loi géométrique de paramètre p.

2. Nous allons maintenant procéder au calcul de l'espérance mathématique de X. Nous avons :

$$\mathbb{E}(X) = \sum_{k=1}^{+\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{+\infty} k(1-p)^{k-1}.$$

Posons $S(x)$ la série de terme général kx^{k-1}. La série $S(x)$ est la dérivée de la série géométrique de terme général x^k et elle converge pour $|x| < 1$. La somme est donc égale à la dérivée de $\frac{1}{1-x}$, par conséquent à $\frac{1}{(1-x)^2}$. Nous en déduisons la valeur de l'espérance mathématique :

$$\mathbb{E}(X) = p \times \frac{1}{(1-(1-p))^2} = \frac{1}{p},$$

qui est bien le résultat annoncé dans la fiche de cours.
Pour calculer la variance, il nous reste à calculer $\mathbb{E}\left(X^2\right)$.

$$\mathbb{E}\left(X^2\right) = \sum_{k=1}^{+\infty} k^2 p(1-p)^{k-1} = \sum_{k=1}^{+\infty} k(k-1)p(1-p)^{k-1} + \sum_{k=1}^{+\infty} kp(1-p)^{k-1}.$$

La première somme est égale à

$$p(1-p) = \sum_{k=1}^{+\infty} k(k-1)(1-p)^{k-2}.$$

En dérivant $S(x)$ pour tout $|x| < 1$, nous obtenons

$$S'(x) = \frac{2}{(1-x)^3} = \sum_{k=1}^{+\infty} k(k-1)(1-p)^{k-2}.$$

Ainsi, nous obtenons

$$\mathbb{E}\left(X^2\right) = \frac{2-p}{p^2}.$$

Nous en déduisons

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2},$$

qui est bien le résultat annoncé dans la fiche de cours.

Les accidents mortels de la route

Dans le département du Bas-Rhin, en 2001, nous enregistrons, d'après les données du ministère de l'intérieur et l'INSEE, 89 personnes tuées sur la route. Calculer la probabilité d'avoir l'année suivante un nombre de personnes tuées sur la route strictement plus grand que 88.
Nous considérons que le décès d'un individu suite à un accident de la route est un phénomène rare qui peut se modéliser par une variable aléatoire X dont la loi de probabilité est une loi de Poisson de paramètre λ. Compte tenu du peu d'information disponible, nous prenons $\lambda = \mathbb{E}(X) = 89$.

Solution

La probabilité recherchée est la suivante :

$$\mathbb{P}(X > 88) = 1 - \exp(-89)\sum_{k=0}^{88}\frac{89^k}{k!} = 1 - 0{,}485903 \approx 0{,}514097.$$

À méditer...

FICHE 13

Variables aléatoires continues

I Définitions

- **Variable aléatoire continue**

Soit X une variable aléatoire réelle (v.a.r.) définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Nous disons que **la variable aléatoire X est continue** s'il existe une fonction f_X définie sur $\mathbb{R}$ telle que

1. $f_X(t) \geqslant 0$ pour tout $t \in \mathbb{R}$;
2. l'ensemble des points de discontinuités de f_X est fini et ces discontinuités sont de première espèce (i.e. la limite à gauche et à droite en chaque point existe) ;
3. pour tout x réel la **fonction de répartition F_X de la variable** X est donnée par

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t.$$

La fonction f_X est une **densité de probabilité de** X.

- **Remarque**

Il faut en particulier que l'intégrale $\int_{-\infty}^{+\infty} f_X(t)\,\mathrm{d}t$ soit convergente et que

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

- ***Propriétés***

Soit X une v.a.r. admettant une densité de probabilité f_X.

1. Pour tout $a \in \mathbb{R}$, $\mathbb{P}([X = a]) = 0$;
2. Pour tout (a, b) avec $-\infty \leqslant a \leqslant b \leqslant +\infty$, nous avons :

$$\mathbb{P}(a < X < b) = \mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f_X(t)\,\mathrm{d}t.$$

- **Fonction d'une variable aléatoire continue**

Soit X une v.a.r. admettant une densité de probabilité f_X et I un intervalle d'extrémités a et b, $(-\infty \leqslant a < b \leqslant +\infty)$, contenant $X(\Omega)$. Soit ϕ de classe $\mathcal{C}^1$ sur I, telle que pour tout x de I, $\phi'(x) > 0$ ou $\phi'(x) < 0$.

1. $\phi(X)$ est une variable aléatoire continue ;

2. une densité de probabilité $g_{\phi(X)}$ de $\phi(X)$ est donnée par :

$$g_{\phi(X)}(x) = \frac{f_X(\phi^{-1}(x))}{|\phi'(\phi^{-1}(x))|} \text{ si } x \in \phi(I), \quad g_{\phi(X)}(x) = 0 \text{ si } x \notin \phi(I).$$

II Moments d'une variable aléatoire continue

Dans tout ce paragraphe, $(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace probabilisé et X une variable aléatoire continue définie sur Ω et de densité de probabilité f_X.

- **Espérance mathématique**

Nous disons que **la variable aléatoire X admet une espérance mathématique** si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t$ converge absolument.

Nous appelons alors **espérance mathématique de** X, la valeur notée $\mathbb{E}(X)$ et définie par

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t.$$

Si deux variables aléatoires réelles X et Y admettent une espérance alors, pour tout $(a, b) \in \mathbb{R}^2$, $\mathbb{E}(aX + bY) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y)$.

- **Espérance d'une fonction d'une variable aléatoire**

Théorème du transfert : Soit I un intervalle d'extrémités a et b, $(-\infty \leqslant a < b \leqslant +\infty)$, contenant $X(\Omega)$. Soit g une fonction à valeurs dans $\mathbb{R}$ définie et continue sur I sauf en un nombre fini de valeurs.

$g(X)$ admet une espérance mathématique $\mathbb{E}(g(X))$ si et seulement si $\displaystyle\int_a^b g(t) f_X(t)\,\mathrm{d}t$ est absolument convergente et nous avons

$$\mathbb{E}(g(X)) = \int_a^b g(t) f_X(t)\,dt.$$

Remarque : Pour calculer $\mathbb{E}(g(X))$, il suffit de changer t en $g(t)$.

- **Moment simple et moment centré d'ordre r**

Soit $r \in \mathbb{N}^*$. Sous réserve d'existence, nous appelons **moment simple d'ordre r de la variable aléatoire** X la valeur

$$\mu_r(X) = \mathbb{E}(X^r) = \int_{-\infty}^{+\infty} t^r f_X(t)\,dt.$$

Remarques

1. Le moment d'ordre 1 de la variable aléatoire X, noté plus simplement μ ou encore μ_X s'il y a plusieurs variables aléatoires à distinguer, est l'espérance mathématique de X.
2. Si une variable aléatoire possède un moment simple d'ordre r, elle possède un moment simple pour tout ordre $s \in [\![1, r]\!]$.

Soit $r \in \mathbb{N}^*$. Sous réserve d'existence, nous appelons **moment centré d'ordre r de** X le réel

$$_\mu m_r(X) = \mathbb{E}\left((X-\mu)^r\right).$$

Remarques

1. Pour $r = 1$, nous avons $_\mu m_1(X) = \mathbb{E}(X-\mu) = 0$.
   Pour $r = 2$, $_\mu m_2(X)$ est appelée la variance de X, voir ci-dessous. Les caractéristiques de forme des distributions sont associées aux moments centrés d'ordres $r = 3$ et $r = 4$.
2. Si une variable aléatoire possède un moment centré d'ordre r, elle possède un moment centré pour tout ordre $s \in [\![1, r]\!]$.
3. Si la variable aléatoire X ne possède pas d'espérance mathématique, alors la variable X ne peut admettre de moment centré.

- **Variance et écart-type**

Nous disons que X **admet une variance** lorsque X admet un moment centré d'ordre 2. Nous appelons alors **variance de X** la valeur

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left((X-\mu)^2\right) = {_\mu m_2}(X).$$

Nous la notons également σ^2 ou σ_X^2 s'il y a plusieurs variables aléatoires à distinguer.
Nous appelons **écart-type de** X la valeur $\sqrt{\mathbb{V}\mathrm{ar}(X)}$, que nous noterons σ ou σ_X selon les cas.

Théorème : X admet une variance si et seulement si X^2 admet une espérance mathématique. La **formule de Huygens** est alors valable :

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X).$$

Propriété : Pour tous réels a et b et pour toute variable aléatoire X admettant une variance, nous avons

$$\mathbb{V}\mathrm{ar}(aX + b) = a^2\mathbb{V}\mathrm{ar}(X).$$

- ***Vocabulaire***

– Si $\mathbb{E}(X) = 0$, alors X **est une variable aléatoire centrée.**
– Si $\mathbb{V}\mathrm{ar}(X) = 1$, alors X **est une variable aléatoire réduite.**
– Si X admet une variance non nulle, la variable $X^* = \dfrac{X - \mu_X}{\sigma_X}$ est appelée **variable centrée et réduite associée à** X.

- **Fonction génératrice des moments**

Considérons la fonction $g_X(t) = \mathbb{E}\left(\exp(tX)\right)$ définie pour l'ensemble des valeurs de t pour lesquelles $\mathbb{E}\left(\exp(tX)\right) < +\infty$. $g_X(t)$ est la **fonction génératrice des moments de la variable aléatoire continue** X lorsqu'elle est définie dans un voisinage de l'origine.

III Indépendance de v.a., cas général

- **Indépendance de deux v.a.** X **et** Y

Deux variables aléatoires X **et** Y, discrètes ou continues, sont dites **indépendantes** lorsque pour tout $(x, y) \in \mathbb{R}^2$

$$\mathbb{P}([X \leqslant x] \cap [Y \leqslant y]) = \mathbb{P}([X \leqslant x])\mathbb{P}([Y \leqslant y]).$$

Propriété : Si X et Y sont indépendantes et g et h sont deux fonctions à valeurs réelles définies et continues sauf en un nombre fini de points respectivement de $X(\Omega)$ et de $Y(\Omega)$, alors $g(X)$ et $h(Y)$ sont deux v.a. indépendantes.

- **Indépendance mutuelle d'une famille de v.a.**

  1. Si l'ensemble K est fini, les v.a. $(X_k)_{k\in K}$, sont dites **mutuellement indépendantes**, lorsque, pour tout $(x_k)_{k\in K} \in \mathbb{R}^K$,

  $$\mathbb{P}\left(\bigcap_{k\in K}[X_k \leqslant x_k]\right) = \prod_{k\in K}\mathbb{P}([X_k \leqslant x_k]).$$

  2. La famille $(X_k)_{k\in K}$ est une **famille infinie de v.a. mutuellement indépendantes** si, pour tout ensemble fini $L \subset K$, les v.a. $(X_l)_{l\in L}$, sont mutuellement indépendantes.

  Théorème : Si les variables aléatoires $X_1,\ldots,X_n$ sont mutuellement indépendantes, alors pour toutes les fonctions ϕ et ψ définies et continues sauf en un nombre fini de points respectivement de $(X_1,\ldots,X_p)(\Omega)$ et de $(X_{p+1},\ldots,X_n)(\Omega)$, les v.a. $\phi(X_1,\ldots,X_p)$ et $\psi(X_{p+1},\ldots,X_n)$ sont indépendantes.

- **Indépendance deux à deux d'une famille de v.a.**

  Les v.a. $(X_k)_{k\in K}$ sont dites **deux à deux indépendantes** si, pour tout $(i,j) \in K$, X_i et X_j sont indépendantes.

  Remarque : L'indépendance mutuelle implique l'indépendance deux à deux. La réciproque n'est pas vraie.

IV Somme et covariance de deux v.a.

Soit X et Y deux v.a. indépendantes et de densité f_X et g_Y. Si la fonction h définie par

$$h(x) = \int_{-\infty}^{\infty} f_X(t)g_Y(x-t)\,\mathrm{d}t.$$

est définie et continue sauf peut-être en un nombre fini de valeurs de $\mathbb{R}$, alors h est une densité de $X+Y$.

Propositions

1. Si deux v.a. X et Y ont une variance alors XY admet une espérance.
2. Si deux v.a. X et Y ont une espérance et sont indépendantes alors XY admet une espérance $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Définition : Soit X et Y deux v.a. ayant une variance. La **covariance de X et de Y** est le nombre réel défini par

$$\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])].$$

Proposition : Soit X et Y deux v.a. ayant une variance.

$$\mathbb{C}\mathrm{ov}(X,X) = \mathbb{V}\mathrm{ar}(X) \quad \text{et} \quad \mathbb{C}\mathrm{ov}(X,Y) = \mathbb{C}\mathrm{ov}(Y,X).$$

Définition : Nous disons que deux v.a. X et Y ayant une variance ne sont pas corrélées si $\mathbb{C}\mathrm{ov}(X,Y) = 0$.

Proposition : Soit X et Y deux v.a. ayant une variance.

1. **Formule de Huygens.**

$$\mathbb{C}\mathrm{ov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Si X et Y sont indépendantes, elles ne sont pas corrélées. La réciproque n'est pas vraie.

2. $X+Y$ admet une variance et

$$\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}(X) + 2\,\mathbb{C}\mathrm{ov}(X,Y) + \mathbb{V}\mathrm{ar}(Y).$$

3. De plus, si X et Y sont indépendantes,

$$\mathbb{V}\mathrm{ar}(X+Y) = \mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y).$$

4. Si $(X_k)_{k\in K}$ est une famille finie de variables aléatoires deux à deux indépendantes et qui admettent une variance, alors $\sum_{k\in K} X_k$ admet une variance et

$$\mathbb{V}\mathrm{ar}\left(\sum_{k\in K} X_k\right) = \sum_{k\in K} \mathbb{V}\mathrm{ar}(X_k).$$

Soit X, Y et Z trois v.a. ayant une variance et a, b, c et d quatre nombres réels.
1. $\mathbb{C}\mathrm{ov}(aX+bY,\, Z) = a\,\mathbb{C}\mathrm{ov}(X,Z) + b\,\mathbb{C}\mathrm{ov}(Y,Z)$.
2. $\mathbb{C}\mathrm{ov}(aX+b,\, cY+d) = ac\,\mathbb{C}\mathrm{ov}(X,Y)$.

Une application directe du cours

Soit f la fonction définie par

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

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

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

Solution

1. a) Pour démontrer que la fonction f est une densité de probabilité, nous devons vérifier les trois points suivants :

1. $f(t) \geqslant 0$ pour tout $t \in \mathbb{R}$;
2. l'ensemble des points de discontinuité de f est fini et ces discontinuités sont de première espèce ;
3. $\int_{\mathbb{R}} f(t)\,\mathrm{d}t = 1$.

Le point 1 est vérifié par définition de la fonction f. Le point 2 est vérifié. En effet, f admet deux points de discontinuité : $t = 0$ et $t = 1$. En $t = 0$, nous avons $\lim_{t\to 0^+} f(t) = \lim_{t\to 0^-} f(t) = 0$. En $t = 1$, nous avons $\lim_{t\to 1^+} f(t) = 0$ et $\lim_{t\to 1^-} f(t) = 2$. Ces deux limites sont effectivement différentes mais elles existent. Le point 3 est vérifié car $\int_{\mathbb{R}} f(t)\,\mathrm{d}t = \int_0^1 2t\,\mathrm{d}t = [t^2]_0^1 = 1$. Donc la fonction f est une densité de probabilité.

b) Démontrons que la loi de probabilité définie par f admet une espérance mathématique. Pour cela, calculons $\mathbb{E}(X)$:

$\mathbb{E}(X) = \int_{\mathbb{R}} tf(t)\,\mathrm{d}t = \int_0^1 2t^2\,\mathrm{d}t = \left[\frac{2}{3}t^3\right]_0^1 = \frac{2}{3} < +\infty.$

Donc la loi de probabilité définie par f admet une espérance mathématique. Il nous reste à montrer que la loi de probabilité définie par f admet une variance. Pour cela, calculons $\mathbb{V}\mathrm{ar}(X)$. Comme $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X)$, il suffit de calculer $\mathbb{E}\left(X^2\right)$:

$\mathbb{E}\left(X^2\right) = \int_{\mathbb{R}} t^2 f(t)\,\mathrm{d}t = \int_0^1 2t^3\,\mathrm{d}t = \left[\frac{1}{2}t^4\right]_0^1 = \frac{1}{2}.$

Par conséquent, nous avons : $\mathbb{V}\mathrm{ar}(X) = 1/2 - 4/9$. Donc la loi de probabilité définie par f admet une variance.

c) La fonction de répartition associée à f, notée F est égale à :

$$F(x) = \begin{cases} 0 & \text{si } x \in]-\infty,0] \\ x^2 & \text{si } x \in [0,1] \\ 1 & \text{si } x \in [1,+\infty[\end{cases}$$

2. Déterminons F_Y. Par définition, nous avons, pour tout réel y,

$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}\left(X^2 + 1 \leqslant y\right) = \mathbb{P}\left(X^2 \leqslant y - 1\right).$

Deux cas se distinguent.

Si $y \leqslant 1$, alors $F_Y(y) = 0$.

Si $y \geqslant 1$, alors $F_Y(y) = \mathbb{P}\left(-\sqrt{y-1} \leqslant X \leqslant \sqrt{y-1}\right) = \mathbb{P}\left(0 \leqslant X \leqslant \sqrt{y-1}\right)$.

À nouveau deux cas se présentent.

Si $\sqrt{y-1} \leqslant 1$ ou encore $1 \leqslant y \leqslant 2$, alors $F_Y(y) = y - 1$.

Si $\sqrt{y-1} \geqslant 1$ ou encore $y \geqslant 2$, alors $F_Y(y) = 1$.

En conclusion, nous avons $F_Y(y) = \begin{cases} 0 & \text{si } y \in]-\infty,1] \\ y-1 & \text{si } y \in [1,2] \\ 1 & \text{si } y \in [2,+\infty[\end{cases}$

FICHE 14

Loi normale ou de Laplace-Gauss

I La loi normale centrée-réduite

- ***Définition***

Une variable aléatoire X à valeurs dans $\mathbb{R}$ suit une **loi normale centrée-réduite**, notée $\mathcal{N}(0,1)$, si X est une variable continue et admet pour densité de probabilité la fonction f_X suivante

$$f_X(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right), \quad \text{pour tout } t \in \mathbb{R}.$$

- ***Propriétés de la loi normale centrée-réduite***

1. Le graphe de f_X a l'allure d'une courbe en cloche assez aplatie. Voir la courbe noire de la figure 14.1.
2. La fonction de répartition de la loi normale $\mathcal{N}(0,1)$, notée généralement Φ, est égale à

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t, \quad \text{pour tout } x \in \mathbb{R}.$$

Le graphe de Φ a l'allure d'une courbe en S assez étalée et est symétrique par rapport au point $(0,1/2)$ et la pente de la tangente en ce point est $1/\sqrt{2\pi}$. Voir la courbe noire de la figure 14.2.

3. Soit X une variable aléatoire qui suit la loi normale $\mathcal{N}(0,1)$. Nous avons : $\mathbb{E}(X) = 0$ et $\mathbb{V}\mathrm{ar}(X) = 1$. C'est la raison pour laquelle la loi est appelée centrée-réduite et est notée $\mathcal{N}(0,1)$.
4. La fonction génératrice des moments de la loi normale $\mathcal{N}(0,1)$ est égale, pour tout $t \in \mathbb{R}$ à

$$g_X(t) = \mathbb{E}(\exp(tX)) = \exp\left(\frac{t^2}{2}\right).$$

Remarques

1. g_X est une fonction paire, ce qui a pour conséquence que les coefficients d'ordre impair dans un développement en série entière sont nuls.
2. g_X est utilisée pour calculer les moments d'ordre n. En effet, g_X est développable en série entière pour tout réel t. Ainsi le coefficient de $t^n/n!$ dans le développement de g_X autour de 0 donne le moment d'ordre n.

5. Nous avons, pour tout $n \in \mathbb{N}$: $\mathbb{E}(X^{2n}) = \dfrac{(2n)!}{2^n\, n!}$ et $\mathbb{E}(X^{2n+1}) = 0$.
6. En particulier, le coefficient d'asymétrie de la loi normale $\mathcal{N}(0,1)$ est nul et le coefficient d'aplatissement de la loi normale $\mathcal{N}(0,1)$ est égal à 3.

II La loi normale de paramètres μ et σ

- ***Définition***

Une variable aléatoire X à valeurs dans $\mathbb{R}$ suit une loi **normale de paramètres** $\mu \in \mathbb{R}$ **et** $\sigma > 0$, notée $\mathcal{N}(\mu,\sigma)$, si X est une variable continue et admet pour densité de probabilité la fonction $f_{\mu,\sigma}$ suivante

$$f_{\mu,\sigma}(t) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right), \quad \text{pour tout } t \in \mathbb{R}.$$

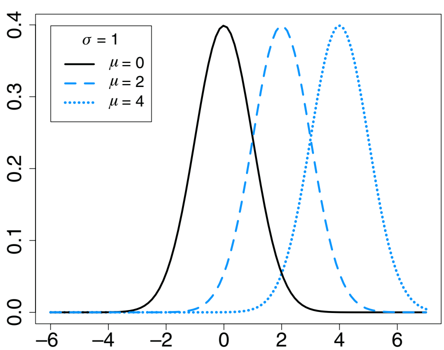


Figure 14.1 Densités de lois normales pour σ fixé

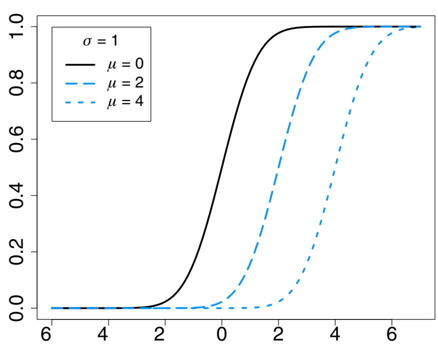


Figure 14.2 Fonctions de répartition de lois normales pour σ fixé

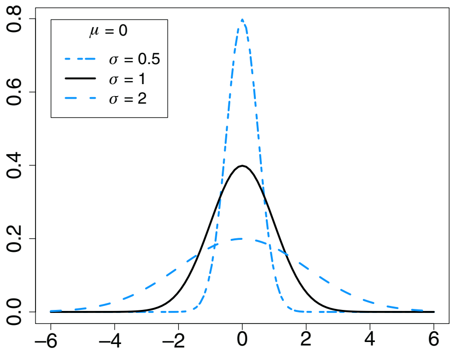


Figure 14.3 Densités de lois normales pour μ fixé

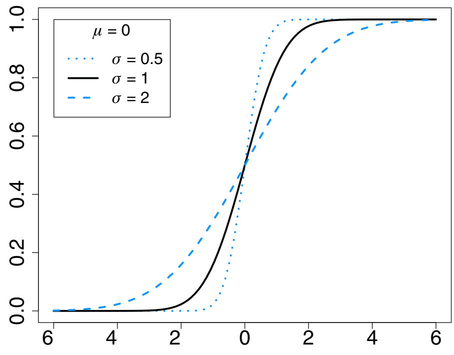


Figure 14.4 Fonctions de répartition de lois normales pour μ fixé

Remarque : Il est à noter que nous notons ici la loi normale $\mathcal{N}(\mu,\sigma)$ et non $\mathcal{N}(\mu,\sigma^2)$. En effet, en statistique, nous aimons avoir des quantités comparables et par conséquent nous utilisons l'écart-type σ et non la variance σ^2. Ainsi les deux paramètres sont exprimés dans la même unité. Il faudra donc faire attention à la notation utilisée dans les livres et dans les logiciels.

