SY01, Automne 2024
Eléments de Probabilités

Equipe de mathématiques, LMAC

Université de Technologie de Compiègne
Laboratoire de Mathématiques Appliquées de Compiègne

Table des matières

Chapitre 1

Ensembles, Evénéments et probabilités

1.1 Introduction

Pour aborder un cours de probabilité, il est normal de vouloir donner un sens au mot clé du cours : **probabilité**. Aussi pouvons-nous commencer par une définition naïve de ce mot.

Définition 1.1.1 *Une probabilité est un nombre compris entre 0 et 1, associé à un événement et indiquant les chances de réalisation de cet événement au cours d'une expérience aléatoire.*

Cette définition laisse entrevoir que pour parler de probabilités il faut trois éléments essentiels :

i- une expérience aléatoire ;

ii- des événements ;

iii- des nombres associés aux événements.

A ces trois éléments nous substituerons respectivement les notions mathématiques d'**univers**, **tribu** et **mesure de probabilité**. Ces trois notions sont liées par des **axiomes** [1] qui permettent de mener des calculs.

Si l'on revient à la définition du mot probabilité on peut s'apercevoir que l'essentiel, c'est-à-dire le sens intrinsèque du mot *probabilité*, est transféré sur le mot *chances.*

Dans ce cours, une probabilité sera toujours un nombre compris entre 0 et 1 non assujetti à notre conception du hasard. Il n'en reste pas moins intéressant de voir comment ont évolué les opinions sur ce sujet d'ordre philosophique. Deux grands

1. Propositions primitives que l'on renonce à démontrer et sur lesquelles est basée une science.

courants s'affrontent : les *objectivistes* et les *subjectivistes*.

Conception objectiviste

Les objectivistes partent du postulat suivant : *la probabilité d'un événement peut être déterminée de manière unique.*
Vision classique. Héritage des jeux de hasard (Chevalier de Méré,...). En général l'ensemble Ω des éventualités (résultat des jeux) est un ensemble fini et des raisons de *symétrie* conduisent à donner la même probabilité à chaque éventualité ; donc 1/2 au jeu de *pile ou face*, 1/6 pour le lancer d'un dé, etc.

Dans ce cas le calcul des probabilités n'est affaire que de dénombrement et si A est un événement, la probabilité $P(A)$ de A est donnée par la célèbre formule :

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

On parle alors d'équiprobabilité. Ainsi, la probabilité d'obtenir un nombre pair en lançant un dé est 3/6=1/2.
Paradoxe de Bertrand. On considère un triangle équilatéral et son cercle circonscrit. On choisit une corde au hasard. Quelle est la probabilité que sa longueur soit supérieure à celle du côté du triangle ?

Solution Pour des raisons de symétrie on fixe un point A sur le cercle puis on choisit l'autre extrémité B de la corde sur le cercle. Comme on peut le voir sur la

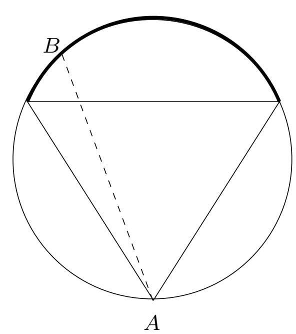


Figure 1.1 – en pointillé une corde d'extrémités A et B

figure 1.1 le point B doit être choisi sur l'arc de cercle qui apparaît en gras. La proportion de tels points sur le cercle étant de 1/3 on conclut que la probabilité cherchée est de 1/3 !

Solution Cette fois-ci on considère que pour déterminer une corde il suffit de choisir son centre A (il la détermine de manière unique). Encore une fois, pour des raisons de symétrie, on considère fixée la direction dans laquelle on choisit le point A. Comme on peut le voir sur la figure 1.2 les points admissibles sont ceux qui apparaissent sur le segment en gras ; ceux-ci représentent donc la moitié des points du diamètre et on conclut que la probabilité cherchée est de 1/2 !

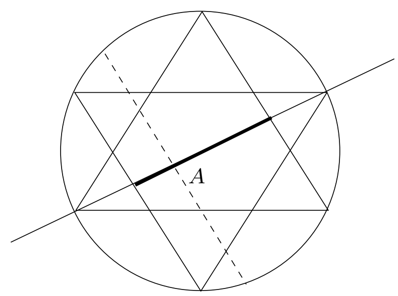


Figure 1.2 – en pointillé une corde de centre A

Les solutions ci-dessus sont exactes ! Et il en existe d'autres. . .le problème est simplement mal posé dans le sens où l'on ne précise pas ce que l'on entend par "choisir une corde au hasard".
Vision fréquentiste. Elle repose sur la "loi des grands nombres". Une seule expérience ne suffit pas à déterminer la probabilité d'un événement alors on répète l'expérience et on appelle probabilité de l'événement la fréquence limite de sa réalisation. Par exemple, si on note A l'événement "obtenir un 6" en lançant un dé et α_n le nombre de 6 obtenus au cours de n lancés, on a :

$$P(A) = \lim_{n\to+\infty} \frac{\alpha_n}{n} = \frac{1}{6}.$$

Cette méthode, bien que très intuitive (nous l'utiliserons d'ailleurs dans le cours), est très limitée en pratique puisqu'elle nécessite une infinité d'expériences. Nous verrons en fait que la loi des grands nombres n'est qu'une conséquence de la définition des probabilités.

Conception subjectiviste

Postulat : "les probabilité n'existent pas !" de Finetti.
Mesure d'incertitude. La probabilité objective n'existe pas, elle n'est donc pas une grandeur mesurable analogue à la masse d'un corps, elle est seulement une *mesure d'incertitude* pouvant varier avec les circonstances et l'observateur, donc subjective ! La seule exigence est qu'elle satisfasse les axiomes du calcul des probabilités. Pour les subjectivistes, ce sont nos incertitudes sur l'expérience du jet d'une pièce qui nous conduisent à évaluer à 1/2 la probabilité d'obtenir pile ou face.
L'approche Bayésienne. Les Bayésiens vont plus loin ! Voici un bref exemple illustrant leur démarche. Dans l'expérience aléatoire de pile ou face, les Bayésiens considèrent que la probabilité p d'obtenir pile (donc $1 - p$ d'obtenir face) est elle-même aléatoire. Comme $p \in [0, 1]$ ils donnent une loi dite *a priori* à p et, après avoir réalisé des expériences, en déduisent une loi dite *a posteriori* pour p tenant compte des résultats de l'expérience.

1.2 Ensembles, événements

1.2.1 Ensembles

Dans ce paragraphe nous allons rappeler un certain nombre de notions et résultats élémentaires dont l'utilisation sera très souvent implicite.

Famille des parties. On appelle $\mathcal{P}(E)$ la famille (l'ensemble) des parties de l'ensemble E. On a la caractérisation suivante :

$$A \in \mathcal{P}(E) \Leftrightarrow A \subset E \Leftrightarrow (\forall x \in A,\ x \in E).$$

Double inclusion. Soient E et F deux ensembles :

$$E = F \Leftrightarrow (E \subset F \text{ et } F \subset E).$$

Opérations élémentaires. Soit E un ensemble et A et B deux parties de E.

$$\begin{array}{rcll}
\bar{A} & = & \{x \in E | x \notin A\} & \text{complémentaire de } A \text{ dans } E \\
A \cup B & = & \{x \in E | x \in A \text{ ou } x \in B\} & \text{réunion de } A \text{ et } B \\
A \cap B & = & \{x \in E | x \in A \text{ et } x \in B\} & \text{intersection } A \text{ et } B \\
A \backslash B & = & \{x \in E | x \in A \text{ et } x \notin B\} & \text{différence de } A \text{ et } B \\
A \Delta B & = & \{x \in E | x \in A \backslash B \text{ ou } x \in B \backslash A\} & \text{différence symétrique de } A \text{ et } B
\end{array}$$

$\bar{A} = E \backslash A$, le complémentaire de A, est aussi noté A^c ou encore C_E^A pour bien spécifier qu'il s'agit du complémentaire de A **dans** E.
Soit I un ensemble d'indices et $(A_i)_{i \in I}$ une famille d'éléments de $\mathcal{P}(E)$ alors :

$$\begin{array}{rcll}
\bigcup_{i \in I} A_i & = & \{x \in E | \exists i \in I ; x \in A_i\} & (\text{réunion des } A_i) \\
\bigcap_{i \in I} A_i & = & \{x \in E | \forall i \in I ; x \in A_i\} & (\text{intersection des } A_i)
\end{array}$$

Partition. Soit I un ensemble d'indices et $(A_i)_{i \in I}$ une famille d'éléments de $\mathcal{P}(E)$ alors $(A_i)_{i \in I}$ est une partition de E si les conditions suivantes sont satisfaites :

— les A_i sont deux à deux disjoints ($A_i \cap A_j = \emptyset$ si $i \neq j$) ;

— la réunion des A_i est E ($\cup_{i \in I} A_i = E$) ;

Produit cartésien. Le *produit cartésien* de deux ensembles E et F, noté $E \times F$, est défini ainsi :

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

Ensembles finis, dénombrables et non dénombrables. Si E est un ensemble comptant n éléments ($n \in \mathbb{N}$) alors E est dit *fini* (si $n = 0$ alors $E = \emptyset$) ; n est appelé *cardinal* de E et plusieurs notations sont utilisées : $n = \mathrm{Card}(E) = \#E = |E|$. Soient E et F deux ensembles finis :

$$|E| = |F| \Leftrightarrow \text{il existe une bijection } f \text{ de } E \text{ dans } F.$$

Un ensemble est dit *dénombrable* s'il peut être mis en bijection avec une partie de $\mathbb{N}$. Tout ensemble fini est donc dénombrable, $\mathbb{N}$ est infini dénombrable ($\mathbb{N}$ est en bijection avec lui-même par l'application identité), $\mathbb{Z}$ est infini dénombrable : l'application φ de $\mathbb{N}$ sur $\mathbb{Z}$ définie par

$$\varphi(x) = \begin{cases} -x/2 & \text{si} \quad x \text{ est pair,} \\ (x+1)/2 & \text{si} \quad x \text{ est impair,} \end{cases}$$

est bijective. En fait on peut montrer que *toute réunion dénombrable d'ensembles dénombrables est encore dénombrable*. Par exemple Q est dénombrable, en effet :

$$\mathrm{Q} = \bigcup_{\mathrm{i} \in \mathbb{N}^*} \mathbb{Z}_{\mathrm{i}} \text{ où } \mathbb{Z}_{\mathrm{i}} = \{\ldots, -\frac{2}{\mathrm{i}}, -\frac{1}{\mathrm{i}}, 0, \frac{1}{\mathrm{i}}, \frac{2}{\mathrm{i}}, \ldots\},$$

et les $\mathbb{Z}_i$ sont dénombrables.
Enfin notons que tous les ensembles ne sont pas dénombrables ! $\mathbb{R}$ en est un exemple, il a la *puissance du continu*. Il en est de même pour tout intervalle $[a, b]$ $(a < b)$ de $\mathbb{R}$.

1.2.2 Application

Soit E et F deux ensembles non vides, une application de E dans F associe à tout élément $x \in E$ un **unique** élément $y = f(x) \in F$ (image de x). Si $y \in F$ tout élément $x \in E$ tel que $y = f(x)$ est appelé antécédent de y par f.

— Si tout élément de F possède au moins un antécédent par f, l'application f est dite *surjective* :
$$f \text{ surjective } \Leftrightarrow f(E) = F$$

— Si tout élément de F possède au plus un antécédent par f, l'application f est dite *injective* :
$$f \text{ injective } \Leftrightarrow (\forall (x, x') \in E^2, x \neq x' \Rightarrow f(x) \neq f(x')).$$

— Si tout élément de F possède un et un seul antécédent par f, l'application f est dite *bijective* :
$$f \text{ bijective } \Leftrightarrow f \text{ injective et surjective}$$

Image et image réciproque. Soit f une application de E dans F, $A \subset E$ et $B \subset F$.
$$f(A) = \{f(x) | x \in A\}$$
est appelé *image de A par f*.
$$f^{-1}(B) = \{x \in E | f(x) \in B\}$$
est appelé *image réciproque de B par f*.
Fonctions indicatrices. Soit E un ensemble et A une partie de E, on appelle

fonction indicatrice de A (parfois appelée fonction caractéristique), notée 1_A, l'application de E dans $\{0, 1\}$ définie par :

$$1_A(x) = \begin{cases} 1 & \text{si} \quad x \in A, \\ 0 & \text{si} \quad x \notin A. \end{cases},$$

Si A et B sont deux parties de E on a les propriétés élémentaires suivantes :

$$\begin{aligned} 1_A = 1_B &\Leftrightarrow A = B, \\ 1_{A\cap B} &= 1_A 1_B, \\ 1_{\bar{A}} &= 1_E - 1_A, \\ 1_{A\cup B} &= 1_A + 1_B - 1_{A\cap B}, \\ 1_{A\cup B} = 1_A + 1_B &\Leftrightarrow A \cap B = \emptyset. \end{aligned}$$

1.2.3 Univers et événements

Univers, événements

Nous allons voir dans ce qui suit que la notion d'événement est entièrement liée à la notion d'ensemble ; seul le vocabulaire fait la différence entre les notions *ensemblistes* et *probabilistes.*

Univers. On appelle *univers* ou *espace fondamental* l'ensemble Ω des issues envisagées d'une expérience aléatoire. Les éléments de Ω sont appelés *issues* ou encore *événements élémentaires.*

Evénement. On appelle événement tout sous-ensemble de Ω. L'ensemble vide $\emptyset$ désigne l'événement *impossible* alors que Ω désigne l'événement *certain.*

Exemple 1.2.1 *Nous décrivons 4 expériences et les univers associés.*

(i) Jeu de pile ou face : $\Omega =\{pile,face\}$.

(ii) Lancer de deux dés : $\Omega = \{1, \ldots, 6\}^2$.

(iii) Nombre d'utilisations d'une voiture avant panne : $\Omega = \mathbb{N}$.

(iv) Durée de bon fonctionnement d'un appareil : $\Omega = \mathbb{R}^+$.

Ci-dessous nous décrivons quelques événements.

Expérience	*événement*	*description*
(i)	*"obtenir pile"*	$\{pile\}$
(ii)	*"la somme est inférieure à 3"*	$\{(1, 1), (1, 2), (2, 1)\}$
(ii)	*"somme est égale à 5"*	$\{(4, 1), (1, 4), (2, 3), (3, 2)\}$
(iii)	*"panne après 3 utilisations"*	$\{4, 5, \ldots\}$
(iv)	*"l'appareil est tombé en panne après 8h de marche"*	$]8, +\infty[$

On notera que dans ces expériences aléatoires, l'univers est fini pour (i) et (ii), infini dénombrable pour (iii) et infini non dénombrable pour (iv).

1.2.4 Tribu

Pour une expérience aléatoire donnée, nous ne considérons pas toujours tous les événements mais seulement ceux qui sont disponibles ou intéressants, d'où la notion de *tribu*.

Définition 1.2.1 *Soit Ω l'univers associé à une expérience aléatoire. On appelle tribu ou σ-algèbre toute famille $\mathcal{F}$ de parties de Ω satisfaisant :*

i- $\Omega \in \mathcal{F}$ *;*

ii- *si $A \in \mathcal{F}$ alors $\bar{A} \in \mathcal{F}$ (stabilité par passage au complémentaire) ;*

iii- *si $(A_i)_{i \in I}$ est une famille dénombrable d'événements alors :*

$$\bigcup_{i \in I} A_i \in \mathcal{F} \quad \textit{(stabilité par réunion dénombrable)}$$

Remarque 1.2.1 *Attention ! $\mathcal{F}$ est une famille de parties de Ω. Il faut donc écrire $A \in \mathcal{F}$ et non pas $A \subset \mathcal{F}$.*

Proposition 1.2.1 *Une tribu est stable par intersection dénombrable.*

Démonstration. En exercice.

Proposition 1.2.2 *Soit $\mathcal{C}$ une famille de parties de Ω. Alors il existe une et une seule plus petite tribu (au sens de l'inclusion) contenant $\mathcal{C}$, on l'appelle la tribu engendrée par $\mathcal{C}$ et on la note $\mathcal{F}(\mathcal{C})$.*

Démonstration. Il suffit de montrer que l'intersection de toutes les tribus contenant $\mathcal{C}$ est encore une tribu ; elle est alors la plus petite au sens de l'inclusion.

Exemple 1.2.2 *$\mathcal{P}(\Omega)$ est une tribu dite exhaustive (elle contient tous les événements).*

Exemple 1.2.3 *La tribu engendrée par $\{\Omega\}$ est la tribu triviale $\{\emptyset, \Omega\}$*

Exemple 1.2.4 *Pour le jeu de pile (p) ou face (f) la tribu exhaustive est*

$$\{\emptyset, \{p\}, \{f\}, \{p, f\}\}.$$

Exemple 1.2.5 *Soit*

$$\mathcal{C} = \{[a_1, b_1] \times \cdots \times [a_p, b_p] \subset \mathbb{R}^p | -\infty < a_i < b_i < +\infty, i = 1, \ldots, p\}.$$

La tribu $\mathcal{F}(\mathcal{C})$ engendrée par $\mathcal{C}$ est appelée tribu des Boréliens, elle est notée $\mathcal{B}$. Cette tribu joue un rôle particulier dans la définition des variables aléatoires réelles (chapitre 3).

Définition 1.2.2 *Soit Ω l'univers associé à une expérience aléatoire et $\mathcal{F}$ une tribu d'événements de Ω. Alors le couple $(\Omega, \mathcal{F})$ est appelé espace probabilisable.*

1.3 Espace probabilisé - Probabilité

Définition 1.3.1 *Soit* $(\Omega, \mathcal{F})$ *un espace probabilisable. On appelle probabilité (ou mesure de probabilité) l'application* P*, définie sur* $\mathcal{F}$ *à valeurs dans* $[0, 1]$ *qui vérifie :*

i- $P(\Omega) = 1$.

ii- *pour toute famille finie ou infinie dénombrable* $(A_i)_{i\in I}$ *d'événements deux à deux incompatibles (i.e.* $A_i \cap A_j = \emptyset$ *si* $i \neq j$*) :*

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

Le triplet $(\Omega, \mathcal{F}, P)$ *est alors appelé espace probabilisé.*

Proposition 1.3.1 *Soit* $(\Omega, \mathcal{F}, P)$ *un espace probabilisé et* A *et* B *deux éléments de* $\mathcal{F}$*, alors :*

i- $P(\emptyset) = 0$.

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

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

iv- $P(A \cup B) \leq P(A) + P(B)$.

v- $(A \subset B) \Rightarrow (P(A) \leq P(B))$.

Evénement négligeable ou presque sûr. On dit que $A \in \mathcal{F}$ est négligeable (resp. presque sûr) si $P(A) = 0$ (resp. $P(A) = 1$). Attention ! un événement négligeable (resp. presque sûr) n'est pas nécessairement l'événement impossible (resp. certain).
Suite monotone d'événements. On dit d'une suite $(A_i)_{i\in\mathbb{N}^*}$ d'événements qu'elle est croissante (resp. décroissante) si pour tout $i \geq 1$ on a $A_i \subset A_{i+1}$ (resp. $A_{i+1} \subset A_i$).

Proposition 1.3.2 *Soit* $(A_i)_{i\in\mathbb{N}^*}$ *une suite croissante (resp. décroissante) d'événements, alors :*

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \lim_{n\to+\infty} P(A_n) \quad \left(\textit{resp. } P\left(\bigcap_{i=1}^{+\infty} A_i\right) = \lim_{n\to+\infty} P(A_n)\right).$$

Démonstration. Voir le document attaché 1.7.1.

1.3.1 Equiprobabilité

Parmi les mesures de probabilité une se distingue particulièrement. Il s'agit de l'équiprobabilité qui correspond au cas où Ω est fini non vide, $\mathcal{F} = \mathcal{P}(\Omega)$ et chaque issue a la même probabilité p. Ainsi on a :

$$P(\Omega) = 1 \Leftrightarrow \sum_{\omega\in\Omega} p = 1 \Leftrightarrow p = \frac{1}{|\Omega|},$$

et pour tout $A \subset \Omega$:

$$P(A) = \sum_{\omega \in A} p = \frac{|A|}{|\Omega|}.$$

Par conséquent, en situation d'équiprobabilité le calcul d'une probabilité se ramène systématiquement au calcul du rapport :

$$\frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

ce qui conduit systématiquement à dénombrer les cas favorables et les cas possibles.

Exemple 1.3.1 *Lorsqu'on jette 1 dé, l'univers Ω est constitué des éléments $\{1, \dots, 6\}$. Si A est l'événement "obtenir un nombre pair" alors $A = \{2, 4, 6\}$, par conséquent $P(A) = 3/6 = 1/2$.*

1.3.2 Dénombrement, arrangements, permutations, combinaisons

***p*-liste.** Soit E un ensemble à n éléments. On appelle *p-liste* d'éléments de E tout p-uplet $(x_1, \dots, x_p)$ composé d'éléments de E.

Exemple 1.3.2 *Lorsqu'on jette 3 dés l'univers Ω est constitué des 3-listes d'éléments de $\{1, \dots, 6\}$.*

Proposition 1.3.3 *Le nombre de p-listes d'un ensemble à n éléments est n^p ; c'est aussi le nombre d'applications allant d'un ensemble à p éléments vers un ensemble à n éléments.*

Démonstration. $n^p = n$ choix pour le premier élément de la liste $\times n$ choix pour le second $\times \cdots \times n$ choix pour le dernier.

Définition 1.3.2 *On appelle arrangement de p éléments parmi n toute p-liste d'éléments distincts ordonnés d'un ensemble à n éléments.*

Proposition 1.3.4 *Le nombre d'arrangements de p éléments parmi n est noté A_n^p et vaut :*

$$\begin{cases} \frac{n!}{(n-p)!} & \text{si} \quad 0 \leq p \leq n, \\ 0 & \text{sinon.} \end{cases}$$

Démonstration. n choix pour le premier élément de la liste $\times(n-1)$ choix pour le second $\times \cdots \times (n-(p-1))$ choix pour le dernier $= n!/(n-p)!$.

Définition 1.3.3 *Un arrangement de n éléments parmi n est appelé une permutation d'un ensemble à n éléments.*

Remarque 1.3.1 *Les permutations d'un ensemble à n éléments correspondent aux différentes façons d'ordonner les n éléments d'un ensemble.*

Proposition 1.3.5 *Le nombre de permutations d'un ensemble à n éléments est $n!$.*

Démonstration. Il y en a exactement $A_n^n = n!$.

Définition 1.3.4 *L'ensemble des p-listes non-ordonnées que l'on peut former avec un arrangement de p éléments parmi n est appelé combinaison de p éléments parmi n.*

Proposition 1.3.6 *Le nombre de combinaisons à p élé- ments d'un ensemble à n éléments est noté C_n^p et vaut :*

$$\begin{cases} \frac{n!}{p!(n-p)!} & si \quad 0 \leq p \leq n, \\ 0 \quad sinon. \end{cases}$$

Démonstration. Par définition on a : $C_n^p \times p! = A_n^p$.

Exemple 1.3.3

- *Dans une course de 20 chevaux la probabilité de gagner le tiercé dans l'ordre est de $1/A_{20}^3$.*
- *Au loto la probabilité d'avoir les 6 bons numéros est de $1/C_{49}^6$.*
- *Le nombre de façons de disposer n convives autour d'une table qui compte n places est $n!$.*

1.3.3 Probabilités géomètriques, probabilités uniformes

Nous nous sommes jusque là essentiellement contentés d'exemples où l'univers est un ensemble fini. La notion d'équiprobabilité est reservée aux univers finis, mais on peut toutefois envisager une notion équivalente pour les parties mesurables de $\mathbb{R}^p$ de la manière suivante.

Soit $\Omega \subset \mathbb{R}^p$ un ensemble tel que :

$$0 < \int_\Omega dx_1 \ldots dx_p = \text{mes}(\Omega) < +\infty.$$

Remarque 1.3.2 *Ici "mes" peut être une longueur, une surface, un volume, une intégrale curviligne, etc.*

On appelle mesure de probabilité uniforme sur Ω ou encore probabilité géométrique, la mesure de probabilité définie par :

$$\forall B \in \mathcal{B}(\mathbb{R}^p), \quad P(B) = \frac{\text{mes}(B \cap \Omega)}{\text{mes}(\Omega)}.$$

Exemple 1.3.4 *Si $\Omega = [a, b]$ avec $-\infty < a < b < +\infty$ alors la probabilité uniforme sur Ω est définie par :*

$$P(A) = \frac{longueur(A \cap [a, b])}{longueur([a, b])} = \frac{longueur(A \cap [a, b])}{b - a}.$$

Exemple 1.3.5 *Si $\Omega = C(O, R)$ (cercle de centre O et de rayon $R > 0$) et si P est la probabilité uniforme sur Ω alors par exemple , en notant A l'arc de cercle défini par les angles $0 \leq \theta_1 < \theta_2 < 2\pi$, on a :*

$$P(A) = \frac{\theta_2 - \theta_1}{2\pi}.$$

Exemple 1.3.6 *Si $\Omega = D(O, R)$ (disque de centre O et de rayon R) et si P est la probabilité uniforme sur Ω, on a par exemple :*

$$P(D(O, r)) = \begin{cases} \frac{r^2}{R^2} & si \quad r \in [0, R], \\ 1 & si \quad r > R. \end{cases}$$

Paradoxe de Bertrand. Le paradoxe de Bertrand évoqué dans l'introduction correspond à deux choix distincts de probabilités uniformes pour traiter un même problème...

1.4 Conditionnement et indépendance

1.4.1 Probabilités conditionnelle

On a affaire à une population de n individus dont $n_B > 0$ ont les yeux bleus, n_M sont myopes et $n_{B \cap M}$ sont myopes aux yeux bleus. Posons-nous cette question : on choisit un individu au hasard dans la population, l'individu choisi a les yeux bleus, quelle est la probabilité qu'il soit myope ?