- ***Propriétés de la loi normale $\mathcal{N}(\mu,\sigma)$***

1. Le graphe de $f_{\mu,\sigma}$ a l'allure d'une courbe en cloche symétrique par rapport à $x = \mu$, très pointue pour σ petit, très aplatie pour σ grand. Voir la figure 14.1 pour $\sigma = 1$ et différentes valeurs de μ et 14.3 pour $\mu = 0$ et différentes valeurs de σ.
2. La fonction de densité d'une loi normale $\mathcal{N}(\mu,\sigma)$ vérifie

$$f_{\mu,\sigma}(\mu + u) = f_{\mu,\sigma}(\mu - u).$$

3. La fonction de répartition d'une loi normale $\mathcal{N}(\mu,\sigma)$ est égale à

$$F_{\mu,\sigma}(x) = \mathbb{P}(X \leqslant x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Son graphe a l'allure d'une courbe en S, très pentue pour σ petit, très étirée pour σ grand. Voir la figure 14.2 pour $\sigma = 1$ et différentes valeurs de μ et 14.4 pour $\mu = 0$ et différentes valeurs de σ.

4. La fonction de répartition d'une loi normale $\mathcal{N}(\mu,\sigma)$ vérifie

$$F_{\mu,\sigma}(\mu - x) = 1 - F_{\mu,\sigma}(\mu + x).$$

Remarque : Cette propriété est mise à profit dans les tables de la loi-normale centrée-réduite où elle permet de ne mentionner que les valeurs-de F_X correspondant aux valeurs positives de x.

5. Soit X une variable aléatoire qui suit une loi normale $\mathcal{N}(\mu,\sigma)$. Nous avons : $\mathbb{E}(X) = \mu$ et $\mathbb{V}\mathrm{ar}(X) = \sigma^2$.
6. La fonction génératrice des moments d'une loi normale $\mathcal{N}(\mu,\sigma)$ est égale, pour tout $t \in \mathbb{R}$ à

$$g_{\mu,\sigma}(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

7. Il en résulte de la forme de la fonction génératrice des moments que si (X_1,X_2) est un couple de variables aléatoires indépendantes de lois respectives $\mathcal{N}(\mu_1,\sigma_1)$ et $\mathcal{N}(\mu_2,\sigma_2)$, avec μ_1,μ_2 des réels et $\sigma_1 > 0,\sigma_2 > 0$, alors la somme $X = X_1 + X_2$ suit une loi normale $\mathcal{N}(\mu,\sigma)$ où $\mu = \mu_1 + \mu_2$ et $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Pour commencer avec la loi normale

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

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

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

a. Quelle est la loi de la variable aléatoire $Y - 2X$?

b. Déterminez μ sachant que $\mathbb{P}(Y \geqslant 2X) = 0{,}516$.

Solution

1. Calculons $\mathbb{P}(X \geqslant 4 | X \leqslant 5{,}2)$. Remarquons que cette probabilité est égale à $\mathbb{P}(4 \leqslant X \leqslant 5{,}2)/\mathbb{P}(X \leqslant 5{,}2)$. Or, d'après l'énoncé, la variable X suit une loi normale $\mathcal{N}(5{,}2;\, 0{,}8)$.
Il ne reste donc plus qu'à centrer et réduire X pour se ramener à une loi normale centrée-réduite. Ainsi, nous avons

$$\mathbb{P}(X \geqslant 4 | X \leqslant 5{,}2) = \mathbb{P}\left(\frac{-1{,}2}{0{,}8} \leqslant \frac{X-5{,}2}{0{,}8} \leqslant 0\right) \Big/ \mathbb{P}\left(\frac{X-5{,}2}{0{,}8} \leqslant 0\right)$$
$$= \big(\Phi(0) - \Phi(-1{,}5)\big) \big/ \Phi(0).$$

Il ne reste plus qu'à lire ces deux valeurs dans une table de la loi normale centrée-réduite. Le seul problème est le calcul de la valeur $\Phi(-1{,}5)$. En effet, les valeurs qui se trouvent dans la table sont toutes positives. Il faut donc utiliser l'égalité suivante : pour tout $x \geqslant 0$, $\Phi(-x) + \Phi(x) = 1$.
Ainsi il nous reste plus qu'à calculer : $\Phi(0) + \Phi(1{,}5) - 1$.
En utilisant la table de la fonction de répartition de la loi normale centrée-réduite, nous trouvons que $\Phi(0) = 0{,}5$ et $\Phi(1{,}5) = 0{,}9332$.
Ainsi, nous en déduisons que $\mathbb{P}(X \geqslant 4 | X \leqslant 5{,}2) = 0{,}4332 \times 2 = 0{,}8664$.

2. a. Comme les variables Y et X sont indépendantes et suivent toutes les deux une loi normale, la variable $Y - 2X$ suit aussi une loi normale. Il ne reste plus qu'à déterminer la moyenne et l'écart-type de cette loi. Pour cela, calculons $\mathbb{E}(Y - 2X)$. Nous avons $\mathbb{E}(Y-2X) = \mathbb{E}(Y) - 2\mathbb{E}(X) = \mu - 2 \times 5{,}2 = \mu - 10{,}4$. Calculons maintenant $\mathbb{V}\mathrm{ar}(Y - 2X)$. Nous avons $\mathbb{V}\mathrm{ar}(Y-2X) = \mathbb{V}\mathrm{ar}(Y) + 4\mathbb{V}\mathrm{ar}(X) = 0{,}8 + 4 \times 0{,}8 = 4{,}0$. En conclusion, $Y - 2X$ suit une loi normale $\mathcal{N}(\mu - 10{,}4; 2)$.

b. Déterminons μ tel que $\mathbb{P}(Y \geqslant 2X) = 0{,}516$. Remarquons que $\mathbb{P}(Y \geqslant 2X) = \mathbb{P}(Y - 2X \geqslant 0)$. Nous sommes ramenés à étudier la variable aléatoire précédente $Y - 2X$ qui suit une loi normale non centrée et non réduite. Nous allons donc centrer et réduire. Ainsi, nous avons

$$\begin{aligned}\mathbb{P}(Y \geqslant 2X) &= \mathbb{P}\left(\frac{Y - 2X - (\mu - 10{,}4)}{2} \geqslant \frac{-(\mu - 10{,}4)}{2}\right)\\ &= 1 - \Phi\left(\frac{-(\mu-10{,}4)}{2}\right) = 0{,}516.\end{aligned}$$

Par conséquent, nous cherchons μ tel que $\Phi\left(\dfrac{\mu - 10{,}4}{2}\right) = 0{,}516$. En lisant dans la table de la fonction de répartition de la loi normale centrée-réduite, nous en déduisons que $\mu = 10{,}48$.

Un système à résoudre

Deux variables aléatoires X et Y suivent respectivement les lois $\mathcal{N}(150; \sigma)$ et $\mathcal{N}(100; \sigma)$. Il existe un réel a tel que $\mathbb{P}(X \leqslant a) = 0{,}017$ et $\mathbb{P}(Y > a) = 0{,}005$. Déterminez σ.

Solution

Tout au long de cette correction, nous supposerons que σ est non nul. En utilisant la première égalité et en centrant et réduisant X, nous obtenons que $\Phi\left(\frac{a-150}{\sigma}\right) = 0{,}017$ ou encore $1 - \Phi\left(-\frac{a-150}{\sigma}\right) = 0{,}017$. Ainsi, en utilisant la table de la fonction de répartition de la loi normale centrée-réduite, nous en déduisons que $\frac{150-a}{\sigma} = 2{,}12$.
En utilisant la seconde égalité, en remarquant que $\mathbb{P}(Y > a) = 1 - \mathbb{P}(Y \leqslant a)$ et en centrant et réduisant Y, nous obtenons que $\Phi\left(\frac{a-100}{\sigma}\right) = 0{,}996$. Ainsi, en utilisant la table de la fonction de répartition de la loi normale centrée-réduite, nous en déduisons que $\frac{a-100}{\sigma} = 2{,}65$.
Par conséquent, nous sommes ramenés à résoudre un système de deux équations à deux inconnues, où seul σ nous intéresse.

$$\left\{\begin{array}{lcl} \dfrac{150-a}{\sigma} & = & 2{,}12 \\ \dfrac{a-100}{\sigma} & = & 2{,}65 \end{array}\right.$$

En le résolvant, nous trouvons que $\sigma = 5000/477 \approx 10{,}48$.
Ce qui confirme bien l'hypothèse que nous avions fait au début de cette correction à savoir $\sigma \neq 0$.

FICHE 15

Lois dérivées de la loi normale

I Loi du Khi-deux $\chi^2(p)$

• ***Définition***

Soit p un entier positif. Une variable aléatoire X suit une **loi de Pearson ou loi du Khi-deux** à p degrés de liberté, notée $\chi^2(p)$, si X est une variable continue et admet pour densité de probabilité la fonction f_X suivante

$$f_X(t) = \begin{cases} \dfrac{1}{2^{\frac{p}{2}}\Gamma\left(\frac{p}{2}\right)} t^{\frac{p}{2}-1}\exp\left(-\dfrac{t}{2}\right) & \text{pour } t \geqslant 0 \\ 0 & \text{pour } t < 0 \end{cases}$$

où $\Gamma(r) = \displaystyle\int_0^{+\infty} t^{r-1}\exp(-t)\,\mathrm{d}t$ est la fonction gamma d'Euler.

Remarque : La figure 15.1 montre le graphe de la densité de probabilité f_X pour différentes valeurs du paramètre p.

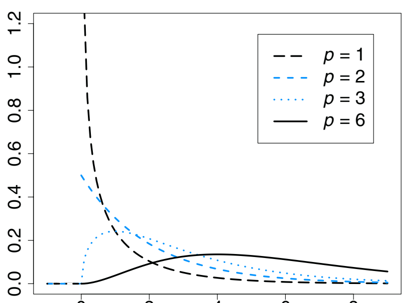


Figure 15.1. Densités

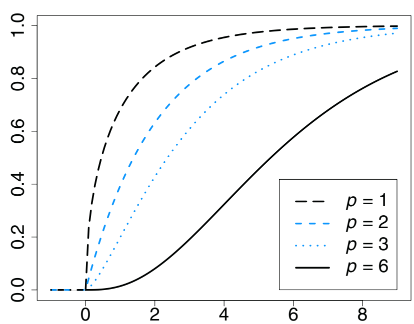


Figure 15.2. Fonctions de répartition

- ***Propriétés***

1. La fonction de répartition ne s'explicite pas. Cependant, il existe des tables de la fonction de répartition. La figure 15.2 montre le graphe de la fonction de répartition pour différentes valeurs du paramètre p.
2. Soit X une variable aléatoire qui suit une loi du Khi-deux $\chi^2(p)$. Nous avons : $\mathbb{E}(X) = p \quad \text{et} \quad \mathbb{V}\text{ar}(X) = 2p$.
3. Soit $X_1, \ldots, X_n$ une suite de v.a. indépendantes et identiquement distribuées qui suit la loi normale $\mathcal{N}(0,1)$. Alors la v.a. $\displaystyle\sum_{i=1}^{n} X_i^2$ suit la loi du Khi-deux $\chi^2(n)$.
4. Soit une v.a. X_1 qui suit une loi du Khi-deux $\chi^2(n_1)$ et X_2 une v.a. qui suit indépendamment de X_1 une loi du Khi-deux $\chi^2(n_2)$. Alors la v.a. $X_1 + X_2$ suit une loi du Khi-deux $\chi^2(n_1 + n_2)$.

II Loi de Student $t(n)$

- ***Définition***

Soit n un entier strictement positif. Une variable aléatoire X suit une **loi de Student** à n degrés de liberté, notée $t(n)$, si X est une variable continue et admet pour densité de probabilité la fonction f_X suivante

$$f_X(t) = \begin{cases} \dfrac{1}{\sqrt{n\pi}} \dfrac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)} \dfrac{1}{\left(1 + \dfrac{t^2}{n}\right)^{\frac{n+1}{2}}} & \text{pour } t \geqslant 0 \\ 0 & \text{pour } t < 0 \end{cases}$$

La figure 15.3 montre le graphe de la densité de probabilité f_X pour différentes valeurs du paramètre n.

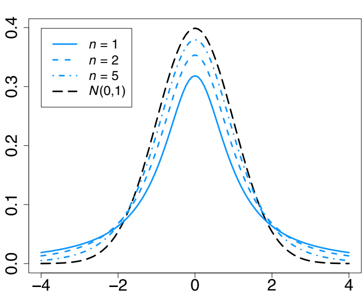


Figure 15.3. Densités

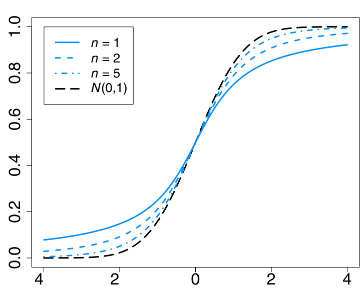


Figure 15.4. Fonctions de répartition

- ***Propriétés***

1. La fonction de répartition ne s'explicite pas. Cependant, il existe des tables de la fonction de répartition. La figure 15.4 montre le graphe de la fonction de répartition pour différentes valeurs du paramètre n.
2. Soit X une variable aléatoire qui suit une loi de Student $t(n)$. Nous avons :

$$\begin{array}{lll} \text{si} & n \geqslant 2 & \mathbb{E}(X) = 0 \\ \text{si} & n \geqslant 3 & \mathbb{V}\text{ar}(X) = \dfrac{n}{n-2}\cdot \end{array}$$

3. Soit une v.a. U qui suit une loi normale $\mathcal{N}(0,\ 1)$ et X une v.a. qui suit indépendamment de U une loi du Khi-deux $\chi^2(n)$. Alors la variable aléatoire $T_n = \dfrac{U}{\sqrt{X/n}}$ suit une loi de Student $t(n)$.
4. Pour $n = 1$, la loi est la **loi de Cauchy**. La loi de Cauchy est en fait la loi du rapport de deux variables qui suivent chacune indépendamment la loi normale $\mathcal{N}(0,1)$.

III Loi de Fisher-Snedecor $F(n,p)$

- ***Définition***

Soit n et p deux entiers positifs. Une variable aléatoire X suit une **loi de Fisher-Snedecor** (ou loi de Fisher) à n et p degrés de liberté, notée $F(n,p)$, si X est une variable continue et admet pour densité de probabilité la fonction f_X suivante

$$f_X(t) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n+p}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{p}{2}\right)} \left(\dfrac{n}{p}\right)^{\frac{n}{2}} \dfrac{t^{\frac{n-2}{2}}}{\left(1+\dfrac{n}{p}t\right)^{\frac{n+p}{2}}} & \text{pour } t \geqslant 0 \\ 0 & \text{pour } t < 0 \end{cases}$$

Les figures 15.5 et 15.7 montrent le graphe de la densité de probabilité f_X pour différentes valeurs des paramètres n et p.

- ***Propriétés***

1. La fonction de répartition ne s'explicite pas. Cependant, il existe des tables de la fonction de répartition. Les figures 15.6 et 15.8 montrent le graphe de la fonction de répartition pour différentes valeurs des paramètres n et p.

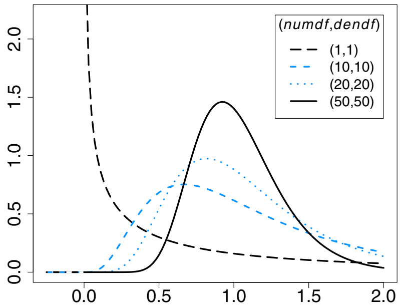


Figure 15.5. Densités

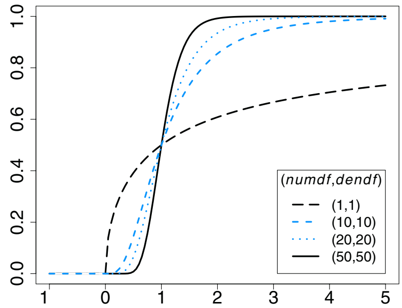


Figure 15.6. Fonctions de répartition

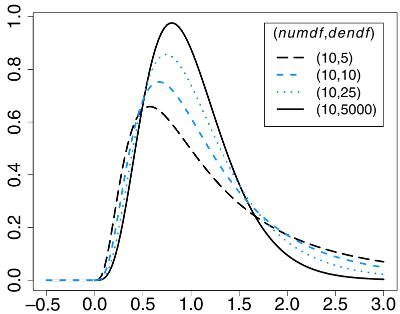


Figure 15.7. Densités

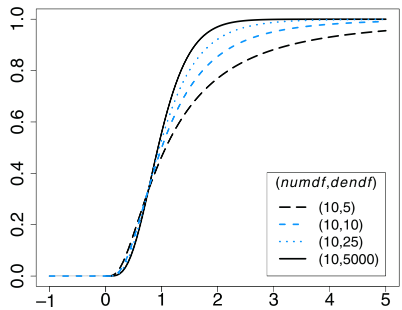


Figure 15.8. Fonctions de répartition

2. Soit X une variable aléatoire qui suit une loi de Fisher $F(n,p)$. Nous avons :

$$\text{si } p \geqslant 3 \quad \mathbb{E}(X) = \frac{p}{p-2}$$

$$\text{si } p \geqslant 5 \quad \mathbb{V}\text{ar}(X) = \frac{2p^2(n+p-2)}{n(p-2)^2(p-4)}.$$

3. Soit X et Y étant deux variables aléatoires suivant indépendamment des lois du χ^2 à n et p degrés de liberté respectivement. Alors la variable aléatoire $\dfrac{X/n}{Y/p}$ suit une loi de Fisher-Snedecor à n degrés de liberté au numérateur et p degrés de liberté au dénominateur.

D'après un sujet de concours de Pharmacie

Soit X la variable aléatoire d'ensemble fondamental $\mathbb{R}^+$ et dont une densité de probabilité f_X est égale à

$$f_X(t) = \frac{t^n \exp\left(-\dfrac{t}{2}\right)}{2^{n+1} n!}, \quad t \geqslant 0.$$

1. Identifiez la variable aléatoire X.

2. Calculez les moments d'ordre 1 et 2 de X.

3. Nous choisissons $n = 5$. Déterminez les deux nombres x_1 et x_2 tels que : $\mathbb{P}_X([0,x_1]) = \mathbb{P}_X([x_2,+\infty[) = 0{,}050$. Que vaut $\mathbb{P}(]x_1,x_2[)$?

Solution

1. Il résulte de la définition que X suit une loi du Khi-deux à $2n+2$ degrés de liberté.

2. – Le moment d'ordre 1 de X n'est rien d'autre que l'espérance mathématique et par conséquent est égal à $2n+2$.

– Pour calculer le moment d'ordre 2 de X, nous allons nous servir de la relation $\mathbb{V}\text{ar}\left(X\right) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2\left(X\right)$. Par conséquent, nous avons

$$\mathbb{E}\left(X^2\right) = 2(2n+2) + (2n+2)^2 = (2n+2)(2n+4) = 4(n+1)(n+2).$$

3. Si $n = 5$, alors $p = 2 \times 5 + 2 = 12$. Nous utiliserons donc la table du χ^2 à 12 degrés de liberté. Déterminons x_1 et x_2. D'après l'énoncé, nous savons que :

$$\mathbb{P}_X([0,x_1]) = 0{,}050.$$

En lisant la table de la loi du χ^2 régions unilatérales, nous obtenons : $x_1 = 5{,}2260$. D'autre part, nous savons que :

$$\mathbb{P}_X([0,x_2]) = 1 - \mathbb{P}_X([x_2,+\infty[) = 1 - 0{,}050 = 0{,}950.$$

En lisant une fois de plus la table de la loi du χ^2 régions unilatérales, nous obtenons : $x_2 = 21{,}0261$.

Calculons $\mathbb{P}(]x_1,x_2[)$. Nous avons :

$$\begin{aligned}\mathbb{P}(]x_1,x_2[&= 1 - \mathbb{P}_X([0,x_1]) - \mathbb{P}_X([x_2,+\infty[) \\ &= 1 - 2 \times 0{,}05 = 1 - 0{,}10 = 0{,}90.\end{aligned}$$

FICHE 16

Lois continues

I Loi uniforme continue $\mathcal{U}[a,b]$

- ***Définition***

Une variable aléatoire$X$ à valeurs dans $[a, b]$, où a et b sont deux réels tels que $a < b$ suit une **loi uniforme continue** sur $[a, b]$, notée $\mathcal{U}[a, b]$, si X est une variable continue et admet pour densité de probabilité la fonction f_X suivante

$$f_X(t) = \begin{cases} \dfrac{1}{b-a} & \text{si} \quad a \leqslant t \leqslant b \\ 0 & \text{sinon} \end{cases}$$

- ***Propriétés***

1. La fonction de répartition d'une loi uniforme $\mathcal{U}[a, b]$ est égale à

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < a \\ \dfrac{x-a}{b-a} & \text{si} \quad a \leqslant x \leqslant b \\ 1 & \text{si} \quad x > b \end{cases}$$

2. Soit X une variable aléatoire qui suit une loi uniforme $\mathcal{U}[a, b]$. Nous avons : $\mathbb{E}(X) = \dfrac{a+b}{2}$ et $\mathbb{V}\text{ar}(X) = \dfrac{(b-a)^2}{12}$.

- **Remarque**

La loi uniforme continue de référence est la loi $\mathcal{U}[0, 1]$, correspondant aux générateurs de nombres au hasard des logiciels. À partir d'un tel générateur, des nombres au hasard sur $[a, b]$ sont produits par la transformation $y = (b-a)x + a$. Nous verrons à la fiche 17 comment simuler une loi quelconque à partir de ces « nombres au hasard ».

II Loi exponentielle $\mathcal{E}(\lambda)$

- ***Définition***

Une variable aléatoire X à valeurs dans $[0,+\infty[$ suit une **loi exponentielle de paramètre** λ ($\lambda > 0$), notée $\mathcal{E}(\lambda)$, si X est une variable continue et admet pour densité de probabilité la fonction f_X suivante

$$f_X(t) = \begin{cases} \lambda \exp(-\lambda t) & \text{pour} \quad t \geqslant 0 \\ 0 & \text{pour} \quad t < 0 \end{cases}.$$

- ***Propriétés***

1. La fonction de répartition d'une loi exponentielle $\mathcal{E}(\lambda)$ est égale à

$$F_X(x) = \begin{cases} 1 - \exp(-\lambda x) & \text{si} \quad x \geqslant 0 \\ 0 & \text{si} \quad x \leqslant 0 \end{cases}$$

Remarque : Il est préférable de travailler avec la **fonction de survie** qui est définie par $r(x) = 1 - F_X(x) = \mathbb{P}(X > x)$.

2. Soit X une variable aléatoire qui suit une loi exponentielle $\mathcal{E}(\lambda)$. Nous avons :

$$\mathbb{E}(X) = \frac{1}{\lambda} \quad \text{et} \quad \mathbb{V}\text{ar}(X) = \frac{1}{\lambda^2}.$$

III Loi gamma $\Gamma(r,\lambda)$

La loi exponentielle est un cas particulier d'une famille de lois appelées lois gamma.

- ***Définition***

Une variable aléatoire X à valeurs dans $[0,+\infty[$ suit une **loi gamma de paramètres** r **et** λ ($r > 0$, $\lambda > 0$), notée $\Gamma(r,\lambda)$, si X est une variable continue et admet pour densité de probabilité (voir figures 16.1 et 16.3) la fonction f_X suivante

$$f_X(t) = \begin{cases} \dfrac{\lambda^r}{\Gamma(r)} t^{r-1} \exp(-\lambda t) & \text{pour} \quad t \geqslant 0 \\ 0 & \text{pour} \quad t < 0, \end{cases}$$

où $\Gamma(r) = \displaystyle\int_0^{+\infty} t^{r-1} \exp(-t)\,\mathrm{d}t$ est la fonction gamma d'Euler.

Remarque : Pour $r = 1$, $\Gamma(1,\lambda)$ coïncide avec $\mathcal{E}(\lambda)$.

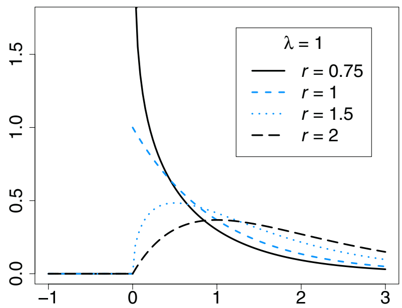


Figure 16.1. Densités

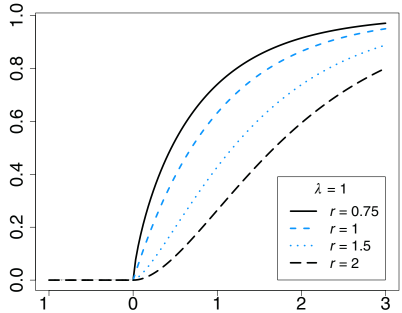


Figure 16.2. Fonctions de répartition

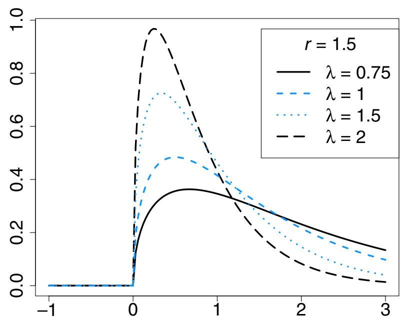


Figure 16.3. Densités

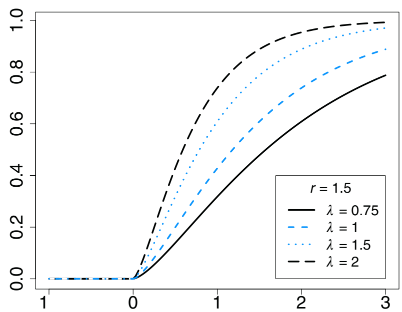


Figure 16.4. Fonctions de répartition

- ***Propriétés***

1. La fonction de répartition d'une loi gamma $\Gamma(r, \lambda)$ n'a pas de forme explicite. Par exemple, les figures 16.2 et 16.4 montrent le graphe de la fonction de répartitions de diverses lois gammas.
2. Soit X une variable aléatoire qui suit une loi gamma $\Gamma(r, \lambda)$. Nous avons : $\mathbb{E}(X) = \dfrac{r}{\lambda}$ et $\mathbb{V}\mathrm{ar}(X) = \dfrac{r}{\lambda^2}$.
3. Si (X_1, X_2) est un couple de variables aléatoires indépendantes de lois respectivement gamma $\Gamma(r_1, \lambda)$ et gamma $\Gamma(r_2, \lambda)$ alors la somme $X_1 + X_2$ suit une loi gamma $\Gamma(r_1 + r_2, \lambda)$.
4. La loi gamma $\Gamma(n, \lambda)$ (n entier > 0, $\lambda > 0$) coïncide avec la loi de la somme de n variables aléatoires indépendantes suivant toutes la loi exponentielle $\mathcal{E}(\lambda)$.
5. Soit Y une variable aléatoire qui suit la loi normale $\mathcal{N}(0,1)$. Alors $X = Y^2$ a pour loi $\Gamma(1/2, 1/2)$.

6. Soit $(Y_1, \ldots, Y_n)$ un système de n variables aléatoires indépendantes suivant la loi normale $\mathcal{N}(0,1)$. Alors $Y_1^2 + \ldots + Y_n^2$ suit une loi gamma $\Gamma(n/2, 1/2)$. Cette loi n'est rien d'autre que la loi du khi-deux à n degrés de liberté (voir fiche 15).

IV Lois bêta $Beta(n,p)$

- ***Définition***

Une variable aléatoire X à valeurs dans [0, 1] suit une **loi bêta de paramètres n et p** ($n > 0$, $p > 0$), notée $Beta(n, p)$, si X est une variable continue et admet pour densité de probabilité (voir figure 16.5 et 16.7) la fonction f_X suivante

$$f_X(t) = \begin{cases} \dfrac{1}{B(n,p)} t^{n-1}(1-t)^{p-1} & \text{pour} \quad 0 \leqslant t \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

où $B(n,p) = \dfrac{\Gamma(n)\Gamma(p)}{\Gamma(n+p)}$ qui est la fonction bêta.

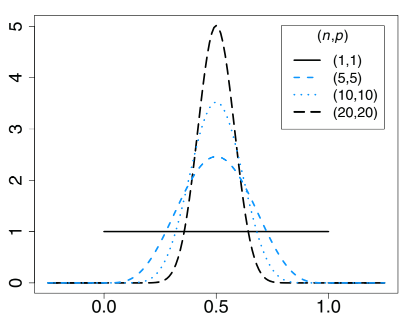


Figure 16.5. Densités

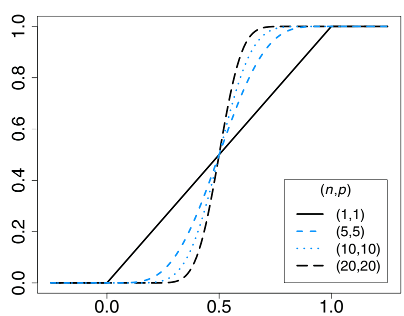


Figure 16.6. Fonctions de répartition

- ***Propriétés***

1. La fonction de répartition d'une loi bêta $Beta(n, p)$ n'a pas de forme explicite. Par exemple, les figures 16.6 et 16.8 montre le graphe de la fonction de répartitions de diverses lois bêtas.
2. Soit X une variable aléatoire qui suit une loi bêta $Beta(n, p)$. Nous avons :

$$\mathbb{E}(X) = \frac{n}{n+p} \quad \text{et} \quad \mathbb{V}\text{ar}(X) = \frac{np}{(n+p+1)(n+p)^2}.$$

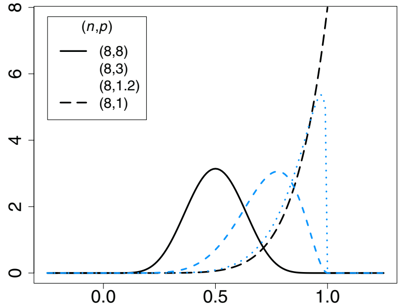


Figure 16.7. Densités

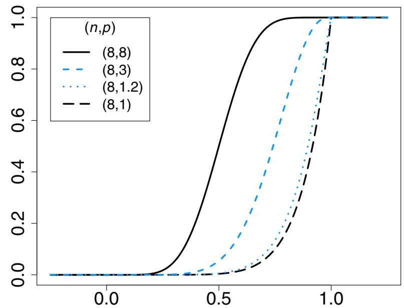


Figure 16.8. Fonctions de répartition

La loi lognormale $\mathcal{LN}(\mu, \sigma)$

Soit X une v.a. à valeurs dans $\mathbb{R}^{+*}$. X suit une loi lognormale de paramètres μ et σ^2, notée $\mathcal{LN}(\mu, \sigma)$, si $\ln(X)$ suit une loi normale $\mathcal{N}(\mu, \sigma)$.

1. Calculez la densité de probabilité f_X associée à X.

2. Calculez l'espérance mathématique et la variance de X.

Pour cela, vous utiliserez le fait que $\mathbb{E}(X^k) = \mathbb{E}(\exp(k \ln X)) = \mathbb{E}(\exp(kY))$ $= g_Y(k) = \exp\left(\mu k + \dfrac{\sigma^2 k^2}{2}\right)$, où Y suit une loi normale $\mathcal{N}(\mu, \sigma)$ et pour la dernière égalité, voir la fiche 14.

3. D'après Michel Lejeune. Application : Grâce à une étude épidémiologique nous constatons que la distribution des poids des individus dans une population adulte suit une loi lognormale. Considérant que le poids moyen est de 70 kg et que l'écart-type des poids est de 12 kg, résolvez les deux équations permettant de déterminer les valeurs des paramètres μ et σ de la loi lognormale.

Solution

Nous supposerons dans tout l'exercice que σ est non nul.

1. Calculons f_X. Pour cela, calculons la fonction de répartition associée à X. Puis dérivons-la et nous obtiendrons ainsi f_X.

Nous avons $F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(\ln X \leqslant \ln x) = \Phi\left(\dfrac{\ln x - \mu}{\sigma}\right)$.

En dérivant cette dernière égalité, nous obtenons,

pour tout $x > 0$ $f_X(x) = \frac{1}{x} \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \left(\frac{\ln x - \mu}{\sigma} \right)^2 \right)$ et

pour $x \leqslant 0, f_X(x) = 0$.

2. Calculons $\mathbb{E}(X)$. En se servant de l'indication, pour $k = 1$, nous avons $\mathbb{E}(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$. Calculons maintenant $\mathbb{V}\mathrm{ar}(X)$. Pour cela, nous utilisons encore une fois le théorème de Huygens : $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X)$. Il reste donc à calculer $\mathbb{E}\left(X^2\right)$. Nous nous servons à nouveau de l'indication donnée dans l'énoncé. Pour $k = 2$, nous avons $\mathbb{E}\left(X^2\right) = \exp\left(2\mu + 2\sigma^2\right)$. Donc, nous en déduisons que :

$$\mathbb{V}\mathrm{ar}(X) = \exp\left(2\mu + 2\sigma^2\right) - \left(\exp\left(\mu + \frac{\sigma^2}{2}\right)\right)^2 = \exp\left(2\mu + \sigma^2\right)\left(\exp(\sigma^2) - 1\right).$$

3. Comme le poids moyen est de 70 kg, nous en déduisons que $\mathbb{E}(X) = 70$ kg. De plus, comme l'écart-type des poids est de 12 kg, nous en déduisons que $\mathbb{V}\mathrm{ar}(X) = 144$ kg². Ainsi le système s'écrit :

$$\left\{ \begin{array}{rcl} \exp\left(\mu + \frac{\sigma^2}{2}\right) & = & 70 \\ \exp\left(2\mu + \sigma^2\right)\left(\exp(\sigma^2) - 1\right) & = & 144 \end{array} \right.$$

ou encore

$$\left\{ \begin{array}{rcl} \mu + \frac{\sigma^2}{2} & = & \ln(70) \\ 2\left(\mu + \frac{\sigma^2}{2}\right) + \ln\left(\exp(\sigma^2) - 1\right) & = & \ln(144) \end{array} \right.$$

Après calculs, nous obtenons $\mu = \ln\left(\frac{2450}{\sqrt{1261}}\right)$ et $\sigma = \sqrt{\ln\left(\frac{1261}{1225}\right)}$.

La loi de Pareto

Soit Y une v.a. exponentielle de loi $\mathcal{E}(\lambda)(\lambda > 0)$. La v.a. $X = \exp(Y)$ suit une loi de Pareto de paramètres λ et 1, notée $\mathcal{PA}(\lambda, 1)$. Habituellement, le deuxième paramètre de la loi de Pareto n'est pas égal à 1, mais est quelconque. C'est un paramètre de seuil, noté par exemple x_{min}.

1. Calculez la densité de probabilité f_X associée à X.

2. Calculez l'espérance mathématique et la variance de X.

Solution

1. Calculons f_X. Pour cela, calculons la fonction de répartition associée à X. Puis dérivons-la et nous obtiendrons ainsi f_X. Nous avons, pour $x \geqslant 1$, $F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(\exp(Y) \leqslant x) = \mathbb{P}(Y \leqslant \ln(x)) = 1 - x^{-\lambda}$, en se souvenant que la v.a. Y suit une loi exponentielle $\mathcal{E}(\lambda)$. En dérivant cette dernière égalité, nous obtenons, pour tout $x \geqslant 1$, $f_X(x) = \lambda \left(\frac{1}{x}\right)^{\lambda+1}$. Pour $x < 1$, $f_X(x) = 0$.

2. Calculons $\mathbb{E}(X)$. Nous avons, pour tout $\lambda > 1$ (ce qui assure la convergence de l'intégrale qui va suivre),

$$\mathbb{E}(X) = \int_1^{+\infty} x \times \lambda \left(\frac{1}{x}\right)^{\lambda+1} \mathrm{d}x = \frac{\lambda}{\lambda - 1}\cdot$$

Calculons maintenant $\mathbb{V}\mathrm{ar}(X)$. Pour cela, nous utilisons encore une fois le théorème de Huygens : $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}^2(X)$. Il reste donc à calculer $\mathbb{E}\left(X^2\right)$. Nous avons, pour tout $\lambda > 2$ (ce qui assure la convergence de l'intégrale qui va suivre),

$$\mathbb{E}\left(X^2\right) = \int_1^{+\infty} x^2 \times \lambda \left(\frac{1}{x}\right)^{\lambda+1} \mathrm{d}x = \frac{\lambda}{\lambda - 2}\cdot$$

Donc, nous en déduisons que

$$\mathbb{V}\mathrm{ar}(X) = \frac{\lambda}{\lambda - 2} - \left(\frac{\lambda}{\lambda - 1}\right)^2 = \frac{\lambda}{(\lambda - 2)(\lambda - 1)^2}\cdot$$

FICHE 17

Simulation d'une expérience aléatoire

I Définition de la simulation statistique

Dans son *Dictionnaire encyclopédique de la Statistique* (Dunod 1997), Yadolah Dodge donne la **définition de la simulation** que nous reproduisons ici : *La simulation est la méthode statistique permettant la reconstitution fictive de l'évolution d'un phénomène. C'est une expérimentation qui suppose la constitution d'un modèle théorique présentant une similitude de propriétés ou de relations avec le phénomène faisant l'objet de l'étude.*
En effet, il n'est pas toujours possible d'étudier le comportement de modèles, d'estimateurs (voir fiche 25) en raison de leur complexité. Dans ces cas, nous aurons alors recourt à des **simulations d'échantillons** pour combler l'absence d'éléments théoriques.

II Les outils informatiques

Actuellement, tous les logiciels de calcul possèdent un **générateur de nombres aléatoires** qui en général correspondent à des observations provenant d'une loi uniforme continue sur l'intervalle $[0,1]$ (fiche 16). Ce sont ces observations que nous qualifions de **nombres pseudo-aléatoires** puisqu'elles sont engendrées par un mécanisme complétement déterministe. Les deux raisons principales pour lesquelles nous nous contentons d'un rendu pseudo-aléatoire sont :

1. il est très difficile d'obtenir de « vrais » nombres aléatoires et que, dans certaines situations, il est possible d'utiliser des nombres pseudo-aléatoires, en lieu et place de vrais nombres aléatoires ;
2. ce sont des générateurs particulièrement adaptés à une implémentation informatique, donc plus facilement et plus efficacement utilisables.

III Le processus de simulation

Nous souhaitons générer des réalisations d'une loi discrète ou continue. Nous commencerons par résoudre le problème pour une loi continue puis pour une loi discrète.

- **Cas d'une loi continue**

La fonction inverse de la fonction de répartition d'une loi continue de densité $f_X > 0$, notée F_X, existe d'après le théorème de la bijection, puisque la fonction de répartition F_X est continue et strictement croissante. Supposons que nous disposons, au moins de façon numérique, de la fonction inverse. Les logiciels de calcul proposent, en général, les fonctions inverses des fonctions de répartition des lois continues classiques. Le processus de simulation est le suivant :
étant donné une variable aléatoire, notée U, qui suit la loi uniforme sur l'intervalle $[0,1]$. Considérons la variable aléatoire $X = F_X^{-1}(U)$ et déterminons sa fonction de répartition. Nous avons

$$\mathbb{P}(X \leqslant x) = \mathbb{P}(F_X^{-1}(U) \leqslant x) = \mathbb{P}(U \leqslant F_X(x)) = F_X(x),$$

puisque U suit la loi uniforme. Donc la fonction de répartition de la variable aléatoire X est égale à F_X.
Ainsi, à partir de **nombres pseudo-aléatoires** $u_1,\ldots,u_n$, simulés suivant une loi uniforme sur l'intervalle $[0,1]$, nous obtenons, en posant $x_i = F_X^{-1}(u_i)$, des **nombres pseudo-aléatoires** $x_1,\ldots,x_n$ simulant les réalisations d'une v.a. de fonction de répartition F_X.

- **Cas d'une loi discrète finie**

Pour une loi discrète, la méthode ci-dessus n'est plus utilisable car la fonction réciproque de F_X n'existe pas puisque que F_X est une fonction en escalier. Dans le cas où le nombre de valeurs possibles $b_1 < b_2 < \ldots < b_p$ est fini, nous pouvons l'adapter ainsi :
si $u_i \in [0, F_X(b_1)]$, alors générer $x_i = b_1$
$\ldots$
si $u_i \in [F_X(b_{j-1}), F_X(b_j)]$, alors générer $x_i = b_j$
$\ldots$
si $u_i \in [F_X(b_{p-1}), 1]$, alors générer $x_i = b_p$.

Remarque : Il existe des méthodes de génération adaptées et plus efficaces pour chaque loi que nous pouvons trouver dans des ouvrages spécialisés.

FICHE 18

Convergences. Théorèmes limites. Approximations

I Convergences

- **Convergence en loi**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ et F_n la suite des fonctions de répartition correspondantes. Soit X une variable aléatoire définie sur le même espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ et F_X sa fonction de répartition. Nous disons que la suite $(X_n)_{n\geqslant 1}$ **converge en loi** vers X si, en tout point x de continuité de F_X, nous avons $\lim\limits_{n\to+\infty} F_n(x) = F_X(x)$.

Remarques

1. Pour des variables discrètes à valeurs entières, la convergence en loi vers une variable discrète à valeurs entières s'exprime par

$$\forall k \in \mathbb{N},\ \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k).$$

2. La suite $(X_n)_{n\geqslant 1}$ converge en loi vers X si et seulement si pour tout couple de points a et b de continuité de F_X, nous avons

$$\lim_{n\to+\infty} \mathbb{P}(a < X_n \leqslant b) = \mathbb{P}(a < X \leqslant b).$$

- **Convergence en probabilité**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ et X une variable aléatoire définie sur le même espace.
Nous disons que la suite $(X_n)_{n\geqslant 1}$ **converge en probabilité** vers X si, pour tout $\varepsilon > 0$, $\lim\limits_{n\to+\infty} \mathbb{P}(|X_n - X| \geqslant \varepsilon) = 0$.

Remarques

1. La convergence en probabilité implique la convergence en loi.
2. La convergence en loi est la plus utilisée en pratique car elle permet d'approcher la fonction de répartition de X_n par celle de X lorsque n est « grand ».

II Théorèmes limites

- **Inégalité de Markov**

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ admettant un moment d'ordre 1. Alors nous avons, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(|X| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{E}\left(|X|\right)}{\varepsilon}.$$

- **Inégalité de Bienaymé-Tchébychev**

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ admettant un moment d'ordre 2. Alors nous avons, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}\mathrm{ar}(X)}{\varepsilon^2}.$$

Cas particulier : variables aléatoires de Bernoulli

Soit $X_1,\ldots,X_n$, n variables aléatoires indépendantes et identiquement distribuées qui suivent une loi de Bernoulli $\mathcal{B}(p)$.

Soit $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Alors nous avons, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(|\overline{X}_n - p| \geqslant \varepsilon\right) \leqslant \frac{p(1-p)}{n\varepsilon^2} \leqslant \frac{1}{4n\varepsilon^2}.$$

La dernière inégalité provient du fait que la fonction de $[0,1]$ dans $\mathbb{R}$ qui à p associe $p(1-p)$ est maximale pour $p = 1/2$ et vaut alors $1/4$.

- **Loi faible des grands nombres**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires de même loi sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, deux à deux indépendantes, ayant une espérance μ et une variance σ^2.

Posons $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
Alors la suite $\left(\overline{X}_n\right)_{n\geqslant 1}$ converge en probabilité vers la variable aléatoire constante et égale à μ.

- **Théorème de la limite centrée (ou théorème central-limite)**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées, ayant une espérance μ et une variance σ^2.
Posons $Y_n = \dfrac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$.
Alors la suite $(Y_n)_{n\geqslant 1}$ converge en loi vers une variable aléatoire de loi normale $\mathcal{N}(0,1)$.

III Approximations

- **Approximation de la loi hypergéométrique par la loi binomiale**

Soit E un ensemble de N éléments, dont une proportion p de type A. Nous effectuons n tirages **sans remise** dans E. Soit X_N le nombre d'éléments de type A obtenus. La variable aléatoire X_N suit alors la loi hypergéométrique $\mathcal{H}(N,n,p)$.
Alors la suite $(X_N)_{N\geqslant 1}$ converge en loi vers une variable aléatoire de loi binomiale $\mathcal{B}(n,p)$.

Remarque : Ainsi, lorsque N devient très grand, n et p restant fixes, effectuer des tirages sans remise revient à effectuer des tirages avec remise.

En pratique : La loi hypergéométrique $\mathcal{H}(N,n,p)$ peut être approchée par la loi binomiale $\mathcal{B}(n,p)$ lorsque $N \geqslant 10n$.

- **Approximation de la loi binomiale par la loi de Poisson**

Soit λ un nombre réel fixé dans $[0,1]$. Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires de loi binomiale $\mathcal{B}(n,p_n)$ telle que $\lim\limits_{n\to+\infty} np_n = \lambda$.
Alors la suite $(X_n)_{n\geqslant 1}$ converge en loi vers une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda)$.