Formellement. Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé, $B \in \mathcal{F}$ un événement de probabilité non nulle et $M \in \mathcal{F}$ un événement quelconque. Comment calculer la probabilité de M lorsque l'on sait que B s'est réalisé ?

Approche fréquentiste. On réalise n (choix de n individus) expériences aléatoires. On note n_B le nombre de fois où un événement B (avoir les yeux bleus) de $\mathcal{F}$ se réalise au cours des n expériences. Ce qui nous intéresse ici est le nombre de fois où $M \cap B$ (myopes aux yeux bleus) se réalise relativement au nombre de fois où B se réalise puisque l'on tient pour certitude le fait que B a lieu. Donc la probabilité qui nous intéresse est donnée par :

$$\frac{n_{M \cap B}}{n_B} = \frac{n_{M \cap B}}{n} \times \frac{n}{n_B} \approx \frac{P(M \cap B)}{P(B)}.$$

Il apparaît donc naturel de définir la **probabilité de M sachant B**, notée $P(M|B)$ ou $P_B(M)$, par $P(M \cap B)/P(B)$.

Définition 1.4.1 *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et B un événement de probabilité non nulle. Soit $A \in \mathcal{F}$ on appelle probabilité de A sachant B, notée $P(A|B)$ ou $P_B(A)$, la probabilité que A se réalise sachant que B s'est réalisé et on a :*

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

Propriétés.
- si A et B sont incompatibles alors $P(A|B) = 0$;
- si $A \subset B$ alors $P(A|B) = P(A)/P(B)$;
- si $B \subset A$ alors $P(A|B) = 1$.

Exemple 1.4.1 *On lance un dé. Il retombe ! Alors les issues possibles de cette expérience sont dans $\Omega = \{1, \dots, 6\}$ et P correspond à l'équiprobabilité. Soit B l'événement $\{$obtenir un nombre pair$\}$. On a*

$$P(\{x\}|B) = \frac{P(\{x\} \cap \{2,4,6\})}{P(\{2,4,6\})} = \frac{|\{x\} \cap \{2,4,6\}|}{|\{2,4,6\}|} = \begin{cases} 1/3 & \text{si } x \text{ est pair,} \\ 0 & \text{sinon.} \end{cases}$$

1.4.2 Probabilités totales, formule de Bayes, probabilités composées

Proposition 1.4.1 *Si $(A_n)_{n \in \mathbb{N}^*}$ est une suite d'événements deux à deux incompatibles et B non négligeable alors :*

$$P\left(\bigcup_{n \in \mathbb{N}^*} A_n | B\right) = \sum_{n \in \mathbb{N}^*} P(A_n|B).$$

Démonstration.

$$P\left(\bigcup_{n \in \mathbb{N}^*} A_n | B\right) = P\left(\left(\bigcup_{n \in \mathbb{N}^*} A_n\right) \cap B\right) /P(B)$$

$$= P\left(\bigcup_{n \in \mathbb{N}^*} (A_n \cap B)\right) /P(B) = \sum_{n \in \mathbb{N}^*} P\left(A_n \cap B\right) /P(B) = \sum_{n \in \mathbb{N}^*} P\left(A_n | B\right).$$

La troisième égalité provient du fait que $(A_n \cap B)_{n \in \mathbb{N}^*}$ est une suite d'événements deux à deux incompatibles.

Proposition 1.4.2 (formule des probabilités totales) *Soient $(\Omega, \mathcal{F}, P)$ un espace probabilisé et $(A_n)_{n\in I}$ un système complet d'événements de probabilités strictement positives. Soit $A \in \mathcal{F}$ alors*

$$P(A) = \sum_{n\in I} P(A|A_n)P(A_n).$$

Démonstration. $\Omega = \cup_{n\in I} A_n$ et les A_n sont deux à deux disjoints, donc

$$\begin{aligned} P(A) = P(A\cap\Omega) &= P\left(A \cap \left(\bigcup_{n\in I} A_n\right)\right) \\ &= P\left(\bigcup_{n\in I}(A\cap A_n)\right) = \sum_{n\in I} P(A\cap A_n) \\ &= \sum_{n\in I} P(A|A_n)P(A_n). \end{aligned}$$

Proposition 1.4.3 (formule de Bayes) *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé, $(A_n)_{n\in I}$ un système complet d'événements de probabilités strictement positives et A un événement non négligeable. Alors pour tout $n \in I$ on a*

$$P(A_n|A) = \frac{P(A|A_n)P(A_n)}{\sum_{i\in I} P(A|A_i)P(A_i)}.$$

Preuve. Par définition d'une probabilité conditionnelle on a

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

En appliquant la formule des probabilités totales à $P(A)$ on obtient la formule de Bayes.

Proposition 1.4.4 (formule des probabilités composées) *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et $A_1, \dots, A_n$ n événements de $\mathcal{F}$. On a*

$$P\left(\bigcap_{k=1}^{n} A_k\right) = P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)\dots P(A_n|A_1\cap\cdots\cap A_{n-1}),$$

si $A_1 \cap \cdots \cap A_{n-1}$ est non négligeable.

Preuve. Par définition d'une probabilité conditionnelle la formule est vraie pour $n = 2$. Supposons la formule vraie au rang $n-1$, alors en notant $A = \cap_{i=1}^{n-1} A_i$ on a par définition d'une probabilité conditionnelle :

$$P(A_n \cap A) = P(A_n|A)P(A).$$

L'hypothèse de récurrence conduit immédiatement au résultat.

1.4.3 Indépendance, formule de Poincaré

Définition 1.4.2 *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et A et B deux événements de $\mathcal{F}$. A et B sont dits indépendants en probabilité ou stochastiquement indépendants si $P(A \cap B) = P(A)P(B)$.*

Remarque 1.4.1 *La définition ci-dessus prend tout son sens lorsqu'on la réécrit sous la forme $P(A|B) = P(A)$ ou $P(B|A) = P(B)$; ce qui signifie que la réalisation d'un des deux événements n'influence en rien la probabilité de réalisation de l'autre.*

Exemple 1.4.2 *Monsieur et Madame X ont dix enfants et seulement des filles. Madame X attend un autre enfant. Quelle est la probabilité que ce soit un garçon ? Bien sûr c'est 1/2 !*

Définition 1.4.3 *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. Soit $I \subset \mathbb{N}$ et $(A_n)_{n\in I}$ une famille d'événements de $\mathcal{F}$. Les événements de $(A_n)_{n\in I}$ sont stochastiquement indépendants si*

$$\forall J \subset I \quad P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j).$$

Exemple 1.4.3 *On jette deux fois un dé non pipé. On définit les événements*

— $A = \{$*on obtient un nombre pair au premier coup*$\}$;
— $B = \{$*on obtient un nombre impair au second coup*$\}$;
— $C = \{$*on obtient deux résultats de même parité*$\}$.

Ces événements sont indépendants deux à deux mais $\{A, B, C\}$ n'est pas une famille d'événements indépendants.

Proposition 1.4.5 (formule de Poincaré) *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et $A_1, \dots, A_n$ n événements de $\mathcal{F}$. On a*

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1\le i_1<\cdots<i_k\le n} P(A_{i_1} \cap \cdots \cap A_{i_k}).$$

Démonstration. Voir le document attaché 1.7.2.

Exemple 1.4.4 *Pour $n = 3$ la formule de Poincaré s'écrit :*

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$$

$$-P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3).$$

Ce résultat est facile à admettre avec un dessin.

1.5 Exercices de cours

EC-1-1 Montrer que si le centre de la corde est choisi au hasard sur le disque, alors la probabilité cherchée est de 1/4.

EC-1-2 Donner une forme simplifiée des ensembles :

a- $(A \cup B) \cap (A \cup C)$;

b- $(A \cup B) \cap (\overline{A} \cup B)$.

EC-1-3 Représenter sur le plan l'application $f : \mathbb{R}^2 \to \{0,1\}$ définie par $f(x,y) = 1_{[0,1]}(x)1_{[0,x]}(y)$. Montrer que $f(x,y) = 1_{[0,1]}(y)1_{[y,1]}(x) = 1_T(x,y)$ où $T = \{(x,y) \in \mathbb{R}^2; 0 \leq y \leq x \leq 1\}$.

EC-1-4 Exprimer les événements suivants à l'aide des événements A, B et C et des opérations intersection, réunion et complémentaire.

a- les trois événements ont lieu ;

b- A et B ont lieu, mais pas C ;

c- au moins un des événements A, B et C a lieu ;

d- au plus un des événements A, B et C est réalisé ;

e- exactement un des événements A, B ou C est réalisé.

EC-1-5 Soit $(\Omega, \mathcal{F})$ un espace probabilisable et $(A_n)_{n\geq 1}$ une suite dénombrable d'événements de $\mathcal{F}$. Montrer que :

$$\bigcap_{n\geq 1} A_n \in \mathcal{F}.$$

Aide : Noter $B_n = \bar{A}_n$ et considérer $\cup_{n\geq 1} B_n$.

EC-1-6 Y a-t-il une contradiction entre les probabilités : $P(A) = 0,8$, $P(B) = 0,2$ et $P(A \cup \overline{B}) = 0,7$?

EC-1-7 A l'oral d'un examen, un étudiant doit répondre à 8 questions sur un total de 10.

a- Combien de choix possibles y a-t-il ?

b- Combien de choix y a-t-il s'il doit répondre aux trois premières questions ?

c- Combien de choix y a-t-il s'il doit répondre au moins à 4 des 5 premières questions ?

EC-1-8 Combien y a-t-il de pièces dans un jeu de dominos sachant que sur chaque pièce figurent deux symboles choisis parmi {blanc, 1, 2, 3, 4, 5, 6} ?

EC-1-9 Dans un stand de tir à l'arc la cible est circulaire, située à 100 mètres du pas de tir, d'un rayon de 50 cm. Sur cette cible sont dessinés des cercles concentriques de rayons 10, 20, 30 et 40 cm qui forment cinq couronnes. Un impact dans la plus petite couronne rapporte 50 points, 40 points pour la couronne immédiatement supérieure, etc. La probabilité qu'un tireur rate la cible est 10 fois l'inverse de la distance qui sépare le tireur de la cible ; un tel résultat entraîne une pénalité de 10 points. Chaque tireur tire une fois et, lorsque la

cible est atteinte, la probabilité qu'une zone donnée de la cible soit touchée est uniforme.
Quels scores peut réaliser un tireur et quelles sont les probabilités correspondantes ? Même question lorsqu'un tireur tire deux fois.

EC-1-10 On a décelé dans un élevage de moutons, une probabilité $0,3$ pour qu'un animal soit atteint par une maladie M. La probabilité qu'un mouton qui n'est pas atteint par M ait une réaction négative à un test T est $0,9$. S'il est atteint par M, la probabilité qu'il ait une réaction positive à T est $0,8$. Quelle est la probabilité qu'un mouton pris au hasard et ayant une réaction positive soit atteint par M ?

EC-1-11 On cherche un parapluie qui, avec la probabilité $p/7$, se trouve dans l'un quelconque des 7 étages d'un immeuble. On a exploré en vain les 6 premiers étages. Quelle est la probabilité que le parapluie se trouve au 7^e étage ?

EC-1-12 Soient deux événements A et B définis sur le même espace de probabilité.

a- Si A est négligeable, montrer que A et B sont indépendants ;

b- Même chose si A est presque sûr.

EC-1-13 On lance 2 dés non pipés, un dé noir et un dé blanc. Soit A l'événement "le chiffre du dé noir est pair", B l'événement "le chiffre du dé blanc est impair", C l'événement "les 2 chiffres ont même parité". Montrer que A et C, A et B, B et C sont indépendants, mais que les trois événements A , B et C ne le sont pas.

EC-1-14 Le pourcentage d'étudiants qui réussissent les u.v. A, B et C, sont : A : 50%, B : 40%, C : 30%, A et B : 35%, B et C : 20%, C et A : 20% et 15% réussissent les trois u.v. Quel est le pourcentage d'étudiants qui obtiennent au moins l'une des trois u.v ?

EC-1-15 Soient A et B deux événements indépendants. Montrer que $\overline{A}$ et B, A et $\overline{B}$, $\overline{A}$ et $\overline{B}$ sont indépendants.

1.6 Exercices de travaux dirigés

ETD-1-1 Soient A, B et C des parties d'un ensemble E.

a- simplifier : $A \cap (\overline{A} \cup B) \cap (\overline{A} \cup \overline{B} \cup C)$;

b- simplifier : $(A \cup (\overline{A} \cap B)) \cup (\overline{A} \cap \overline{B} \cap C)$;

c- montrer que : $(A \cup B \cup C) \cap (\overline{A} \cup \overline{B} \cup C) = (A \Delta B) \cup C$.

ETD-1-2 Dire, en justifiant votre réponse, si les ensembles suivants sont dénombrables ou pas.

a- Les entiers relatifs pairs ;

b- Les nombres rationnels ($\mathbb{Q}$) ;

c- Le segment $[0, 1]$ de $\mathbb{R}$.

ETD-1-3 Soit $\mathcal{F}$ une σ-algèbre de sous-ensembles de Ω et soit $B \in \mathcal{F}$. Montrer que $\mathcal{R} = \{A \cap B \ : \ A \in \mathcal{F}\}$ est une σ-algèbre de sous-ensembles de B.

ETD-1-4 Soient $(\Omega, \mathcal{F}_1)$ et $(\Omega, \mathcal{F}_2)$ deux espaces probabilisables.

a- Montrer que $(\Omega, \mathcal{F}_1 \cap \mathcal{F}_2)$ est un espace probabilisable.

b- Montrer qu'en général, $(\Omega, \mathcal{F}_1 \cup \mathcal{F}_2)$ n'est pas un espace probabilisable (prendre $\Omega = \{a, b, c\}$).

ETD-1-5 On jette trois dés. Soient A l'événement *"obtenir au moins un 1"*, B l'événement *"2 faces montrent le même résultat au moins"*, C l'événement *"la somme des faces est paire"* et D l'événement $B \cap C$.

a- Quel est l'espace fondamental ?

b- Donner l'expression d'un événement élémentaire et de sa probabilité.

c- Calculer les probabilités des événements A, B, C et D.

ETD-1-6 Le programme d'une épreuve d'examen comporte 100 sujets. Trois d'entre eux, tirés au sort, sont proposés à chaque candidat. Un candidat n'ayant étudié que le quart des sujets du programme subit l'épreuve. Quelle est la probabilité que ce candidat ait étudié :

(a) les trois sujets proposés ?

(b) deux de ces sujets ?

(c) aucun des trois ?

(d) au moins l'un des trois ?

ETD-1-7 A un examen se présentent 100 individus dont n sont des filles. On a observé que parmi les filles il y a 10% de mentions et seulement 5% chez les garçons.

i- 7% des étudiants ont eu une mention. Quelle est le nombre de filles ayant passé l'examen ?

ii- Un étudiant a eu une mention. Quelle est la probabilité que ce soit un garçon ?

ETD-1-8 Dans une entreprise pharmaceutique, la production d'une variété d'ampoules est assurée par n machines, $M_1, M_2, \ldots, M_n$ $n \geq 2$, dans des proportions respectives $\alpha_1, \alpha_2, \ldots, \alpha_n$. On sait d'autre part que les proportions respectives d'ampoules défectueuses fabriquées par ces machines sont respectivement $p_1, p_2, \ldots, p_n$. On choisit au hasard dans la production de ces n machines une ampoule et on constate qu'elle n'est pas défectueuse. Calculer la probabilité que cette ampoule ait été fabriquée par la machine M_k $(1 \leq k \leq n)$.

ETD-1-9 Montrer que le nombre de façons de distribuer r balles identiques dans n cases est égal à C^r_{n+r-1}.

ETD-1-10 Soit $\mathcal{C}$ le cercle circonscrit au carré $\mathcal{K}$ de côté de longueur $R > 0$. On choisit un point au hasard uniformément dans $\mathcal{C}$, montrer que la probabilité que ce point soit dans $\mathcal{K}$ est de $2/\pi$.

ETD-1-11 Combien faut-il mettre de raisins dans un kilogramme de pâte pour que, en mangeant une part de gâteau de 50 g, on ait au moins 99 chances sur 100 de manger du raisin ?
Indication : négliger la masse des raisins !

ETD-1-12 Des individus ne transmettent que deux informations (V et F) dont l'une est vraie (V) et l'autre fausse (F). Un individu I_1 reçoit l'information correcte qu'il transmet à un individu I_2. I_2 la transmet à son tour à un individu I_3 qui la transmet alors à I_4 et ainsi de suite. Un individu quelconque a une probabilité $p \in]0, 1[$ de transmettre correctement l'information qu'il a reçue. On note p_n la probabilité que I_n reçoive une information correcte.

a- Calculer p_1, p_2 et p_3.

b- Soit $(u_n)_{n\geq 1}$ la suite numérique définie par $u_{n+1} = au_n + b$ où $(a, b) \in (\mathbb{R} \setminus \{1\}) \times \mathbb{R}$. Démontrer que pour $n \geq 1$ on a :

$$u_{n+1} = a^n u_1 + b \left(\frac{1 - a^n}{1 - a} \right).$$

c- Établir une relation entre p_{n+1} et p_n pour $n \geq 1$. En déduire p_{n+1} en fonction de n et p.

d- Calculer $\lim_{n\to+\infty} p_n$. Que peut-on en conclure ?

ETD-1-13 Un monsieur distrait écrit n lettres différentes à n personnes distinctes et ferme les enveloppes avant d'avoir écrit les adresses, qu'il inscrit ensuite au hasard. Quelle est la probabilité qu'un destinataire au moins reçoive la lettre qui lui était destinée ?

ETD-1-14 On effectue une expérimentation sur le comportement des rats, consistant à les faire choisir entre deux portes A et B. Si le rat choisit la porte A, il reçoit une décharge électrique et la porte reste fermée ; s'il choisit la porte B, elle s'ouvre et le rat sort.
On constate expérimentalement qu'il y a deux types de rats : la probabilité

conditionnelle pour qu'un rat sorte par la porte B au k^e essai, sachant qu'il a échoué $k-1$ fois à la porte A, est $p_k = 1/2$ pour les rats du type I (rats sans mémoire) et est $q_k = k/(k+1)$ pour les rats de type II (rats avec mémoire).

a- Calculer, pour le type I et le type II, les probabilités P_n et Q_n qu'un rat sorte au n^e essai.

b- On choisit au hasard un rat dans une population contenant 60% de rats du type I et 40% de rats du type II. Calculer la probabilité conditionnelle que le rat soit du type II sachant qu'il est sorti au n^e essai, pour les valeurs suivantes $n = 1$, 2, 3 et 4.

ETD-1-15 Une école d'ingénieurs a N élèves dont n_1 élèves en première année, n_2 élèves en deuxième année et n_3 élèves en troisième année. On tire au sort deux élèves parmis N. L'un dit qu'il est plus ancien que l'autre dans l'école.

a- Quelle est la probabilité qu'il soit en deuxième année ?

b- Quelle est la probabilité qu'il soit en troisième année ?

ETD-1-16 La mercière de Saint-Martin-sur-Yvette joue avec ses boutons. Dans une boîte à ouvrage, elle a mélangé v boutons verts avec r boutons rouges. Elle en prend un au hasard, puis le remet dans la boîte avec c boutons de la même couleur. Elle recommence n fois ce processus. Montrer que la probabilité de saisir un bouton vert la n-ième fois est indépendante de n.

ETD-1-17 Dans une bibliothèque n livres sont exposés sur une étagère rectiligne et répartis au hasard. Parmi ces n livres, k sont du même auteur A, les autres étant d'auteurs tous différents. Calculer la probabilité qu'au moins p livres de A se retrouvent côte-à-côte dans les cas suivants :

a- $n = 20$, $k = 3$, $p = 3$;

b- $n = 20$, $k = 5$, $p = 2$.

ETD-1-18 Soit f une application de $(\Omega_1, \mathcal{F}_1)$ dans $(\Omega_2, \mathcal{F}_2)$ (deux espaces probabilisables). Soit $\mathcal{C} = \{A \in \mathcal{F}_2 \ : \ f^{-1}(A) \in \mathcal{F}_1\}$. Montrer que $\mathcal{C}$ est une σ-algèbre.

ETD-1-19 Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. On considère une suite $(B_n)_{n\geq 1}$ d'événements de $\mathcal{F}$ qui vérifie $B_{n+1} \subset B_n$ pour tout $n \geq 1$ (suite décroissante d'événements) et $\cap_{n\geq 1} B_n = \emptyset$.

a- On pose $A_n = B_n \cap \overline{B_{n+1}}$, $n \geq 1$. Montrer que $A_n \cap A_m = \emptyset$ si $n \neq m$.

b- Montrer que $B_m = \cup_{n\geq m} A_n$ pour tout $m \geq 1$.

c- En déduire que $\lim_{n\to+\infty} P(B_n) = 0$.

Solution. Voir la preuve de la proposition 1.3.2 en document 1.7.1.

ETD-1-20 Une expérience aléatoire consiste à jeter un dé non truqué jusqu'à l'obtention d'un as. Définir l'univers associé à cette expérience et calculer la probabilité que l'expérience se termine au n-ième coup.

ETD-1-21 Un conducteur sobre a 1 chance sur 1000 d'avoir un accident de voiture au cours d'une période. Un conducteur ivre a une chance sur 50 d'avoir un accident

de voiture au cours de la même période. On admet qu'un conducteur sur 100 conduit en état d'ivresse.
Quelle est la probabilité pour qu'il y ait un accident et que le conducteur soit ivre ?

ETD-1-22 On note S_n^p le nombre d'applications surjectives d'un ensemble à n éléments dans un ensemble à p éléments ($p \leq n$).

a- Déterminer S_n^1 et S_n^2.

b- Montrer que $\sum_{k=1}^{p} C_p^k S_n^k = p^n$.

c- Application : au pays de Lilliput, il y a trois types de comportement devant les œufs. Une personne sur trois les mange durs alors que deux sur trois les mangent à la coque. Parmi les amateurs d'œufs à la coque, il y a autant de grosboutistes qui les mangent par le gros bout, que de petitsboutistes qui les mangent par le petit bout. Considérons un échantillon de n personnes, choisies indépendamment et avec remise. Calculer la taille minimale de l'échantillon pour qu'il contienne au moins une personne de chaque type avec une probabilité supérieure à 0,95.

ETD-1-23 Un distributeur automatique de billes reçoit ses ordres sous forme d'une suite de "bits" 0 ou 1 qu'il lit de gauche à droite. Il est positionné au départ au-dessus du premier tiroir T_1 d'une suite de n tiroirs.
Tout 0 déclenche la chute d'une bille, tout 1 provoque le déplacement, sans chute de bille, de l'automate au-dessus du tiroir suivant. Les billes sont indiscernables.

I- Quel nombre N de bits 0 ou 1 doit contenir le code donné à la machine pour qu'en fin d'exécution elle soit positionnée sur le $n^{\text{ème}}$ tiroir et ait distribué exactement p billes ?

II- En supposant que tous ces codes à N bits dÃ©crits Ã la question I puissent être choisis avec la même probabilité, calculer la probabilité des événements suivants :

a- A_k : "il y a k billes dans T_1 et $p-k$ billes dans T_n" pour $k \in \{1, 2, \cdots, p-1\}$,

b- B : "tous les tiroirs sont vides sauf T_1 et T_n entre lesquels sont réparties les p billes",

c- C : "tous les tiroirs sont vides sauf deux entre lesquels sont réparties les p billes",

d- D : "il y a au moins une bille dans chacun des n tiroirs" (ceci pour $p \geq n$).

ETD-1-24 Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. Soient $A_1, \ldots, A_n$ n événements de $\mathcal{F}$, non négligeables, deux à deux incompatibles et de réunion égale à Ω.

a- Redémontrer la formule des probabilités totales : pour tout $A \in \mathcal{F}$, $P(A) = \sum_{k=1}^{n} P(A|A_k)P(A_k)$.

b- Redémontrer la formule de Bayes : si A est non négligeable et si $1 \leq j \leq n$ alors

$$P(A_j|A) = \frac{P(A|A_j)P(A_j)}{\sum_{k=1}^{n} P(A|A_k)P(A_k)}.$$

ETD-1-25 N personnes dont a hommes attendent à l'entrée d'un dispensaire. Le médecin ne peut recevoir que n d'entre elles (on suppose $n \leq a$ et $n \leq N - a$). Pour tout entier k de $\{0, 1, \cdots, n\}$ quelle est la probabilité qu'il y ait exactement k hommes parmi les n personnes reçues. En déduire

$$\sum_{k=0}^{n} C_a^k C_{N-a}^{n-k}$$

en fonction de n et N.

ETD-1-26 La `Française des Jeux` s'interroge sur les probabilités d'obtenir k bons numéros plus le complémentaire au jeu du loto (pour $k = 3$, 4 ou 5). Deux personnes X et Y sont chargées d'effectuer les calculs. Afin d'éviter les erreurs on décide que X et Y adopteront des points de vue différents :

a- X adopte le point de vue de la Française des Jeux, c'est-à-dire que 6 numéros sont fixés sur la grille et elle tire 6+1 boules sans remise dans une urne qui en contient 49.

b- Y adopte le point de vue du joueur, c'est-à-dire qu'elle considère les 7 numéros choisis par la Française des Jeux fixés et elle en choisit 6 au hasard sur la grille ;

Mener les calculs de X et Y et comparer les résultats obtenus.

Solution.

a- On tire 7 numéros distincts dans $\{1, 2, \ldots, 49\}$. Pour tenir compte du fait que le dernier numéro a une place particulière (il s'agit du complémentaire) on tient compte de l'ordre. Par conséquent l'univers Ω est l'ensemble des 7-listes d'éléments distincts et ordonnés de $\{1, 2, \ldots, 49\}$. Nous somme dans une situation d'équiprobabilité, par conséquent, pour tout $A \subset \Omega$ on a :

$$P(A) = \frac{|A|}{|\Omega|} \quad \text{où} \quad |\Omega| = A_{49}^7.$$

Soit A_k $(k = 3, 4, 5)$ l'événement *"obtenir k bons numéros plus le complémentaire"*.