En pratique : La loi binomiale $\mathcal{B}(n,p)$ peut être approchée par la loi de Poisson $\mathcal{P}(np)$ lorsque :

$$\begin{cases} p & \leqslant & 0{,}1 \\ n & \geqslant & 30 \\ np & < & 15 \end{cases}$$

ou lorsque d'autres conditions données par l'énoncé sont vérifiées.
Attention, dans certains cas, l'astuce sera d'approcher la loi binomiale $\mathcal{B}(n,1-p)$ et non la loi binomiale $\mathcal{B}(n,p)$.

- **Approximation de la loi binomiale par la loi normale**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes de loi de Bernoulli $\mathcal{B}(p)$, où $p \in]0,1[$. $S_n = \sum_{k=1}^{n} X_k$ suit la loi binomiale $\mathcal{B}(n,p)$. Alors la suite $\left(\frac{S_n - np}{\sqrt{np(1-p)}}\right)_{n\geqslant 1}$ converge en loi vers une variable aléatoire de loi normale $\mathcal{N}(0,1)$.

En pratique : La loi binomiale $\mathcal{B}(n,p)$ peut être approchée par la loi normale $\mathcal{N}\left(np,\sqrt{np(1-p)}\right)$ lorsque :

$$\begin{cases} n & \geqslant & 30 \\ np & \geqslant & 15 \\ np(1-p) & > & 5 \end{cases}$$

ou lorsque d'autres conditions données par l'énoncé sont vérifiées.

Correction de continuité

Si la variable aléatoire X suit la loi binomiale $\mathcal{B}(n,p)$, alors la variable aléatoire X prend des valeurs entières positives entre 0 et n. Remplacer la loi binomiale $\mathcal{B}(n,p)$ par la loi normale $\mathcal{N}\left(np,\sqrt{np(1-p)}\right)$ revient à considérer la variable aléatoire X comme une variable qui prend donc toutes les valeurs réelles.
L'intervalle $[k-0{,}5; k+0{,}5[$ est l'ensemble des nombres réels qui s'arrondissent à k, c'est-à-dire pour $k \in [\![1,n-1]\!]$, nous remplacerons $\mathbb{P}(X=k)$ par $\mathbb{P}(k-0{,}5 \leqslant X < k+0{,}5)$.

Remarque : Pour que la somme des valeurs approchées des $\mathbb{P}(X=k)$, k variant de 0 à n, soit égale à 1, nous remplacerons $\mathbb{P}(X=0)$ par $\mathbb{P}(X<0{,}5)$ et $\mathbb{P}(X=n)$ par $\mathbb{P}(n-0{,}5 \leqslant X)$.

Les camping-cars

En France, la proportion des camping-cars par rapport à l'ensemble des véhicules est égale à 0,07.

1. Soit X le nombre des camping-cars parmi 100 véhicules choisis au hasard sur un parking contenant 2000 véhicules. Calculez $\mathbb{P}(X \geqslant 5)$.

2. Soit Y le nombre de camping-cars parmi 1000 véhicules circulant sur le boulevard périphérique d'une grande ville à une heure d'affluence. Nous supposerons que $N \geqslant 20\,000$. Calculez $\mathbb{P}(65 \leqslant Y \leqslant 75)$.

3. Nous choisissons n véhicules au hasard. Déterminez pour quelle valeur de n nous pouvons affirmer que la proportion des camping-cars parmi ces n véhicules est comprise entre 0,06 et 0,08 avec un risque d'erreur inférieur à 0,05.

Solution

Soit N le nombre total de véhicules. Soit S_n le nombre de camping-cars parmi n véhicules choisis au hasard. La variable S_n suit la loi hypergéométrique $\mathcal{H}(N; n; 0{,}07)$.

1. Dans cette question $n = 100$. Pour calculer $\mathbb{P}(X \geqslant 5)$, nous allons remplacer la loi hypergéométrique par une loi dont la probabilité est plus facile à calculer. Comme $N = 2000$, l'inégalité $N \geqslant 10n$ est vérifiée. Dans ce cas nous pouvons remplacer la loi hypergéométrique par la loi binomiale de paramètres $n = 100$ et $p = 0{,}07$. Il nous reste donc à calculer $\mathbb{P}(X_1 \geqslant 5)$ où X_1 suit une loi binomiale $\mathcal{B}(100; 0{,}07)$. Ce calcul est très long à réaliser. Nous cherchons à remplacer la loi binomiale $\mathcal{B}(100; 0{,}07)$ par une autre loi et en particulier par la loi de Poisson. Pour cela, il faut nous assurer que les conditions sont bien vérifiées. En effet, nous avons $p = 0{,}07 \leqslant 0{,}1$, $n = 100 \geqslant 30$ et $np = 100 \times 0{,}07 = 7 < 15$. Nous pouvons donc approcher la loi binomiale $\mathcal{B}(100; 0{,}07)$ par la loi de Poisson $\mathcal{P}(7)$. Ainsi, nous obtenons $\mathbb{P}(X \geqslant 5) \approx 1 - \sum_{k=0}^{4} \mathbb{P}(X_2 = k) = 1 - \sum_{k=0}^{4} \exp(-7)\frac{7^k}{k!} \approx 0{,}827$, où X_2 suit la loi de Poisson $\mathcal{P}(7)$.

2. Dans cette question $n = 1000$. Pour calculer $\mathbb{P}(65 \leqslant Y \leqslant 75)$, nous allons remplacer la loi hypergéométrique par une loi dont la probabilité est plus facile à calculer. Comme $N \geqslant 20\,000$, l'inégalité $N \geqslant 10n$ est vérifiée. Dans ce cas nous pouvons remplacer la loi hypergéométrique par la loi binomiale de paramètres $n = 1000$ et $p = 0{,}07$. Il nous reste donc à calculer $\mathbb{P}(65 \leqslant Y_1 \leqslant 75)$ où Y_1 suit une loi binomiale $\mathcal{B}(1000; 0{,}07)$. Ce calcul est très long à réaliser. Nous cherchons à remplacer la loi binomiale $\mathcal{B}(1000; 0{,}07)$ par une autre loi et en particulier par la loi normale. Pour cela, il faut nous assurer que les conditions sont bien vérifiées. En effet, nous avons $n = 1000 \geqslant 30$ et $np = 1000 \times 0{,}07 = 70 \geqslant 15$ et $np(1-p) = 1000 \times 0{,}07 \times 0{,}93 = 65{,}1 > 5$.

Nous pouvons donc approcher la loi binomiale $\mathcal{B}(1000; 0{,}07)$ par la loi normale $\mathcal{N}(70; \sqrt{65{,}1})$. Ainsi, nous obtenons $\mathbb{P}(65 \leqslant Y \leqslant 75) \approx \mathbb{P}\left(\frac{65-70}{\sqrt{65{,}1}} \leqslant Y_2^* \leqslant \frac{75-70}{\sqrt{65{,}1}}\right)$, où Y_2^* est la variable centrée et réduite calculée à partir de la variable Y_2 qui suit la loi nomale $\mathcal{N}(70; \sqrt{65{,}1})$.

Donc, $\mathbb{P}(65 \leqslant Y \leqslant 75) \approx \Phi\left(\frac{75{,}5-70}{\sqrt{65{,}1}}\right) - \Phi\left(\frac{64{,}5-70}{\sqrt{65{,}1}}\right) = 2 \times \Phi(0{,}68) - 1 = 0{,}5034$.

3. Soit $F_n = \dfrac{S_n}{n}$ la proportion de camping-cars parmi n véhicules choisis au hasard. Nous supposons que $N \geqslant 10n$. En supposant que $n \geqslant 30$, $0{,}07 \times n \geqslant 15$ et $0{,}07 \times 0{,}93 \times n > 5$, c'est-à-dire $n \geqslant 215$, nous pouvons approcher la loi binomiale $\mathcal{B}(n; 0{,}07)$ par la loi normale $\mathcal{N}(0{,}07 \times n; \sqrt{0{,}0651 \times n})$. Par conséquent F_n suit approximativement la loi normale $\mathcal{N}\left(0{,}07; \sqrt{\frac{0{,}0651}{n}}\right)$ et $F_n^* = \frac{F_n - 0{,}07}{\sqrt{\frac{0{,}0651}{n}}}$ suit approximativement la loi normale $\mathcal{N}(0,1)$. Donc, nous obtenons

$$\mathbb{P}(0{,}06 \leqslant F_n \leqslant 0{,}08) \quad = \quad \mathbb{P}\left(\frac{0{,}06 - 0{,}07}{\sqrt{\frac{0{,}0651}{n}}} \leqslant F_n^* \leqslant \frac{0{,}08 - 0{,}07}{\sqrt{\frac{0{,}0651}{n}}}\right)$$

ou encore $\mathbb{P}(0{,}06 \leqslant F_n \leqslant 0{,}08) = 2\Phi\left(\frac{0{,}01}{\sqrt{\frac{0{,}0651}{n}}}\right) - 1$. Par conséquent, $\mathbb{P}(0{,}06 \leqslant F_n \leqslant 0{,}08) \geqslant 0{,}95 \Leftrightarrow \Phi\left(\frac{0{,}01}{\sqrt{\frac{0{,}0651}{n}}}\right) \geqslant 0{,}975 = \Phi(1{,}96)$ ou encore $n \geqslant 0{,}0651\left(\frac{1{,}96}{0{,}01}\right)^2$. Donc $\mathbb{P}(0{,}06 \leqslant F_n \leqslant 0{,}08) \geqslant 0{,}95 \Leftrightarrow n \geqslant 2501$, valeur qui est bien supérieure à 215.

Vecteurs aléatoires. Vecteurs gaussiens

FICHE 19

I Vecteurs aléatoires

Définition : Un **vecteur aléatoire X** est une application de $(\Omega,\ \mathcal{F},\ \mathbb{P})$ dans un espace vectoriel réel, en général $\mathbb{R}^p$ muni de sa tribu borélienne. En pratique, $\mathbb{R}^p$ est muni de sa base canonique et nous identifierons $\mathbf{X}$ au p-uple de variables aléatoires formé par ses composantes sur cette base $\mathbf{X} = (X_1, X_2, \ldots, X_p)$.

Définition : **La fonction de répartition** F est une application de $\mathbb{R}^p$ dans $\mathbb{R}^p$ définie par $F(x_1, \ldots, x_p) = \mathbb{P}\left[X_1 \leqslant x_1, \ldots, X_p \leqslant x_p\right]$.

Définition : La **densité** f si elle existe est définie, en tout point $(x_1, \ldots, x_p)$ de $\mathbb{R}^p$, par l'égalité $f(x_1, \ldots, x_p) = \dfrac{\partial^p F}{\partial x_1 \ldots \partial x_p}(x_1, \ldots, x_p)$.

Théorème de Cramer-Wold : La loi du vecteur aléatoire **X** est entièrement déterminée par celles de toutes les combinaisons linéaires de ses composantes.

Définition : Soit $\mathbf{X} = \left(X_1,\ X_2, \ldots, X_p\right)$ un vecteur aléatoire tel que chacune de ses composantes X_i, admette une espérance. L'**espérance de X** $= \left(X_1,\ X_2, \ldots, X_p\right)$, est le vecteur :

$$\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu} = \left(\mathbb{E}(X_1), \mathbb{E}(X_2), \cdots, \mathbb{E}(X_p)\right).$$

Définition : Soit $\mathbf{X} = \left(X_1,\ X_2, \ldots, X_p\right)$ un vecteur aléatoire tel que chacun des couples de ses composantes $(X_i,\ X_j)$, admette une covariance. Nous appelons **matrice de variance-covariance de X** la matrice

$$\mathbb{V}\mathrm{ar}(\mathbf{X}) = \boldsymbol{\Sigma} = \begin{bmatrix} \mathbb{V}\mathrm{ar}(X_1) & \mathbb{C}\mathrm{ov}(X_1, X_2) & \ldots & \mathbb{C}\mathrm{ov}(X_1, X_p) \\ \mathbb{C}\mathrm{ov}(X_2, X_1) & \mathbb{V}\mathrm{ar}(X_2) & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \mathbb{C}\mathrm{ov}(X_p, X_1) & \cdot & \ldots & \mathbb{V}\mathrm{ar}(X_p) \end{bmatrix}.$$

Remarques

1. Si $A \in \mathcal{M}_p(\mathbb{R})$ et $B \in \mathbb{R}^p$, $\mathbb{E}(AX + B) = A\mathbb{E}(X) + B$.
2. $\mathbf{\Sigma}$ est une matrice carrée symétrique réelle d'ordre p.
3. $\mathbf{\Sigma} = \mathbb{E}({}^t\mathbf{XX}) - {}^t\,\boldsymbol{\mu}\boldsymbol{\mu}$.
4. Si $A \in \mathcal{M}_p(\mathbb{R})$ et $B \in \mathbb{R}^p$, $\mathbb{V}\mathrm{ar}(A\mathbf{X} + B) = A\mathbb{V}\mathrm{ar}(\mathbf{X})^t A$.
5. Si les variables X_i sont réduites, $\mathbf{\Sigma}$ s'identifie avec la matrice des corrélations linéaires

$$\begin{bmatrix} 1 & R(X_1,X_2) & \dots & R(X_1,X_p) \\ \cdot & 1 & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & 1 \end{bmatrix}.$$

II Vecteurs aléatoires gaussiens

Définition : $\mathbf{X}$ est **un vecteur gaussien à p dimensions** si toute combinaison linéaire de ses composantes suit une loi normale à une dimension.
Attention, la normalité de chaque composante ne suffit nullement à définir un vecteur gaussien.

Théorème : Les composantes d'un vecteur gaussien $\mathbf{X}$ sont indépendantes si et seulement si $\mathbf{\Sigma}$, est diagonale, c'est-à-dire si les composantes sont non corrélées.

Notation : Nous notons $\mathcal{N}_p\,(\boldsymbol{\mu},\ \mathbf{\Sigma})$ la loi normale à $p \geqslant 2$ dimensions d'espérance $\boldsymbol{\mu}$ et de matrice de variance-covariance $\mathbf{\Sigma}$.

Théorème : Si la matrice $\mathbf{\Sigma}$ est inversible, alors le vecteur aléatoire $\mathbf{X}$ admet pour densité

$$f(x_1,x_2,\dots,x_p) = \frac{1}{(2\pi)^{p/2}(\det \mathbf{\Sigma})^{p/2}} \exp\left(-\frac{1}{2}{}^t(\mathbf{x} - \boldsymbol{\mu})\mathbf{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right).$$

Manipulation de vecteurs gaussiens

Soit $(X,\ Y,\ Z)$ un vecteur aléatoire de loi $\mathcal{N}_3(\boldsymbol{\mu},\ \mathbf{\Sigma})$ avec $\boldsymbol{\mu} = (0,\ 2,\ 1)$ et $\mathbf{\Sigma} = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 4 & 2 \\ 3 & 2 & 3 \end{pmatrix}$.

1. Quelles sont les lois de X, de Y, de Z, de $-X + 2Z$?
2. Le vecteur $(X,\ Y + Z)$ est-il un vecteur gaussien ? Déterminez sa loi.

Solution

1. Le vecteur aléatoire (X, Y, Z) est un vecteur gaussien donc, par définition, chacune des combinaisons linéaires $\alpha X + \beta Y + \gamma Z$, pour $(\alpha, \beta, \gamma) \in \mathbb{R}^3$, suit une loi normale. En choisissant $(\alpha, \beta, \gamma) = (1, 0, 0)$, nous obtenons que la variable aléatoire X suit une loi normale. L'espérance de X est donnée par la première composante de $\boldsymbol{\mu}_1 = 0$ et sa variance par le coefficient $\boldsymbol{\Sigma}_{1,1} = 1$. Ainsi X suit une loi normale $\mathcal{N}(0, 4)$. De la même manière Y et Z suivent une loi normale $\mathcal{N}(2, 4)$ et $\mathcal{N}(1, 3)$.
Pour $(\alpha, \beta, \gamma) = (-1, 0, 2)$, nous déduisons que $-X + 2Z$ suit une loi normale. Calculons son espérance : $\mathbb{E}(-X + 2Z) = -\mathbb{E}(X) + 2\mathbb{E}(Z) = 2$.
Pour la variance, $\mathbb{V}\mathrm{ar}(-X + 2Z) = \mathbb{V}\mathrm{ar}(-X) + 2\mathbb{C}\mathrm{ov}\,(-X, 2Z) + \mathbb{V}\mathrm{ar}(2Z) = \mathbb{V}\mathrm{ar}(X)$ $-4\mathbb{C}\mathrm{ov}\,(X, Z) + 4\mathbb{V}\mathrm{ar}(Z) = 4 - 4 \times 3 + 4 \times 3 = 4$.

2. Le vecteur $\mathbf{V} = (X, Y + Z)$ est gaussien si chacune des combinaisons linéaires $\alpha X + \beta(Y + Z)$, pour $(\alpha, \beta) \in \mathbb{R}^2$, suit une loi normale, ce qui est bien le cas puisque (X, Y, Z) est un vecteur gaussien et que donc, par définition, chacune des combinaisons linéaires $\alpha X + \beta Y + \gamma Z$, pour $(\alpha, \beta, \gamma) \in \mathbb{R}^3$, suit une loi normale.
Calculons l'espérance de $\mathbf{V}$: $\mathbb{E}(\mathbf{V}) = (\mathbb{E}(X), \mathbb{E}(Y + Z)) = (0, 3)$. Puis déterminons sa matrice de variance-covariance $\mathbb{V}\mathrm{ar}(\mathbf{V})$:

$$\mathbb{V}\mathrm{ar}(\mathbf{V}) = \begin{pmatrix} \mathbb{V}\mathrm{ar}(X) & \mathbb{C}\mathrm{ov}(X, Y + Z) \\ \mathbb{C}\mathrm{ov}(Y + Z, X) & \mathbb{V}\mathrm{ar}(Y + Z) \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 5 & 11 \end{pmatrix}.$$

Vecteur non gaussien à composantes gaussiennes

Soit X une variable aléatoire de loi $\mathcal{N}(0, 1)$ et T indépendante de X telle que $\mathbb{P}[T = 1] = \mathbb{P}[T = -1] = 1/2$.

1. Montrez que $Y = TX$ suit une loi $\mathcal{N}(0, 1)$.

2. Montrez que le vecteur aléatoire (X, Y) n'est pas un vecteur gaussien.

Solution

1. Soit $y \in \mathbb{R}$. Calculons $\mathbb{P}[Y \leqslant y]$.

$$\begin{aligned} \mathbb{P}[Y \leqslant y] &= \mathbb{P}[Y \leqslant y \text{ et } T = 1] + \mathbb{P}[Y \leqslant y \text{ et } T = -1] \\ &= \mathbb{P}[X \leqslant y \text{ et } T = 1] + \mathbb{P}[-X \leqslant y \text{ et } T = -1] \\ &= \mathbb{P}[X \leqslant y]\mathbb{P}[T = 1] + \mathbb{P}[-X \leqslant y]\mathbb{P}[T = -1]. \end{aligned}$$

La dernière égalité provient du fait que les deux variables aléatoires X et T sont indépendantes.

Comme X suit une loi normale centrée et réduite, $-X$ suit également une loi normale centrée et réduite par symétrie de cette loi par rapport à 0. Ainsi les probabilités $\mathbb{P}[X \leqslant y]$ et $\mathbb{P}[-X \leqslant y]$ sont égales.

$$\begin{aligned}\mathbb{P}[Y \leqslant y] &= \mathbb{P}[X \leqslant y]\mathbb{P}[T = 1] + \mathbb{P}[-X \leqslant y]\mathbb{P}[T = -1]\\ &= \mathbb{P}[X \leqslant y](\mathbb{P}[T = 1] + \mathbb{P}[T = -1])\\ &= \mathbb{P}[X \leqslant y].\end{aligned}$$

Par conséquent, la loi de Y est une loi normale centrée et réduite.

2. Si le vecteur aléatoire (X, Y) est un vecteur gaussien alors, par définition, chacune des combinaisons linéaires $\alpha X + \beta Y$, pour $(\alpha, \beta) \in \mathbb{R}^2$, doit suivre une loi normale. En particulier $X + Y$ doit suivre une loi normale et alors nécessairement $\mathbb{P}[X + Y = 0] = 0$. Or, si nous calculons cette probabilité :

$$\begin{aligned}\mathbb{P}[X + Y = 0] &= \mathbb{P}[X + XT = 0]\\ &= \mathbb{P}[X(1 + T) = 0]\\ &= \mathbb{P}[X = 0 \text{ ou } 1 + T = 0]\\ &= \mathbb{P}[X = 0] + \mathbb{P}[T = -1] - \mathbb{P}[X = 0 \text{ et } T = -1]\\ &= \mathbb{P}[X = 0] + \mathbb{P}[T = -1] - \mathbb{P}[X = 0]\mathbb{P}[T = -1]\\ &= 1/2\\ &\neq 0.\end{aligned}$$

Vocabulaire de la statistique

- **Statistique**

  La **statistique** désigne à la fois un ensemble de données et l'ensemble des activités consistant à collecter ces données, à les traiter et à les interpréter.

- **Statistiques**

  Les **statistiques**, l'ensemble des données numériques, interviennent pratiquement dans tous les domaines d'activité : gestion financière (états, banques, assurances, entreprises...), démographie, contrôles de qualité, études de marché, sciences expérimentales (biologie, psychologie...)...

- **Statistique descriptive**

  Le traitement des données, pour en dégager un certain nombre de renseignements qualitatifs ou quantitatifs à des fins de comparaison, s'appelle la **statistique descriptive** (fiches 21, 22 et 23).

- **Statistique inférentielle**

  Un autre but de la statistique consiste à extrapoler à partir d'un échantillon de la population à étudier, le comportement de la population dans son ensemble (sondages, contrôle de qualité comportant un test destructif...) C'est la **statistique inductive** ou encore appelée **statistique inférentielle** (fiches 24, 25, 26, 27, 28, 29 et 30).

- **Population**

  L'ensemble sur lequel porte l'activité statistique s'appelle **la population**. Elle est généralement notée Ω pour rappeler la notation des probabilités mais par exemple dans la théorie des sondages elle est notée U, U comme Univers.

- **Individus**

  Les éléments qui constituent la population sont appelés les **individus** ou encore les **unités statistiques**.

Remarque : Ces « individus » peuvent être de natures très diverses : ensemble de personnes, mois d'une année, pièces produites par une usine, résultats d'expériences répétées un certain nombre de fois.