Pour trouver la probabilité $P(A_k)$ il suffit de calculer $|A_k|$. On a :

$\lvert A_k\rvert =$	C_6^k	(k no. tirés dans l'urne ont été choisis par X)
$\times$	A_6^k	(placement des k bons no. dans les 6 premières places de la 7-liste)
$\times$	C_{43}^{6-k}	($6-k$ no. sont tirés parmi les 43 non choisis par X)
$\times$	A_{6-k}^{6-k}	(placement des $6-k$ no. parmi les $6-k$ places restantes des 6 premières places de la 7-liste)
$\times$	C_{6-k}^1	(le complémentaire est choisi parmi les $6-k$ no. restants)
$\times$	A_1^1	(on met le complémentaire à la 7-ième place de la 7-liste).

On trouve alors :

$$P(A_k) = \frac{(6!)^2 42! 43!}{k!(5-k)!(6-k)!(37+k)!49!}.$$

b- Dans la situation de Y on tire 6 numéros parmi 49 ; dans ce cas l'univers Ω est composé des 6-listes d'éléments de $\{1, \ldots, 49\}$ non ordonnées c'est-à-dire des combinaisons de 6 éléments parmi 49. On a donc toujours équiprobabilité avec $|\Omega| = C_{49}^6$. Pour trouver la probabilité $P(A_k)$ il suffit de calculer $|A_k|$. On a :

$\lvert A_k\rvert =$	C_6^k	(k no. choisis parmi les 6 bons no. de la Française des Jeux (F.J.))
$\times$	C_{42}^{5-k}	($6-(k+1)$ mauvais no. choisis parmi ceux non selectionnés par la F.J.)
$\times$	C_1^1	(le no. complémentaire pour lequel il n'y a qu'un choix)

On trouve alors :

$$P(A_k) = \frac{(6!)^2 42! 43!}{k!(5-k)!(6-k)!(37+k)!49!}.$$

ETD-1-27 On dispose de six boîtes dont les compositions respectives sont : B_i contient i boules rouges et $6-i$ boules noires pour $i = 1, \ldots, 6$. On note $\mathcal{R}$ l'événement : {la boule tirée est rouge}, $\mathcal{B}_i$ l'événement {on tire dans B_i} (pour le deuxième tirage on les notera $\mathcal{R}_2$ et $\mathcal{B}_{i,2}$).

I- On lance un dé non pipé puis on tire une boule dans la boîte dont l'indice est le numéro de la face obtenue.

a- Calculer la probabilité que la boule tirée soit rouge.

b- Si on a tiré une boule rouge quelle est la probabilité que ce soit dans la boîte B_6 ?

II- On remet en place la première boule tirée. Si elle était rouge on tire une boule dans B_6, sinon, on choisit au hasard l'une des cinq autres boîtes, et on y tire une boule. Calculer $P(\mathcal{R}_2|\bar{\mathcal{R}})$ puis $P(\mathcal{R}_2)$.

ETD-1-28 Une urne contient initialement b boules blanches et r boules rouges. On effectue des tirages successifs d'une boule de cette urne selon le protocole suivant : si à un rang quelconque on obtient une boule rouge, celle-ci est remise dans l'urne avant le tirage suivant et si à un rang quelconque on obtient une boule blanche, on la mange !

a- Quelle est la probabilité de tirer une boule blanche au cours des n premiers tirages ?

b- Quelle est la probabilité de manger au moins une boule blanche au cours des n premiers tirages ?

c- Sachant qu'au cours des n premiers tirages on a tiré exactement une boule blanche, quelle est la probabilité qu'elle ait été tirée en dernier ?

ETD-1-29 Soit X_n un ensemble de cardinal n. On note π_n le nombre de partitions de X_n, ce nombre s'appelle le nombre de *Bell* d'indice n. On conviendra que $\pi_0 = 1$.

a- Caculer π_1, π_2 et π_3.

b- Soit X_n un ensemble de cardinal n ($n \geq 1$) et soit a un élément fixé de X_n. Soit $P = \{A_1, \ldots, A_k\}$ une partition quelconque de X_n, on suppose que A_1 est l'unique terme de P contenant a. En raisonnant sur le cardinal de A_1 , démontrer la relation :

$$\forall n \in \mathbb{N}^*, \quad \pi_n = \sum_{p=1}^{n} C_{n-1}^{p-1} \pi_{n-p} = \sum_{k=0}^{n-1} C_{n-1}^{k} \pi_k.$$

c- Vérifier les résultats de la question a- et calculer π_6.

ETD-1-30 Soient $(\Omega, \mathcal{F}, P)$ un espace probabilisé, A et B deux événements réalisables non indépendants tels que $A \cap B$ et $A \cap \overline{B}$ sont non négligeables.

I- On pose $\alpha = P(A|\overline{B})/P(A|B)$. Montrer que

$$P(B|A) = \frac{P(B)}{\alpha + (1-\alpha)P(B)}.$$

En déduire que $P(B|A)$ est une fonction strictement croissante de $P(B)$.

II- Application : un test est destiné à déceler la présence d'une maladie dans une population.

a- Lorsque la maladie est présente, le test la révèle dans 99% des cas.

b- Lorsque la maladie est absente, le test révèle cette absence dans 99% des cas.

On convient de dire que le test est efficace si, révélant la présence de la maladie chez une personne choisie au hasard dans la population, il y a une probabilité au moins égale à 0,95 que cette personne soit effectivement atteinte de la maladie.
Déterminer à partir de quel pourcentage de personnes atteintes de la maladie dans la population le test est efficace.

ETD-1-31 Démontrer que pour deux événements A et B de $(\Omega, \mathcal{F}, P)$, on a l'équivalence entre A et B sont indépendants et $P(A \cap B)P(\overline{A} \cap \overline{B}) = P(A \cap \overline{B})P(\overline{A} \cap B)$.

ETD-1-32 Dans le repère orthonormé xOy on considère le triangle T de sommets A, B et C de coordonnées respectives $(0, 0)$, $(1, 1)$ et $(2, 0)$. Un point P est choisi de manière uniforme dans T, puis il est projeté en Q appartenant au segment $[A, B]$ suivant la direction Oy. Soit D le point de coordonnées $(x, 0)$, calculer la probabilité que Q appartienne au segment $[A, D]$. Faire un dessin.

ETD-1-33 On rappelle que $(x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k}$. Soit Ω un ensemble fini. Soit $\mathcal{P}(\Omega)$ la famille des sous-ensembles de Ω et $\mathcal{P}_k(\Omega)$ la famille des sous-ensembles de Ω à k éléments $(0 \leq k \leq \text{Card}(\Omega))$.

a- Calculer $\text{Card}(\mathcal{P}_k(\Omega))$.

b- Quelle relation lie $\mathcal{P}(\Omega)$ aux $\mathcal{P}_k(\Omega)$?

c- En déduire $\text{Card}(\mathcal{P}(\Omega)) = 2^{\text{Card}(\Omega)}$.

ETD-1-34 On considère l'épreuve : jet d'une pièce de monnaie jusqu'à ce que face apparaisse. On définit par w_i l'événement "Face apparaît au i-ème jet". On suppose que $P(w_i) = 1/2^i$.

(a) L'espace fondamental est-il fini ? infini ? dénombrable ? Justifier la réponse et donner explicitemment l'espace fondamental.

(b) Montrer que $\Sigma_i P(w_i) = 1$.

ETD-1-35 Sur n boîtes de conserves deux sont avariées. On ouvre successivement ces n boîtes au hasard. Soit k un entier naturel non nul inférieur ou égal à n. On note : A_k l'événement "aucune des k premières boîtes ouvertes n'est avariée".

(a) Evaluer (en justifiant la réponse) : $p_{n-1} = P(A_{n-1})$ et $p_n = P(A_n)$.

(b) Quelle relation satisfont entre eux les événements A_{k-1} et A_k ?

(c) Quelle est la probabilité qu'aucune des k premières boîtes ouvertes ne soit avariée ?

ETD-1-36 a- Déterminer Ω dans les cas suivants :

i- tirer une pièce trois fois.

ii- deux balles sont tirées sans remise d'une urne contenant deux balles rouges et deux balles bleues.

iii- lancer une pièce jusqu'à l'obtention de pile.

b- Montrer que si A et B sont deux événements d'une même tribu alors $A \cap B$, $A \cap \bar{B}$ et $A \Delta B$ sont aussi dans la tribu.

c- Soit A un événement, montrer que si $P(A)$=0 ou 1 alors A est indépendant de tous les autres événements.

ETD-1-37 Un voyageur arrive à un carrefour. Il sait qu'à cet endroit il va trouver deux routes : un cul de sac et la bonne route.
Il y a trois frères à ce carrefour : F_1, F_2 et F_3.
F_1 dit une fois sur dix la vérité, F_2 dit cinq fois sur dix la vérité et F_3 dit neuf

fois sur dix la vérité.
Le voyageur s'adresse "au hasard" à un et à un seul des trois frères. Il demande son chemin, et s'aperçoit par la suite que ce chemin est le bon. Quelle est la probabilité qu'il se soit adressé à F_1 ?

Solution. Soit V l' événement "dire la vérité" et F_i l' événement "le voyageur s'est adressé à F_i.
Si le voyageur a pris le bon chemin, c'est que le frère auquel il s'est adressé lui a dit la vérité, or

$$P(V|F_1) = \frac{1}{10},\ P(V|F_2) = \frac{5}{10},\ P(V|F_3) = \frac{9}{10}.$$

D'autre part le voyageur a choisi au hasard son informateur, donc :

$$P(F_1) = P(F_2) = P(F_3) = \frac{1}{3}$$

on peut donc utiliser le Théorème de Bayes :

$$P(F_1/V) = \frac{P(V|F_1)P(F_1)}{P(V|F_1)P(F_1) + P(V|F_2)P(F_2) + P(V|F_3)P(F_3)} = \frac{1}{15}$$

ETD-1-38 Dans un champ contenant n fleurs, un papillon vole au hasard de l'une à l'autre. On suppose que lorsqu'il quitte une fleur, il choisit de façon équiprobable la suivante parmi les autres.

a- Quelle est la probabilité pour qu'il revienne pour la première fois au point de départ après $r \ \leq n-1$ changements de fleur exactement ?

b- Quelle est la probabilité pour qu'après r changements de fleur, il ne se soit jamais posé deux fois sur la même ?

Solution

a- A chaque changement de fleur, le papillon a $n-1$ choix possibles. Les choix successifs sont des événements indépendants. : soit A_i l'événement "au i-ème changement de fleur, le papillon passe d'une fleur différente de la première à celle ci".

$$P(\bar{A}_i) = \frac{n-2}{n-1}\ (i > 1)$$

La probabilité cherchée est :

$$P(\bar{A}_2 \cdots \bar{A}_{r-1} A_r) = P(\bar{A}_2) \cdots P(\bar{A}_{r-1}) P(A_r) = \frac{n-2}{n-1}\frac{n-2}{n-1} \cdots \frac{n-2}{n-1}\frac{1}{n-1}$$

$$= \frac{(n-2)^{r-2}}{(n-1)^{r-1}}$$

b- nombre de cas possibles : $(n-1)^r$
Nombre de cas favorables : après s' étre posé sur la i-ème fleur, le papillon doit choisir la suivante parmi les $(n-i)$ fleurs qu'il n'a pas encore visitées. Il y a donc :

$$(n-1)(n-2)\dots(n-r) = \frac{(n-1)!}{(n-r-1)!}$$

cas favorables. C'est évidemment le nombre d'arrangements de r objets pris parmi $n-1$. D'où la probabilité cherchée :

$$\frac{(n-1)!}{(n-1)^r(n-r-1)!}.$$

ETD-1-39 Soit un dé à six faces dont deux sont peintes en rouge, deux en jaune et deux en bleu.

a- L'expérience consiste à "lancer deux fois le dé" (ou de façon équivalente "on lance deux dés qui sont identiques"). Décrire l'ensemble fondamental associé à cette expérience et donner la probabilité associée aux issues de cette expérience ?

b- On lance ce dé jusqu'à l'obtention d'une face rouge. Décrire un ensemble fondamental associé à cette expérience.
Quelle distribution (loi) de probabilités lui attribuez-vous ?
On posera $\bar{R} = \{$ face non rouge$\}$.

c- Deux personnes A et B jouent dans cet ordre avec la règle précédente (cf b). Le premier qui obtient une face rouge a gagné.
Quelles sont les probabilités de gain pour chacun des joueurs (le jeu peut durer indéfiniment) ?

Solution.

a- Pour un dé $\Omega = \{R, J, B\}$ avec $P(R) = P(J) = P(B) = 1/3$.
Pour deux dés

$$\Omega = \{(R,R),(R,J),(R,B),(J,R),(J,J),(J,B),(B,R),(B,J),(B,B)\}.$$

La probabilité de chaque couple étant 1/9.

b- Désignons par $\bar{R}$ l'événement "non rouge" : $\bar{R} = J \cup B$.

$$\Omega = \{R, \bar{R}R, \bar{R}\bar{R}R, \bar{R}\bar{R}\bar{R}R, \dots\}.$$

$P(\bar{R}\bar{R}\bar{R}\dots\bar{R}R) = (\frac{2}{3})^{n-1}\frac{1}{3}$, où n est le nombre de lancers pour obtenir la première face rouge.

c-

$$\begin{aligned} P(Agagne) &= P(R) + P(\bar{R}\bar{R}R) + P(\bar{R}\bar{R}\bar{R}\bar{R}R) + \dots \\ &= \frac{1}{3} + (\frac{2}{3})^2\frac{1}{3} + (\frac{2}{3})^4\frac{1}{3} + \dots \\ &= \sum_{k=0}^{\infty}(\frac{2}{3})^{2k}\frac{1}{3} = \frac{1}{3}\frac{1}{1-(2/3)^2}. \end{aligned}$$

$P(Bgagne) = \sum_{k=0}^{\infty}(\frac{2}{3})^{2k+1}\frac{1}{3} = \frac{2}{9}\frac{1}{1-(2/3)^2}$.

ETD-1-40 Un voyageur traversant une frontière passe en général successivement les contrôles suivants :
police du pays qu'il quitte - douane du pays où il entre - police du pays où il entre. Considèrons les événements suivants :
A : "le voyageur est contrôlé au premier poste de police"
B : "le voyageur est contrôlé à la douane "
C : "le voyageur est contrôlé au deuxième poste de police"
On suppose connues les probabilités de A, B, C, $A \cap B$, $A \cap C$, $B \cap C$ et $A \cap B \cap C$. Quelle est la probabilité que le voyageur soit contrôlé au moins une fois ?

ETD-1-41 Il y a un concours de pronostics sur la finale d'un tournoi d'escrime qui oppose un français et un américain. Parmi les participants la moitié sont des anglais, un tiers de français et un sixième des suisses.
Parmi les anglais, 60% prévoient l'américain comme gagnant de la finale et 40% prévoient le français. Chez les français, 30% prévoient l'américain et 70% prévoient le français, enfin chez les suisses 90% prévoient l'américain et 10% prévoient le français. On tire une personne au hasard parmi les participants. Elle prévoit le français gagnant. Quelle est la probabilité qu'il s'agisse d'un citoyen anglais ?

ETD-1-42 Soient A, B, C trois événements de $\mathcal{P}(\Omega)$. On suppose que chacun d'eux est non négligeable [2] et non certain [3].

i- Montrer que si A est indépendant de $B \cap C$ et de $B \cap \bar{C}$ alors A est indépendant de B.

ii- Montrer que A et B ne peuvent pas être simultanément incompatibles et indépendants.

iii- Si $\Omega = A \cup B$, A et B peuvent -ils être indépendants ?

ETD-1-43 Dans une urne il y a dix boules numérotées de 0 à 9.
On tire deux boules au hasard en même temps.

a- Décrire Ω l'espace fondamental associé à cette expérience aléatoire.

2. un événement est négligeable si sa probabilité est nulle.
3. un événement est certain si sa probabilité est égale à un.

b- Soit X la variable aléatoire donnant le plus petit des nombres portés par les deux boules. Donner $E = X(\Omega)$. Soit $x \in E$, donner $X^{-1}(x)$.

1.7 documents

1.7.1 Démonstration de la proposition 1.3.2

Espace probabilisé Soit $(A_i)_{i\in\mathbb{N}^*}$ une suite croissante d'événements, la suite $(B_i)_{i\in\mathbb{N}^*}$ définie par :

$$\begin{cases} B_1 = A_1, \\ B_n = A_n \backslash A_{n-1} \quad \text{pour} \quad n \geq 2, \end{cases}$$

est une suite d'événements deux à deux incompatibles telle que

$$\bigcup_{n\in\mathbb{N}^*} A_n = \bigcup_{n\in\mathbb{N}^*} B_n.$$

Par conséquent

$$p\left(\bigcup_{n\in\mathbb{N}^*} A_n\right) = P\left(\bigcup_{n\in\mathbb{N}^*} B_n\right) = \sum_{n=1}^{+\infty} P(B_n).$$

Or, $P(B_1) = P(A_1)$ et pour tout $n \geq 1$, $P(B_n) = P(A_n) - P(A_{n-1})$ donc

$$P(A_n) = \sum_{k=1}^{n} P(B_k),$$

$$\sum_{n=1}^{+\infty} P(B_n) = \lim_{n\to+\infty} \sum_{k=1}^{n} P(B_k) = \lim_{n\to+\infty} P(A_n).$$

Ainsi $P(\cup_{n=1}^{+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$.

Si $(A_i)_{i\in\mathbb{N}^*}$ une suite décroissante d'événements, alors la suite $(\overline{A_i})_{i\in\mathbb{N}^*}$ est une suite croissante d'événements donc $P(\cup_{n=1}^{+\infty} \overline{A_n}) = \lim_{n\to+\infty} P(\overline{A_n})$. Or,

$$P\left(\bigcap_{n=1}^{+\infty} A_n\right) = P\left(\overline{\bigcup_{n=1}^{+\infty} \overline{A_n}}\right) = 1 - P\left(\bigcup_{n=1}^{+\infty} \overline{A_n}\right),$$

d'où :

$$P\left(\bigcap_{n=1}^{+\infty} A_n\right) = 1 - \lim_{n\to+\infty} P(\overline{A_n}) = \lim_{n\to+\infty} (1 - P(\overline{A_n})) = \lim_{n\to+\infty} P(A_n).$$

1.7.2 Démonstration de la proposition 1.4.5

Indépendance stochastique En vertu de la proposition 1.3.1 la proposition est vraie pour $n = 2$. Faisons l'hypothèse de récurrence qu'elle est vraie au rang $n-1$ et posons $A = \cup_{k=1}^{n-1} A_k$. En appliquant à nouveau la proposition 1.3.1 on a

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

$$= P\left(\bigcup_{k=1}^{n-1} A_k\right) - P\left(\bigcup_{k=1}^{n-1} (A_k \cap A_n)\right) + P(A_n).$$

L'hypothèse de récurrence nous permet d'appliquer la formule de Poincaré au rang $n-1$ aux deux premières probabilités, ce qui donne

$$P(A \cup A_n) = P(A_n) +$$

$$\sum_{k=1}^{n-1} (-1)^{k+1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} (P(A_{i_1} \cap \cdots \cap A_{i_k}) - P(A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n)) \right)$$

$$= \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} P(A_{i_1} \cap \cdots \cap A_{i_k}),$$

il suffit de remarquer que

$$\{(i_1, \ldots, i_k) \; ; \; 1 \le i_1 < \cdots < i_k \le n\}$$

est réunion disjointe de

$$\{(i_1, \ldots, i_k) \; ; \; 1 \le i_1 < \cdots < i_k < n\}$$

et

$$\{(i_1, \ldots, i_{k-1}, n) \; ; \; 1 \le i_1 < \cdots < i_{k-1} < n\}.$$

Chapitre 2

Variables aléatoires discrètes

2.1 Variable aléatoire et probabilités

On considère deux jets indépendants d'un même dé. A cette expérience aléatoire on associe l'ensemble fondamental $\Omega = \Omega_1 \times \Omega_2$, constitué des couples (ω_1, ω_2) décrivant les résultats successifs des deux épreuves.

Si X réprésente la somme des deux nombres successifs, X peut prendre des valeurs variables $(2, 3, \ldots, 12)$ qui dépendent du hasard.

La relation $X = n$ est vraie pour un certain nombre d'éventualités seulement. Elle définit donc un événement à savoir la partie de Ω constituée par les couples (ω_1, ω_2) tels que $\omega_1 + \omega_2 = n$.

Par exemple, la probabilité de l'événement $\{X = 3\}$ est égale à la probabilité de l'événement $\{(1, 2), (2, 1)\}$ soit $2/36 = 1/18$.

2.1.1 Variable aléatoire

Définition 2.1.1 *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. On appelle variable aléatoire discrète, définie sur $(\Omega, \mathcal{F}, P)$ à valeurs dans un ensemble fini ou infini dénombrable E, toute application X de Ω dans E, telle que pour tout x de E on ait :*

$$X^{-1}(\{x\}) = \{w \in \Omega;\ X(w) = x\} \in \mathcal{F}.$$

Si $E = \mathbb{Z}$ (resp. $\mathbb{N}$) on dit qu'on a une variable aléatoire entière (resp. entière positive).

Remarque 2.1.1 *Si $\mathcal{B}(E)$ est une tribu de parties de E, alors on a aussi pour tout $B \in \mathcal{B}(E)$:*

$$X^{-1}(B) = \bigcup_{x \in B} X^{-1}(\{x\}) \in \mathcal{F}.$$

Les événements $X^{-1}(\{x\})$ et $X^{-1}(B)$ sont encore notés $\{X = x\}$ et $\{X \in B\}$.

2.1.2 Loi de probabilité

Proposition 2.1.1 *Si X est une variable aléatoire sur $(\Omega, \mathcal{F}, P)$, l'application p_X définie par $p_X(B) = P(\{w,\ X(w) \in B\})$ pour tout $B \in \mathcal{B}$ est une probabilité sur $(E, \mathcal{B})$, où $\mathcal{B}$ est une tribu sur E.*

Démonstration . Vérifier en utilisant la définition 1.2.3 (chapitre 1) que p_X est une probabilité sur $(E, \mathcal{B})$.

Définition 2.1.2 *Etant donné une famille $\{p_X(x);\ x \in E\}$ vérifiant $p_X(x) \geq 0$, $\sum_{x\in E} p_X(x) = 1$, on définit une probabilité sur E en posant :*

$$p_X(A) = \sum_{x \in A} p_X(x),$$

pour tout $A \in \mathcal{B}$.

Remarque 2.1.2 *Si P est une probabilité sur $(\Omega, \mathcal{F})$, la probabilité d'événements tels que $\{w \in \Omega;\ |X(w)| \geq a\}$ sera écrite indifféremment : $P(\{w \in \Omega;\ |X(w)| \geq a\})$, $P(\{|X(w)| \geq a\})$ ou $P(|X| \geq a)$.*

Définition 2.1.3 *L'application p_X définie sur $\mathcal{B}$ par :*

$$p_X(B) = P\left(X^{-1}(B)\right) = P(\{w \in \Omega;\ X(w) \in B\}),\ \forall B \in \mathcal{B},$$

est appelée mesure de probabilité image de X ou loi de probabilité de la v.a. X.

Remarque 2.1.3 *L'intérêt principal des variables aléatoires est de permettre de calculer facilement les probabilités de certains événements. Dans l'exemple introductif de ce chapitre, si on note B l'événement "obtenir une somme paire" et si la loi de X est connue alors on a*

$$p_X(B) = p_X(\{2,4,6,8,10,12\}) = \sum_{x\in\{2,4,6,8,10,12\}} p_X(x)$$

$$= \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36} = \frac{1}{2}.$$

Il eut été beaucoup plus pénible de dénombrer $\{(x,y) \in \{1,\ldots,6\}^2\ ;\ x+y \text{ est pair}\}$!

Cas particuliers :

- X est une v.a. **certaine** si et seulement si elle est constante quel que soit le résultat de l'épreuve c'est-à-dire $X = a$. Sa loi est alors définie par la relation unique $P(X = a) = 1$.
- soit $A \subset \mathcal{F}$ et $\Omega = A \cup \overline{A}$. Définissons $X = 1$, si A est réalisé, $X = 0$ sinon, alors $P(X = 1) = P(A)$ et $P(X = 0) = 1 - P(A)$. On note $X(w) = 1_A(w)$; X est alors appelée v.a. indicatrice.

2.1.3 Indépendance, fonctions de v.a.

Proposition 2.1.2 *Soit X une v.a. discrète définie sur $(\Omega, \mathcal{F}, P)$, à valeurs dans E. Si f est une application de E dans $\mathbb{R}$, alors $f(X)$ est une v.a. discrète à valeurs dans $F = f(E)$.*

Démonstration. Soit $Y = f(X)$. Y est à valeurs dans $f(E)$, qui est au plus infini dénombrable car $f(E) = \{f(x); x \in E\}$ et E est au plus infini dénombrable. De plus, si $y \in f(E)$, alors :

$$\{Y = y\} = \{f(X) = y\} = \{X \in \{x \in E; y = f(x)\}\} = \cup_{\{x\in E; y=f(x)\}}\{X = x\},$$

il s'agit d'une réunion dénombrable d'événements de $\mathcal{F}$ donc d'un élément de $\mathcal{F}$.

Définition 2.1.4 *Soient $X_1, \ldots, X_n$ des v.a. discrètes à valeurs dans $E_1, \ldots, E_n$, des ensembles dénombrables. Ces variables sont dites indépendantes si quel que soit $A_i \subset E_i$*

$$P\left(X_1 \in A_1; \ldots; X_n \in A_n\right) = P(X_1 \in A_1) \ldots P(X_n \in A_n).$$

Remarque 2.1.4 *Ceci est équivalent à (uniquement dans le cas discret !)*

$$P(X_1 = x_1; \ldots; X_n = x_n) = P(X_1 = x_1) \ldots P(X_n = x_n).$$

Ici, les "point-virgules" (;) jouent le rôle de l'intersection ($\cap$).

Proposition 2.1.3 *Toute v.a. discrète X à valeurs dans un ensemble au plus infini dénombrable $E = \{x_i; i \in I\}$, est identique en loi à la v.a. discrète :*

$$Y = \sum_{i\in I} x_i 1_{A_i}$$

où $\{A_i, i \in \mathbb{N}\}$ forme un système exhaustif de Ω, tel que $A_i = \{X = x_i\}$.

Démonstration. Il est clair que Y est à valeurs dans E. De plus, pour tout $j \in I$, $P(Y = x_j) = P(\sum_{i\in I} x_i 1_{A_i} = x_j) = P(1_{A_j} = 1) = P(X = x_i)$ car $\{A_i, i \in \mathbb{N}\}$ forme un système exhaustif de Ω.

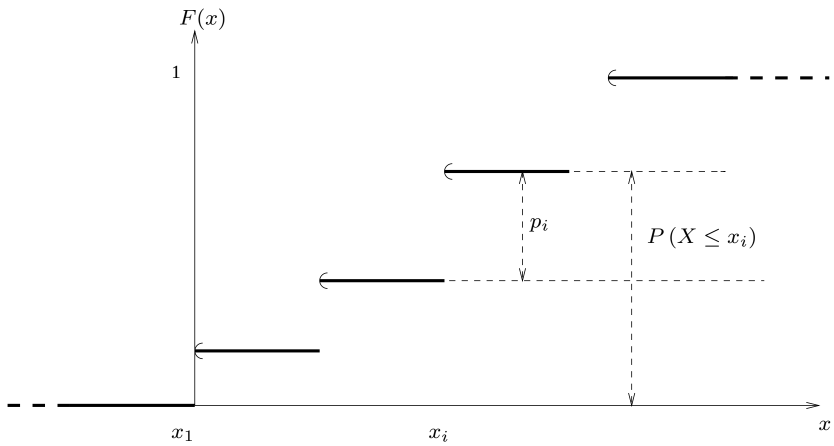
2.1.4 Fonction de répartition

Définition 2.1.5 *Soit X une v.a. de loi $p_i = P\left(X = x_i\right),\ i \in I$. On appelle fonction de répartition de la loi de X, la fonction réelle $F_X : \mathbb{R} \to [0, 1]$, définie par*

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

on peut écrire $F_X(x) = \sum_{\{i: x_i \leq x\}} p_i$.

Valeurs possibles x de X	Probabilités correspondantes
x_1	p_1
$\dots$	$\dots$
x_i	p_i
$\dots$	$\dots$
$\bigcup_{i\in I} x_i$	1



Propriétés

— La fonction de répartition est définie sur tout $\mathbb{R}$.
— $\lim_{x\to-\infty} F_X(x) = 0$; $\lim_{x\to+\infty} F_X(x) = 1$.
— Si $x_2 > x_1$, alors $F_X(x_2) \geq F_X(x_1)$.
— La fonction de répartition est continue à droite.

Remarque 2.1.5 *La donnée de la fonction de répartition permet de calculer la probabilité que la v.a. appartienne à un intervalle, c'est-à-dire :*
— $P(a \leq X \leq b) = F_X(b) - F_X(a-)$,
— $P(a < X \leq b) = F_X(b) - F_X(a)$,
— $P(a < X < b) = F_X(b-) - F_X(a)$,
— $P(X = a) = F_X(a) - F_X(a-)$,
où $F_X(a-) = \lim_{\varepsilon\searrow 0} F_X(a-\varepsilon)$.

2.1.5 Propriétés des lois

— **Loi de Bernoulli.** Une v.a. X suit une loi de Bernoulli si $E = \{0,1\}$ avec $P(X=1) = p$ et $P(X=0) = 1-p$, avec $p \in]0,1[$ et on note $X \sim B(p)$. Toute expérience aléatoire dont l'issue est soit un succès $(X=1)$, soit un échec $(X=0)$ est appelée épreuve de Bernoulli.
— **Loi Binomiale.** Lorsqu'on répète indépendamment n épreuves de Bernoulli on obtient un nombre X de succès qui s'écrit :

$$X = X_1 + \cdots + X_n,$$

où les v.a. X_i sont indépendantes et suivent la loi $B(p)$. Il est alors évident que $E = \{0, \ldots, n\}$. La loi de X est appelée loi binomiale et est notée $B(n, p)$. Elle est définie par

$$p_X(k) = P(X = k) = C_n^k p^k (1-p)^{n-k}, \quad 0 \leq k \leq n.$$

En effet,

$$\begin{aligned}\{X = k\} &= \{X_1 + \cdots + X_n = k\} \\ &= \bigcup_{1 \leq i_1 < \cdots < i_k \leq n} \{X_{i_1} = 1 \; ; \ldots ; \; X_{i_k} = 1 \; ; \text{ autres } X_i = 0\}.\end{aligned}$$

Il s'agit d'une réunion d'événements incompatibles donc

$$\begin{aligned}P(X = k) &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} P(X_{i_1} = 1 \; ; \ldots ; \; X_{i_k} = 1 \; ; \text{ autres } X_i = 0) \\ &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} p^k (1-p)^{n-k} \\ &= p^k (1-p)^{n-k} |\{(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k \; ; \; 1 \leq i_1 < \cdots < i_k \leq n\}| \\ &= C_n^k p^k (1-p)^{n-k}.\end{aligned}$$

On remarque que

$$\sum_{x \in E} p_X(x) = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

— **Loi de Poisson.** On appelle loi de Poisson de paramètre λ et on note $P(\lambda)$ la loi sur $\mathbb{N}$, définie par :

$$P(X = x) = \exp(-\lambda) \frac{\lambda^x}{x!},$$

où $x \in \mathbb{N}, \; \lambda \in \mathbb{R}^{*+}$. Notation : $X \sim P(\lambda)$

— **Loi Géométrique.** Soit X le nombre de fois qu'il faut exécuter de manière indépendante des épreuves de Bernoulli jusqu'à l'obtention du premier succès. La loi d'une telle v.a. est dite géométrique et est notée $G(p)$ si les épreuves de Bernoulli ont pour loi $B(p), \; p \in]0, 1[$.

Bien sûr, X est à valeurs dans $\mathbb{N}^*$ et

$$P(Y = k) = P(X_1 = 0; \ldots ; X_{k-1} = 0; X_k = 1) = p(1-p)^{k-1}, \quad k \in \mathbb{N}^*,$$

où les var X_i sont indépendantes et identiquement distribuées (i.i.d.) suivant la loi $B(p)$.

— **Loi Hypergéométrique.** On pose pour $m, \; n \in \mathbb{N}$ et si $m \leq n$:

$$\begin{pmatrix} n \\ m \end{pmatrix} = \frac{n!}{m!(n-m)!} = C_n^m.$$

On considère une urne contenant N boules, M blanches et $N-M$ noires. On en tire n au hasard sans replacement (i.e. remise). Soit X la v.a. nombre de boules blanches tirées, alors :

$$P(X=m)=\frac{\begin{pmatrix} M \\ m \end{pmatrix}\begin{pmatrix} N-M \\ n-m \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}},$$

pour $m=0,1,\ldots,M$. La loi hypergéométrique est notée : $H(N,M,n)$.

2.1.6 Espérance, variance, moments

Définition 2.1.6 *Soit X une v.a. à valeurs dans E. Si $\sum_{x\in E}|x|\,P\left(X=x\right)<\infty$, on appelle espérance de X et on note $\mathbb{E}\left[X\right]$ la quantité $\sum_{x\in E}xP\left(X=x\right).$*

Proposition 2.1.4 *Soient X et Y deux v.a. discrètes définies sur un espace probabilisé $(\Omega,\mathcal{F},P)$ à valeurs dans E_X et dans E_Y respectivement. On a*

1. *$\forall(a,b)\in\mathbb{R}^2$, $\mathbb{E}[aX+bY]=a\mathbb{E}[X]+b\mathbb{E}[Y]$;*
2. *Si $X\leq Y$ avec probabilité 1, alors $\mathbb{E}[X]\leq\mathbb{E}[Y]$.*

Démonstration. Voir document 2.4.1.

Théorème 2.1.1 *Soient X et Y deux v.a. discrètes définies sur un espace probabilisé $(\Omega,\mathcal{F},P)$ à valeurs dans E_X et dans E_Y respectivement. Si X et Y sont indépendantes $\mathbb{E}\left[XY\right]=\mathbb{E}\left[X\right]\mathbb{E}\left[Y\right].$*

Démonstration. Voir document 2.4.2.

Proposition 2.1.5 *Soit X une v.a. à valeurs dans E dénombrable. X a pour loi $\{p(x);x\in E\}$ si, et seulement si, pour tout $f:\mathbb{R}\to\mathbb{R}$ telle que :*

$$\sum_{x\in E}|f(x)|\,p(x)<+\infty,$$

$$\mathbb{E}\left[f(X)\right]=\sum_{x\in E}f(x)p(x). \tag{2.1}$$

Démonstration. Voir document 2.4.3.

Remarque 2.1.6 *L'intégrale de Riemman-Stieltjès introduite en 2.4.4 permet d'adopter la notation $\int_{-\infty}^{+\infty}f(x)dF(x)$ pour la quantité définie en (2.1).*

Remarque 2.1.7 *Lorsque $f(x)=x$ et $\sum_{x\in E}|x|\,p(x)<\infty$ alors $\mathbb{E}\left[X\right]=\sum_{x\in E}xp(x)$. Plus généralement, on définit pour une telle v.a. le moment d'ordre n comme $\mathbb{E}\left[X^n\right]$ si $\mathbb{E}\left[|X^n|\right]<\infty$ et on a $\mathbb{E}\left[X^n\right]=\sum_{x\in E}x^np(x)$.*

Définition 2.1.7 *Soit X une v.a. discrète telle que $\mathbb{E}\left[X^2\right] < \infty$, on appelle variance de X, la quantité :*

$$\operatorname{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sum_{x \in E} (x-m)^2 p(x).$$

De plus, la quantité $\sigma(X) = \sqrt{Var(X)}$ est appelée écart-type de X.

Définition 2.1.8 *On appelle moment centré d'ordre $r \in \mathbb{N}^*$ de X le nombre réel, s'il existe, noté μ_r et défini par :*

$$\mu_r = \mathbb{E}[(X - \mathbb{E}(X))^r] = \sum_{x \in E} (x - \mathbb{E}[X])^r p(x).$$

Propriétés.

- $\operatorname{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
- $\operatorname{Var}(aX) = a^2 \operatorname{Var}(X)$, $a \in \mathbb{R}$.
- $\operatorname{Var}(X + b) = \operatorname{Var}(X)$, $b \in \mathbb{R}$.
- Si $X_1, \ldots, X_n$ sont indépendantes alors

$$\operatorname{Var}(X_1 + \cdots + X_n) = \operatorname{Var}(X_1) + \cdots + \operatorname{Var}(X_n).$$

TABLE 2.1 – caractéristiques des lois usuelles.

Lois	Espérance	Variance
Bernoulli $B(p)$	p	$p(1-p)$
Binomiale $B(n,p)$	np	$np(1-p)$
Poisson $P(\lambda)$	λ	λ
Géométrique $G(p)$	$1/p$	$(1-p)/p^2$

2.1.7 Fonctions génératrices, moments

Définition 2.1.9 *Soit X une v.a. à valeurs dans $\mathbb{N}$ (entière) et soit $p_n = P(X = n)$. La fonction définie par*

$$g(u) = \mathbb{E}\left[u^X\right] = \sum_{n \in \mathbb{N}} p_n u^n,$$

$(-1 \leq u \leq 1)$ s'appelle la fonction génératrice de la v.a. X.

Remarque 2.1.8 *$g(u)$ est indéfiniment dérivable sur $]-1, 1[$ et $g^{(n)}(0) = p_n n!$, ceci montre que la fonction génératrice détermine la loi de X.*

Remarque 2.1.9 *Si* $0 \leq u < 1$, *on a :*

$$g'(u) = \mathbb{E}\left[Xu^{X-1}\right] = \sum_{n=1}^{+\infty} np_n u^{n-1},$$

mais si $u \uparrow 1$, $g'(u)$ *converge soit vers une limite finie* $g'(1)$, *soit diverge vers* $+\infty$ *et on convient alors que* $g'(1) = +\infty$. *Comme par ailleurs* $Xu^{X-1} \uparrow X$, *on a* $\mathbb{E}[X] = g'(1) = \lim_{u\uparrow 1} g'(u)$ *(finie ou infinie).*

De même on établit que si $\mathbb{E}[X] < \infty$, $\mathbb{E}[X(X-1)] = g''(1) = \lim_{u\uparrow 1} g''(u)$ *finie ou infinie et on en déduit que* $\mathrm{Var}(X) = g''(1) + g'(1) - (g'(1))^2$.

Proposition 2.1.6 *La fonction génératrice de la somme de deux variables entières indépendantes est le produit des fonctions génératrices.*

Démonstration.

$$g_{X+Y}(u) = \mathbb{E}[u^{X+Y}] = \mathbb{E}[u^X u^Y] = \mathbb{E}[u^X]\mathbb{E}[u^Y] = g_X(u)g_Y(u),$$

en vertu du fait que X et Y sont indépendantes.

TABLE 2.2 – Fonctions génératrices des lois usuelles.

Lois	Fonction génératrice
Bernoulli $B(p)$	$1 - p + pu$
Binomiale $B(n, p)$	$(1 - p + pu)^n$
Poisson $P(\lambda)$	$\exp(-\lambda + \lambda u)$
Géométrique $G(p)$	$\dfrac{pu}{1 - (1-p)u}$

2.1.8 Convolution, somme de v.a.

Définition 2.1.10 *On appelle convolution des suites numériques* $(a_n)_{n\geq 0}$ *et* $(b_n)_{n\geq 0}$ *la suite numérique* $(c_n)_{n\geq 0}$, *notée* $(a_n * b_n)$ *dont le terme général est défini par :*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Proposition 2.1.7 *Soient* X *et* Y *deux v.a. entières indépendantes de lois respectives* p_X *et* p_Y. *Alors, la loi de la somme est notée* p_{X+Y} *et vaut* $p_X * p_Y$.

Démonstration. Il est clair que $E_{X+Y} = \mathbb{N}$ et que p_{X+Y} est définie (pour $n \in \mathbb{N}$), par :

$$p_{X+Y}(n) = P(X+Y=n)$$
$$P(X+Y=n\ ;\ \bigcup_{k=0}^{n}\{Y=k\})$$
$$= \sum_{k=0}^{n} P(X=n-k\ ;\ Y=k) = \sum_{k=0}^{n} p_X(n-k)p_Y(k),$$

d'où le résultat.

2.1.9 Inégalité

Renvois | $EC2-17$

Proposition 2.1.8 (inégalité de Bienaymé-Čebyčev) *Soit X une v.a. discrète admettant une espérance mathématique $\mathbb{E}(X) = m$ et une variance $\mathrm{Var}(X) = \sigma^2$. Alors pour tout $\varepsilon > 0$, on a :*

$$P(|X-\mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Démonstration.

$$\begin{aligned} P(|X-m| \geq \varepsilon) &= \sum_{\{x\in E;\ |x-m|\geq\varepsilon\}} p_X(x) \\ &\leq \sum_{\{x\in E;\ |x-m|\geq\varepsilon\}} \left(\frac{|x-m|}{\varepsilon}\right)^2 p_X(x) \\ &\leq \frac{1}{\varepsilon^2}\sum_{x\in E} (|x-m|)^2 p_X(x) \\ &= \frac{\mathrm{Var}(X)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}. \end{aligned}$$

Interprétation. En prenant $\varepsilon = k\sqrt{\mathrm{Var}(X)} = k\sigma$ où $k > 0$ on obtient

$$P(|X-\mathbb{E}[X]| > k\sigma(X)) \leq \frac{1}{k^2}.$$

2.2 Exercices de cours

EC-2-1 Décrire les événements $\{X = 7\}$ et $\{X \geq 10\}$ et donner leurs probabilités.

EC-2-2 On lance deux dés et on s'intéresse aux résultats obtenus. Définir l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ correspondant à cette expérience aléatoire. Montrer que les applications suivantes définissent des v.a.r. discrètes.

a- $X(\omega_1, \omega_2) = \omega_1 + \omega_2$;

b- $Y(\omega_1, \omega_2) = \omega_1 - \omega_2$;

c- $Z(\omega_1, \omega_2) = (\omega_1 - \omega_2)^2$.

EC-2-3 Soit $\Omega = \{1, \ldots, 6\}$ un espace fondamental. On considère les familles d'ensembles $\mathcal{F}_1 = \mathcal{P}(\Omega)$ et $\mathcal{F}_2 = \{\emptyset, \{1,3,5\}, \{2,4,6\}, \Omega\}$.

a- Soit X l'application identité définie sur Ω. X est-elle une v.a.r. sur $(\Omega, \mathcal{F}_1)$? Sur $(\Omega, \mathcal{F}_2)$?

b- Soit Y l'application définie sur Ω par $Y(\omega) = 1$ si ω est impair et $Y(\omega) = 0$ si ω est pair. Y est-elle une v.a.r. sur $(\Omega, \mathcal{F}_1)$? Sur $(\Omega, \mathcal{F}_2)$?

EC-2-4 Donner les lois des v.a.r. définies dans les exercices 2.2 et 2.2.

EC-2-5 Les v.a. X et Y de l'exercice 2.2 sont-elles indépendantes ?

EC-2-6 Donner les f.d.r. des lois des v.a.r. définies dans les exercices 2.2 et 2.2.

EC-2-7 Soit F la fonction définie par :

$$F(x) = \begin{cases} 0 & \text{si} \quad x < -1, \\ 0,29 & \text{si} \quad -1 \leq x < 1, \\ 0,37 & \text{si} \quad 1 \leq x < 7, \\ 0,69 & \text{si} \quad 7 \leq x < 11, \\ 1 & \text{si} \quad x \geq 11. \end{cases}$$

a- Vérifier que F est bien une f.d.r.

b- Soit X la v.a. admettant F pour f.d.r. ; quelle est la loi de X ?

EC-2-8 Parmi les lois usuelles, quelle est celle qui modélise le mieux :

a- Le nombre d'étudiants ayant réussi l'U.V. SY01 ?

b- Le nombre de grilles de loto jouées pour gagner pour la première fois ?

EC-2-9 Une corbeille de fruits contient 3 pêches jaunes et 5 blanches. Avant de partir en cours Pierre prend 2 pêches au hasard dans la corbeille. Quelle est la probabilité que Pierre goûte les deux types de pêches ?

EC-2-10 Retrouver les données de la table 3.1.

EC-2-11 Soit X une v.a. de loi de Poisson de paramètre $\lambda > 0$. Calculer $\mathbb{E}\left[(1+X)^{-1}\right]$.

EC-2-12 Montrer que $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2$.

EC-2-13 Le nombre X de kg de tomates récoltés dans un jardin en une semaine, est une variable aléatoire dont la loi de probabilité est la suivante :

x	0	1	2	3
$p_X(x)$	0,1	0,5	0,3	0,1

a- Quelle est l'espérance mathématique de X et quelle est sa variance ?

b- Pendant les six semaines de la saison de récolte, la distribution de probabilité ne varie pas. Calculer l'espérance mathématique et la variance de la variable aléatoire Y donnant *la récolte totale en kg durant les six semaines.*

EC-2-14 Démontrer que si X est une v.a.r. entière admettant un moment d'ordre 2, alors :

a- $\mathbb{E}(X) = g'_X(1)$;

b- $\mathrm{Var}(X) = g''_X(1) + g'_X(1) - (g'_X(1))^2$.

EC-2-15 Calculer les fonctions génératrices données dans la table 2.2 et retrouver les moyennes et variances des lois considérées.

EC-2-16 Montrer que si X et Y sont deux v.a. indépendantes de lois respectives $B(n, p)$ et $B(m, p)$ alors $X + Y \sim B(n + m, p)$. Utiliser les fonctions génératrices, puis la convolution.

EC-2-17 Soit $X \sim P(1)$. Calculer $P(|X - 1| > 2)$ de manière exacte puis en donner une majoration.

2.3 Exercices de travaux dirigés

ETD-2-1 Soit X une variable aléatoire à valeurs dans $\{-2, -1, 0, 1, 2\}$, telle que $p_X(-2) = p_1$, $p_X(-1) = p_2$, $p_X(0) = p_3$, $p_X(1) = p_4$ et $p_X(2) = p_5$.

a- Que doivent vérifier p_1, p_2, p_3, p_4 et p_5 ?

b- On suppose pour la suite que $p_1 = p_2 = p_3 = p_4 = p_5$. Calculer l'espérance et la variance de X.

c- Soit $Y = X^2$. Y est-elle une variable aléatoire discrète ? Si oui, donner sa loi.

d- Tracer F_X et F_Y (si elle existe).

e- Soit Z une v.a. de Fdr $F_Z(x) = \frac{1}{2}1_{[0,+\infty[}(x) + \frac{1}{3}1_{[1,+\infty[}(x) + \frac{1}{6}1_{[4,+\infty[}(x)$. Donner la loi de Z.

ETD-2-2 Soient X et Y deux v.a. indépendantes de lois respectives :

$$P(X = 1) = p_1 \ ; \ P(X = 2) = p_2,$$

$$P(Y = 1) = q_1 \ ; \ P(Y = 2) = q_2,$$

avec $0 < p_1 < 1$; $0 < p_2 < 1$; $0 < q_1 < 1$; $0 < q_2 < 1$, satisfaisant de plus $p_1+p_2 = 1$ et $q_1+q_2 = 1$. On pose : $Z_1 = \min(X, Y)$; $Z_2 = \max(X, Y)$; $Z_3 = Z_2 - Z_1$; $Z_4 = Z_1 + Z_2$.

a- Déterminer les lois de probabilités de Z_1, de Z_2, du couple (Z_1, Z_2) de Z_3 et de Z_4.

b- Les v.a. Z_1 et Z_2 sont-elles indépendantes ?

ETD-2-3 L'oral d'un examen comporte vingt sujets possibles. Le candidat tire trois sujets au hasard, parmi ces trois sujets il choisit le sujet qu'il désire traiter. Ce candidat a révisé seulement douze sujets. On considère la v.a. X, nombre de sujets révisés parmi les trois sujets tirés.

a- Quelle est la loi de X ?

b- Quelle est la probabilité que le candidat obtienne au moins un sujet révisé ?

ETD-2-4 Un lot contient 3% de pièces défectueuses. On prélève au hasard un échantillon de 10 pièces. Les pièces étant très nombreuses, on admet que le tirage peut être considéré comme fait au hasard et avec remise. Soit X la variable aléatoire "nombre de pièces défectueuses dans l'échantillon".

a- Quelle est la loi de X ? Que valent $\mathbb{E}[X]$ et $\mathrm{Var}[X]$?

b- Calculer $P(X = 0)$ et $P(X \geq 1)$.

ETD-2-5 On considère deux avions, un biréacteur B et un quadriréacteur Q. On suppose que tous les réacteurs de ces avions ont la même probabilité p de tomber en panne et qu'ils sont indépendants les uns des autres. Soient X et Y les v.a. indiquant respectivement le nombre de réacteurs de B et Q tombant en panne au cours d'un même vol.

a- Etablir les lois de probabilité de X et Y.

b- On estime qu'un avion ne peut achever son vol que si la moitié au moins de ses réacteurs fonctionne normalement. Soient P_B et P_Q les probabilités d'un vol réussi respectivement par un biréacteur et par un quadriréacteur. Donner P_B et P_Q en fonction de p. Indiquer selon les valeurs de p l'avion qui offre le plus de sécurité.

ETD-2-6 Notons p l'application suivante :

$$\begin{array}{rcl} \mathbb{N}^* & \to & \mathbb{R} \\ k & \mapsto & p(k) = \theta(1-\theta)^{k-1}, \end{array}$$

pour $\theta \in \mathbb{R}$. On rappelle que la somme des $(n+1)$ premiers termes d'une suite géométrique de raison $x \neq 1$, est donnée par :

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1-x}.$$

On en déduit que :

$$f(x) = \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}, \quad \text{sur} \quad]-1, 1[.$$

a- Pour quelles valeurs de θ, l'application p définit-elle une loi de probabilité ?

b- Soit X une variable aléatoire réelle d'ensemble fondamental $\mathbb{N}^*$ et de loi p.

   i- Trouver la valeur de θ pour laquelle $P(X \leq 1) = P(X > 1)$. Notons θ_0 cette valeur.

   ii- Calculer $P(X > 2 | X > 1)$ pour $\theta = \theta_0$.

   iii- On note F_X la fonction de répartition de X. Calculer $F_X(x)$ pour $x \in \mathbb{R}$ et θ quelconque.

c- La loi de X est-elle une loi de probabilité connue ?

d- Donner un exemple concret d'une variable aléatoire X qui suit cette loi.

e- En considérant la fonction f calculer l'espérance mathématique et donner (sans calcul) la variance de la variable aléatoire X.

ETD-2-7 Soit $(X_n)_{n\geq 1}$ une suite de v.a. telle que $X_n \sim B(n; p_n)$ et $\lim_{n\to+\infty} np_n = \lambda \in]0, +\infty[$. Montrer que lorsque $n \to +\infty$, la loi de X_n converge vers une loi de Poisson de paramètre λ.