- **Caractères**

Les caractéristiques étudiées sur les individus d'une population sont appelées les **caractères**. Un **caractère** est une application χ d'un ensemble Ω fini de cardinal N (la population) dans un ensemble C (l'**ensemble des valeurs possibles du caractère**), qui associe à chaque individu ω de Ω la valeur $\chi(\omega)$ que prend ce caractère sur l'individu ω. Nous considérons plusieurs types de caractères :

1. Les **caractères qualitatifs**.
Exemples : profession, adresse, sexe, numéro de téléphone...
2. Les **caractères quantitatifs** : leur détermination produit un nombre ou une suite de nombres. Nous distinguons
   (i) les **caractères simples** ou **univariés** : leur mesure sur un individu produit un seul nombre. L'ensemble de leurs valeurs est donc $\mathbb{R}$ ou une partie de $\mathbb{R}$.
   Exemples : taille, poids, salaire, température...
   (ii) les **caractères multiples** : leur mesure sur un individu produit une suite finie de nombres. L'ensemble de leurs valeurs est donc $\mathbb{R}^n$ ou une partie de $\mathbb{R}^n$.
   Exemples : relevé de notes d'un(e) étudiant(e), fiche de salaire...

Remarques

1. Les caractères qualitatifs peuvent toujours être transformés en quantitatifs par codage. C'est ce qui se fait le plus généralement. Mais un tel codage est purement conventionnel et n'a pas vraiment un sens quantitatif.
   Exemple : Nous ne pouvons pas calculer le sexe moyen.
2. Certains caractères qualitatifs s'expriment à l'aide de nombres.
   Exemple : Un numéro de téléphone.
   Mais ils n'ont pas non plus de sens quantitatif.
   Exemple : Calculer un numéro de téléphone moyen n'est pas pertinent.

- **Données brutes**

La suite des valeurs $\chi(\omega)$ prises par χ s'appelle les **données brutes**. C'est une suite finie $(X_1,\ldots,X_N)$ de l'ensemble C.

FICHE 21

Statistique descriptive univariée. Représentations graphiques

I Description d'une série statistique

Si X est un caractère quantitatif simple, la suite finie $X(\Omega) = (X_1, \ldots, X_N)$ des valeurs atteintes par le caractère (ou données brutes) est un ensemble fini $\{x_1, \ldots, x_p\}$. Nous supposons que ces valeurs sont ordonnées :

$$x_1 < \ldots < x_p.$$

Le fait que telle valeur soit relative à tel individu est un renseignement qui n'intéresse pas le statisticien. Seul l'ensemble des valeurs atteintes et le nombre de fois que chacune d'elle est atteinte sont utiles.

- **Distribution statistique discrète**

  – ***Définitions***

  Soit une série statistique $(x_i, n_i)_{i=1,\ldots,p}$. Nous appelons

  1. **effectif de la valeur** x_i : le nombre n_i de fois que la valeur x_i est prise, c'est-à-dire le cardinal de l'ensemble $X^{-1}(x_i)$;
  2. **effectif cumulé en** x_i : la somme $\displaystyle\sum_{j=1}^{i} n_j$;
  3. **fréquence de la valeur** x_i : le rapport $f_i = \dfrac{n_i}{N}$ de l'effectif de x_i à l'effectif total N de la population, c'est-à-dire le cardinal de Ω ou encore la somme des n_i ;
  4. **fréquence cumulée en** x_i : la somme $\displaystyle\sum_{j=1}^{i} f_j$.

– **Distribution statistique discrète**

La série statistique $(x_i, n_i)_{i=1,\ldots,p}$ ou $(x_i, f_i)_{i=1,\ldots,p}$ est appelée **distribution statistique discrète**.

- **Distribution statistique groupée**

Lorsque le caractère quantitatif discret ou continu comprend un grand nombre de valeurs, nous préférons regrouper les valeurs en intervalles appelées **classes** pour rendre la statistique plus lisible. Nous partageons alors l'ensemble C des valeurs du caractère en classes $]a_i, a_{i+1}]$ avec $a_i < a_{i+1}$.

– ***Définitions***

Soit une série statistique $(]a_i, a_{i+1}], n_i)_{i=1,\ldots,p}$. Nous appelons

1. **effectif de** $]a_i, a_{i+1}]$: le nombre n_i de valeurs prises dans $]a_i, a_{i+1}]$, c'est-à-dire $X^{-1}(]a_i, a_{i+1}])$;
2. **effectif cumulé en** a_i : le nombre de valeurs prises dans l'intervalle $]-\infty, a_i]$;
3. **fréquence de** $]a_i, a_{i+1}]$: le rapport $f_i = \dfrac{n_i}{N}$;
4. **fréquence cumulée en** a_i : la somme $\displaystyle\sum_{j=1}^{i} f_j$.

– **Distribution statistique groupée**

La série statistique $(]a_i, a_{i+1}], n_i)_{i=1,\ldots,p}$ ou $(]a_i, a_{i+1}], f_i)_{i=1,\ldots,p}$ est appelée **distribution statistique groupée** ou **continue**.

II Représentations graphiques

- **Représentations graphiques pour une distribution statistique discrète**

– **Diagramme en bâtons**

Le **diagramme en bâtons des effectifs (resp. des fréquences)** d'une distribution statistique discrète est constitué d'une suite de segments verticaux d'abscisses x_i dont la longueur est proportionnelle à l'effectif (resp. la fréquence) de x_i.

– **Polygone des effectifs ou des fréquences**

Le **polygone des effectifs (resp. des fréquences)** d'une distribution statistique discrète est obtenu à partir du diagramme en bâtons des effectifs (resp. des fréquences) en joignant par un segment les sommets des bâtons.

– **Polygone des effectifs (resp. fréquences) cumulés**

En remplaçant dans la définition précédente le mot effectifs (resp. fréquences) par effectifs cumulés (resp. fréquences cumulées), nous obtenons le **polygone des effectifs cumulés (resp. des fréquences cumulées)**.

- **Représentations graphiques pour une distribution statistique groupée**

– **Histogramme**

L'**histogramme** est la représentation graphique d'une distribution statistique groupée. Deux cas se distinguent :

1. Dans le cas où les amplitudes des classes sont égales, cet histogramme est constitué d'un ensemble de rectangles dont la largeur est égale à a, l'amplitude de la classe, et la hauteur égale à $K \times n_j$ où n_j est l'effectif de la classe et K est un coefficient arbitraire (choix d'une échelle), de sorte que l'aire totale sous l'histogramme est égale à $K \times N \times a$ où N est l'effectif total.
2. Dans le cas de classes d'amplitudes $k_j \times a$ inégales, multiples entiers de l'une d'entre elles a, nous convenons, pour conserver le résultat précédent, de prendre pour hauteur du rectangle de la classe numéro j le quotient $(K \times n_j)/k_j$.

Remarque : Lorsqu'il est utilisé pour estimer une densité, un histogramme délimite une aire totale égale à 1.

– **Polygone des effectifs ou des fréquences**

Le **polygone des effectifs** ou **des fréquences** d'une distribution statistique groupée est obtenu en joignant dans l'histogramme de cette distribution les milieux des côtés horizontaux supérieurs.

– **Polygone des fréquences cumulées**

Le **polygone des fréquences cumulées d'une distribution statistique groupée** est la représentation graphique de la fonction définie sur chaque intervalle $]a_i, a_{i+1}]$, $1 \leqslant i \leqslant p$ par $f(x) = \sum_{j=1}^{i-1} f_j + \dfrac{x - a_i}{a_{i+1} - a_i} f_i$.

Application directe du cours

Représentez la distribution statistique discrète ((1,2), (2,3), (3,4), (4,1), (5,6), (6,5), (7,2), (8,3), (9,1), (10,1)) :

1. à l'aide du diagramme en bâtons des effectifs ;

2. en superposant le diagramme en bâtons et le polygone des fréquences ;
3. en superposant le diagramme en bâtons et le polygone des effectifs cumulés.

Solution

1. La figure 21.1 correspond au diagramme en bâtons des effectifs.

2. La figure 21.2 a été obtenue en superposant le diagramme en bâtons et le polygone des fréquences.

3. La figure 21.3 a été obtenue en superposant le diagramme en bâtons et le polygone des effectifs cumulés.

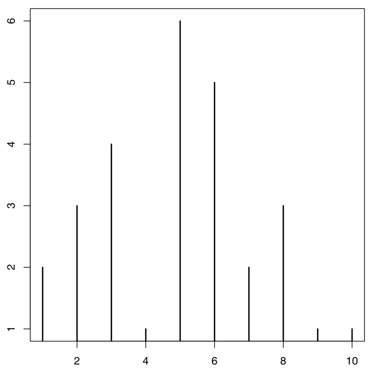


Figure 21.1

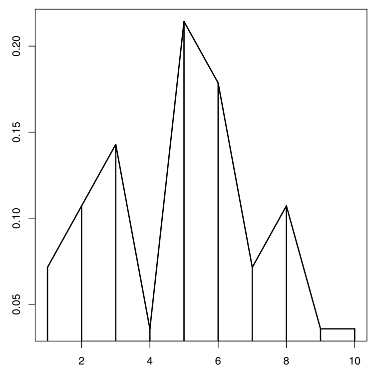


Figure 21.2

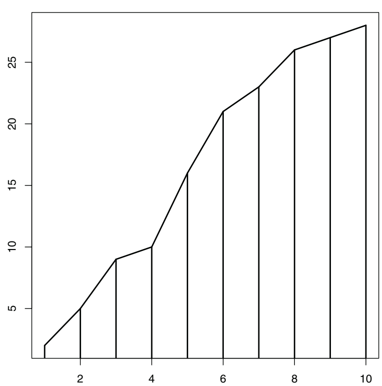


Figure 21.3

Application du cours

Représentez la distribution statistique groupée suivante ((]1,3],4), (]3,4],8), (]4,5.5],10), (]5.5,6],14), (]6,8],20), (]8,10],12), (]10,11],9), (]11,12.5],3)) :

1. à l'aide d'un histogramme des fréquences ;
2. en superposant l'histogramme des fréquences et son polygone des fréquences ;
3. à l'aide du polygone des fréquences cumulées.

Solution

1. L'histogramme des fréquences demandé est la figure 21.4.
2. La figure 21.5 superpose l'histogramme des fréquences et le polygone des fréquences.
3. Nous obtenons le polygone des fréquences cumulées à la figure 21.6.

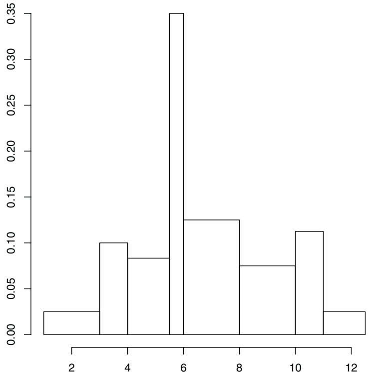


Figure 21.4

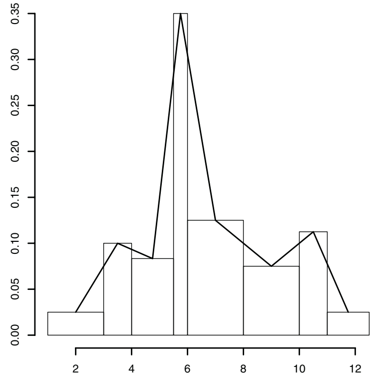


Figure 21.5

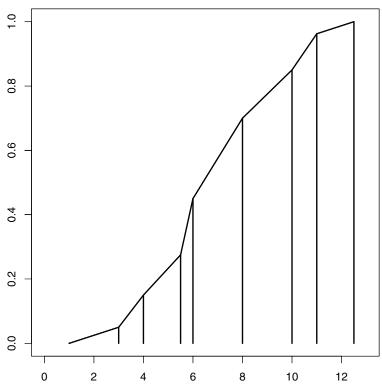


Figure 21.6

FICHE 22

Diverses caractéristiques

I Caractéristiques de position

- **Mode et classe modale**

Un **mode**, $\mathrm{Mo}(x)$, d'une distribution statistique discrète est l'une des valeurs $x_1, x_2, \ldots, x_p$ dont la fréquence est maximale. Pour une distribution statistique groupée, une **classe modale**, $\mathrm{Mo}(x)$, est une classe de densité, c'est-à-dire de rapport fréquence/longueur, maximale. La distribution est **unimodale** si elle a un seul mode. Si elle en a plusieurs, elle est **plurimodale** (**bimodale**, **trimodale**, ...).

- **Médiane**

Soit m et d les parties entière et décimale de $(N+1)/2$.
La **médiane**, $Q_2(x)$ ou $Q_{0,5}(x)$, d'une série statistique est définie par $Q_2(x) = x_{(m)} + d(x_{(m+1)} - x_{(m)})$, où $x_{(m)}$ signifie la m-ième valeur lorsque la série statistique est classée par ordre croissant. $x_{(m)}$ est aussi appelée la m-ième **statistique d'ordre**.

> **Remarque :** Il existe beaucoup de définitions pour la médiane. Celle que nous avons décidé d'adopter ne permet pas de discussion au niveau du calcul.

- **Quartiles**

Soit m et d les parties entière et décimale de $(N+1)/4$.
Le **premier quartile**, $Q_1(x)$ ou $Q_{0,25}(x)$, d'une série statistique est défini par $Q_1(x) = x_{(m)} + d(x_{(m+1)} - x_{(m)})$.
Soit m' et d' les parties entière et décimale de $3(N+1)/4$.
Le **troisième quartile**, $Q_3(x)$ ou $Q_{0,75}(x)$, d'une série statistique est défini par $Q_3(x) = x_{(m')} + d'(x_{(m'+1)} - x_{(m')})$.

- **Moyenne arithmétique**

La **moyenne**, $\mu(x)$ ou parfois notée simplement μ, d'une distribution statistique discrète est le nombre réel défini par $\mu(x) = \sum_{i=1}^{p} x_i f_i = \frac{1}{N} \sum_{i=1}^{p} x_i n_i$. Pour une statistique groupée, la moyenne se calcule par $\mu(x) = \sum_{i=1}^{p} \frac{a_i + a_{i+1}}{2} f_i$.

II Caractéristiques de dispersion

- **Étendue**

L'**étendue**, notée $\mathrm{e}(x)$, est la différence entre la plus grande et la plus petite des valeurs prises, donc $\mathrm{e}(x) = \max(x) - \min(x)$.

- **Étendue interquartile**

L'**étendue interquartile**, notée $\mathrm{EIQ}(x)$, est la différence entre les quartiles $\mathrm{Q}_{0,75}(x)$ et $\mathrm{Q}_{0,25}(x)$, donc $\mathrm{EIQ}(x) = \mathrm{Q}_{0,75}(x) - \mathrm{Q}_{0,25}(x)$.

- **Variance et écart-type**

La **variance** d'une distribution statistique discrète est définie par

$$\mathbb{V}\mathrm{ar}\,(x) = \sigma^2(x) = \sum_{i=1}^{p} (x_i - \mu(x))^2 f_i.$$

L'**écart-type** est égal à $\sqrt{\mathbb{V}\mathrm{ar}\,(x)}$.
Formule de Huygens : $\mathbb{V}\mathrm{ar}(x) = \mu(x^2) - \mu^2(x)$, où $\mu(x^2)$ est la moyenne du carré des valeurs de la distribution.

- **Moment centré d'ordre r**

$$_{\mu}m_r(x) = \sum_{i=1}^{p} (x_i - \mu(x))^r f_i.$$

- **Coefficient de variation (cas d'une variable positive)**

$$\mathrm{CV}(x) = \frac{\sigma(x)}{\mu(x)}.$$

III Caractéristiques de forme

- **Coefficient d'asymétrie de Fisher**

$$\gamma_1(x) = \frac{{}_\mu m_3(x)}{\sigma^3(x)} = \frac{{}_\mu m_3(x)}{({}_\mu m_2(x))^{3/2}}.$$

- **Coefficient d'asymétrie de Pearson**

$$\beta_1(x) = \frac{({}_\mu m_3(x))^2}{(\sigma^2(x))^3} = \frac{({}_\mu m_3(x))^2}{({}_\mu m_2(x))^3} = \gamma_1^2(x).$$

- **Coefficient d'aplatissement de Fisher**

$$\gamma_2(x) = \frac{{}_\mu m_4(x)}{({}_\mu m_2(x))^2} - 3.$$

- **Coefficient d'aplatissement de Pearson**

$$\beta_2(x) = \frac{{}_\mu m_4(x)}{({}_\mu m_2(x))^2} = \frac{{}_\mu m_4(x)}{\sigma^4(x)}.$$

Temps de travail en Europe en 2006

Établissez, pour le tableau ci-dessous (Nombre d'heures travaillées par semaine des personnes ayant un emploi à plein temps dans les États membres)

Pays	Durée (heures)	Pays	Durée (heures)
Allemagne	41,70	Lettonie	43,00
Autriche	44,10	Lituanie	39,80
Belgique	41,00	Luxembourg	40,90
Chypre	41,80	Malte	41,20
Danemark	40,50	Pays-Bas	40,80
Espagne	42,20	Pologne	42,90
Estonie	41,50	Portugal	41,60
Finlande	40,50	République Tchèque	42,70
France	**41,00**	Royaume-Uni	43,10
Grèce	44,10	Slovaquie	41,60
Hongrie	41,00	Slovénie	42,50
Irlande	40,70	Suède	41,10
Italie	41,30		

Source : Eurostat.

1. la moyenne, la médiane, le minimum, le maximum, le premier et troisième quartile,

2. l'étendue, l'étendue interquartile, l'écart-type,

3. les coefficients d'asymétrie et les coefficients d'aplatissement,

4. et la **boîte de dispersion**, appelée également la boîte à moustaches, qui se représente en suivant les instructions suivantes :

(i) Nous traçons un rectangle de largeur fixée à priori et de longueur $\text{EIQ} = \text{Q}_{0,75} - \text{Q}_{0,25}$.

(ii) Nous y situons la médiane par un segment positionné à la valeur $\text{Q}_{0,5}$, par rapport à $\text{Q}_{0,75}$ et $\text{Q}_{0,25}$. Parfois nous y ajoutons aussi la moyenne.

(iii) Nous calculons $\left(\text{Q}_{0,75} + 1{,}5 \times \text{EIQ}\right)$ et nous cherchons la dernière observation x_h en deçà de la limite $\left(\text{Q}_{0,75} + 1{,}5 \times \text{EIQ}\right)$, soit $x_h = \max\left\{x_i / x_i \leqslant \text{Q}_{0,75} + 1{,}5 \times \text{EIQ}\right\}$.

(iv) Nous calculons $\left(\text{Q}_{0,25} - 1{,}5 \times \text{EIQ}\right)$ et nous cherchons la première observation x_b au delà de la limite $\left(\text{Q}_{0,25} - 1{,}5 \times \text{EIQ}\right)$, soit $x_b = \min\left\{x_i / x_i \geqslant \text{Q}_{0,25} - 1{,}5 \times \text{EIQ}\right\}$.

(v) Nous traçons deux lignes allant des milieux des largeurs du rectangle aux valeurs x_b et x_h.

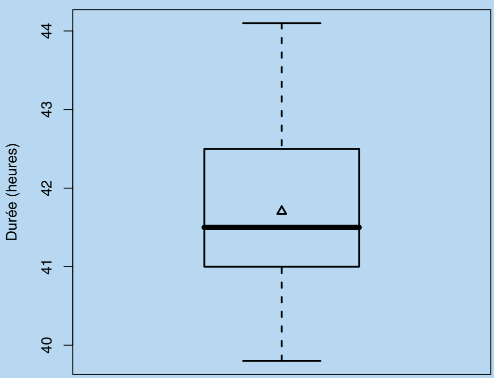


Figure 22.1. Boîte de dispersion associée à l'exemple du tableau « Le temps de travail dans les États membres », 2006, Eurostat.

Solution

Donnons les valeurs des différentes caractéristiques

1. la moyenne $\mu(x) = 41{,}70h$, la médiane $\text{Q}_{0,5}(x) = 41{,}50h$, le minimum $\min(x) = 39{,}80h$, le maximum $\max(x) = 44{,}10h$, le premier quartile $\text{Q}_{0,25}(x) = 40{,}95h$ et le troisième quartile $\text{Q}_{0,75}(x) = 42{,}60h$,

2. l'étendue $\mathrm{e}(x) = 4{,}30h$, l'étendue interquartile $\mathrm{EIQ}(x) = 1{,}75h$, l'écart-type $\sqrt{\mathbb{V}\mathrm{ar}(x)} = 1{,}09h$,

3. les coefficients d'asymétrie de Pearson $\beta_1^{(x)} = 0{,}44$ et de Fisher $\gamma_1^{(x)} = 0{,}67$ et les coefficients d'aplatissement de Pearson $\beta_2^{(x)} = 2{,}72$ et de Fisher $\gamma_2^{(x)} = -0{,}28$,

4. et la **boîte de dispersion** représentée à la figure 22.1.

FICHE 23

Statistique descriptive bivariée

Soit (X,Y) une distribution statistique d'un couple de caractères sur une population d'effectif N. Notons $(x_1, x_2, \ldots, x_p)$, et $(y_1, y_2, \ldots, y_q)$ les valeurs distinctes observées pour X et Y ordonnées dans l'ordre croissant.

I Distribution conjointe

- **Effectif du couple**

  L'**effectif du couple** (x_i, y_j) est égal au nombre $n_{i,j}$ de couples de valeurs égaux à (x_i, y_j), c'est-à-dire le cardinal de $(X,Y)^{-1}(x_i, y_j)$.

- **Fréquence du couple**

  La **fréquence du couple** (x_i, y_j) est égal au nombre $f_{i,j} = n_{i,j}/N$.

- **Distributions des effectifs et des fréquences**

  $\big((x_i, y_j), n_{i,j}\big)_{\substack{i=1,\ldots,p \\ j=1,\ldots,q}}$ et $\big((x_i, y_j), f_{i,j}\big)_{\substack{i=1,\ldots,p \\ j=1,\ldots,q}}$ sont la **distribution des effectifs** et la **distribution des fréquences** du couple (X,Y).

II Distributions marginales

- **Effectif et fréquences marginaux**

  L'**effectif marginal** de x_i (respectivement de y_j) est égal au nombre $n_{i,\bullet} = \sum_{j=1}^{q} n_{i,j}$ (respectivement $n_{\bullet,j} = \sum_{i=1}^{p} n_{i,j}$) et la **fréquence marginale** de x_i (respectivement y_j) au nombre $f_{i,\bullet} = \sum_{j=1}^{q} f_{i,j}$ (respectivement $f_{\bullet,j} = \sum_{i=1}^{p} f_{i,j}$).

- **Distributions marginales des effectifs et des fréquences**

  $(x_i, n_{i,\bullet})_{i=1,\dots,p}$ et $(y_j, n_{\bullet,j})_{j=1,\dots,q}$ (respectivement $(x_i, f_{i,\bullet})_{i=1,\dots,p}$ et $(y_j, f_{\bullet,j})_{j=1,\dots,q}$) sont appelées les **distributions marginales des effectifs** (respectivement des **fréquences**) de X et de Y.

III Indépendance et distributions conditionnelles

- **Indépendance**

  Y est **indépendant** de X si la distribution des fréquences de Y dans la sous-population Ω_i des individus pour lesquels la valeur de X est x_i ne dépend pas de i.

- **Fréquence conditionnelle**

  La fraction $\dfrac{n_{i,j}}{n_{i,\bullet}} = \dfrac{f_{i,j}}{f_{i,\bullet}}$ représente la **fréquence conditionnelle** de la valeur y_j sachant x_i.