ETD-2-8 Paris est interdit à la circulation pour laisser le champ libre aux voitures officielles. Entre l'Etoile et Orly, il y a 13 feux tricolores qui fonctionnent de manière indépendante. Chacun est au rouge un tiers du temps. Soit X le nombre de feux rouges qu'une escorte de motards ignore sur son passage.

a- Déterminer la loi de probabilité de X, son espérance et sa variance.

b- Le cortège qui conduit un ministre à l'aéroport est lancé à 120 km/h sur les Champs Elysées. Le chauffeur du ministre s'arrête brusquement au premier feu rouge qu'il rencontre en créant un carambolage avec les véhicules chargés de le protéger. On note Y le nombre de feux franchis avant l'accident. Déterminer la loi de Y et son espérance.

ETD-2-9 Soit X une v.a. de loi uniforme sur $\{1, \ldots, n\}$ ($P(X = k) = 1/n$ pour $1 \leq k \leq n$). Quelle est la fonction de répartition de X ? Calculer $\mathbb{E}[X]$ et $\mathrm{Var}[X]$.

ETD-2-10 Soit X de loi $P(\lambda)$ et Y de loi $P(\mu)$ indépendantes. Quelle est la loi de $X+Y$? *Indication* : utiliser la convolution et les fonctions génératrices.

ETD-2-11 Soient X et Y deux variables aléatoires réelles indépendantes à valeurs dans $\{-1, 0, 1\}$. On suppose que la variable aléatoire X satisfait :

$$P(X = -1) = p_1, \ P(X = 1) = p_2,$$

avec $p_1 \in [0, \frac{1}{2}]$ et $p_2 \in [0, \frac{1}{2}]$ et que la loi de Y est uniforme sur $\{-1, 0, 1\}$.

a- Donner les lois de X et Y.

b- Déterminer la loi de $S = X + Y$.

c- On définit la variable aléatoire Z par :

$$Z = \begin{cases} \theta_1, & \text{si} \qquad X + Y = 0, \\ \theta_2, & \text{sinon,} \end{cases}$$

où θ_1 et θ_2 sont deux entiers distincts fixés.

i- Montrer que la loi de Z ne dépend pas de p_1 et p_2.

ii- Déterminer les valeurs de θ_1 et θ_2 de sorte que $\mathbb{E}(Z) = 3$ et $\mathrm{Var}(Z) = 2$.

iii- Retrouver $\mathbb{E}(Z)$ et $\mathrm{Var}(Z)$ à l'aide de la fonction génératrice de Z.

ETD-2-12 On dispose d'une boîte contenant initialement a boules rouges et b boules blanches ($a \in \mathbb{N}^*$ et $b \in \mathbb{N}^*$). On tire une boule au hasard puis on la remet dans la boîte, ainsi qu'une autre boule de la même couleur. Puis on recommence. On note R_i l'événement {la $i^{\text{ème}}$ boule tirée est rouge}, B_i l'événement {la $i^{\text{ème}}$ boule tirée est blanche}. On considère la variable aléatoire :

$$X_i = \begin{cases} 1, & \text{si} \quad R_i, \\ 0, & \text{si} \quad B_i. \end{cases}$$

a- Quelle est la loi $\mathcal{L}$ de X_1 ? Calculer son espérance.

b- Calculer $P(R_2|R_1)$ et $P(R_2|B_1)$. Quelle est la loi de X_2 ?

c- On pose $S_n = \sum_{i=1}^{n} X_i$.

i- Dans quel ensemble $\mathcal{S}_n$ la variable aléatoire S_n prend-elle ses valeurs ? Soit $k \in \mathcal{S}_n$. Si $S_n = k$, quelle est la composition de la boîte après le $n^{\text{ème}}$ tirage ? En déduire $P(X_{n+1} = 1|S_n = k)$.

ii- Montrer que :

$$P(X_{n+1} = 1) = \sum_{k=0}^{n} \frac{a+k}{a+b+n} P(S_n = k).$$

iii- En déduire que :

$$P(X_{n+1} = 1) = \frac{a}{a+b+n} + \frac{\mathbb{E}[S_n]}{a+b+n}.$$

d- On fait l'hypothèse de récurrence :

$$\mathcal{P}_n \; : \; (X_1, \cdots, X_n \text{ suivent la loi } \mathcal{L} \text{ du a}).$$

i- Si $\mathcal{P}_n$ est vraie calculer $\mathbb{E}[S_n]$.

ii- Montrer que $\mathcal{P}_n$ est vraie pour tout n.

ETD-2-13 Une entreprise cherche à organiser sa production. Elle produit des machines spécialisées de type A toutes différentes. Elle a une probabilité 0,6 de vendre immédiatement chacune des machines de type A qu'elle produit. On suppose qu'elle en produit trois.

a- Quelle est la probabilité qu'elle vende les trois ?

b- Quelle est la probabilité qu'elle en vende au moins deux ?

c- Indiquer la loi de probabilité de la variable aléatoire X indiquant le nombre de machines vendues.

d- Quelle est l'espérance mathématique de X ? Préciser sa signification.

ETD-2-14 a- Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{N}$. Déterminer la loi de X sachant que :
$X(\Omega) = \mathbb{N}^*$ et qu'il existe $p \in]0,1[$, tel que $\forall n \in \mathbb{N}^*, \; P(X = n) = pP(X \geq n)$.

b- Soient X et Y deux v.a. indépendantes de même loi géométrique $G(a)$. Déterminer $P(X = Y)$ et $P(2Y \leq X)$.

ETD-2-15 Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ de loi : $P(X = k) = (1 - \theta)\theta^{k-1}$, $\forall k \in \mathbb{N}^*$, et $\theta \in]0,1[$.Pour $n_0 \in \mathbb{N}^*$, on pose $Y = \min\{n_0, X\}$ et $Z = \max\{n_0, X\}$.

a- Déterminer la loi de Y.

b- Déterminer la loi de Z.

Montrer que si X est une v.a. entière alors : $\mathbb{E}[X] = \sum_{k=1}^{\infty} P(X \geq k)$.

ETD-2-16 Considèrons une suite de v.a. de Bernoulli indépendantes $X_1, X_2, \cdots$ de paramètre p où $p \in]0,1[$. Soient les événements : $A_1 = \{X_1 = 1\}, \cdots, A_n = \{X_1 = 0, X_2 = 0, \cdots, X_n = 1\}$ pour $n \geq 2$ et $E = \{\text{nous n'observons aucun as}\}$.

a- Calculer la probabilité de A_n.

b- Exprimer E à l'aide des A_n. Calculer la probabilité de E.

c- Soit T_1 le premier instant où 1 apparaît dans la suite de réalisations de X_k. Déterminer la loi de T_1.

d- Soit T_2 le premier instant après T_1 où 1 apparaît de nouveau. Montrer que les v.a. $T_2 - T_1$ et T_1 sont indépendantes et de même loi.

e- Quelle est la loi de T_2 ?

ETD-2-17 I- Dans l'algorithme qui suit la fonction `ALEA` renvoie un nombre aléatoire compris entre 0 et 1 et uniformément distribué sur $[0, 1]$. De plus, les résultats des appels successifs d'`ALEA` sont indépendants. Soit $p \in]0, 1[$, on considère l'algorithme suivant :

```
X=1
Tant que (ALEA > p)
  Faire
    X = X + 1
  Fin de Faire
```

X étant le résultat obtenu en fin d'algorithme, trouver la loi de la variable aléatoire X. De quelle loi connue s'agit-il ?

II- Soit X une variable aléatoire suivant une loi géométrique de paramètre p. On rappelle que si $x \neq 1$ et $n \in \mathbb{N}$ alors :

$$\sum_{j=0}^{n} x^j = 1 + x + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

a- Déterminer g la fonction génératrice de la variable aléatoire X. En déduire $\mathbb{E}(X)$ et $\mathrm{Var}(X)$.

b- Pour $k \in \mathbb{N}$ et $n \in \mathbb{N}$, montrer que $P(X > k + n | X > n) = P(X > k)$.

c- Soient Y et Z deux variables aléatoires indépendantes de même loi géométrique de paramètre p.

i- Soit S la variable aléatoire définie par $S = Y + Z$. Lorsque $S = s$, où $s \geq 2$, quelles sont les valeurs possibles des variables aléatoires Y et Z ?

ii- Déterminer la loi de $S = Y + Z$.

iii- Calculer $P(Y = j | S = s)$ et préciser les valeurs possibles de j. Quel fait surprenant fait apparaître ce résultat ?

ETD-2-18 Soit X de loi géométrique $G(p)$. Calculer $\mathbb{E}[X]$ et $\mathrm{Var}[X]$.

ETD-2-19 Ma fille Camille achète régulièrement chez le libraire des images (1 F l'unité) du film *Le Roi Lion*. Elle colle chaque image, qui n'est pas un double d'une image précédente, dans un petit album où se trouvent 200 places numérotées. Je me demande combien va me coûter en moyenne le remplissage de l'album. On suppose que le libraire dispose en même quantité de chaque modèle d'image ; on note Y_{200} le nombre d'images à acquérir pour avoir les 200 images différentes. On note X_k ($k = 1, \dots, 200$) le nombre d'images à acquérir pour avoir k images différentes sachant qu'on a déjà $k-1$ images différentes.

a- Calculer $\mathbb{E}(Y_{200})$ en fonction de $\mathbb{E}(X_1), \dots, \mathbb{E}(X_{200})$.

b- Calculer $P(X_k = n)$ $(k = 1, \dots, 200)$; en déduire que X_k suit une loi géométrique dont on précisera le paramètre.

c- En déduire $\mathbb{E}(Y_{200})$ sachant que $\sum_{n=1}^{200} \frac{1}{n} \approx \ln 200 + \gamma$ (où γ, la constante d'Euler, est à peu près égale à $0,58$ et $\ln 200 \approx 5,30$).

ETD-2-20 Pour l'expérience aléatoire qui consiste à jeter *ad infimum* et de manière indépendante une pièce (pas forcément équilibrée), définissons les v.a. $X_1, X_2, \dots,$ telles que :

$$X_i = \begin{cases} +1, & \text{si la pièce donne pile au } i\text{-ème lancer,} \\ -1, & \text{si la pièce donne face au } i\text{-ème lancer.} \end{cases}$$

Considérons les v.a. $S_1, S_2, \dots, S_n$ $(n \geq 1)$, définies par :

$$S_k = X_1 + \cdots + X_k, \quad 1 \leq k \leq n.$$

a- Donner la loi de X_1 et calculer $\mathbb{E}[X_1]$ et $\mathrm{Var}[X_1]$.

b- Calculer $P(S_n = n)$, $\mathbb{E}[S_n]$ et $\mathrm{Var}[S_n]$. Sous quelle condition avons-nous $\mathbb{E}[S_n] < 0$?

c- Calculer $P(S_n \geq 0)$.

d- Calculer la loi de S_n.

e- Si $p = 1/2$, montrer[1] que $P(S_{2k} = 0) \sim 1/\sqrt{\pi k}$ lorsque $k \to \infty$.

Aide : utiliser la formule de Stirling : $n! \sim \sqrt{2\pi n}\; e^{-n} n^n$ lorsque $n \to \infty$.

ETD-2-21 Une banque a connu quelques difficultés de trésorerie ; son conseil d'administration décide d'organiser la gestion de manière à ce qu'il y ait 999 chances sur 1000 de toujours pouvoir faire face aux demandes de retrait des déposants (on suppose qu'il n'y a aucun mouvement de panique parmi les déposants ; les habitants de ce pays mènent une vie paisible et ne se préoccupent pas les uns des autres). La banque a 1000 clients et le dépôt de chaque client est de 1000 F. La probabilité qu'un client retire son argent un jour donné est 0,001. Dans ces conditions, combien la banque doit-elle conserver de liquidités journalières pour obéir au principe de gestion qui a été posé ?

ETD-2-22 Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ de loi : $P(X = k) = \theta(1-\theta)^{k-1}$, $\forall k \in \mathbb{N}^*$, et $\theta \in]0,1[$ et Y une variable aléatoire définie sur le même espace probabilisé $(\Omega, \mathcal{F}, P)$ que X, indépendante de X et suivant la même loi que X. On considère $Z = X + Y$.

1. La notation $u_n \sim v_n$ signifie que $u_n/v_n \to 1$ lorsque $n \to \infty$, on dit que les suites $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 1}$ sont équivalentes au voisinage de l'infini.

a- Quel est l'ensemble des valeurs prises par Z. Si Z est une variable aléatoire réelle, de quel type est-elle ?

b- Montrer que si $A \in \mathcal{F}$ et $B \in \mathcal{F}$, alors $A \cap B \in \mathcal{F}$.

c- Montrer que Z est une variable aléatoire réelle.

d- Calculer la loi de Z.

e- Donner l'espérance mathématique et la variance de Z.

f- Calculer la fonction génératrice g_Z de Z ; retrouver, à l'aide de g_Z, la moyenne et la variance de Z.

Solution.

a- $Z \in \{2, 3, \dots\}$ donc à valeurs dans un ensemble dénombrable donc discrète.

b- $\overline{A \cap B} = \overline{A} \cup \overline{B}$ or $\overline{A}$ et $\overline{B}$ sont dans $\mathcal{F}$ puisque A et B y sont. Donc comme $\mathcal{F}$ est stable par réunion dénombrable $\overline{A} \cup \overline{B}$ est dans $\mathcal{F}$ donc son complémentaire aussi, c'est-à-dire $A \cap B$.

c- Soit $k \in \{2, 3, \dots\}$ il suffit de montrer que $\{Z = k\}$ est dans $\mathcal{F}$. Or $\{Z = k\} = \{X + Y = k\} = \cup_{n=1}^{k-1} \{X = n\} \cap \{Y = k - n\}$. Maintenant, pour tout n $\{X = n\} \cap \{Y = k - n\}$ est dans $\mathcal{F}$ puisque X et Y sont des v.a.r. discrètes et d'après le **b**, donc la réunion finie est aussi dans $\mathcal{F}$ ce qui finit la preuve.

d- Nous venons de voir que pour $k \in \{2, 3, \dots\}$:

$$\{Z = k\} = \{X + Y = k\} = \cup_{n=1}^{k-1} \{X = n\} \cap \{Y = k - n\}.$$

Donc

$$P(Z = k) = \sum_{n=1}^{k-1} P(X = n) P(Y = k - n)$$

$$= \sum_{n=1}^{k-1} \theta^2 (1-\theta)^{n-1} (1-\theta)^{k-n-1} = \theta^2 \sum_{n=1}^{k-1} (1-\theta)^{k-2} = (k-1)\theta^2 (1-\theta)^{k-2}.$$

e- $\mathbb{E}(Z) = \mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) = 2/\theta$; $\mathrm{Var}(Z) = \mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 2(1-\theta)/\theta^2$.

f- Puisque X et Y sont indépendantes on a, pour $-1 < u < 1$:

$$g_Z(u) = g_X(u) g_Y(u) = \left(\frac{\theta u}{1 - (1-\theta)u} \right)^2$$

g- On rappelle que $\mathbb{E}(Z) = g'_Z(1)$. Or

$$g'_Z(u) = \frac{2\theta^2 u (1 - (1-\theta)(u - u^2))}{(1 - (1-\theta)u)^3}$$

ce qui conduit à : $g'_Z(1) = 2/\theta$. Pour la variance il suffit d'utiliser la relation $\mathrm{Var}(Z) = g''_Z(1) + g'_Z(1) - (g'_Z(1))^2$.

ETD-2-23 Considérons une machine qui au début d'une journée est soit en panne soit en état de marche. Si la machine est en panne au début de la n-ième journée, la probabilité qu'elle soit réparée et opérationnelle au début de la journée suivante est p. De façon analogue, si au début de la n-ième journée la machine est en état de bon fonctionnement, la probabilité qu'elle soit en panne au début de la journée suivante est q.

Notons X_n la variable aléatoire qui représente l'état de la machine au début de la n-ième journée ; X_n prend la valeur 1 si la machine fonctionne et 0 sinon.

a- Exprimer en fonction de p et q les probabilités suivantes : $P(X_{n+1} = 1|X_n = 0)$, $P(X_{n+1} = 0|X_n = 1)$, $P(X_{n+1} = 0|X_n = 0)$ et $P(X_{n+1} = 1|X_n = 1)$.

b- On note $\pi_n = P(X_n = 0)$ pour $n \in \mathbb{N}$. Donner une expression de π_{n+1} en fonction de p, q et π_n.

c- Supposons que $0 < p, q < 1$. Exprimer π_n en fonction de p, q et π_0 (on calculera π_1, π_2 puis on utilisera la relation : $1 + x + x^2 + \cdots + x^n = (1 - x^{n+1})/(1 - x)$ lorsque $x \neq 1$).

d- Calculer, pour $0 < p, q < 1$, $\lim_{n\to+\infty} \pi_n$. Comment interpréter ce résultat ?

ETD-2-24 A et B jouent aux dés. Une partie se déroule ainsi : A lance le dé, s'il obtient un 1 ou un 2 il gagne la partie, sinon B lance le dé et gagne la partie s'il obtient 3, 4 ou 5, sinon A lance à nouveau le dé, etc.

a- Pour $k \in \mathbb{N}^*$ calculer les probabilités des événements :

- $A_{2k-1} = \{\text{A gagne au } 2k-1\text{-ième coup}\}$;
- $B_{2k} = \{\text{B gagne au } 2k\text{-ième coup}\}$.

b- La partie pouvant continuer indéfiniment, calculer les probabilités des événements $G_A = \{\text{A gagne la partie}\}$ et $G_B = \{\text{B gagne la partie}\}$.

c- Soit X la variable aléatoire égale au nombre de coups joués. Déterminer la loi de X et son espérance mathématique $\mathbb{E}(X)$.

Solution.

(a) $A_{2k-1} = \bar{A}_1 \cap \bar{B}_2 \cap \ldots \bar{B}_{2k-2} \cap A_{2k-1}$ alors, d'après la formule des probabilités composées :

$$P(A_{2k-1}) =$$

$$P(A_{2k-1}|\bar{A}_1\cap\bar{B}_2\cap\cdots\cap\bar{B}_{2k-2})\times P(\bar{B}_{2k-2}|\bar{A}_1\cap\bar{B}_2\cap\cdots\cap\bar{A}_{2k-3})\times\cdots\times P(\bar{B}_2|\bar{A}_1)\times P(\bar{A}_1)$$

$$= \frac{2}{6} \times \left(\frac{4}{6}\right)^{k-1} \times \left(\frac{3}{6}\right)^{k-1} = \frac{1}{3^k}.$$

Remarque : les événements $A_1, B_2, A_3, \ldots$ ne sont pas indépendants ! Par exemple si A_1 est réalisé, B_2 n'est pas libre de se réalisér.

De même, on a :

$$P(B_{2k}) =$$

$$P(B_{2k}|\bar{A}_1\cap\bar{B}_2\cap\cdots\cap\bar{A}_{2k-1})\times P(\bar{A}_{2k-1}|\bar{A}_1\cap\bar{B}_2\cap\cdots\cap\bar{B}_{2k-2})\times\cdots\times P(\bar{B}_2|\bar{A}_1)\times P(\bar{A}_1)$$
$$= \frac{3}{6}\times\left(\frac{4}{6}\right)^k\times\left(\frac{3}{6}\right)^{k-1} = \frac{1}{3^k}.$$

(b) On a :
$$G_A = \bigcup_{k=1}^{+\infty} A_{2k-1}$$
or les $(A_{2k-1})_{k=1,2,\ldots}$ sont 2 à 2 incompatibles, donc :
$$P(G_A) = P\left(\bigcup_{k=1}^{+\infty} A_{2k-1}\right) = \sum_{k=1}^{+\infty} P(A_{2k-1}) = \sum_{k=1}^{+\infty}\frac{1}{3^k} = \frac{1}{3}\sum_{k=1}^{+\infty}\frac{1}{3^{k-1}}$$
$$= \frac{1}{3}\frac{1}{(1-1/3)} = \frac{1}{2}.$$
De même :
$$P(G_B) = P\left(\bigcup_{k=1}^{+\infty} B_{2k}\right) = \sum_{k=1}^{+\infty}\frac{1}{3^k} = \frac{1}{2}.$$
On a donc $P(G_A)+P(G_B) = 1$ ce qui n'est pas évident car il est possible, a priori, qu'une partie dure indéfiniment. Le résultat ci-dessus nous indique que la probabilité que cette situation ait lieu est nulle.

(c) X est le nombre de coups joués dans une partie. Donc $X \in \mathbb{N}^*$ et on a pour $n \in \mathbb{N}^*$:
$$\{X = n\} = \begin{cases} A_n & \text{si} \quad n = 2k-1, \\ B_n & \text{si} \quad n = 2k. \end{cases}$$
Il vient :
$$P(X = n) = 1/3^k \quad \text{si} \quad n = 2k-1 \quad \text{ou} \quad n = 2k,$$
Pour les puristes, on peut écrire :
$$P(X = n) = \frac{1}{3^{[\frac{n+1}{2}]}},$$
où $[\cdot]$ désigne la partie entière. Maintenant :
$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} nP(X = n) = \sum_{k=1}^{+\infty}(2k-1)(1/3)^k + \sum_{k=1}^{+\infty} 2k(1/3)^k$$
$$= 4/3\sum_{k=1}^{+\infty} k(1/3)^{k-1} - \sum_{k=1}^{+\infty}(1/3)^k$$
$$= 4/3\left(\sum_{k=1}^{+\infty} x^k\right)'_{|x=1/3} - 1/2 = 4/3\,(1/(1-x))'_{|x=1/3} - 1/2$$
$$= (4/3)\times(1-1/3)^2 - 1/2 = 5/2.$$

ETD-2-25 Soient X et Y deux variables aléatoires indépendantes et uniformément distribuées sur $\{1, 2, \dots, n\}$ ($n \in \mathbb{N}^*$).

a- Donner la fonction génératrice g_X de X.

b- Calculer $g_X^{(k)}(0)/k!$ (où $g_X^{(k)}(0)$ désigne la dérivée k-ième de la fonction g_X en 0) pour $k \in \mathbb{N}$. Quelle propriété des fonctions génératrices retrouvez-vous ?

c- On suppose que $n = 3$ et on définit $Z = X + Y$.

(i) Quel est l'ensemble E des valeurs possibles de Z.

(ii) Donner g_Z.

(iii) Trouver la loi de Z en utilisant g_Z.

d- Trouver un exemple concret d'expérience illustrant X, Y et Z (n étant à préciser).

Solution.

(a) $g_X(u) = \mathbb{E}(u^X) = \dfrac{u + u^2 + \cdots + u^n}{n}$.

(b) Pour $1 \leq k \leq n$ on a :

$$g_X^{(k)}(0) = k!/n,$$

donc $g_X^{(k)}(0)/k! = 1/n = P(X = k)$, ce qui est une propriété (vue en cours !) des fonctions génératrices.

(c) (i) $Z \in \{2, 3, 4, 5, 6\}$;

(ii) $g_Z = g_X \times g_Y$ car X et Y sont indépendantes, donc :

$$g_Z(u) = \frac{1}{9}(u + u^2 + u^3)^2 = \frac{1}{9}(u^2 + 2u^3 + 3u^4 + 2u^5 + u^6).$$

(iii) En utilisant la propriété $g_X^{(k)}(0)/k! = P(X = k)$ pour $k = 2, 3, \dots, 6$, on obtient :

$$P(Z = 2) = \frac{2/9}{2!} = 1/9 \ ; \ P(Z = 3) = \frac{2 \times 3 \times 2/9}{3!} = 2/9 \dots$$

Finalement on a :

k	2	3	4	5	6
$P(Z = k)$	1/9	2/9	3/9	2/9	1/9

(d) On lance deux dés non pipés et on note X et Y les résultats obtenus, alors la somme Z est bien la somme de deux v.a. indépendantes de loi uniforme sur $\{1, 2, 3, 4, 5, 6\}$.

ETD-2-26 Soient X_1 et X_2 deux variables aléatoires indépendantes (v.a.) de même loi de Poisson de paramètre $\lambda > 0$. Définisons les v.a. $Z = X_1 + X_2$ et $U = aX_1 + b$ où $a > 0$ et b sont des paramètres réels.

(a) Calculer explicitement la loi de la v.a.Z.

(b) Donner l'espérance et la variance de Z.

(c) Calculer l'espérance et la variance de U.

(d) Peut on déterminer $a > 0$ et b tels que Z et U aient la même loi. (Indication On pourra utiliser 2. et 3.).

ETD-2-27 Soit Y une v.a. de loi de Poisson de paramètre $\mu > 0$ et soit $S_n = X_1 + \cdots + X_n$ où les X_i sont des v.a. indépendantes et équidistribuées (i.i.d.) et suivent la loi de Poisson de paramètre $\lambda > 0$.

(a) Donner en justifiant votre réponse la loi de S_n.

(b) Pour quelle valeur du paramètre λ, Y et S_n suivent la même loi ?

(c) Comment peut-on donc représenter la loi de Poisson de paramètre μ ?

ETD-2-28 Dans une banque, un guichet automatique permet de retirer des billets de 50 Euros. Le nombre de clients utilisant l'appareil dans un intervalle de temps de 5 minutes est une variable aléatoire N telle que :

$$P(N = 0) = 0.3;\ P(N = 1) = 0.1;\ P(N = 2) = 0.4;\ P(N = 3) = 0.2.$$

a- Déterminer l'espérance de la v.a. N.

b- On suppose que les nombre des clients utilisant l'appareil sur deux périodes de 5 minutes ne se chevauchant pas sont indépendants. Soit C la variable aléatoire "nombre de clients utilisant l'appareil en une heure". C peut s'écrire :

$$C = \sum_{i=1}^{12} N_i$$

où N_i désigne le nombre de clients utilisant l'appareil au cours du i-ème intervalle de 5 minutes, lorsque l'on découpe l'heure en 12 intervalles de 5 minutes ; chaque N_i suit la même loi que N.

i- Quelles sont les valeurs possibles de C. Calculer E(C).

ii- Calculer P(C=0).

ETD-2-29 Un garage loue des voitures sans chauffeur à la journée. Toute voiture louée lui laisse un bénéfice de 50 euros pour une journée, soit 50% du tarif de location. Il dispose de 2 véhicules, mais chaque véhicule indépendamment de l'autre, peut être immobilisé pour entretien, avec une probabilité $p = 0,1$. On note Y le nombre de voitures disponibles à la location un jour donné. Le nombre X de clients désirant louer une voiture pour une journée donnée est une variable aléatoire de fonction de répartition F donnée par :

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ 0,2 & \text{si} \quad 0 \leq x < 1, \\ 0,5 & \text{si} \quad 1 \leq x < 2, \\ 0,8 & \text{si} \quad 2 \leq x < 3, \\ 1 & \text{si} \quad x \geq 3. \end{cases}$$

a- Quelle est la loi de Y ? Quelle est sa moyenne ?

b- Quelle est la loi de X ? Quelle est sa moyenne ?

c- Soit Z la variable aléatoire indiquant le nombre de clients qui louent effectivement une voiture dans ce garage un jour donné.

i- Exprimer Z en fonction de X et Y.

ii- Calculer la loi de Z (arrondir les probabilités au premier chiffre après la virgule).

iii- Calculer l'espérance de Z.

iv- Quel est le bénéfice moyen réalisé par le garage en une journée ? en une semaine (7 jours ouvrables) ?

d- Le garage ne dispose plus que d'une voiture. Quel tarif de location doit-il proposer pour réaliser le même bénéfice moyen journalier ?

ETD-2-30 Soit X une variable aléatoire de fonction génératrice g définie par :

$$g(x) = 1/4 + x/4 + x^2/4 + x^3/4.$$

a- Donner la loi de X.

b- Calculer la moyenne et la variance de X.

ETD-2-31 Une variable aléatoire à valeurs dans $\mathbb{N}$ peut-elle avoir une loi uniforme ?

ETD-2-32 Soit X et Y deux variables aléatoires à valeurs dans $\{1, 2, 3\}$ telles que $P(X = x, Y = y) = 1/9$ pour tout $(x, y) \in \{1, 2, 3\}^2$.

a- X et Y sont-elles indépendantes ?

b- Calculer la loi de $Z = X + Y$.

ETD-2-33 Soit T le domaine triangle défini par $T = \{(x, y)/0 \leq x \leq y \leq 1\}$. On choisit un point P au hasard dans T suivant une probabilité uniforme. On note X l'abscisse de P.

a- Représenter en fonction de T l'événement $\{X \leq x\}$ pour $x \in [0, 1]$.

b- En déduire la fonction de répartition de X.

2.4 Documents

2.4.1 Démonstration de la Proposition 2.1.4

Démontrons 1.

$$\mathbb{E}[aX+bY] = \sum_{(x,y)\in E_X\times E_Y} (ax+by)P(X=x\ ;\ Y=y)$$

$$= \sum_{x\in E_X}\sum_{y\in E_Y} axP(X=x\ ;\ Y=y) + \sum_{x\in E_X}\sum_{y\in E_Y} byP(X=x\ ;\ Y=y)$$

$$= a\sum_{x\in E_X} xP(X=x\ ;\ Y\in E_Y) + b\sum_{y\in E_Y} yP(X\in E_X\ ;\ Y=y)$$

$$= a\sum_{x\in E_X} xP(X=x) + b\sum_{y\in E_Y} yP(Y=y) = a\mathbb{E}[X]+b\mathbb{E}[Y]$$

Démontrons 2. Si $X\geq 0$, c'est-à-dire si $E_X\subset\mathbb{R}^+$ alors il est évident que

$$\mathbb{E}[X] = \sum_{x\in E_X} xP(X=x)\geq 0.$$

Maintenant si $X\leq Y$ alors $Z=Y-X$ est une v.a. positive, ce qui entraîne $\mathbb{E}[Z]\geq 0$, or d'après 1, on a

$$\mathbb{E}[Z]=\mathbb{E}[Y-X]=\mathbb{E}[Y]-\mathbb{E}[X]\geq 0.$$

2.4.2 Démonstration du Théorème 2.1.1

Notons $Z=XY$ alors $E_Z=\{z=xy\ ;\ (x,y)\in E_X\times E_Y\}$ donc

$$\{Z=z\} = \bigcup_{(x,y)\in E_X\times E_Y\ ;\ xy=z} \{X=x\}\cap\{Y=y\}\in\mathcal{F},$$

donc Z est une v.a. discrète. De plus

$$\mathbb{E}[Z] = \sum_{z\in E_Z} zP(Z=z) = \sum_{(x,y)\in E_X\times E_Y} xyP(X=x\ ;\ Y=y)$$

$$= \sum_{x\in E_X}\sum_{y\in E_Y} xyP(X=x)P(Y=y)$$

$$= \left(\sum_{x\in E_X} xP(X=x)\right)\left(\sum_{y\in E_Y} yP(X=y)\right) = \mathbb{E}[X]\mathbb{E}[Y].$$

2.4.3 Démonstration de la Proposition 2.1.5

D'après la proposition 3.1.5 $f(X)$ est une v.a.r. discrète. Si X admet pour loi $\{p(x); x \in E\}$, on a alors :

$$\begin{aligned}
\mathbb{E}[f(X)] &= \sum_{y \in F} y P(f(X) = y) \\
&= \sum_{y \in F} y \sum_{x : f(x) = y} P(X = x) \\
&= \sum_{y \in F} \sum_{x : f(x) = y} f(x) P(X = x) \\
& \sum_{x \in E} f(x) P(X = x),
\end{aligned}$$

où $F = f(E)$.

Réciproquement, si $y \in E$ et $f(x) = 1_{\{y\}}(x)$, on a : $\mathbb{E}[f(X)] = P(X = y) = p(y)$. Donc si 2.1 est vraie pour toute application f, la loi de X est donnée par $\{p(x); x \in E\}$.

2.4.4 Complément d'intégration : Int. de Riemann-Stieltjès

Dans ce paragraphe nous allons présenter une généralisation de l'intégrale de Riemann (vue en MT22) qui nous permettra de traiter d'une manière homogène les v.a. discrètes et absolument continues du chapitre suivant.

Soit g une fonction définie sur un intervalle $[a, b]$ de $\mathbb{R}$. On note

$$\mathcal{P}_n(a, b) = \{a = x_0 < x_1 < \cdots < x_n = b\}$$

une partition de $[a, b]$ et

$$\tau_n(a, b) = \max\{x_i - x_{i-1} \ ; \ 1 \leq i \leq n\}$$

le plus grand intervalle de la partition $\mathcal{P}_n(a, b)$.

De plus on note pour $1 \leq i \leq n$

$$m_i = \inf\{g(x) \ ; \ x \in [x_{i-1}, x_i]\},$$

et

$$M_i = \sup\{g(x) \ ; \ x \in [x_{i-1}, x_i]\}.$$

Soit F une fonction croissante sur $[a, b]$. On définit

$$\overline{S}_n = \sum_{i=1}^{n} M_i (F(x_i) - F(x_{i-1}))$$

et

$$\underline{S}_n = \sum_{i=1}^{n} m_i(F(x_i) - F(x_{i-1})).$$

Définition 2.4.1 *Si $\overline{S}_n$ et $\underline{S}_n$ admettent une même limite lorsque $\tau_n(a,b)$ tend vers 0, alors cette limite définit l'intégrale de Riemann-Stieltjès de g par rapport à F sur $[a,b]$. On la note*

$$\int_{[a,b]} g(x)dF(x) \quad \textit{ou encore} \quad \int_a^b gdF.$$

Proposition 2.4.1 (non démontrée) *Soit $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}$ et g_1 et g_2 deux fonctions définies sur $[a,b]$ admettant une intégrale de Riemann-Stieltjès par rapport à deux fonctions F_1 et F_2 croissantes sur $[a,b]$. Alors*

— *[i] $\int_a^b (\alpha g_1 + \beta g_2)dF_1 = \alpha \int_a^b g_1 dF_1 + \beta \int_a^b g_2 dF_1$;*

— *[ii] $\int_a^b g_1 d(\alpha F_1 + \beta F_2) = \alpha \int_a^b g_1 dF_1 + \beta \int_a^b g_1 dF_2$;*

— *[iii] si $a < c < b$ alors*

$$\int_a^b g_1 dF_1 = \int_a^c g_1 dF_1 + \int_c^b g_1 dF_1 \ ;$$

— *[iv] si $g_1 \leq g_2$ sur $[a,b]$ alors*

$$\int_a^b g_1 dF_1 \leq \int_a^b g_2 dF_1.$$

Définition 2.4.2 *soient g définie sur $\mathbb{R}$ et F une fonction croissante définie sur $\mathbb{R}$. On appelle intégrale de Riemann-Stieltjès généralisée de g par rapport à F, lorsqu'elle existe, la quantité*

$$\int_{\mathbb{R}} gdF = \lim_{a \to -\infty} \lim_{b \to +\infty} \int_a^b gdF.$$

Nous l'avons déjà dit, notre intention est d'utiliser de telles intégrales avec des fonctions de répartition F de variables aléatoires discrètes ou absolument continues.

Exemple. Soit F la fonction de répartition associée à une variable aléatoire prenant la valeur c avec probabilité 1. Une telle variable aléatoire a une loi dite de Dirac, notée δ_c. Alors il est clair que

$$F(x) = \begin{cases} 0 & \text{si} \quad x < c, \\ 1 & \text{si} \quad x \geq c. \end{cases}$$

Soit g une fonction continue sur $\mathbb{R}$; calculons $\int_a^b gdF$.

Si $c \notin [a,b]$ alors pour toute partition $\mathcal{P}_n(a,b)$ on a : $F(x_i) - F(x_{i-1}+) = 0$ pour $1 \leq i \leq n$, donc $\overline{S}_n = \underline{S}_n = 0$ pour tout $n \geq 1$. Par passage à la limite on a donc

$$\int_a^b gdF = 0.$$

Si $c \in [a,b]$, d'après la proposition (2.4.1) [iii] et le résultat ci-dessus on a

$$\int_a^b g dF = \int_a^c g dF \quad \text{car} \quad \int_c^d g dF = 0.$$

Soit $\mathcal{P}_n(a,b) = \{x_j = a + (c-a)j/n \; ; \; 0 \leq j \leq n\}$, comme F est constante en dehors de c on a

$$\overline{S}_n = M_n(F(c) - F(x_{n-1})) = M_n$$

et

$$\underline{S}_n = m_n(F(c) - F(x_{n-1})) = m_n.$$

Or $\lim_{\tau_n(a,b)\to 0} m_n = \lim_{x_{n-1}\to c} \inf\{g(x) \; ; \; x \in [x_{n-1}, c]\} = g(c)$ car g est continue, de même on a $\lim_{\tau_n(a,b)\to 0} M_n = g(c)$, d'où

$$\int_a^b g dF = g(c).$$

En fait il est facile de voir que pour une variable aléatoire discrète X à valeurs dans $E_X = \{x_j \in \mathbb{R} \; ; \; j \in J \subset \mathbb{N}\}$ on a

$$F_X(x) = \sum_{j \in J} p_j F(x - x_j), \quad \forall x \in \mathbb{R},$$

où pour tout $j \in J$ $p_j = P_X(x_j)$.

D'après l'exemple précédent et la proposition (2.4.1) [ii] on a

$$\int_{-\infty}^{+\infty} g dF_X = \sum_{j \in J} g(x_j) p_j.$$

Par exemple, le moment d'ordre r ($r \in \mathbb{N}$) de X est donné par

$$\mathbb{E}[X^r] = \int_{-\infty}^{+\infty} x^r dF(x) = \sum_{j \in J} x_j^r p_j.$$

Exemple. Soit F une fonction telle qu'il existe une fonction positive f avec

$$F(x) = \int_{-\infty}^{x} f(u) du.$$

Soit $(a,b) \in \mathbb{R}^2$ tels que $a < b$ et $F(b) < +\infty$. Soit g une fonction continue définie sur $[a,b]$ et $\mathcal{P}_n(a,b)$ une partition de $[a,b]$, calculons $\int_a^b g dF$.

$$\overline{S}_n = \sum_{i=1}^{n} M_i (F(x_i) - F(x_{i-1})) = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} M_i f(x) dx$$

et

$$\underline{S}_n = \sum_{i=1}^{n} m_i(F(x_i) - F(x_{i=1})) = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} m_i f(x)dx.$$

Par conséquent

$$\underline{S}_n \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} g(x)f(x)dx = \int_a^b g(x)f(x)dx \leq \overline{S}_n.$$

Donc lorsque cette intégrale de Riemann-Stieltjès existe on a

$$\int_{]a,b]} g(x)dF(x) = \int_a^b g(x)f(x)dx.$$

Chapitre 3

Variables aléatoires continues

3.1 Variable aléatoire continue, loi de probabilité

3.1.1 Introduction

Lorsque les valeurs susceptibles d'être prises par une variable aléatoire forment un intervalle (ou une réunion d'intervalles) de $\mathbb{R}$, on dit que la variable est continue, ou que le modèle est continu.

Du point de vue mathématique, il y a donc une différence essentielle avec les modèles discrets (finis ou infinis). Cette différence se manifeste par le fait que pour une v.a. discrète, la probabilité totale (égale à un) se repartit en des points bien précis, alors que pour une v.a. continue il ne peut y avoir de probabilité non nulle en un point.

3.1.2 Variables aléatoires réelles

On appelle tribu des boréliens de $\mathbb{R}$, notée $\mathcal{B}$, la tribu engendrée par les ensembles ouverts de $\mathbb{R}$. Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé associé à une expérience aléatoire.

Définition 3.1.1 *On appelle variable aléatoire réelle (v.a.r.) définie sur $(\Omega, \mathcal{F}, P)$ toute application X de Ω dans $\mathbb{R}$ qui vérifie :*

$$\forall B \ \in \mathcal{B} \quad X^{-1}(B) = \{\omega \in \Omega \ ; \ X(\omega) \in B\} \in \mathcal{F}.$$

Proposition 3.1.1 (non démontrée) *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé associé à une expérience aléatoire et X une application de Ω dans $\mathbb{R}$. Si*

$$\forall x \ \in \mathbb{R} \quad X^{-1}(]-\infty, x]) \in \mathcal{F}$$

alors X est une v.a.r.

Remarque 3.1.1 *Les variable aléatoires discrètes vues au chapitre 2 répondent à cette définition ; elles ne sont en fait qu'un cas particulier des variables aléatoires réelles.*

3.1.3 Loi de probabilité

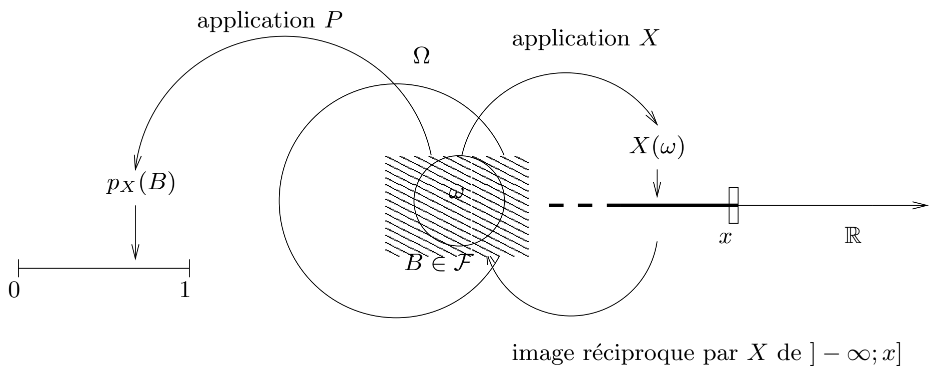
Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et X une v.a.r. définie sur Ω.

Théorème 3.1.1 *L'application p_X définie sur $\mathcal{B}$ par :*

$$p_X(B) = P(X^{-1}(B)) = P(\{\omega \in \Omega \;;\; X(\omega) \in B\}), \quad \forall B \in \mathcal{B},$$

est une mesure de probabilité sur l'espace probabilisable $(\mathbb{R}, \mathcal{B})$. Cette application est appelée mesure de probabilité image de X ou loi de probabilité de la v.a.r. X.

Démonstration. Voir document 3.5.1.



3.1.4 Caractéristiques des lois

Définition 3.1.2 *L'application F_X définie sur $\mathbb{R}$, à valeurs dans $[0,1]$, par :*

$$F_X(x) = p_X(]-\infty, x]), \quad \forall x \in \mathbb{R},$$

est appelée fonction de répartition (f.d.r.) de la variable aléatoire X.

Théorème 3.1.2 *La fonction de répartition d'une v.a.r. X est croissante, nulle en $-\infty$, égale à 1 en $+\infty$ et est continue à droite en chacun de ses points de discontinuité.*

Définition 3.1.3 *Une v.a. est dite continue si sa f.d.r. est continue.*

Proposition 3.1.2 *Le saut de F_X en un point x_0, c'est-à-dire la quantité $F(x_0) - F(x_0-) = F(x_0) - \lim_{x \nearrow x_0} F(x)$, est la probabilité que X prenne la valeur x_0 que l'on note $p_X(\{x_0\})$ ou $P(X = x_0)$ cette probabilité est nulle lorsque la variable aléatoire est continue.*

Démonstration. Voir document 3.5.2

Remarque 3.1.2 *Dans la suite de ce cours nous noterons $P\,(a < X < b)$ au lieu de $P(\{\omega \in \Omega; a < X(\omega) < b\})$.*

Théorème 3.1.3 (non démontré) *Soit F une fonction définie sur $\mathbb{R}$ à valeurs dans $[0,1]$ qui vérifie les propriétés suivantes :*
(i) F est croissante ;
(ii) $\lim_{-\infty} F = 0$ et $\lim_{+\infty} F = 1$;
(iii) F est continue à droite en chacun de ses points de discontinuité.
Alors il existe une v.a.r. admettant F pour f.d.r.

Remarque 3.1.3 *Ce dernier théorème montre qu'une v.a.r. X peut être décrite seulement par sa f.d.r. Bien souvent, l'espace fondamental sous-jacent $(\Omega, \mathcal{F}, P)$ n'interviendra pas, seule comptera la loi de probabilité p_X de X, ou, plus simplement, sa f.d.r. F_X.*

Proposition 3.1.3 *Soit X une v.a.r. de f.d.r. F, et a et b ($a < b$) des points de continuité de F alors :*

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b).$$

Remarque 3.1.4 *D'une manière générale pour une v.a.r. X de f.d.r. F et pour $a < b$ on a :*

$$F(b) - F(a) = P(X \in]-\infty, b]) - P(X \in]-\infty, a]) = P(a < X \leq b).$$

3.1.5 Densité

Définition 3.1.4 *On dit d'une v.a.r. X qu'elle est absolument continue (ou à densité) s'il existe une fonction f_X positive satisfaisant*

$$p_X(B) = \int_B f_X(x)dx$$

pour tout élément $B \in \mathcal{B}$. Cette fonction f_X, lorsqu'elle existe, est appelée densité de probabilité de X.

Proposition 3.1.4 *Toute densité de probabilité f satisfait les propriétés suivantes :*
a- $\forall x \in \mathbb{R}, \quad f(x) \geq 0$;
b- $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1.$
Réciproquement, si f satisfait les propriétés a et b ci-dessus, alors il existe une v.a.r. admettant f pour densité.

Remarque 3.1.5 *Bien sûr, lorsque $B =]-\infty,\ x]$, nous obtenons :*

$$P(X \in]-\infty,\ x]) = F_X(x) = \int_{-\infty}^{x} f_X(y)dy.$$

Si x est un point de continuité de f_X alors on a :

$$F'_X(x) = \frac{dF_X}{dx}(x) = f_X(x).$$

Remarque 3.1.6 *Une densité peut prendre des valeurs supérieures à 1 !*

3.1.6 Espérance

Définition 3.1.5 *Soit X une v.a.r. absolument continue de densité f_X. Si*

$$\int_{-\infty}^{+\infty} |x| f_X(x)dx < +\infty,$$

alors X admet une espérance $\mathbb{E}[X]$ définie par :

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)dx.$$

Proposition 3.1.5 *Soit X une v.a.r. et $\phi : \mathbb{R} \to \mathbb{R}$ une application telle que $\phi(X)$ soit une v.a.r. (voir 3.5.5). L'espérance de $\phi(X)$, lorsqu'elle existe, est égale à :*

$$\mathbb{E}[\phi(X)] = \int_{-\infty}^{+\infty} \phi(x) f_X(x)dx,$$

où f_X est la densité de X.

Remarque 3.1.7 *Une probabilité est une espérance dans le sens où pour tout $A \subset \mathbb{R}$ (borélien) :*

$$P(X \in A) = \int_A f_X(x)dx = \int_{-\infty}^{+\infty} 1_A(x) f_X(x)dx = \mathbb{E}[1_A(X)].$$

Remarque 3.1.8 *D'après la proposition 3.1.5 et la remarque précédente, lorsque ϕ est strictement croissante, on a :*

$$F_{\phi(X)}(y) = \mathbb{E}[1_{]-\infty,y]}(\phi(X))] = \int_{-\infty}^{\phi^{-1}(y)} f_X(x)dx.$$

3.1.7 Indépendance

Définition 3.1.6 *Les v.a.r. X et Y sont indépendantes si, et seulement si, pour tous $B_1, B_2 \in \mathcal{B}$:*

$$P(X \in B_1, Y \in B_2) = P(X \in B_1)P(Y \in B_2).$$

La proposition qui suit donne une condition d'indépendance plus simple à vérifier.

Proposition 3.1.6 *X et Y sont indépendantes si et seulement si :*

$$\forall (x,y) \in \mathbb{R}^2, P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y).$$

Définition 3.1.7 *Les v.a.r. $X_1, \ldots, X_n$ sont indépendantes si, et seulement si, on a :*

$$P(X_1 \in B_1, \ldots, X_n \in B_n) = P(X_1 \in B_1) \ldots P(X_n \in B_n),$$

où $B_1, \ldots, B_n \in \mathcal{B}$.

Proposition 3.1.7 *Les v.a.r. $X_1, \ldots, X_n$ sont indépendantes si, et seulement si :*

$$\mathbb{E}[h_1(X_1) \times \cdots \times h_n(X_n)] = \mathbb{E}[h_1(X_1)] \times \cdots \times \mathbb{E}[h_n(X_n)] \tag{3.1}$$

pour toutes les fonctions $h_i : \mathbb{R} \to \mathbb{R}$, telles que (3.1) ait un sens.

3.1.8 Moments, médiane

Définition 3.1.8 *Soit X une v.a.r. absolument continue de densité f_X. On appelle moment d'ordre $r \in \mathbb{N}^*$ de X la quantité, lorsqu'elle existe, définie par $\mathbb{E}[X^r]$.*

Définition 3.1.9 *On appelle moment centré d'ordre r de X, s'il existe, le réel μ_r défini par $\mu_r = \mathbb{E}[(X - \mathbb{E}[X])^r]$. μ_2 est appelé variance de X, elle est notée $\mathrm{Var}(X)$ (parfois $\sigma^2(X)$), sa racine carrée, notée $\sigma(X)$ ou σ_X est appelée écart-type de X.*

Remarque 3.1.9 *L'écart-type, comme la variance, mesure la tendance à s'écarter de la moyenne. De plus l'écart-type est une grandeur homogène 'a X.*

Propriétés. Soient X et Y deux v.a. absolument continues. On a :

a- $\forall (a,b) \in \mathbb{R}^2$, $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$;

b- Si $X \leq Y$ avec probabilité 1, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

c- $\mathrm{Var}\,(aX) = a^2 \mathrm{Var}\,(X)$, $a \in \mathbb{R}$.

d- $\mathrm{Var}\,(X + b) = \mathrm{Var}\,(X)$, $b \in \mathbb{R}$.

e- Si $X_1, \ldots, X_n$ sont des v.a.r. indépendantes alors :

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}\,(X_1) + \cdots + \mathrm{Var}\,(X_n)\,.$$

Définition 3.1.10 *On appelle médiane de la v.a.r. X tout réel M qui satisfait :*

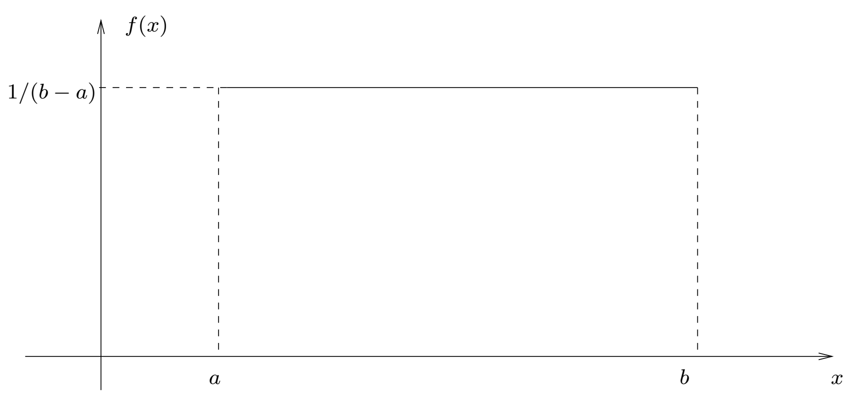
$$P(X \leq M) \geq \frac{1}{2} \quad et \quad P(X \geq M) \geq \frac{1}{2}.$$

Remarque 3.1.10 *Lorsque X admet une f.d.r. continue et strictement croissante F la médiane est unique et est solution de $F(M) = 1/2$.*

3.1.9 Lois usuelles

Loi Uniforme. Soit X une v.a.r. à valeurs dans $I = [a, b]$ $(a < b)$ un intervalle de $\mathbb{R}$, alors X suit une loi uniforme sur I si elle admet une densité f définie par :

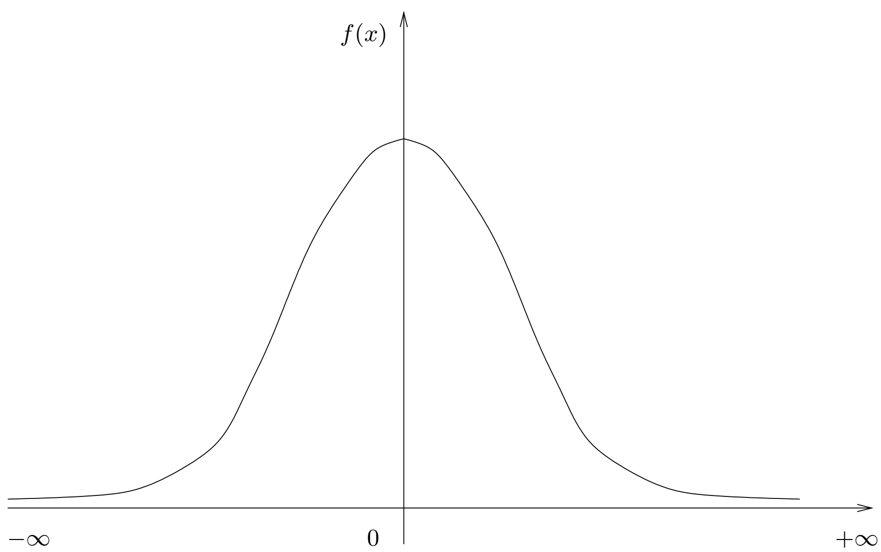
$$f_X(x) = \frac{1}{b-a} 1_{[a,b]}(x).$$



Un calcul simple montre que : $\mathbb{E}[X] = \frac{a+b}{2}$, $\mathbb{E}[X^2] = \frac{a^2+ab+b^2}{3}$ et $\mathrm{Var}(X) = \frac{(a-b)^2}{12}$.