- **Distribution conditionnelle des fréquences**

  La distribution $\left(y_j, \dfrac{f_{i,j}}{f_{i,\bullet}}\right)_{j=1,\dots,q}$ est appelée la **distribution conditionnelle des fréquences** de Y sachant que $X = x_i$.

- **Remarques**

  1. Nous définissons d'une manière analogue la **distribution conditionnelle des fréquences** de X sachant que $Y = y_j$.
  2. Pour que Y soit indépendant de X il faut que la distribution conditionnelle des fréquences de Y sachant que $X = x_i$ ne dépende pas de i.

IV Covariance et coefficient de corrélation linéaire

- **Covariance du couple** (X,Y)

$$\mathbb{C}\mathrm{ov}(X,Y) = \sum_{i,j} (x_i - \mu(X))(y_j - \mu(Y)) f_{i,j}.$$

Remarques

1. Si X et Y sont indépendants, alors la covariance du couple (X,Y) est nulle.
2. $\mathbb{C}\text{ov}(X,Y) = \mu(XY) - \mu(X)\mu(Y)$ où $\mu(XY) = \sum_{i,j} x_i y_j f_{i,j}$.

- **Coefficient de corrélation linéaire**

$$r(X,Y) = \frac{\mathbb{C}\text{ov}(X,Y)}{\sigma(X)\sigma(Y)}.$$

Remarque : Nous avons $|r(X,Y)| \leqslant 1$.

Les oraux du CAPES de mathématiques

À l'oral du CAPES de mathématiques, chaque candidat est interrogé à l'épreuve d'exposé où il obtient la note X et à l'épreuve sur dossier où il obtient la note Y (notes sur 20). Les résultats obtenus par 104 candidats, données fictives, sont donnés dans le tableau ci-dessous :

X \ Y	[0,4[	[4,8[	[8,12[	[12,16[	[16,20[
[0,4[	1	1			
[4,8[	1	3	5	11	
[8,12[	2	10	10	28	
[12,16[		1	3	9	11
[16,20[			2	4	2

1. Calculez la moyenne de X, de Y, l'écart-type de X et de Y.
2. Donnez la distribution marginale des effectifs de X et de Y.
3. Calculez la fréquence conditionnelle de $X \in [12,16[$ sachant $Y \in [4,8[$.
4. Calculez la covariance de X et de Y puis déduisez-en le coefficient de corrélation.

Solution

Nous travaillons avec une distribution statistique groupée, la valeur associée à chaque classe sera la valeur égale au point milieu de la classe. Ces valeurs sont notées x_i et y_j. Nous allons introduire un tableau qui nous permettra d'obtenir les résultats demandés.

X \ Y	[0,4[	[4,8[	[8,12[	[12,16[	[16,20[	x_i	$n_{i,\bullet}$	$x_i n_{i,\bullet}$	$x_i^2 n_{i,\bullet}$
[0,4[	1	1				2	2	4	8
[4,8[	1	3	5	11		6	20	120	720
[8,12[	2	10	10	28		10	50	500	5000
[12,16[		1	3	9	11	14	24	336	4704
[16,20[			2	4	2	18	8	144	2592
y_j	2	6	10	14	18				
$n_{\bullet,j}$	4	15	20	52	13				
$y_j n_{\bullet,j}$	8	90	200	728	23				
$y_j^2 n_{\bullet,j}$	16	540	2000	10192	4212				

1. Calculons la moyenne de X. Nous avons

$$\mu(X) = \frac{1}{104}(4 + 120 + 500 + 336 + 144) = 11{,}04.$$

En faisant le même calcul, nous obtenons $\mu(Y) = 12{,}60$. Calculons l'écart-type de X. Pour cela, calculons la variance de X. Nous avons, en utilisant la formule de Huygens,

$$\mathbb{V}\mathrm{ar}(X) = \frac{1}{104}(8 + 720 + 5000 + 4704 + 2592) - 11{,}04^2 = 8{,}36.$$

Ainsi, nous avons $\sigma(X) = 2{,}89$. En faisant le même calcul, nous obtenons $\mathbb{V}\mathrm{ar}(Y) = 10{,}84$. Ainsi, nous en déduisons $\sigma(Y) = 3{,}30$.

2. La distribution marginale des effectifs est

	[0,4[	[4,8[	[8,12[	[12,16[	[16,20[
X	2	20	50	24	8
Y	4	15	20	52	13

3. La fréquence conditionnelle de $X \in [12,16[$ sachant que $Y \in [4,8[$ est égale à

$$\frac{n_{4,2}}{n_{\bullet,2}} = \frac{1}{15}.$$

4. La covariance de X et de Y s'exprime, avec la formule de Huygens, de la manière suivante :

$$\mathbb{C}\mathrm{ov}(X,Y) = \mu(XY) - \mu(X)\mu(Y) = 139{,}76 - 11{,}04 \times 12{,}60 = 0{,}66.$$

La valeur du coefficient de corrélation linéaire entre X et Y est alors :

$$r(X,Y) = \frac{\mathbb{C}\mathrm{ov}(X,Y)}{\sigma(X)\sigma(Y)} = \frac{0{,}66}{2{,}89 \times 3{,}30} = 0{,}07.$$

FICHE 24

Échantillonnage. Modèles statistiques

I Introduction

Le problème de l'estimation est l'impossibilité de connaître exactement la valeur d'un **paramètre inconnu** noté θ. Ce problème est très général et a des aspects distincts. Les observations, par exemple obtenues à partir d'une méthode d'échantillonnage, permettent de construire **une estimation de** θ. Ainsi chaque observation est la valeur d'une variable aléatoire X dont la loi dépend de θ. Cela revient à doter l'**état de la nature** inconnu d'un **modèle probabiliste**. Ce dernier est complété par un **modèle d'échantillonnage** décrivant la manière dont les observations sont recueillies.

II Échantillonnage

- **Échantillon aléatoire**

Soit X une variable aléatoire réelle. Un **échantillon aléatoire** d'effectif $n \geqslant 1$ est un vecteur aléatoire $\mathbf{X_n} = (X_1, X_2, \ldots, X_n)$ à n composantes qui sont n variables aléatoires indépendantes suivant la même loi que X, appelée **variable aléatoire parente**.

> **Remarque :** Pour des raisons de commodité, nous avons supposé que les X_i sont mutuellement indépendantes. Dans certains cas, l'indépendance deux à deux sera suffisante.

- **Statistique de l'échantillon**

Toute variable aléatoire T, fonction de l'échantillon aléatoire $\mathbf{X_n} = (X_1, X_2, \ldots, X_n)$, est appelée **statistique de l'échantillon**.

Remarques

1. Une statistique peut être à valeurs dans $\mathbb{R}$ ou $\mathbb{R}^p$. Dans le dernier cas, nous parlerons de statistique vectorielle.
2. La difficulté de cette notion est la suivante : nous avons une double conception, qui est la base de la statistique mathématique. Les valeurs observées $(x_1, x_2, \ldots, x_n)$ (noter que ce sont des minuscules) constituent n réalisations indépendantes d'une variable aléatoire X ou encore, une réalisation unique du vecteur aléatoire $\mathbf{X_n} = (X_1, X_2, \ldots, X_n)$ à n composantes où les X_i sont n variables aléatoires indépendantes et de même loi.

La théorie de l'échantillonnage se propose d'étudier les propriétés du vecteur aléatoire à n composantes et des caractéristiques le résumant, encore appelées statistiques, à partir de la distribution supposée connue de la variable parente X, et d'étudier en particulier ce qui se passe lorsque la taille de l'échantillon est de plus en plus élevée. C'est généralement ce qui préoccupe les statisticiens bien que depuis quelques années des théories concernant les petits échantillons se développent également.

III Modèles

- **Modèle statistique**

Un **modèle statistique** est défini par la donnée d'une caractéristique vectorielle $\mathbf{X}$ et d'une famille de lois de probabilité de $\mathbf{X}$ notée $(\mathbb{P}_\theta)_{\theta\in\Theta}$. Il est généralement noté $(D_\mathbf{X}, \mathbb{P}_\theta)_{\theta\in\Theta}$.

- **Modèles paramétrique et non paramétrique**

Lorsque la famille de lois de probabilité $(\mathbb{P}_\theta)_{\theta\in\Theta}$ peut être indexée par un paramètre δ dont l'ensemble des valeurs possibles, noté Δ, est un sous-ensemble de $\mathbb{R}^k$, le modèle est appelé **modèle paramétrique**. Dans le cas contraire, le modèle est appelé **modèle non paramétrique**.

FICHE 25

Estimateur et propriétés d'un estimateur

I Estimateur et estimation

- **Estimateur**

Si $(X_1,\ldots,X_n)$ est un échantillon aléatoire d'effectif n de loi parente la loi de X, alors nous appelons **estimateur** du paramètre θ toute fonction h_n de l'échantillon aléatoire $(X_1,\ldots,X_n)$, noté $\widehat{\theta}_n$:

$$\widehat{\theta}_n = h_n(X_1,\ldots,X_n).$$

Remarques

1. À priori l'estimateur $\widehat{\theta}_n$ est à valeurs dans un ensemble Θ, contenant l'ensemble des valeurs possibles du paramètre θ.
2. $\widehat{\theta}_n$ est une **v.a.** de loi de probabilité qui dépend du paramètre θ.
3. $\widehat{\theta}_n$ peut être univarié ou multivarié.

- **Estimation**

Une fois l'échantillon prélevé, nous disposons de n valeurs observées $x_1,\ldots,x_n$, ce qui nous fournit une valeur $h_n(x_1,\ldots,x_n)$ qui est une réalisation de $\widehat{\theta}_n$ et que nous appelons **estimation**.

Remarques

1. Nous distinguons la variable aléatoire $\widehat{\theta}_n$ de sa valeur observée, notée $\widehat{\theta}_n(x_1,\ldots,x_n)$.
2. Nous utiliserons les notations suivantes :
   (i) $(X_1,\ldots,X_n)$ désigne l'échantillon aléatoire de taille n et les n observations ne sont pas encore à disposition.
   (ii) $(x_1,\ldots,x_n)$ désigne une réalisation de l'échantillon aléatoire et les n observations sont à disposition.
3. Il faut systématiquement se demander : « suis-je entrain de manipuler une variable aléatoire ou l'une de ses réalisations ? »

II Propriétés d'un estimateur

Le choix d'un estimateur va reposer sur **ses qualités**. Le premier défaut possible concerne la possibilité de comporter un **biais**.

- **Biais d'un estimateur**

  Le **biais** de $\widehat{\theta}_n$ se définit par $B\left(\widehat{\theta}_n\right) = \mathbb{E}\left(\widehat{\theta}_n\right) - \theta$.

- **Estimateur sans biais**

  $\widehat{\theta}_n$ est un **estimateur sans biais** (ou non biaisé) du paramètre θ si $B\left(\widehat{\theta}_n\right) = 0$, c'est-à-dire si $\mathbb{E}\left(\widehat{\theta}_n\right) = \theta$.

- **Estimateur asymptotiquement sans biais**

  Un estimateur $\widehat{\theta}_n$ est **asymptotiquement sans biais** pour θ si $\lim\limits_{n \to +\infty} \mathbb{E}\left(\widehat{\theta}_n\right) = \theta$.

- **Écart quadratique moyen**

  Si $\widehat{\theta}_n$ est un estimateur de θ, nous mesurons la précision de $\widehat{\theta}_n$ par l'**écart quadratique moyen**, noté EQM :

$$EQM\left(\widehat{\theta}_n\right) = \mathbb{E}\left(\left(\widehat{\theta}_n - \theta\right)^2\right) = \mathbb{V}\mathrm{ar}\left(\widehat{\theta}_n\right) + B\left(\widehat{\theta}_n\right)^2.$$

> **Remarque :** Si $\widehat{\theta}_n$ est un estimateur sans biais, c'est-à-dire si $B\left(\widehat{\theta}_n\right) = 0$, alors : $EQM\left(\widehat{\theta}_n\right) = \mathbb{V}\mathrm{ar}\left(\widehat{\theta}_n\right)$.

  Propriété : Entre deux estimateurs de θ, nous choisissons celui dont l'écart quadratique moyen ou le risque est le plus faible.

- **Estimateur relativement plus efficace**

  Un estimateur $\widehat{\theta_n^1}$ est **relativement plus efficace** qu'un estimateur $\widehat{\theta_n^2}$ s'il est plus précis que le second, c'est-à-dire si :

$$EQM\left(\widehat{\theta_n^1}\right) \leqslant EQM\left(\widehat{\theta_n^2}\right).$$

- **Estimateur sans biais optimal**

  Nous appelons **estimateur sans biais optimal** parmi les estimateurs sans biais, un estimateur $\widehat{\theta}_n$ préférable à tout autre au sens de la variance c'est-à-dire l'estimateur le plus efficace parmi tous les estimateurs sans biais.

- **Estimateur convergent**

  Un estimateur $\widehat{\theta}_n$ est un **estimateur convergent** s'il converge en probabilité vers θ quand n tend vers l'infini.

  Propriété : Si un estimateur est sans biais et que sa variance tend vers zéro quand n tend vers l'infini, alors cet estimateur est convergent.

III Trois exemples

Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire de loi parente la loi de X.

- **Estimateur de la moyenne**

  L'estimateur $\widehat{\mu}_n$ est égal à

$$\widehat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

  Propriété : Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ, $\widehat{\mu}_n$ est un **estimateur sans biais** de la moyenne μ, c'est-à-dire $\mathbb{E}(\widehat{\mu}_n) = \mu$. Lorsque la loi parente admet une variance, notée σ^2, la variance de l'estimateur $\widehat{\mu}_n$ est égale à $\mathbb{V}\mathrm{ar}\left(\widehat{\mu}_n\right) = \sigma^2/n$ et $\widehat{\mu}_n$ est un **estimateur convergent** de la moyenne μ.

- **Estimateur de la variance**

  L'estimateur S_n^2 est égal à

$$S_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \widehat{\mu}_n)^2.$$

  Propriété : Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ et une variance notée σ^2, S_n^2 est un **estimateur biaisé** de la variance σ^2 et le biais $B\left(S_n^2\right)$ est égal à $-\sigma^2/n$.

  S_n^2 est donc un **estimateur asymptotiquement sans biais**.

- **Estimateur corrigé de la variance**

  L'estimateur corrigé de la variance $S_{n,c}^2$ est égal à

$$S_{n,c}^2 = \frac{nS_n^2}{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \widehat{\mu}_n)^2.$$

  Propriété : Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ et une variance notée σ^2, $S_{n,c}^2$ est un **estimateur sans biais** de la variance σ^2.

D'après un exercice d'examen de l'Université de Strasbourg

Soit T_1 et T_2 deux estimateurs sans biais et indépendants, d'un paramètre θ, de variances respectives σ_1^2, σ_2^2, deux nombres réels strictement positifs. Soit α un réel appartenant à $[0,\ 1]$.

1. Montrez que la variable aléatoire $T = \alpha T_1 + (1-\alpha)T_2$ est aussi un estimateur sans biais de θ.

2. Pour quelle valeur du paramètre α la variance de la variable aléatoire T est-elle minimale ?

Solution

1. T est une variable aléatoire réelle comme combinaison linéaire de deux variables aléatoires réelles T_1 et T_2. Comme T_1 et T_2 sont des estimateurs sans biais de θ, T_1 et T_2 admettent une espérance. Nous avons, pour $\alpha \in [0,1]$

$$\begin{aligned}\mathbb{E}(T) &= \mathbb{E}\left(\alpha T_1 + (1-\alpha)T_2\right) = \alpha\mathbb{E}(T_1) + (1-\alpha)\mathbb{E}(T_2)\\ &= \alpha\theta + (1-\alpha)\theta = \theta\end{aligned}$$

Donc T est un estimateur de θ sans biais.

2. Calculons la variance de T. Comme T_1 et T_2 sont indépendants, nous avons

$$\begin{aligned}\mathbb{V}\text{ar}\,(T) &= \alpha^2\mathbb{V}\text{ar}\,(T_1) + (1-\alpha)^2\mathbb{V}\text{ar}\,(T_2)\\ &= \sigma_1^2\alpha^2 + \sigma_2^2(1-\alpha)^2\\ &= \left(\sigma_1^2+\sigma_2^2\right)\alpha^2 - 2\sigma_2^2\alpha + \sigma_2^2.\end{aligned}$$

Posons, pour $\alpha \in [0,1]$, $f(\alpha) = \left(\sigma_1^2+\sigma_2^2\right)\alpha^2 - 2\sigma_2^2\alpha + \sigma_2^2$. f est un polynôme du second degré dont le coefficient du terme de degré 2 est strictement positif.
Pour tout $\alpha \in [0,\ 1]$, nous avons $f'(\alpha) = 2\left(\sigma_1^2+\sigma_2^2\right)\alpha - 2\sigma_2^2$.
$f'(\alpha_0) = 0$ équivaut à $\alpha_0 = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\cdot$ Cette valeur α_0 appartient à l'intervalle $]0,\ 1[$ et par conséquent la variance de T est minimale pour $\alpha = \alpha_0 = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\cdot$

Nous avons alors l'égalité $\mathbb{V}\text{ar}(T) = \dfrac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$: la variance de T est la moyenne harmonique des variances de T_1 et T_2.

FICHE 26

Méthode du maximum de vraisemblance

I Vraisemblance d'un échantillon

La **vraisemblance des observations** $\mathbf{x} = (x_1,\ldots,x_n)$ d'un échantillon aléatoire de loi parente la loi de X est définie de la façon suivante :

- Si X est une variable aléatoire continue :

$$\theta \in \Theta \mapsto L(x_1,\ldots,x_n|\theta) = \prod_{i=1}^{n} f(x_i,\theta),$$

  où Θ est l'ensemble des valeurs possibles du paramètre θ.
- Si X est une variable aléatoire discrète :

$$\theta \in \Theta \mapsto L(x_1,\ldots,x_n|\theta) = \prod_{i=1}^{n} \mathbb{P}_\theta(X_i = x_i).$$

Remarque : Les expressions des vraisemblances ci-dessus ne sont valables que parce que les variables aléatoires $X_1,\ldots,X_n$ sont indépendantes par définition d'un **échantillon aléatoire**.

II Estimateur du maximum de vraisemblance

Définition

Un **estimateur du maximum de vraisemblance** (EMV) du paramètre θ est une statistique de l'échantillon :

$$\begin{array}{rcl} \widehat{\theta}_n : D^n_{\mathbf{X}} & \to & \Theta \\ \mathbf{x} = (x_1,\ldots,x_n) & \mapsto & \widehat{\theta}_n(x_1,\ldots,x_n) \end{array}$$

telle que $\forall \theta \in \Theta,\ L(x_1,\ldots,x_n|\widehat{\theta}_n) \geqslant L(x_1,\ldots,x_n|\theta)$.

Remarques

1. $L(\mathbf{x}|\theta)$ n'a aucune raison d'être différentiable en θ.
2. Il n'y a aucune raison pour qu'un EMV soit sans biais.
3. Un EMV n'a aucune raison d'être unique.

Remarque : Le principe de vraisemblance, à la base de la procédure d'estimation du maximum de vraisemblance, revient à rechercher la valeur de θ, fonction des observations $(x_1,\ldots,x_n)$, qui assure la plus grande probabilité d'obtenir ces observations.
La recherche d'un maximum de la vraisemblance n'est pas forcément réduite à un simple calcul des zéros de la dérivée de L.
Cependant, ce cas étant le plus fréquent, il est logique de poser les deux hypothèses suivantes :

1. $D_{\mathbf{X}}$ ne dépend pas de θ.
2. $\forall \mathbf{x} \in D_{\mathbf{X}}^n$, $\forall \theta \in \Theta$, L est deux fois continûment dérivable par rapport à θ.

Alors $\widehat{\theta}_n$, EMV, est solution du système d'équations en θ suivant :

$$(I) \Leftrightarrow \left\{ \begin{array}{lcll} \dfrac{\partial L}{\partial \theta}(\mathbf{x}|\theta) & = & 0 & (1) \\ \dfrac{\partial^2 L}{\partial \theta^2}(\mathbf{x}|\theta) & < & 0 & (2) \end{array} \right.$$

En fait, vu la forme des densités des lois usuelles de probabilité, il est aussi aisé d'utiliser le logarithme de la vraisemblance, $\log L(x_1,\ldots,x_n|\theta)$, si $f(\mathbf{x}, \theta) > 0$, pour tout $\mathbf{x} \in D_{\mathbf{X}}^n$, pour tout $\theta \in \Theta$, ce que nous supposerons :

$$(1) \Leftrightarrow (1') \quad \frac{\partial \log L}{\partial \theta}(\mathbf{x}|\theta) = 0$$

et

$$(2) \Leftrightarrow (2') \quad \frac{\partial^2 \log L}{\partial \theta^2}(\mathbf{x}|\theta) < 0.$$

Remarque : L'équation $(1')$ s'appelle l'équation de vraisemblance.

Estimation par intervalle de confiance

FICHE 27

I Introduction

Lorsque nous procédons à une estimation, il est plus pertinent de fournir un intervalle $\theta_1 < \theta < \theta_2$ plutôt que simplement une estimation ponctuelle, c'est-à-dire d'écrire uniquement $\widehat{\theta} = c$. Un tel intervalle $]\theta_1, \theta_2[$ s'appelle une **estimation par intervalle de confiance** du paramètre θ ou **estimation ensembliste** du paramètre θ.

II Espérance d'une loi normale

Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire de loi parente une loi normale d'espérance μ et d'écart-type σ.

- **L'écart-type σ est connu**

  $\widehat{\mu}_n$ suit une loi $\mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

  – **Intervalle de probabilité**

  L'intervalle de probabilité pour $\widehat{\mu}_n$ à $1-\alpha$ est égal à

$$\mu - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \widehat{\mu}_n < \mu + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},$$

  où $u_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ pour la loi normale centrée et réduite.

  – **Intervalle de confiance**

  L'intervalle de confiance pour μ à $1-\alpha$ est égal à

$$\widehat{\mu}_n(obs) - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \widehat{\mu}_n(obs) + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Remarque : Souvent $1-\alpha$ sera fixé à 0,95 et par conséquent $u_{1-\alpha/2}$ sera égal à 1,96.

- **L'écart-type σ est inconnu**

  – ***Théorème de Fisher***

Soit $X_1,\ldots,X_n$ n variables aléatoires indépendantes de loi une loi normale $\mathcal{N}(\mu,\sigma)$. $\widehat{\mu}_n$ et $\dfrac{nS_n^2}{\sigma^2}$ sont deux variables aléatoires indépendantes et $\dfrac{\widehat{\mu}_n-\mu}{\sigma/\sqrt{n}}$ suit une loi normale $\mathcal{N}(0,1)$ et $\dfrac{nS_n^2}{\sigma^2}$ suit une loi du Khi-deux $\chi^2(n-1)$. $\dfrac{\widehat{\mu}_n-\mu}{S_n/\sqrt{n-1}}$ suit une loi de Student $t(n-1)$.