Loi Normale ou de Gauss. Soit X une v.a.r. On dit que X suit une loi normale centrée ($\mathbb{E}[X] = 0$) réduite ($\mathrm{Var}(X) = 1$) et on note $X \sim N(0,1)$ si X admet une densité f définie par :

$$\forall\, x \in \mathbb{R}, \quad f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$



Soit Y une v.a.r. telle que $Y = \sigma X + \mu$ ($\mu \in \mathbb{R}$ et $\sigma \in]0, +\infty[$). Alors

$$\mathbb{E}[Y] = \sigma \mathbb{E}[X] + \mu = \mu \quad \text{et} \quad \mathrm{Var}[Y] = \sigma^2 \mathrm{Var}[X] = \sigma^2.$$

De plus

$$F_Y(x) = P(Y \leq x) = P\left(X \leq \frac{x-\mu}{\sigma}\right) = F_X\left(\frac{x-\mu}{\sigma}\right).$$

Donc

$$f_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

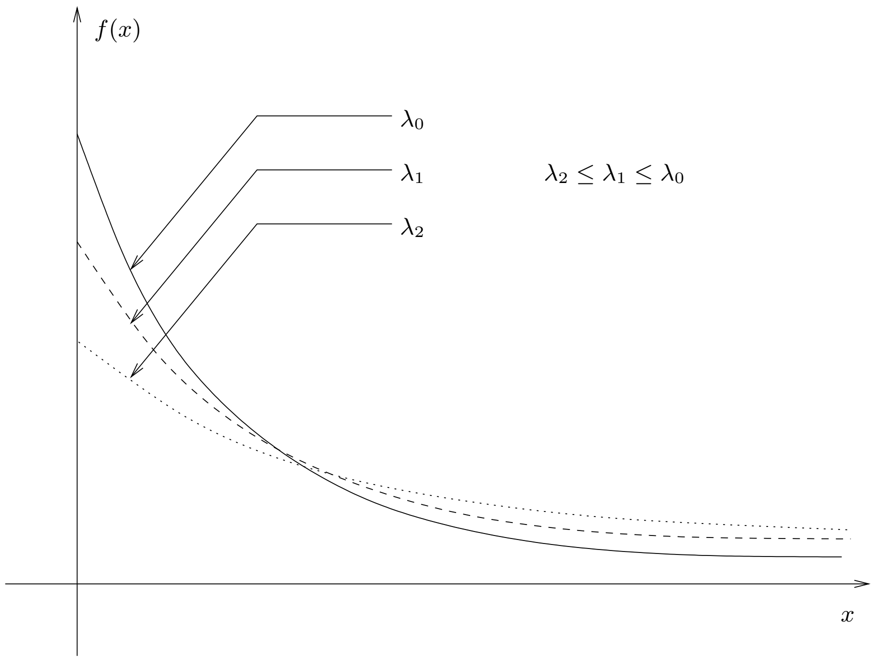
Toute v.a.r. Y admettant la densité f_Y définie ci-dessus est dite normale de moyenne μ et de variance σ^2 ; on note alors $Y \sim N(\mu, \sigma^2)$. Réciproquement si $Y \sim N(\mu, \sigma^2)$ alors $X = (Y - \mu)/\sigma \sim N(0,1)$.
Remarque : les paramètres de la loi de Gauss sont sa moyenne μ et sa variance σ^2.

Proposition 3.1.8 *Si $X \sim N(\mu_1, \sigma_1^2)$ et $Y \sim N(\mu_2, \sigma_2^2)$ sont indépendantes alors $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.*

Loi exponentielle. Soit X une v.a.r. admettant pour densité f où f est définie par

$$\forall\ x \in \mathbb{R}, \quad f_X(x) = \lambda \exp(-\lambda x) 1_{[0,+\infty[}(x).$$

Alors on dit que X suit une loi exponentielle de paramètre $\lambda > 0$, et on note $X \sim E(\lambda)$.



On a $\mathbb{E}[X] = 1/\lambda$ et $\mathrm{Var}[X] = 1/\lambda^2$ et la f.d.r. F est définie par :

$$F(x) = (1 - \exp(-\lambda x)) 1_{[0,+\infty[}(x).$$

Loi Gamma. Une v.a.r. X suit une loi gamma de paramètres $\alpha \in \mathbb{R}_+^*$ et $\lambda \in \mathbb{R}_+^*$ si elle admet une densité f définie par :

$$\forall\ x \in \mathbb{R}, \quad f(x) = \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda \exp(-\lambda x) 1_{[0,+\infty[}(x),$$

où

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} \exp(-x) dx.$$

La fonction $\Gamma(\alpha)$ satisfait pour $\alpha > 0$, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.
On note $X \sim \gamma(\alpha, \lambda)$ et on a $\mathbb{E}[X] = \alpha/\lambda$ et $\mathrm{Var}[X] = \alpha/\lambda^2$. De plus, si $\alpha \in \mathbb{N}^*$ alors $\Gamma(\alpha) = (\alpha - 1)!$; on parle dans ce cas de loi d'Erlang.

Proposition 3.1.9 *Soient $X_1, \ldots, X_n$ n v.a.r. indépendantes et de même loi $E(\lambda)$. Alors $Y = X_1 + \cdots + X_n \sim \gamma(n, \lambda)$.*

Loi du chi-deux. On dit qu'une variable aléatoire X suit une loi du χ^2 à n degrés de liberté si $X \sim \gamma(n/2, 1/2)$ et on note $X \sim \chi^2(n)$. En fait, une v.a.r. X suit une telle loi lorsqu'elle s'écrit :

$$X = X_1^2 + \cdots + X_n^2,$$

où les v.a.r. $X_1, \ldots, X_n$ sont i.i.d. de loi $N(0,1)$. On a alors : $\mathbb{E}[X] = n$ et $\mathrm{Var}(X) = 2n$.

3.1.10 Fonction génératrice des moments

Définition 3.1.11 *On appelle fonction génératrice des moments d'une v.a.r. X, la fonction M_X définie par :*

$$M_X(t) = \mathbb{E}[\exp(tX)] = \int_{\mathbb{R}} \exp(tx) f_X(x) dx,$$

sur $\{t \in \mathbb{R} \ ; \ M_X(t) < +\infty\}$. Elle est définie sur un intervalle qui contient l'origine.

Remarque 3.1.11 *M_X est appelée fonction génératrice des moments (f.g.m.) à cause de la relation :*

$$M_X^{(n)}(0) = \mathbb{E}[X^n], \quad \textit{pour} \quad n \in \mathbb{N}.$$

Cette transformation de f_X en M_X est aussi connue sous le nom de transformée de Laplace de la fonction f_X.

Exemple 3.1.1 (loi $E(\lambda)$) *Soit $X \sim E(\lambda)$ alors :*

$$M_X(t) = \int_0^{+\infty} \lambda \exp(-(\lambda - t)x) dx = \frac{\lambda}{\lambda - t} \quad si \quad \lambda > t.$$

Proposition 3.1.10 *Si M_X est définie sur $[-a, a]$ avec $a > 0$ alors :*

a. *X possède des moments finis de tous ordres ;*

b. $M_X(t) = \displaystyle\sum_{k=0}^{+\infty} \frac{t^k}{k!} E[X^k].$

3.1.11 Somme de deux variables aléatoires

Soient X et Y deux v.a.r. définies sur $(\Omega, \mathcal{F}, P)$ de f.d.r. (resp. densités) respectives F_X et F_Y (resp. f_X et f_Y).

Proposition 3.1.11 *La fonction de répartition F_Z de la variable aléatoire $Z = X + Y$, est donnée par le produit de convolution $F_X * F_Y$ de F_X et F_Y, c'est-à-dire :*

$$F_Z(x) = (F_X * F_Y)(x) = \int_{\mathbb{R}} F_X(x - y) f_Y(y) dy = \int_{\mathbb{R}} F_Y(x - y) f_X(y) dy.$$

TABLE 3.1 – f.g.m. des lois usuelles.

Lois	f.g.m.
Uniforme $U(a,b)$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Exponentielle $E(\lambda)$	$\frac{\lambda}{\lambda-t}$
Normale $N(\mu,\sigma^2)$	$e^{\frac{t^2\sigma^2}{2}+\mu t}$
Gamma $\gamma(\alpha,\lambda)$	$\left(\frac{\lambda}{\lambda-t}\right)^\alpha$

De plus, $Z = X + Y$ admet une densité f_Z donnée par $f_X * f_Y$, c'est-à-dire pour $x \in \mathbb{R}$:

$$f_Z(x) = (f_X * f_Y)(x) = \int_{\mathbb{R}} f_X(x-y)f_Y(y)dy = \int_{\mathbb{R}} f_Y(x-y)f_X(y)dy.$$

Démonstration. Voir le document 3.5.3.

Corollaire 3.1.1 *Si les v.a.r. $X_1, \dots, X_n$ sont indépendantes, la loi de $X_1+\cdots+X_n$ est égale au produit de convolution des lois de X_i. La réciproque est fausse.*

Exemple 3.1.2 *Soient X et Y deux variables aléatoires indépendantes et de même loi Uniforme sur $[-1,1]$. Déterminer la fonction de répartition de la variable $X+Y$. La fonction de répartition de $X+Y$ est donnée par :*

$$F_{X+Y}(t) = \int_R F_Y(t-x)f_X(x)dx = \int_R \frac{t-x+1}{2} 1_{[-1,\ 1]}(t-x)\frac{1}{2}1_{[-1,\ 1]}(x)dx$$

$$= \frac{1}{4}\int_R (t-x+1)1_{[-1+t,\ 1+t]}(x)1_{[-1,\ 1]}(x)dx = \begin{cases} 0, & si\ t < -2, \\ \frac{1}{4}\int_{-1}^{t+1}(t-x+1)dx = \frac{t}{2}+\frac{t^2}{8}+\frac{1}{2}, & si\ t \in [-2,\ 0], \\ \frac{1}{2}+\frac{1}{4}\int_{t-1}^{1}(t-x+1)dx = \frac{t}{2}-\frac{t^2}{8}+\frac{1}{2}, & si\ t \in]0,\ 2], \\ 1, & si\ t > 2. \end{cases}$$

Exemple 3.1.3 *Soit $X \sim U([0,1])$ et $Y \sim E(1)$ indépendantes ; alors la densité de $Z = X+Y$ est donnée par*

$$f_Z(x) = \int_{\mathbb{R}} 1_{[0,1]}(x-y)\exp(-y)1_{[0,+\infty[}(y)dy.$$

Remarquons que :

$$1_{[0,1]}(x-y)1_{[0,+\infty[}(y) = 1_{]-\infty,0[}(x) + 1_{[0,x]}(y)1_{[0,1]}(x) + 1_{]1,+\infty[}(x)1_{[x-1,x]}(y).$$

Il vient donc

$$f_Z(x) = \begin{cases} 0 & si \quad x \in]-\infty, 0[, \\ \int_0^x \exp(-y)dy & si \quad x \in [0,1], \\ \int_{x-1}^x \exp(-y)dy & si \quad x \in]1, +\infty[, \end{cases} = \begin{cases} 0 & si \quad x \in]-\infty, 0[, \\ 1 - \exp(-x) & si \quad x \in [0,1], \\ \exp(1-x) - \exp(-x) & si \quad x \in]1, +\infty[. \end{cases}$$

Proposition 3.1.12 *Si les v.a.r. $X_1, \dots, X_n$ sont indépendantes alors :*

$$M_{X_1+\dots+X_n}(t) = M_{X_1}(t) \dots M_{X_n}(t).$$

Démonstration. Comme les v.a.r. $X_1, \dots, X_n$ sont indépendantes, on a :

$$\mathbb{E}[e^{t(X_1+\dots+X_n)}] = \mathbb{E}[e^{tX_1} \times \dots \times e^{tX_n}] = \mathbb{E}[e^{tX_1}] \times \dots \times \mathbb{E}[e^{tX_n}].$$

d'après la proposition 3.1.

3.1.12 Inégalités

Proposition 3.1.13 (Inégalité de Bienaymé-Čebyčev) *Soit X une v.a.r. absolument continue admettant un moment d'ordre 2. Alors pour tout $\varepsilon > 0$, on a :*

$$P(|X - \mathbb{E}[X]| > \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Démonstration. Voir le document 3.5.4.

Proposition 3.1.14 (Inégalité de Cauchy-Schwarz) *Si les v.a.r. X et Y ont des moments d'ordre 2 (finis donc) alors :*

$$(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2].$$

L'égalité est atteinte si et seulement si il existe une constante c telle que $Y = cX$ avec probabilité 1.

Démonstration. Pour tout $t \in \mathbb{R}$ on a :

$$0 \leq \mathbb{E}[(X + tY)^2] = \mathbb{E}[X^2] + 2t\mathbb{E}[XY] + t^2\mathbb{E}[Y^2].$$

Pour que le polynôme du second degré en t ci-dessus soit positif il est nécessaire que son discriminant Δ soit négatif ou nul. Ce qui conduit à

$$\Delta = 4(\mathbb{E}[XY])^2 - 4\mathbb{E}[X^2]\mathbb{E}[Y^2] \leq 0.$$

Proposition 3.1.15 (Inégalité de Jensen) *Soient X une v.a.r. et φ une fonction convexe sur $\mathbb{R}$. Alors*

$$\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}[X]).$$

Exemple. L'application $x \mapsto x^2$ est convexe sur $\mathbb{R}$; par conséquent $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \geq 0$!

3.2 Exercices de cours

EC-3-1 Montrer que si X est une v.a.r. sur $(\Omega, \mathcal{F}, P)$ et $-\infty < a < b < +\infty$ alors $X^{-1}(]a,b])$, $X^{-1}([a,b])$, $X^{-1}([a,b[)$, $X^{-1}(]-\infty,a[)$, $X^{-1}([a,+\infty[)$ et $X^{-1}(]a,+\infty[)$ sont bien des éléments de $\mathcal{F}$ (ce qui justifie que l'on puisse calculer leurs probabilités !).

EC-3-2 Reprendre le schéma de "fonctionnement" de la loi d'une v.a.r. avec $\Omega = [-1,1]$, P la probabilité géométrique sur Ω et $X(\omega) = \omega^2$ pour trouver p_X.

EC-3-3 Un point P est choisi au hasard uniformément dans un disque centré en O de rayon $\varrho > 0$. Soit R la v.a.r. donnant la distance entre les points O et P. Trouver la f.d.r. de R. Tracer son graphe.

EC-3-4 Soit X l'âge d'un individu. Soit F la fonction définie par :

$$F(x) = (1 - \exp(-x))1_{[0,+\infty[}(x).$$

Vérifier que F est une fonction de répartition. On suppose que F est la f.d.r. de X. Calculer $P(X \leq 0)$ et $P(X \geq s+t | X \geq s)$ pour $s > 0, t > 0$; interpréter les résultats. Montrer qu'il existe une fonction f telle que :

$$\forall x \in \mathbb{R}, \quad F(x) = \int_{-\infty}^{x} f(y)dy.$$

EC-3-5 Soit f définie par $f_X(x) = k(2-x)1_{[0,1]}(x)$.

a- Pour quelle valeur de k cette fonction est-elle une densité de probabilité ? Tracer son graphe.

b- Calculer sa fonction de répartition et tracer son graphe.

EC-3-6 Soit f la fontion définie par $f(x) = kx^2 1_{[-1,1]}(x)$. Pour quelle valeur de k f est-elle une densité ? Calculer la f.d.r. F correspondante. Soit X la v.a. admettant f pour densité. Calculer $\mathbb{E}(X^k)$ $(k \geq 1)$ et $\mathrm{Var}(X)$.

EC-3-7 Calculer la médiane d'une v.a.r. X admettant la densité f définie par $f(x) = \lambda \exp(-\lambda x)1_{[0,+\infty[}(x)$.

EC-3-8 Soient X et Y deux v.a.r. indépendantes et de même loi. On suppose que $\frac{X+Y}{\sqrt{2}}$ est de même loi que X et Y. On suppose que cette loi est de variance finie σ^2.

a- Montrer que la loi de X est centrée.

b- Montrer que si X_1, X_2, Y_1 et Y_2 sont des v.a.r. indépendantes de même loi que X, alors :

$$\frac{X_1 + X_2 + Y_1 + Y_2}{2}$$

a la même loi que X.

EC-3-9 Soit X une v.a. de loi normale centrée réduite.

a- Retrouver en utilisant la densité, son espérance et sa variance.

b- Soit Y une v.a.r. telle que $Y = \sigma X + \mu$ ($\mu \in \mathbb{R}$ et $\sigma \in]0, +\infty[$). Trouver la densité Y. Calculer l'espérance et la variance de Y sans passer par leurs définitions.

EC-3-10 Calculer l'espérance et la variance de la loi d'Erlang.

EC-3-11 Un charpentier est amené à mesurer et couper un tronc de longueur L en son centre O pour en faire une poutre. Montrer que si les erreurs de mesure son uniformément distribuées sur un intervalle centré en O, de $2k \leq L$ cm de longueur, l'espérance est nulle. Pourquoi ce résultat est-il moins rassurant à mesure que k augmente ?

EC-3-12 Supposons que la durée de vie d'une ampoule électrique d'un certain type suit une loi normale de moyenne $m = 180$ heures et d'écart-type $\sigma = 20$ heures. Dans un ensemble de quatre de ces ampoules :

a- quelle est la probabilité que les quatre ampoules aient une durée de vie supérieure à 200 heures ?

b- quelle est la probabilité de pouvoir éclairer une pièce pendant plus de 800 heures avec ces ampoules ?

EC-3-13 Soient X et Y deux v.a.r. indépendantes. Montrer, en utilisant la fonction génératrice des moments, que si X suit la loi $\gamma(\alpha, \beta)$ et Y la loi $\gamma(\mu, \beta)$ alors la v.a.r. $X + Y$ suit la loi $\gamma(\alpha + \mu, \beta)$.

EC-3-14 Soit X et Y deux v.a.r. indépendantes de loi $U(0, 1)$. Calculer la f.d.r. et la densité de $X + Y$.

EC-3-15 Montrer que si $x > 0$, alors :

$$P(X > x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp(-t^2/2)dt \leq \frac{1}{x\sqrt{2\pi}} \exp(-x^2/2).$$

Indication. Remarquer que pour $t \in [x, +\infty[, 1 \leq \frac{t}{x}$.

EC-3-16 Montrer que si $g(x) \geq 0$, pour tout x et $g(x) \geq c > 0$, pour $x \in]\alpha, \beta[$, alors :

$$P(X \in]\alpha, \beta[) \leq c^{-1} \mathbb{E}\left[g(X)\right].$$

3.3 Exercices de travaux dirigés

ETD-3-1 Montrer que l'ensemble des points de discontinuité d'une fonction de répartition est au plus infini dénombrable.

ETD-3-2 Un tireur à l'arc vise le centre d'une cible de rayon $R > 0$. La probabilité qu'il atteigne la cible est p ($0 < p < 1$). Si la cible est touchée, la probabilité qu'une zone donnée de la cible soit atteinte est uniforme. Lors d'un concours, les points sont attribués comme suit :

$-k$ points si la cible n'est pas atteinte ;

$R - r$ points lorsque la cible est atteinte, où r est la distance (aléatoire) séparant le centre de la cible du point d'impact sur la cible.

Le gain du tireur, pour un seul tir, est une v.a.r. X. Donner la fonction de répartition F_r de la v.a.r. r lorsque la cible est atteinte. Donner la fonction de répartition F_X de X. A quels types de v.a.r. a-t-on affaire ?

ETD-3-3 Soient $X_1, \ldots, X_n$ n v.a.r. indépendantes et identiquement distribuées (i.i.d.). On définit $M_n = \max_{1 \leq i \leq n} X_i$ et $N_n = \min_{1 \leq i \leq n} X_i$. Quelles sont les fonctions de répartition de M_n et N_n ? Supposons que la loi des X_i soit uniforme sur $[0, 1]$, vers quoi converge la f.d.r. F_{M_n} lorsque $n \to +\infty$? Quelle interprétation donner à ce résultat ?

ETD-3-4 a- Soit X une variable aléatoire réelle de densité f telle que $\mathbb{E}(|X|) < +\infty$.

- i- Montrer que si f est paire alors $\mathbb{E}(X) = 0$;
- ii- Montrer que si le graphe de f est symétrique par rapport à la droite d'équation $x = x_0$ alors $\mathbb{E}(X) = x_0$;

b- Les fonctions de répartition suivantes ont-elles des densités (si oui, expliciter la densité, si non, justifier la réponse) ?

- i- $F_1(x) = (1 - \exp(-x^3))1_{[0,+\infty[}(x)$;
- ii- $F_2(x) = (1 - \exp(-x))1_{[1,+\infty[}(x)$.

c- Soit g définie sur $\mathbb{R}$ par $g(x) = a/(1 + x^2)$.

- i- A quelle condition g est-elle une densité ?
- ii- Montrer qu'une variable aléatoire X, de densité g, n'a pas de moyenne.

ETD-3-5 Soit X une v.a.r. de f.d.r. F. On suppose F continue. Donner la loi de probabilité de $F(X)$. (On pourra supposer F strictement croissante dans un premier temps.)

ETD-3-6 Un appareil électronique est soumis à des impulsions séparées par des intervalles de temps variables indépendants les uns des autres. On considère que la durée T (exprimée en secondes) séparant deux impulsions successives est une variable aléatoire définie par $T = 2 + 5X$, où X suit la loi $E(1)$.

I- Exprimer la fonction de répartition F_T en fonction de F_X, en déduire la densité de T.

II- Calculer les probabilités des événements suivants :

a- la durée séparant deux impulsions est inférieure à deux secondes.

b- la durée séparant deux impulsions est comprise entre deux et cinq secondes.

III- Calculer la durée moyenne séparant deux impulsions et sa variance.

ETD-3-7 Soit X une v.a. qui suit la loi uniforme sur $[0, 1]$.On pose $Y = X^2$ et $Z = e^X$. Déterminer les densités de probabilités des v.a. Y et Z.

ETD-3-8 Soient X et Y deux v.a.r. indépendantes. Montrer en utilisant le produit de convolution ou la fonction génératrice de moments que si : X suit la loi $N(m,\ \sigma^2)$ et Y la loi $N(l,\ \rho^2)$ alors la v.a.r. $X+Y$ suit la loi $N(m+l,\ \sigma^2+\rho^2)$.

ETD-3-9 Un promoteur organise une petite réunion amicale avec les personnes influentes qui lui ont écarté les obstacles pour construire un immeuble de grand standing sur l'emplacement d'un square, rénové un an auparavent par une de ses filiales (au frais des contribuables). Les serveurs présentent à chacun des cent cinquante convives une coupe de caviar servi à la louche. Une louche contient en moyenne 140 gr. de caviar avec un écart-type de 11 gr. Sachant que le bélouga utilisé revient à 8500F. le Kg., on note X le prix du premier service de caviar.

a- Etablir la loi de X, son espérance et son écart-type.

b- Calculer la probabilité d'avoir servi pour moins de 170000 F de caviar.

c- Reprendre le problème avec les deux cent cinquante ouvriers du chantier, des oeufs de lump à 82 F le kilo servis dans un verre de cantine avec une cuillère de contenance moyenne de 45 gr., avec un écart-type de 7 gr.

d- Calculer la probabilité d'avoir dépensé plus de 1000 F d'ersatz.

ETD-3-10 Une machine fabrique des résistances électriques dont la valeur en ohms est une variable aléatoire R de loi normale $N(100; 9)$. Une seconde machine fabrique des résistances dont la valeur en ohms est une variable aléatoire R' de loi normale $N(200; 16)$.

a- Quelle est la loi suivie par la résistance obtenue en montant en série deux résistances prélevées au hasard dans les productions respectives de la première et de la seconde machine ?

b- Quelle est la probabilité qu'une telle résistance soit comprise entre 290 et 305 ohms ?

ETD-3-11 Soit Y une v.a. de densité : $be^{-x}1_{[2,\infty[}(x)$ où b est une constante positive.

a- Déterminer la valeur de b.

b- On pose $Z = [Y]$. Calculer $P(Z = n)$.

c- On pose $X = Y - [Y]$, où $[Y]$ est la partie entière de Y. Calculer $E(X)$.