– **Intervalle de probabilité**

L'intervalle de probabilité pour $\dfrac{\widehat{\mu}_n-\mu}{S_n/\sqrt{n-1}}$ à $1-\alpha$ est égal à

$$-t_{n-1;1-\alpha/2} < \frac{\widehat{\mu}_n-\mu}{S_n/\sqrt{n-1}} < t_{n-1;1-\alpha/2},$$

où $t_{n-1;1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ pour la loi de Student $t(n-1)$.

– **Intervalle de confiance**

L'intervalle de confiance pour μ à $1-\alpha$ est égal à

$$\widehat{\mu}_n(obs) - t_{n-1;1-\alpha/2}\frac{S_n(obs)}{\sqrt{n-1}} < \mu < \widehat{\mu}_n(obs) + t_{n-1;1-\alpha/2}\frac{S_n(obs)}{\sqrt{n-1}}.$$

III Variance d'une loi normale

- **La moyenne μ est connue**

$\widehat{\sigma^2}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i-\mu)^2$ est un estimateur de σ^2 et $n\widehat{\sigma^2}_n/\sigma^2$ suit une loi du Khi-deux $\chi^2(n)$.

– **Intervalle de probabilité**

k_1 et k_2 sont les bornes de l'intervalle de probabilité pour $n\widehat{\sigma^2}_n/\sigma^2$ si

$$\mathbb{P}\left[k_1 < \frac{n\,\widehat{\sigma^2}_n}{\sigma^2} < k_2\right] = 1-\alpha.$$

Par exemple nous pouvons prendre k_1 égal au quantile d'ordre $\alpha/2$ pour une loi du Khi-deux $\chi^2(n)$ et k_2 égal au quantile d'ordre $1-\alpha/2$ pour une loi du Khi-deux $\chi^2(n)$.

– **Intervalle de confiance**

L'intervalle de confiance pour σ^2 à $1-\alpha$ est égal à

$$\frac{n\,\widehat{\sigma^2}_n(obs)}{k_2} < \sigma^2 < \frac{n\,\widehat{\sigma^2}_n(obs)}{k_1}.$$

- **La moyenne μ est inconnue**

$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \widehat{\mu}_n)^2$ est un estimateur de σ^2 et nS_n^2/σ^2 suit une loi du Khi-deux $x^2(n-1)$.

– **Intervalle de probabilité**

l_1 et l_2 sont les bornes de l'intervalle de probabilité pour nS_n^2/σ^2 si

$$\mathbb{P}\left[l_1 < \frac{nS_n^2}{\sigma^2} < l_2\right] = 1-\alpha.$$

Par exemple nous pouvons prendre l_1 égal au quantile d'ordre $\alpha/2$ pour une loi du Khi-deux $\chi^2(n-1)$ et l_2 égal au quantile d'ordre $1-\alpha/2$ pour une loi du Khi-deux $\chi^2(n-1)$.

– **Intervalle de confiance**

L'intervalle de confiance pour σ^2 à $1-\alpha$ est égal à

$$\frac{nS_n^2(obs)}{l_2} < \sigma^2 < \frac{nS_n^2(obs)}{l_1}.$$

Les plantes marines

Un biologiste étudie un type d'algue qui attaque les plantes marines. La toxine contenue dans cette algue est obtenue sous forme d'une solution organique. Il mesure la quantité de toxine par gramme de solution. Il a obtenu les neuf mesures suivantes, exprimées en milligrammes : 1,2 ; 0,8 ; 0,6 ; 1,1 ; 1,2 ; 0,9 ; 1,5 ; 0,9 ; 1,0. Nous supposerons que ces mesures sont les réalisations de variables aléatoires indépendantes et identiquement distribuées suivant une loi normale d'espérance μ et d'écart-type σ.

1. Donnez une estimation ponctuelle de l'espérance μ et de l'écart-type σ de la quantité de toxine par gramme de solution.

2. Déterminez un intervalle de confiance à 95 % pour l'espérance μ de la quantité de toxine par gramme de solution.

3. Le biochimiste trouve que l'intervalle obtenu n'est pas satisfaisant car trop étendu. Que doit-il faire pour obtenir une estimation plus précise ?

Solution

1. Donnons une estimation ponctuelle de l'espérance μ de la quantité de toxine par gramme de solution.
Nous avons

$$\begin{aligned}\widehat{\mu}_9(obs) &= \frac{1}{9}(1{,}2+0{,}8+0{,}6+1{,}1+1{,}2+0{,}9+1{,}5+0{,}9+1{,}0)\\ &= 9{,}200\,mg.\end{aligned}$$

Nous donnons une estimation ponctuelle de l'écart-type σ de la quantité de toxine par gramme de solution à l'aide de l'estimateur corrigé $S_{9,c}$. Nous avons $S_{9,c}(obs) = 0{,}264\,mg$.

2. L'espérance μ et l'écart-type σ étant inconnus, l'intervalle de confiance pour μ de niveau 95 % s'obtient avec la formule suivante :

$$\begin{aligned}\widehat{\mu}_9(obs) - t_{8;0{,}975}\frac{S_9(obs)}{\sqrt{8}} < \quad &\mu \quad < \widehat{\mu}_9(obs) + t_{8;0{,}975}\frac{S_9(obs)}{\sqrt{8}}\\ 8{,}977\,mg < \quad &\mu \quad < 9{,}403\,mg\end{aligned}$$

où $t_{8;0{,}975}$ est le quantile d'ordre $0{,}975$ pour la loi de Student à huit degrés de liberté $t(8)$.

3. Il doit augmenter la taille de l'échantillon.

Tests d'hypothèse

FICHE 28

I Probabilités d'erreur et hypothèses

Un test statistique est un mécanisme qui permet de trancher entre deux hypothèses à partir des résultats d'un échantillon. Soit $\mathcal{H}_0$ et $\mathcal{H}_1$ ces deux hypothèses, dont une et une seule est vraie. La décision consiste à choisir $\mathcal{H}_0$ ou $\mathcal{H}_1$. Il y a quatre cas qui sont reproduits dans le tableau ci-dessous

	$\mathcal{H}_0$ vraie	$\mathcal{H}_1$ vraie
$\mathcal{H}_0$ décidée	$1 - \alpha$	β
$\mathcal{H}_1$ décidée	α	$1 - \beta$

où α **et** β **sont les probabilités d'erreur de 1re et 2e espèce :**
– α est la probabilité de décider $\mathcal{H}_1$ alors que $\mathcal{H}_0$ est vraie ;
– β est la probabilité de décider $\mathcal{H}_0$ alors que $\mathcal{H}_1$ est vraie.

Remarque : Ces deux probabilités d'erreur sont antagonistes, plus α est grande (respectivement petite), plus β est petite (respectivement grande).

Définition : La **puissance d'un test** est égale à $1 - \beta$.
Dans la pratique des tests statistiques, il est de règle de se fixer α. Les valeurs les plus courantes sont 1 %, 5 % ou 10 %.
$\mathcal{H}_0$ est **l'hypothèse de référence**, aussi appelée **hypothèse nulle**. $\mathcal{H}_1$ est **l'hypothèse alternative**.

II Vocabulaire pour un test statistique

- En fonction des hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ à tester, nous choisissons une **variable de décision**. La loi de cette variable doit être connue dans au moins une hypothèse, le plus souvent $\mathcal{H}_0$, afin de ne pas introduire de nouvelles inconnues dans le problème. Elle est appelée **distribution de référence**.
- La probabilité d'erreur de 1$^{\text{re}}$ espèce α est fixée.
- **Région critique**

  Nous appelons **région critique** l'ensemble des valeurs de la variable de décision qui conduisent à écarter l'hypothèse nulle $\mathcal{H}_0$ au profit de l'hypothèse alternative $\mathcal{H}_1$.
  La région critique ou la **zone de rejet** correspond aux intervalles dans lesquels les différences sont trop grandes pour être le fruit du hasard d'échantillonnage.
- **Région d'acceptation**

  Nous appelons **région d'acceptation** la région complémentaire de la région critique.
  La région ou la **zone d'acceptation** correspond à l'intervalle dans lequel les différences observées entre les réalisations et la théorie sont attribuables aux fluctuations d'échantillonnage.
- La construction d'un test est la détermination de la région critique, cette détermination se fait sans connaître le résultat de l'expérience, donc *a priori*.

III Tests bilatéral et unilatéral

Avant d'appliquer un test statistique, il s'agit de définir le problème posé. En effet, selon les hypothèses formulées, nous appliquons soit un **test bilatéral**, soit un **test unilatéral**.

- Un **test bilatéral** s'applique quand nous cherchons une différence entre les valeurs de deux paramètres, ou entre la valeur d'un paramètre et une valeur donnée sans se préoccuper du signe de la différence. La **zone de rejet** de l'hypothèse de référence peut être l'union de deux zones situées dans chacune des queues de la distribution.
- Un **test unilatéral** s'applique quand nous cherchons à savoir si un paramètre est supérieur (ou inférieur) à un autre paramètre ou à une valeur donnée. La **zone de rejet** de l'hypothèse principale est située d'un seul côté de la distribution de probabilité de référence.

Tests du χ^2

FICHE 29

Le test du χ^2 s'applique à des variables qualitatives à plusieurs modalités.

I Test d'indépendance

Étude de la liaison entre deux caractères qualitatifs X et Y.

$\mathcal{H}_0$: Les variables X et Y sont indépendantes

contre

$\mathcal{H}_1$: Les variables X et Y ne sont pas indépendantes

Nous considérons donc le tableau suivant :

X \ Y	Modalité 1	...	Modalité j	...	Modalité J	Total
Modalité 1	m_{11}	...	m_{1j}	...	m_{1J}	$m_{1\bullet}$
...	...	...	...	...	...	...
Modalité i	m_{i1}	...	m_{ij}	...	m_{iJ}	$m_{i\bullet}$
...	...	...	...	...	...	...
Modalité I	m_{I1}	...	m_{Ij}	...	m_{IJ}	$m_{I\bullet}$
Total	$m_{\bullet 1}$	...	$m_{\bullet j}$	...	$m_{\bullet J}$	$m_{\bullet\bullet} = n$

où m_{ij} correspond aux nombre d'individus observés ayant la modalité i pour X et la modalité j pour Y.

La notation $m_{i\bullet}$ correspond à $\sum_{j=1}^{J} m_{ij}$ et la notation $m_{\bullet j}$ correspond à $\sum_{i=1}^{I} m_{ij}$.

Le principe du test consiste à comparer les effectifs tels que nous les avons, à la répartition que nous aurions si les variables étaient indépendantes. Dans ce cas, en considérant que les marges

$$(m_{1\bullet},\ldots,m_{i\bullet},\ldots,m_{I\bullet},m_{\bullet 1},\ldots,m_{\bullet j},\ldots,m_{\bullet J})$$

sont fixées, nous pouvons calculer cette répartition théorique dans chacun des échantillons. Nous avons alors :

$$c_{ij} = \frac{m_{i\bullet} m_{\bullet j}}{m_{\bullet\bullet}}.$$

Il s'agit donc des effectifs théoriques sous l'hypothèse d'indépendance $\mathcal{H}_0$. Afin d'étudier l'écart entre ces deux répartitions, observée et théorique, nous adoptons l'indice suivant, dû à Pearson

$$\chi^2_{obs} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(m_{ij} - c_{ij})^2}{c_{ij}}.$$

Les conditions d'application du test détaillé ci-dessous sont

$$c_{ij} \geqslant 5 \quad \text{et} \quad n \geqslant 50.$$

Sous l'hypothèse nulle $\mathcal{H}_0$ et lorsque les conditions d'application du test sont remplies, χ^2_{obs} est une réalisation d'une variable aléatoire qui suit approximativement une loi du χ^2 à $(I-1)(J-1)$ degrés de liberté.

Pour un seuil fixé α, les tables de la loi du χ^2 à $(I-1)(J-1)$ degrés de liberté, nous fournissent une valeur critique c telle que $\mathbb{P}\left[\chi^2((I-1)(J-1)) \leqslant c\right] = 1 - \alpha$. Si nous utilisons un logiciel de statistique celui-ci nous fournit une p-valeur.

Règles de décision

Alors nous décidons

$$\begin{cases} \mathcal{H}_0 \text{ est vraie si } \chi^2 < c, \\ \mathcal{H}_1 \text{ est vraie si } \chi^2 \geqslant c. \end{cases}$$

Dans le cas où nous ne pouvons pas rejeter l'hypothèse nulle $\mathcal{H}_0$ et par conséquent nous l'acceptons, nous devrions calculer le risque de seconde espèce du test. Dans le cadre de ce livre, nous ne donnerons pas la formule pour calculer le risque et par conséquent la puissance du test.

1. Lorsque les conditions ne sont pas remplies, il existe des corrections, par exemple celle de Yates, ou le test exact de Fisher dans le cas de deux variables qualitatives à deux modalités.
2. S'il y a plus de deux modalités, nous pouvons essayer d'en regrouper si cela est possible, c'est-à-dire si cela a un sens.
3. Ce test, tel qu'il est exposé, ne peut pas être appliqué à des échantillons appariés.

II Test d'adéquation d'une loi à une loi donnée

Le test suivant est un test adapté pour s'intéresser à la possibilité d'une adéquation d'une distribution à une loi de probabilité donnée. Il est adapté pour des lois de probabilité discrètes et peut être également utilisé pour une loi continue entièrement spécifiée. Nous détaillons son application au cas d'une loi discrète de support fini définie par $(q_{x_k})_{x_k \in \mathcal{X}}$ avec Card $\mathcal{X} = N$.

$$\boxed{\mathcal{H}_0 : (p_{x_k})_{x_k \in \mathcal{X}} = (q_{x_k})_{x_k \in \mathcal{X}}}$$

contre

$$\boxed{\mathcal{H}_1 : (p_{x_k})_{x_k \in \mathcal{X}} \neq (q_{x_k})_{x_k \in \mathcal{X}}}$$

Nous mesurons ensuite la distance entre les effectifs observés m_i et les effectifs théoriques c_i de la même façon que dans le paragraphe précédent

$$\chi^2_{obs} = \sum_{i=1}^{I} \frac{(m_i - c_i)^2}{c_i}.$$

Les conditions d'application du test détaillé ci-dessous sont

$$c_i \geqslant 5 \quad \text{et} \quad n \geqslant 50.$$

Sous l'hypothèse nulle $\mathcal{H}_0$ et lorsque les conditions d'application du test sont remplies, χ^2_{obs} est une réalisation d'une variable aléatoire qui suit approximativement une loi du χ^2 à $(I-1)$ degrés de liberté, où I est le nombre de modalités. La règle de décision est identique au cas précédent.

Les incontournables boules

Nous sélectionnons 160 boules dans une urne à quatre couleurs dont nous ne connaissons pas la répartition. À partir de cet échantillon, nous voulons savoir s'il est possible que la répartition des couleurs dans l'urne soit de 9/16 de boules noires, 3/16 de boules rouges, 3/16 de boules jaunes et 1/16 de boules vertes.
Les données ont été reportées dans le tableau suivant :

Couleurs	Noir	Rouge	Jaune	Vert	Total
Proportion théorique	$\frac{9}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	1
Effectif observé	100	18	24	18	160

Solution

Nous allons calculer quels seraient les effectifs si la répartition théorique était respectée. Ainsi avec les 160 boules, si la répartition théorique était respectée nous aurions $160 \times 9/16 = 90$ boules noires.
Nous obtenons les résultats suivants :

Couleurs	Noir	Rouge	Jaune	Vert	Total
Effectif théorique	90	30	30	10	160

Les effectifs théoriques sont tous supérieurs à 5 et l'effectif total est supérieur à 50. Les conditions d'application du test d'adéquation du χ^2 sont donc bien remplies.
Nous mesurons ensuite la distance entre les effectifs observés et les effectifs théoriques à l'aide de la statistique du khi-carré :

$$\chi^2_{obs} = \sum_{i=1}^{I} \frac{(m_i - c_i)^2}{c_i}.$$

Nous obtenons une valeur $\chi^2_{obs} = 13{,}5111$. Pour $\alpha = 5\%$, la valeur critique d'une loi du χ^2 à 3 degrés de liberté vaut 7,8147 et par conséquent $\chi^2_{obs} > c$. La répartition observée nous permet d'affirmer que la répartition des couleurs au sein de l'urne est significativement différente de la répartition théorique introduite ci-dessus.

Régression linéaire par MCO

I Ajustement linéaire

Nous voulons ajuster la distribution de (X, Y) à une fonction linéaire ou affine $Y = aX + b$. Cet ajustement s'appelle la **régression linéaire de Y en X**. Soit $(X_1, Y_1), \ldots, (X_N, Y_N)$ les données brutes.
Nous allons déterminer (a_0, b_0) de manière à ce que la fonction

$$d(a,b) = \sum_{k=1}^{N} (aX_k + b - Y_k)^2$$

ait un minimum global sur $\mathbb{R}^2$ en (a_0, b_0) (c'est pourquoi cette méthode s'appelle aussi **la méthode des moindres carrés**). Le minimum de cette fonction positive et différentiable sur $\mathbb{R}^2$ est réalisé en un unique point critique. Nous en déduisons les valeurs de a_0 et b_0 :

$$a_0 = \frac{\mathbb{C}\mathrm{ov}(X, Y)}{\sigma(X)^2} = \frac{r(X, Y)\sigma(Y)}{\sigma(X)} \quad \text{et} \quad b_0 = \mu(Y) - a_0\mu(X).$$

Remarques

1. La droite D de régression par la méthode des moindres carrés, d'équation $D : \ y = a_0x + b_0$, passe par le centre de masse $(\mu(X), \mu(Y))$ du nuage de points (X_1, Y_1), ..., (X_N, Y_N).
2. En échangeant les rôles de X et Y nous obtenons la régression linéaire de X en Y. En général les deux droites de régression sont distinctes.

II Variation expliquée et inexpliquée

Le but d'un modèle de régression linéaire est d'expliquer **une partie de la variation** de la variable expliquée Y du fait de sa dépendance linéaire à la variable explicative X.

- Pour mesurer la variation de la variable Y, il est d'usage de considérer les différences entre les observations Y_k et leur moyenne $\mu(Y)$. La somme des carrés $\sum_{k=1}^{N}(Y_k-\mu(Y))^2$ est appelée **la somme des carrés totale** ou **variation totale** et est notée SC_{tot}.
- Si la variable Y dépend de X, et que nous la mesurons sur des individus avec différentes valeurs de X, nous observerons une variation en conséquence. Ici nous ne mesurons que **la variation expliquée** par la régression linéaire de Y en X obtenue par la méthode des moindres carrés. Nous notons $\widehat{Y_k}$ la valeur prédite par ce modèle pour $X = X_k$. Elle est évaluée par la somme des carrés $\sum_{k=1}^{N}\left(\widehat{Y_k}-\mu(Y)\right)^2$ qui est appelée **la somme des carrés due à la régression** ou **variation expliquée par le modèle** et est notée SC_{reg}.
- Généralement, lorsque nous mesurons Y sur des individus avec une même valeur de X, nous observons encore une certaine variation. Il s'agit de **la variation inexpliquée** par la régression linéaire de Y en X obtenue par la méthode des moindres carrés.

  La somme des carrés $\sum_{k=1}^{N}\left(Y_k-\widehat{Y_k}\right)^2$ est appelée **la somme des carrés des résidus** ou **la somme des carrés résiduelle** ou encore **variation inexpliquée par le modèle** et est notée SC_{res}.

Théorème : Nous avons une décomposition des variations suivantes

$$\sum_{k=1}^{N}(Y_k-\mu(Y))^2=\sum_{k=1}^{N}\left(Y_k-\widehat{Y_k}\right)^2+\sum_{k=1}^{N}\left(\widehat{Y_k}-\mu(Y)\right)^2.$$

Le pourcentage de la variation totale qui est expliquée par le modèle est évalué par le **coefficient de détermination**, noté R^2, qui est défini par

$$R^2=\frac{\text{Variation expliquée}}{\text{Variation totale}}=\frac{SC_{reg}}{SC_{tot}}.$$

Remarques :

1. Le coefficient de détermination R^2 prend ses valeurs entre 0 et 1.
2. Plus le coefficient de détermination R^2 est proche de 1, plus les données sont alignées sur la droite de régression.

Table des quantiles de la loi normale centrée et réduite

	0	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0	$-\infty$	-3.090	-2.878	-2.748	-2.652	-2.576	-2.512	-2.457	-2.409	-2.366
0.01	-2.326	-2.290	-2.257	-2.226	-2.197	-2.170	-2.144	-2.120	-2.097	-2.075
0.02	-2.054	-2.034	-2.014	-1.995	-1.977	-1.960	-1.943	-1.927	-1.911	-1.896
0.03	-1.881	-1.866	-1.852	-1.838	-1.825	-1.812	-1.799	-1.787	-1.774	-1.762
0.04	-1.751	-1.739	-1.728	-1.717	-1.706	-1.695	-1.685	-1.675	-1.665	-1.655
0.05	-1.645	-1.635	-1.626	-1.616	-1.607	-1.598	-1.589	-1.581	-1.572	-1.563
0.06	-1.555	-1.546	-1.538	-1.530	-1.522	-1.514	-1.506	-1.498	-1.491	-1.483
0.07	-1.476	-1.468	-1.461	-1.454	-1.447	-1.440	-1.433	-1.425	-1.419	-1.412
0.08	-1.405	-1.398	-1.392	-1.385	-1.379	-1.372	-1.366	-1.359	-1.353	-1.347
0.09	-1.341	-1.335	-1.329	-1.323	-1.316	-1.311	-1.305	-1.299	-1.293	-1.287
0.9	1.282	1.287	1.293	1.299	1.305	1.311	1.316	1.323	1.329	1.335
0.91	1.341	1.347	1.353	1.359	1.366	1.372	1.379	1.385	1.392	1.398
0.92	1.405	1.412	1.419	1.425	1.433	1.440	1.447	1.454	1.461	1.468
0.93	1.476	1.483	1.491	1.498	1.506	1.514	1.522	1.530	1.538	1.546
0.94	1.555	1.563	1.572	1.581	1.589	1.598	1.607	1.616	1.626	1.635
0.95	1.645	1.655	1.665	1.675	1.685	1.695	1.706	1.717	1.728	1.739
0.96	1.751	1.762	1.774	1.787	1.799	1.812	1.825	1.838	1.852	1.866
0.97	1.881	1.896	1.911	1.927	1.943	1.960	1.977	1.995	2.014	2.034
0.98	2.054	2.075	2.097	2.120	2.144	2.170	2.197	2.226	2.257	2.290
0.99	2.326	2.366	2.409	2.457	2.512	2.576	2.652	2.748	2.878	3.090

Table de la fonction de répartition de la loi normale centrée et réduite

	0	.01	.02	.03	.04	.05	.06	.07	.08	.09
0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Table de la loi de Student

ddl	0.900	0.950	0.975	0.990	0.995
1	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.8856	2.9200	4.3027	6.9646	9.9248
3	1.6377	2.3534	3.1824	4.5407	5.8409
4	1.5332	2.1318	2.7764	3.7469	4.6041
5	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.9980	3.4995
8	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.3722	1.8125	2.2281	2.7638	3.1693
11	1.3634	1.7959	2.2010	2.7181	3.1058
12	1.3562	1.7823	2.1788	2.6810	3.0545
13	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.3450	1.7613	2.1448	2.6245	2.9768
15	1.3406	1.7531	2.1314	2.6025	2.9467
16	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.3277	1.7291	2.0930	2.5395	2.8609
20	1.3253	1.7247	2.0860	2.5280	2.8453
21	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.3178	1.7109	2.0639	2.4922	2.7969
25	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.3150	1.7056	2.0555	2.4786	2.7787
27	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.3114	1.6991	2.0452	2.4620	2.7564
30	1.3104	1.6973	2.0423	2.4573	2.7500

Table de la loi du χ^2, régions unilatérales

ddl	0.01	0.05	0.10	0.90	0.95	0.99
1	0.0002	0.0039	0.0158	2.7055	3.8415	6.6349
2	0.0201	0.1026	0.2107	4.6052	5.9915	9.2103
3	0.1148	0.3518	0.5844	6.2514	7.8147	11.3449
4	0.2971	0.7107	1.0636	7.7794	9.4877	13.2767
5	0.5543	1.1455	1.6103	9.2364	11.0705	15.0863
6	0.8721	1.6354	2.2041	10.6446	12.5916	16.8119
7	1.2390	2.1673	2.8331	12.0170	14.0671	18.4753
8	1.6465	2.7326	3.4895	13.3616	15.5073	20.0902
9	2.0879	3.3251	4.1682	14.6837	16.9190	21.6660
10	2.5582	3.9403	4.8652	15.9872	18.3070	23.2093
11	3.0535	4.5748	5.5778	17.2750	19.6751	24.7250
12	3.5706	5.2260	6.3038	18.5493	21.0261	26.2170
13	4.1069	5.8919	7.0415	19.8119	22.3620	27.6882
14	4.6604	6.5706	7.7895	21.0641	23.6848	29.1412
15	5.2293	7.2609	8.5468	22.3071	24.9958	30.5779
16	5.8122	7.9616	9.3122	23.5418	26.2962	31.9999
17	6.4078	8.6718	10.0852	24.7690	27.5871	33.4087
18	7.0149	9.3905	10.8649	25.9894	28.8693	34.8053
19	7.6327	10.1170	11.6509	27.2036	30.1435	36.1909
20	8.2604	10.8508	12.4426	28.4120	31.4104	37.5662
21	8.8972	11.5913	13.2396	29.6151	32.6706	38.9322
22	9.5425	12.3380	14.0415	30.8133	33.9244	40.2894
23	10.1957	13.0905	14.8480	32.0069	35.1725	41.6384
24	10.8564	13.8484	15.6587	33.1962	36.4150	42.9798
25	11.5240	14.6114	16.4734	34.3816	37.6525	44.3141
26	12.1981	15.3792	17.2919	35.5632	38.8851	45.6417
27	12.8785	16.1514	18.1139	36.7412	40.1133	46.9629
28	13.5647	16.9279	18.9392	37.9159	41.3371	48.2782
29	14.2565	17.7084	19.7677	39.0875	42.5570	49.5879
30	14.9535	18.4927	20.5992	40.2560	43.7730	50.8922

Table de la loi du χ^2, régions bilatérales

ddl	0.01	0.05	0.10	0.90	0.95	0.99
1	0.0000	0.0010	0.0039	3.8415	5.0239	7.8794
2	0.0100	0.0506	0.1026	5.9915	7.3778	10.5966
3	0.0717	0.2158	0.3518	7.8147	9.3484	12.8382
4	0.2070	0.4844	0.7107	9.4877	11.1433	14.8603
5	0.4117	0.8312	1.1455	11.0705	12.8325	16.7496
6	0.6757	1.2373	1.6354	12.5916	14.4494	18.5476
7	0.9893	1.6899	2.1673	14.0671	16.0128	20.2777
8	1.3444	2.1797	2.7326	15.5073	17.5345	21.9550
9	1.7349	2.7004	3.3251	16.9190	19.0228	23.5894
10	2.1559	3.2470	3.9403	18.3070	20.4832	25.1882
11	2.6032	3.8157	4.5748	19.6751	21.9200	26.7568
12	3.0738	4.4038	5.2260	21.0261	23.3367	28.2995
13	3.5650	5.0088	5.8919	22.3620	24.7356	29.8195
14	4.0747	5.6287	6.5706	23.6848	26.1189	31.3193
15	4.6009	6.2621	7.2609	24.9958	27.4884	32.8013
16	5.1422	6.9077	7.9616	26.2962	28.8454	34.2672
17	5.6972	7.5642	8.6718	27.5871	30.1910	35.7185
18	6.2648	8.2307	9.3905	28.8693	31.5264	37.1565
19	6.8440	8.9065	10.1170	30.1435	32.8523	38.5823
20	7.4338	9.5908	10.8508	31.4104	34.1696	39.9968
21	8.0337	10.2829	11.5913	32.6706	35.4789	41.4011
22	8.6427	10.9823	12.3380	33.9244	36.7807	42.7957
23	9.2604	11.6886	13.0905	35.1725	38.0756	44.1813
24	9.8862	12.4012	13.8484	36.4150	39.3641	45.5585
25	10.5197	13.1197	14.6114	37.6525	40.6465	46.9279
26	11.1602	13.8439	15.3792	38.8851	41.9232	48.2899
27	11.8076	14.5734	16.1514	40.1133	43.1945	49.6449
28	12.4613	15.3079	16.9279	41.3371	44.4608	50.9934
29	13.1211	16.0471	17.7084	42.5570	45.7223	52.3356
30	13.7867	16.7908	18.4927	43.7730	46.9792	53.6720

Table de la loi de Fisher-Snedecor, $\alpha = 5\ \%$

num	den 1	2	3	4	5	6	7	8	9	10
1	161.4476	18.5128	10.1280	7.7086	6.6079	5.9874	5.5914	5.3177	5.1174	4.9646
2	199.5000	19.0000	9.5521	6.9443	5.7861	5.1433	4.7374	4.4590	4.2565	4.1028
3	215.7073	19.1643	9.2766	6.5914	5.4095	4.7571	4.3468	4.0662	3.8625	3.7083
4	224.5832	19.2468	9.1172	6.3882	5.1922	4.5337	4.1203	3.8379	3.6331	3.4780
5	230.1619	19.2964	9.0135	6.2561	5.0503	4.3874	3.9715	3.6875	3.4817	3.3258
6	233.9860	19.3295	8.9406	6.1631	4.9503	4.2839	3.8660	3.5806	3.3738	3.2172
7	236.7684	19.3532	8.8867	6.0942	4.8759	4.2067	3.7870	3.5005	3.2927	3.1355
8	238.8827	19.3710	8.8452	6.0410	4.8183	4.1468	3.7257	3.4381	3.2296	3.0717
9	240.5433	19.3848	8.8123	5.9988	4.7725	4.0990	3.6767	3.3881	3.1789	3.0204
10	241.8817	19.3959	8.7855	5.9644	4.7351	4.0600	3.6365	3.3472	3.1373	2.9782
11	242.9835	19.4050	8.7633	5.9358	4.7040	4.0274	3.6030	3.3130	3.1025	2.9430
12	243.9060	19.4125	8.7446	5.9117	4.6777	3.9999	3.5747	3.2839	3.0729	2.9130
13	244.6898	19.4189	8.7287	5.8911	4.6552	3.9764	3.5503	3.2590	3.0475	2.8872
14	245.3640	19.4244	8.7149	5.8733	4.6358	3.9559	3.5292	3.2374	3.0255	2.8647
15	245.9499	19.4291	8.7029	5.8578	4.6188	3.9381	3.5107	3.2184	3.0061	2.8450
16	246.4639	19.4333	8.6923	5.8441	4.6038	3.9223	3.4944	3.2016	2.9890	2.8276
17	246.9184	19.4370	8.6829	5.8320	4.5904	3.9083	3.4799	3.1867	2.9737	2.8120
18	247.3232	19.4402	8.6745	5.8211	4.5785	3.8957	3.4669	3.1733	2.9600	2.7980
19	247.6861	19.4431	8.6670	5.8114	4.5678	3.8844	3.4551	3.1613	2.9477	2.7854
20	248.0131	19.4458	8.6602	5.8025	4.5581	3.8742	3.4445	3.1503	2.9365	2.7740
21	248.3094	19.4481	8.6540	5.7945	4.5493	3.8649	3.4349	3.1404	2.9263	2.7636
22	248.5791	19.4503	8.6484	5.7872	4.5413	3.8564	3.4260	3.1313	2.9169	2.7541
23	248.8256	19.4523	8.6432	5.7805	4.5339	3.8486	3.4179	3.1229	2.9084	2.7453
24	249.0518	19.4541	8.6385	5.7744	4.5272	3.8415	3.4105	3.1152	2.9005	2.7372
25	249.2601	19.4558	8.6341	5.7687	4.5209	3.8348	3.4036	3.1081	2.8932	2.7298
26	249.4525	19.4573	8.6301	5.7635	4.5151	3.8287	3.3972	3.1015	2.8864	2.7229
27	249.6309	19.4587	8.6263	5.7586	4.5097	3.8230	3.3913	3.0954	2.8801	2.7164
28	249.7966	19.4600	8.6229	5.7541	4.5047	3.8177	3.3858	3.0897	2.8743	2.7104
29	249.9510	19.4613	8.6196	5.7498	4.5001	3.8128	3.3806	3.0844	2.8688	2.7048
30	250.0951	19.4624	8.6166	5.7459	4.4957	3.8082	3.3758	3.0794	2.8637	2.6996

Table de la loi de Fisher-Snedecor, α = 5 % (suite)

num	den 11	12	13	14	15	16	17	18	19	20
1	4.8443	4.7472	4.6672	4.6001	4.5431	4.4940	4.4513	4.4139	4.3807	4.3512
2	3.9823	3.8853	3.8056	3.7389	3.6823	3.6337	3.5915	3.5546	3.5219	3.4928
3	3.5874	3.4903	3.4105	3.3439	3.2874	3.2389	3.1968	3.1599	3.1274	3.0984
4	3.3567	3.2592	3.1791	3.1122	3.0556	3.0069	2.9647	2.9277	2.8951	2.8661
5	3.2039	3.1059	3.0254	2.9582	2.9013	2.8524	2.8100	2.7729	2.7401	2.7109
6	3.0946	2.9961	2.9153	2.8477	2.7905	2.7413	2.6987	2.6613	2.6283	2.5990
7	3.0123	2.9134	2.8321	2.7642	2.7066	2.6572	2.6143	2.5767	2.5435	2.5140
8	2.9480	2.8486	2.7669	2.6987	2.6408	2.5911	2.5480	2.5102	2.4768	2.4471
9	2.8962	2.7964	2.7144	2.6458	2.5876	2.5377	2.4943	2.4563	2.4227	2.3928
10	2.8536	2.7534	2.6710	2.6022	2.5437	2.4935	2.4499	2.4117	2.3779	2.3479
11	2.8179	2.7173	2.6347	2.5655	2.5068	2.4564	2.4126	2.3742	2.3402	2.3100
12	2.7876	2.6866	2.6037	2.5342	2.4753	2.4247	2.3807	2.3421	2.3080	2.2776
13	2.7614	2.6602	2.5769	2.5073	2.4481	2.3973	2.3531	2.3143	2.2800	2.2495
14	2.7386	2.6371	2.5536	2.4837	2.4244	2.3733	2.3290	2.2900	2.2556	2.2250
15	2.7186	2.6169	2.5331	2.4630	2.4034	2.3522	2.3077	2.2686	2.2341	2.2033
16	2.7009	2.5989	2.5149	2.4446	2.3849	2.3335	2.2888	2.2496	2.2149	2.1840
17	2.6851	2.5828	2.4987	2.4282	2.3683	2.3167	2.2719	2.2325	2.1977	2.1667
18	2.6709	2.5684	2.4841	2.4134	2.3533	2.3016	2.2567	2.2172	2.1823	2.1511
19	2.6581	2.5554	2.4709	2.4000	2.3398	2.2880	2.2429	2.2033	2.1683	2.1370
20	2.6464	2.5436	2.4589	2.3879	2.3275	2.2756	2.2304	2.1906	2.1555	2.1242
21	2.6358	2.5328	2.4479	2.3768	2.3163	2.2642	2.2189	2.1791	2.1438	2.1124
22	2.6261	2.5229	2.4379	2.3667	2.3060	2.2538	2.2084	2.1685	2.1331	2.1016
23	2.6172	2.5139	2.4287	2.3573	2.2966	2.2443	2.1987	2.1587	2.1233	2.0917
24	2.6090	2.5055	2.4202	2.3487	2.2878	2.2354	2.1898	2.1497	2.1141	2.0825
25	2.6014	2.4977	2.4123	2.3407	2.2797	2.2272	2.1815	2.1413	2.1057	2.0739
26	2.5943	2.4905	2.4050	2.3333	2.2722	2.2196	2.1738	2.1335	2.0978	2.0660
27	2.5877	2.4838	2.3982	2.3264	2.2652	2.2125	2.1666	2.1262	2.0905	2.0586
28	2.5816	2.4776	2.3918	2.3199	2.2587	2.2059	2.1599	2.1195	2.0836	2.0517
29	2.5759	2.4718	2.3859	2.3139	2.2525	2.1997	2.1536	2.1131	2.0772	2.0452
30	2.5705	2.4663	2.3803	2.3082	2.2468	2.1938	2.1477	2.1071	2.0712	2.0391

Table de la loi de Fisher-Snedecor, α = 2,5 %

num	den 1	2	3	4	5	6	7	8	9	10
1	647.7890	38.5063	17.4434	12.2179	10.0070	8.8131	8.0727	7.5709	7.2093	6.9367
2	799.5000	39.0000	16.0441	10.6491	8.4336	7.2599	6.5415	6.0595	5.7147	5.4564
3	864.1630	39.1655	15.4392	9.9792	7.7636	6.5988	5.8898	5.4160	5.0781	4.8256
4	899.5833	39.2484	15.1010	9.6045	7.3879	6.2272	5.5226	5.0526	4.7181	4.4683
5	921.8479	39.2982	14.8848	9.3645	7.1464	5.9876	5.2852	4.8173	4.4844	4.2361
6	937.1111	39.3315	14.7347	9.1973	6.9777	5.8198	5.1186	4.6517	4.3197	4.0721
7	948.2169	39.3552	14.6244	9.0741	6.8531	5.6955	4.9949	4.5286	4.1970	3.9498
8	956.6562	39.3730	14.5399	8.9796	6.7572	5.5996	4.8993	4.4333	4.1020	3.8549
9	963.2846	39.3869	14.4731	8.9047	6.6811	5.5234	4.8232	4.3572	4.0260	3.7790
10	968.6274	39.3980	14.4189	8.8439	6.6192	5.4613	4.7611	4.2951	3.9639	3.7168
11	973.0252	39.4071	14.3742	8.7935	6.5678	5.4098	4.7095	4.2434	3.9121	3.6649
12	976.7079	39.4146	14.3366	8.7512	6.5245	5.3662	4.6658	4.1997	3.8682	3.6209
13	979.8368	39.4210	14.3045	8.7150	6.4876	5.3290	4.6285	4.1622	3.8306	3.5832
14	982.5278	39.4265	14.2768	8.6838	6.4556	5.2968	4.5961	4.1297	3.7980	3.5504
15	984.8668	39.4313	14.2527	8.6565	6.4277	5.2687	4.5678	4.1012	3.7694	3.5217
16	986.9187	39.4354	14.2315	8.6326	6.4032	5.2439	4.5428	4.0761	3.7441	3.4963
17	988.7331	39.4391	14.2127	8.6113	6.3814	5.2218	4.5206	4.0538	3.7216	3.4737
18	990.3490	39.4424	14.1960	8.5924	6.3619	5.2021	4.5008	4.0338	3.7015	3.4534
19	991.7973	39.4453	14.1810	8.5753	6.3444	5.1844	4.4829	4.0158	3.6833	3.4351
20	993.1028	39.4479	14.1674	8.5599	6.3286	5.1684	4.4667	3.9995	3.6669	3.4185
21	994.2856	39.4503	14.1551	8.5460	6.3142	5.1538	4.4520	3.9846	3.6520	3.4035
22	995.3622	39.4525	14.1438	8.5332	6.3011	5.1406	4.4386	3.9711	3.6383	3.3897
23	996.3462	39.4544	14.1336	8.5216	6.2891	5.1284	4.4263	3.9587	3.6257	3.3770
24	997.2492	39.4562	14.1241	8.5109	6.2780	5.1172	4.4150	3.9472	3.6142	3.3654
25	998.0808	39.4579	14.1155	8.5010	6.2679	5.1069	4.4045	3.9367	3.6035	3.3546
26	998.8490	39.4594	14.1074	8.4919	6.2584	5.0973	4.3949	3.9269	3.5936	3.3446
27	999.5609	39.4609	14.1000	8.4834	6.2497	5.0884	4.3859	3.9178	3.5845	3.3353
28	1000.2225	39.4622	14.0930	8.4755	6.2416	5.0802	4.3775	3.9093	3.5759	3.3267
29	1000.8388	39.4634	14.0866	8.4681	6.2340	5.0724	4.3697	3.9014	3.5679	3.3186
30	1001.4144	39.4646	14.0805	8.4613	6.2269	5.0652	4.3624	3.8940	3.5604	3.3110

Table de la loi de Fisher-Snedecor, α = 2,5 % (suite)

num	den 11	12	13	14	15	16	17	18	19	20
1	6.7241	6.5538	6.4143	6.2979	6.1995	6.1151	6.0420	5.9781	5.9216	5.8715
2	5.2559	5.0959	4.9653	4.8567	4.7650	4.6867	4.6189	4.5597	4.5075	4.4613
3	4.6300	4.4742	4.3472	4.2417	4.1528	4.0768	4.0112	3.9539	3.9034	3.8587
4	4.2751	4.1212	3.9959	3.8919	3.8043	3.7294	3.6648	3.6083	3.5587	3.5147
5	4.0440	3.8911	3.7667	3.6634	3.5764	3.5021	3.4379	3.3820	3.3327	3.2891
6	3.8807	3.7283	3.6043	3.5014	3.4147	3.3406	3.2767	3.2209	3.1718	3.1283
7	3.7586	3.6065	3.4827	3.3799	3.2934	3.2194	3.1556	3.0999	3.0509	3.0074
8	3.6638	3.5118	3.3880	3.2853	3.1987	3.1248	3.0610	3.0053	2.9563	2.9128
9	3.5879	3.4358	3.3120	3.2093	3.1227	3.0488	2.9849	2.9291	2.8801	2.8365
10	3.5257	3.3736	3.2497	3.1469	3.0602	2.9862	2.9222	2.8664	2.8172	2.7737
11	3.4737	3.3215	3.1975	3.0946	3.0078	2.9337	2.8696	2.8137	2.7645	2.7209
12	3.4296	3.2773	3.1532	3.0502	2.9633	2.8890	2.8249	2.7689	2.7196	2.6758
13	3.3917	3.2393	3.1150	3.0119	2.9249	2.8506	2.7863	2.7302	2.6808	2.6369
14	3.3588	3.2062	3.0819	2.9786	2.8915	2.8170	2.7526	2.6964	2.6469	2.6030
15	3.3299	3.1772	3.0527	2.9493	2.8621	2.7875	2.7230	2.6667	2.6171	2.5731
16	3.3044	3.1515	3.0269	2.9234	2.8360	2.7614	2.6968	2.6404	2.5907	2.5465
17	3.2816	3.1286	3.0039	2.9003	2.8128	2.7380	2.6733	2.6168	2.5670	2.5228
18	3.2612	3.1081	2.9832	2.8795	2.7919	2.7170	2.6522	2.5956	2.5457	2.5014
19	3.2428	3.0896	2.9646	2.8607	2.7730	2.6980	2.6331	2.5764	2.5265	2.4821
20	3.2261	3.0728	2.9477	2.8437	2.7559	2.6808	2.6158	2.5590	2.5089	2.4645
21	3.2109	3.0575	2.9322	2.8282	2.7403	2.6651	2.6000	2.5431	2.4930	2.4484
22	3.1970	3.0434	2.9181	2.8139	2.7260	2.6507	2.5855	2.5285	2.4783	2.4337
23	3.1843	3.0306	2.9052	2.8009	2.7128	2.6374	2.5721	2.5151	2.4648	2.4201
24	3.1725	3.0187	2.8932	2.7888	2.7006	2.6252	2.5598	2.5027	2.4523	2.4076
25	3.1616	3.0077	2.8821	2.7777	2.6894	2.6138	2.5484	2.4912	2.4408	2.3959
26	3.1516	2.9976	2.8719	2.7673	2.6790	2.6033	2.5378	2.4806	2.4300	2.3851
27	3.1422	2.9881	2.8623	2.7577	2.6692	2.5935	2.5280	2.4706	2.4200	2.3751
28	3.1334	2.9793	2.8534	2.7487	2.6602	2.5844	2.5187	2.4613	2.4107	2.3657
29	3.1253	2.9710	2.8451	2.7403	2.6517	2.5758	2.5101	2.4527	2.4019	2.3569
30	3.1176	2.9633	2.8372	2.7324	2.6437	2.5678	2.5020	2.4445	2.3937	2.3486

Index

Z

EXPRESS SCIENCES

Daniel FREDON
Myriam MAUMY-BERTRAND
Frédéric BERTRAND

Mathématiques
Statistique et probabilités
en 30 fiches

Des principes aux applications

Comment aller à l'essentiel, comprendre les méthodes et les démarches avant de les mettre en application ?

Conçue pour faciliter aussi bien l'apprentissage que la révision, la collection « **EXPRESS** » vous propose une présentation simple et concise de la statistique et des probabilités en **30 fiches pédagogiques**.

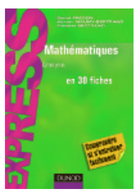
Chaque fiche comporte :

- les **idées clés** à connaître
- la **méthode** à mettre en œuvre
- les **applications** sous forme d'exercices corrigés.

Contenu :

• Analyse combinatoire • Variables aléatoires • Lois discrètes • Statistique descriptive • Échantillonnage • Régression linéaire

Des mêmes auteurs :



Daniel Fredon
était maître de conférences à l'université de Limoges

Myriam Maumy-Bertrand
est maître de conférence à l'université Louis Pasteur (Strasbourg)

Frédéric Bertrand
est maître de conférence à l'université Louis Pasteur (Strasbourg)

- **L1/L2 Mathématiques, Informatique, Sciences physiques, Cycles préparatoires intégrés**
- **IUT**

DUNOD

ISBN 978-2-10-054257-4

www.dunod.com