ETD-3-12 Soient X et Y deux variables aléatoires indépendantes de loi $E(1)$.

a- Calculer la fonction de répartition de $-X$.

b- En déduire la densité de $-X$.

c- Montrer que la densité f_Z de $Z = Y - X$ est définie par :

$$f(x) = \frac{1}{2}\exp(-|x|), \quad x \in \mathbb{R}.$$

d- Calculer $\mathbb{E}(Z)$ et $\mathrm{Var}(Z)$.

ETD-3-13 Soient X une variable aléatoire X de loi $N(0,1)$ et Y une variable aléatoire à valeurs dans $\{-1,1\}$ avec $P(Y=1)=p$, pour $p \in]0,1[$. On suppose que X et Y sont indépendantes. On considère la variable aléatoire $Z = XY$. Montrer que X et Z ont même loi. Sont-elles indépendantes ?

ETD-3-14 Soit Y une v.a. réelle uniformément distribuée sur l'intervalle $[3,6]$. Pour tout $n > 0$, on pose $X_n = 5n^2$, si $3 \leq Y \leq 3 + 4/n^2$ et $X_n = 0$, sinon.

a- Déterminer $E(X_n)$ et $E(X_n^2)$.

b- Calculer $E(X_{n+1}X_{n+2})$.

ETD-3-15 On considère une v.a.r. X centrée et de variance finie σ^2.

a- Montrer que , pour tout $a > 0$,

$$a \leq \mathbb{E}[(a-X)1_{]-\infty,a]}(X)] \leq \{P(X \leq a)\}^{1/2}\{\sigma^2 + a^2\}^{1/2}.$$

En déduire que :

$$P(X > a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

b- Une usine fabrique chaque semaine un nombre aléatoire d'objets Y. On suppose $\mathbb{E}(Y) = 100$ et $\mathrm{Var}(Y) = 400$. Trouver à l'aide de la question précédente un majorant de la probabilité que la production hebdomadaire dépasse 120. Comparer ce résultat avec celui obtenu par l'inégalité de Bienaymé-Tchebychev.

ETD-3-16 Selon les lois de Mendel une certaine variété de pois de senteur a une probabilité 1/4 de fleurir blanche et une probabilité 3/4 de fleurir rose ou rouge. Combien faut-il observer de fleurs de cette espèce pour que la fréquence du nombre de fleurs blanches ne s'écarte pas plus de 0,05 de la fréquence observée (on admet 0,01 d'erreur).

ETD-3-17 Soit X une variable aléatoire dont la densité f est définie par :

$$f_X(x) = a\exp(-x)1_{[0,+\infty[}(x).$$

a- Déterminer a et tracer le graphe de f.

b- Calculer la fonction de répartition de X.

c- Calculer l'espérance et la variance de X.

ETD-3-18 Une tension électrique aléatoire X (mesurée en volts) de densité de probabilité $f_X(x) = \lambda e^{-\lambda x}1_{\mathbb{R}_+}(x), (\lambda > 0)$ passe par un limitateur qui "coupe" toutes les valeurs de la tension inférieures à u_1 et supérieures à $u_2, (0 < u_1 < u_2 < \infty)$, dans le premier cas en l'élevant à u_1 et dans le deuxième cas en la réduisant à u_2. Notons Y la variable aléatoire désignant la tension issue du limiteur.

a- Calculer $P(Y = u_1)$ et $P(Y = u_2)$. Quel est le type de la variable aléatoire Y ? Ecrire Y en fonction de la variable aléatoire X.

b- Donner la fonction de répartition de Y ; tracer son graphe.

c- Calculer l'espérance mathématique et la variance de Y.

ETD-3-19 Soient X et Y deux variables aléatoires positives, indépendantes, suivant une même loi définie par sa densité $f_X(x) = xe^{-x}1_{\mathbb{R}_+}(x)$.

a- Vérifier que f est bien une densité.

b- Calculer $\mathbb{E}[X^n]$ (par récurrence) et la variance de X.

c- Pour $\alpha > -1$, calculer $\mathbb{E}[\exp(-\alpha X)]$.

d- Déterminer la fonction de répartition de X.

e- On pose $T = \inf\{X, Y\}$. Calculer la fonction de répartition de T.

ETD-3-20 Considérons une suite de variables aléatoires $X_1, X_2, \ldots$ indépendantes et identiquement distribuées de loi exponentielle $E(\lambda)$. On note f_1 la densité d'une de ces variables aléatoires.

a- Montrer que la densité de $X_1 + X_2$ est $f_2(t) = \lambda^2 te^{-\lambda t}1_{\mathbb{R}_+}(t)$.

b- Montrer par récurrence que la densité de probabilité f_n de $X_1 + \cdots + X_n$ est définie par :

$$f_n(t) = \frac{\lambda(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}1_{\mathbb{R}_+}(t).$$

c- Soit Y une variable aléatoire indépendante des X_i , de loi géométrique $G(p)$ avec $0 < p < 1$. Calculer la densité de la variable aléatoire $Z = \sum_{i=1}^{Y} X_i$.

d- Donner l'espérance et la variance de la variable aléatoire Z.

ETD-3-21 Soit f la fonction définie sur $\mathbb{R}$ par :

$$f_X(x) = \begin{cases} 0 & \text{si} \quad |x| > 1, \\ x+1 & \text{si} \quad x \in [-1, 0], \\ 1-x & \text{si} \quad x \in [0, 1]. \end{cases}$$

a- Montrer que f est une densité de probabilité. Soit X une v.a.r. de densité f.

b- Déterminer la fonction de répartition de X.

c- Calculer $\mathbb{E}[X]$ et $\mathrm{Var}(X)$. Justifier sans calcul la valeur de $\mathbb{E}[X]$.

d- Déterminer $P(|X| \geq a)$.

e- Déterminer les valeurs de a pour lesquelles $P(|X| \geq a) \leq \dfrac{1}{6a^2}$.

f- Soit Y la variable aléatoire uniforme sur $[-1, 1]$, indépendante de X. Calculer $P((X, Y) \in [-1/2, 1/2] \times [-1/2, 1/2])$.

g- Soit $Z = \min(X, Y)$. Quelle est la loi de Z ?

ETD-3-22 On suppose que la charge de rupture d'un fil est une variable aléatoire X de moyenne $m = 851$ grammes et d'écart-type 71 grammes. D'après l'inégalité de Bienaymé-Tchebychev, que peut-on dire de la probabilité de l'événement $\{X \leq m + 2\sigma\}$? Que signifie cette probabilité ?

ETD-3-23 Soit Y une v.a.r. définie par : $Y = m + \theta$ où la v.a.r. θ suit la loi $N(0,1)$[1] et soit $X = \exp(Y)$.

a- Déterminer la densité f_X de la variable aléatoire X.

b- Calculer $\mathbb{E}(X)$ et $\mathrm{Var}(X)$.

c- Donner la fonction génératrice des moments de la v.a. Y. En déduire $\mathbb{E}(X)$ et $\mathbb{E}(X^2)$.

ETD-3-24 Considèrons une bouche du métro. On suppose que le nombre X_t de voyageurs sortant de cette bouche pendant le temps t, suit une loi de Poisson de paramètre λt, c'est-à-dire pour $n \in \mathbb{N}$ $P(X_t = n) = \exp(-\lambda t)\frac{(\lambda t)^n}{n!}$. Un observateur se poste à la sortie de la bouche du métro et il s'écoule un temps T avant la sortie du premier voyageur. On suppose de plus que le métro fonctionne 24h sur 24h.

a- Que signifie l'événement $\{T > t\}$ où $t \in \mathbb{R}^+$. Calculer sa probabilité et en déduire la densité de T.

b- Montrer que $P(T > t + s | T > s) = P(T > t)$ où $s \in \mathbb{R}^+$.

c- On définit la fonction $G(x) = \ln(1 - F_T(x))$ où $x \in \mathbb{R}$.

i. En utilisant la définition de la probabilité conditionnelle et la question 2, donner l'expression de $\ln(P(T > t+s|T > s))$ en utilisant la fonction G.

ii. Montrer que $G(x) = ax$, satisfait la relation ci-dessus, et dire comment doit être le paramètre a. Retrouver $F_T(t)$.

ETD-3-25 Montrer que si X et Y sont i.i.d., alors pour tout $t > 0$:

$$P(|X - Y| > t) \leq 2P\left(|X| > \frac{t}{2}\right).$$

ETD-3-26 Pour fonctionner, un système utilise une cellule interchangeable. On dispose de la pièce originale et d'une cellule de rechange, chacune d'une durée de vie aléatoire X (en mois). Si la densité f de X est donnée par :

$$f(x) = \begin{cases} cxe^{-x/2} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0, \end{cases}$$

quelle est la fiabilité du système au bout de 5 mois ? Calculer le taux de défaillance de ce système.

Indications. On appelle fiabilité du système la fonction R définie par $R(x) = P(X > x)$. On appelle taux de défaillance du système la fonction λ définie par $\lambda(x) = f(x)/R(x)$.

1. de densité $f_\theta(t) = \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\} 1_{\mathbb{R}}(t)$

ETD-3-27 Soit X une variable aléatoire normale $N(0,2)$ et soit $Y = \exp(-\alpha X^2)$ avec $\alpha > 0$. Donner la loi de Y et son espérance.

ETD-3-28 Calculer l'espérance et la variance d'une loi de Weibull de paramètres (λ, α), dont la densité est $\lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}1_{\mathbb{R}_+}(t)$.

ETD-3-29 Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et X une variable aléatoire définie sur Ω. On note F_X la fonction de répartition de X.

I- Montrer que F_X est croissante.

II- Soit $x \in \mathbb{R}$ fixé et $\varepsilon > 0$. On pose $A_\varepsilon = \{X \in]-\infty, x+\varepsilon]\} \cap \overline{\{X \in]-\infty, x]\}}$.

a- Montrer que si $\varepsilon_1 < \varepsilon_2$ alors $A_{\varepsilon_1} \subset A_{\varepsilon_2}$.

b- Montrer que $\cap_{\varepsilon>0} A_\varepsilon = \emptyset$.

III- Montrer que pour chaque $\varepsilon > 0$ on a : $F_X(x+\varepsilon) - F_X(x) = P(A_\varepsilon)$, $\forall x \in \mathbb{R}$.

IV- Déduire de ce qui précède que F_X est continue à droite en chacun de ses points de discontinuité.

ETD-3-30 Soient X, Y et Z trois v.a.r. indépendantes de même loi uniforme sur $[-1, +1]$. Calculer la loi de $X+Y$, $\frac{X+Y}{2}$, $X+Y+Z$ et $\frac{X+Y+Z}{3}$.

ETD-3-31 Soit f la fonction définie par

$$f(x) = ae^x 1_{]-\infty,-\ln 2]}(x) + be^{-x}1_{[\ln 2,+\infty[}(x).$$

a- Quelles conditions doivent satisfaire a et b pour que f soit une densité ?

b- On suppose que $a = b$, donner l'expression de f et l'allure de son graphe.

c- Soit X une variable aléatoire admettant une densité f. La variable aléatoire X peut-elle être centrée (ne pas calculer $\mathbb{E}(X)$) ? A quelle condition sur a et b, X est-elle centrée (toujours sans calculer $\mathbb{E}(X)$) ?

ETD-3-32 Dans un atelier on assemble 100 ordinateurs par jour. En général, 99% des ordinateurs assemblés fonctionnent lorsqu'on les teste. Soit X le nombre d'ordinateurs en état de marche parmi les ordinateurs assemblés un jour donné.

a- Quelle est la loi de X ?

b- Quelle relation existe-t-il entre les événements $\{X \leq 95\}$ et $\{|X-99| \geq 4\}$?

c- Donner une majoration de $P(|X-99| \geq 4)$, puis en déduire une majoration de $P(X < 95)$.

ETD-3-33 Soit X une v.a. positive de densité f continue et de fonction de répartition (Fdr) F.

a- Soit $s > 0$, montrer que : $\lim_{s\to 0} \frac{1}{s} P(t < X \leq t+s | X \geq t) = \frac{f(t)}{1-F(t)}$, $t \geq 0$. On appelle cette dernière quantité le taux de défaillance de X et on note h cette fonction i.e. $h(t) = \frac{f(t)}{1-F(t)}$, $t \geq 0$.

b- On suppose que la v.a. X suit une loi exponentielle de paramètre $\lambda > 0$, i.e. de densité $f(x) = \lambda \exp\{-\lambda x\} 1_{[0,+\infty[}(x)$.

i- Calculer la Fdr et le taux de défaillance de X.

ii- Soit Y une v.a. positive admettant une densité continue g et un taux de défaillance constant égal à $\mu > 0$. Montrer qu' alors la Fdr G de Y satisfait une équation différentielle. En déduire une expression de g, puis que Y suit une loi exponentielle de paramètre μ.

c- On considère un circuit électronique constitué de r résistances indépendantes ayant des durées de fonctionnement aléatoires $X_1, \cdots, X_n$ de même loi que X.

i- Ces résistances sont montées en série de sorte que le circuit tombe en panne dès que l'une des résistances tombe en panne. Soit T la durée de bon fonctionnement du circuit. Quelle relation lie T aux variables X_i ? En déduire la Fdr de T puis son taux de défaillance.

ii- Ces résistances sont montées en parallèle de sorte que le circuit tombe en panne lorsque toutes les résistances sont tombées en panne. Quelle relation lie T aux variables X_i ? En déduire la Fdr de T puis son taux de défaillance.

ETD-3-34 a- Donner la densité d'une variable aléatoire X de loi $N(0, 1)$ (normale centrée réduite).

b- Soit $Y = X^2$. Exprimer la fonction de répartition de Y en fonction de celle de X.

c- En déduire la densité de Y.

d- En déduire la valeur de $\int_0^{+\infty} \frac{1}{\sqrt{x}} \exp(-x/2)dx$.

ETD-3-35 Soit f la fonction définie par $f(x) = c_1 1_{[-2,-1]}(x) + c_2 1_{[1,2]}(x)$.

a- Quelles conditions doivent vérifier c_1 et c_2 pour que f soit une densité ?

b- Soit X une variable aléatoire de densité f. Soit $Y = 1_{]0,+\infty[}(X) - 1_{]-\infty,0[}(X)$. Exprimer la loi p_Y de Y en fonction de c_1 et c_2.

c- Calculer la variance de Y.

3.4 Exemples

3.4.1 Représentation en base dix d'un nombre

On choisit un nombre "au hasard" dans $[0, 1]$. Pour ne favoriser aucun nombre de $[0, 1]$ on choisit un nombre de la manière suivante. Un nombre $x \in [0, 1]$ est représenté en base dix de la manière suivante :

$$\omega = 0, x_1 x_2 x_3 \ldots \quad \text{où} \quad x_i \in \{0, 1, \ldots, 9\},\ \forall i \in \mathbb{N}^*,$$

où l'équivoque sur le développement décimal est levée en prenant la convention

$$0,1999\cdots \equiv 0,2.$$

Les x_i sont choisis dans $\Theta = \{0, 1, \ldots, 9\}$ de manière indépendante et avec équiprobabilité. L'univers est donc défini par $\Omega = \Theta \times \Theta \times \cdots = \Theta^{\mathbb{N}}$; il s'agit donc d'un produit infini dénombrable de Θ. Soit $(\Omega, \mathcal{F}, P)$ l'espace probabilisé correspondant (dont on suppose l'existance) et $X : \Omega \to [0, 1]$ définie par $X(\omega) = x$ (dont on suppose que c'est une v.a.r.).

Voyons ainsi comment calculer $P(X \leq 0,25)$. Remarquons tout d'abord que

$$\begin{array}{rcl} \{X \leq 0,25\} & = & \{x_1 = 0; x_2 \in \Omega; x_3 \in \Omega; \ldots\} \\ & \cup & \{x_1 = 1; x_2 \in \Omega; x_3 \in \Omega; \ldots\} \\ & \cup & \{x_1 = 2; x_2 \leq 4; x_3 \in \Omega; x_4 \in \Omega; \ldots\} \\ & \cup & \{x_1 = 2; x_2 = 5; x_3 = 0; x_4 = 0; \ldots\}. \end{array}$$

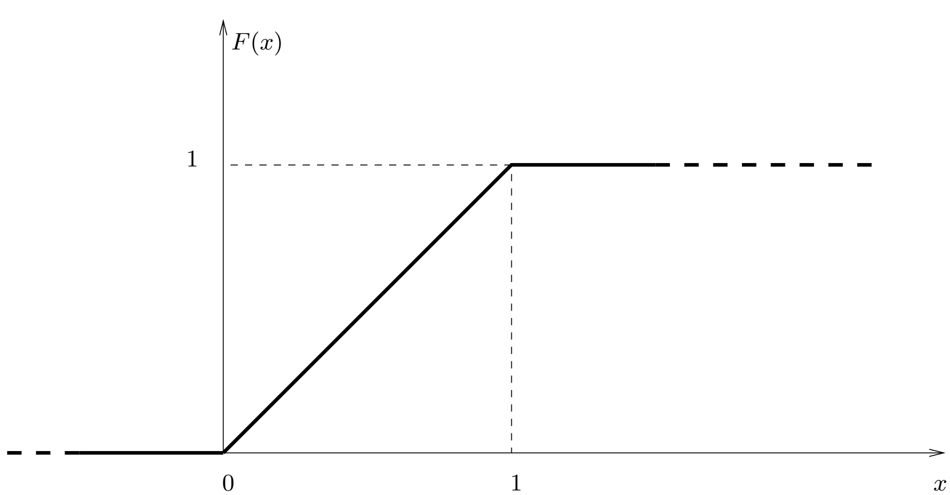
La réunion ci-dessus étant disjointe, on a

$$P(X \leq 0,25) = \left(\frac{1}{10}\right) + \left(\frac{1}{10}\right) + \left(\frac{1}{10} \times \frac{5}{10}\right) + \lim_{n \to +\infty} \left(\frac{1}{10}\right)^n = 0,25.$$

Plus généralement on trouve

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ x & \text{si} \quad x \in [0,1], \\ 1 & \text{si} \quad x > 1. \end{cases}$$

On remarque que la loi de probabilité de X correspond à la mesure de probabilité uniforme sur $[0, 1]$. Il est alors facile de voir que le graphe de F_X est



Paradoxe ? Cette manière de choisir un nombre dans $[0, 1]$ fait que X prendra toujours une valeur. Malgré tout, la probabilité que X prenne une valeur donnée est nulle puisque :

$$P(X = x) = \frac{1}{10} \times \frac{1}{10} \times \cdots = \lim_{n \to +\infty} \left(\frac{1}{10}\right)^n = 0.$$

Ça ne contredit pas les axiomes de probabilité car il serait **faux** d'écrire :

$$P(X \in [0,1]) = \sum_{x \in [0,1]} P(X = x)$$

car ici $[0,1]$ n'est pas dénombrable !

3.4.2 Fonction de répartition

Soit X une v.a. dont la f.d.r. vérifie

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

On définit

$$Y = \max(1/2, X).$$

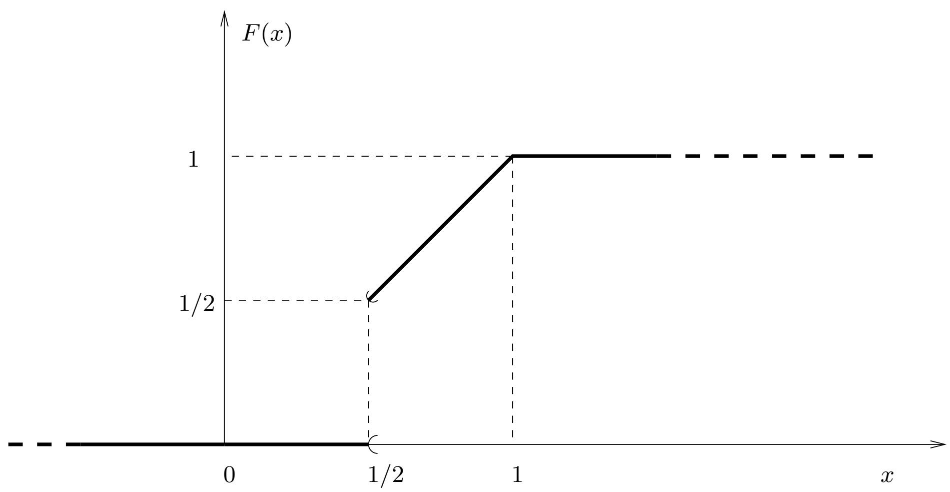
Y est une v.a.r. sur $[0,1]$ à valeurs dans $[1/2, 1]$. En effet

$$\{Y \leq y\} = \begin{cases} \emptyset & \text{si} \quad y < 1/2, \\ \{X \leq y\} & \text{si} \quad y \geq 1/2. \end{cases}$$

On a alors :

$$F_Y(y) = p_Y(]-\infty, y]) = \begin{cases} p_X(\emptyset) & \text{si} \quad y < 1/2 \\ p_X(]-\infty, y]) & \text{si} \quad y \geq 1/2 \end{cases} = y 1_{[1/2,1]}(y) + 1_{]1,+\infty[}(y).$$

Le graphe de F_Y est donc donné par :



F_Y n'est ni constante par morceaux (comme les v.a.r. discrètes) ni continue (comme les v.a.r. à densité). Une telle v.a.r. est dite mixte.

3.5 Documents

3.5.1 Démonstration du théorème 3.1.1

$(\mathbb{R}, \mathcal{B})$ étant probabilisable, il suffit de vérifier que
(i) $P_X(\mathbb{R}) = 1$;
(ii) Pour toute suite $(B_n)_{n\geq 1}$ d'éléments de $\mathcal{B}$ deux à deux disjoints

$$P_X\left(\bigcup_{n=1}^{+\infty} B_n\right) = \sum_{n=1}^{+\infty} P_X(B_n).$$

Pour (i) nous avons $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R}))$. Or $X^{-1}(\mathbb{R}) = \Omega$ donc $P_X(\mathbb{R}) = P(\Omega) = 1$ car $(\Omega, \mathcal{F}, P)$ est un espace probabilisé.

Pour (ii) nous avons

$$P_X\left(\bigcup_{n=1}^{+\infty} B_n\right) = P\left(X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right)\right),$$

où

$$X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right) = \left\{\omega \in \Omega\ ;\ X(\omega) \in \bigcup_{n=1}^{+\infty} B_n\right\}.$$

Notons $A_n = X^{-1}(B_n)$ pour $n \geq 1$ alors $A_n \in \mathcal{F}$ puisque X est une v.a.r. Soit $\omega \in \bigcup_{n=1}^{+\infty} A_n$ alors il existe $n_0 \in \mathbb{N}^*$ tel que $\omega \in A_{n_0}$ donc $X(\omega) \in B_{n_0} \subset \bigcup_{n=1}^{+\infty} B_n$ soit encore $\omega \in X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right)$. Nous avons donc

$$\bigcup_{n=1}^{+\infty} A_n \subset X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right).$$

Maintenant si $\omega \in X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right)$ alors $X(\omega) \in \bigcup_{n=1}^{+\infty} B_n$ donc il existe n_0 tel que $X(\omega) \in B_{n_0}$ soit encore $\omega \in A_{n_0} \subset \bigcup_{n=1}^{+\infty} A_n$. Par conséquent

$$\bigcup_{n=1}^{+\infty} A_n \supset X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right),$$

et finalement

$$\bigcup_{n=1}^{+\infty} A_n = X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right).$$

Par suite

$$P_X\left(X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right)\right) = P\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} P(A_n) = \sum_{n=1}^{+\infty} P_X(B_n)$$

car P est une mesure de probabilité et il est facile de voir que les A_n sont deux à deux disjoints. D'où le résultat du théorème.

3.5.2 Démonstration de la proposition 3.2.1

Pour $a > 0$, nous avons :

$$0 \leq P(X = x) \leq P(x - a < X \leq x),$$

soit : $0 \leq P(X = x) \leq F(x) - F(x - a)$, comme F est une fonction continue :

$$\lim_{a \nearrow 0} F(x) - F(x - a) = 0,$$

d'où $P(X = x) = 0$.

3.5.3 Démonstration de la proposition 3.5.1

Soit $\tau = \{y_i; i \in \mathbb{Z}\}$ avec $y_i < y_{i+1}$ pour tout $i \in \mathbb{Z}$ et $\cup_{i\in\mathbb{Z}}]y_{i-1}, y_i] = \mathbb{R}$. On suppose que pour tout $i \in \mathbb{R}$, $|y_i - y_{i-1}| \leq \Delta$ avec $\Delta > 0$. On note $A_i = \{Y \in]y_{i-1}, y_i]\}$ alors :

$$\{X + Y \leq x\} = \cup_{i\in\mathbb{Z}}\{X + Y \leq x\} \cap A_i,$$

or pour tout $i \in \mathbb{Z}$ on a :

$$\cup_{i\in\mathbb{Z}}\{X \leq x - y_i\} \cap A_i \subset \{X + Y \leq x\} \subset \cup_{i\in\mathbb{Z}}\{X \leq x - y_{i-1}\} \cap A_i.$$

soit

$$\sum_{i\in\mathbb{Z}} P(X \leq x - y_i)P(Y \in]y_{i-1}, y_i]) \leq P(X + Y \leq x) \leq \sum_{i\in\mathbb{Z}} P(X \leq x - y_{i-1})P(Y \in]y_{i-1}, y_i]).$$

Or

$$\sum_{i\in\mathbb{Z}} P(X \leq x - y_i)P(Y \in]y_{i-1}, y_i]) = \sum_{i\in\mathbb{Z}} \int_{y_{i-1}}^{y_i} P(X \leq x - y_i) f_Y(y)dy$$

$$\geq \sum_{i\in\mathbb{Z}} \int_{y_{i-1}}^{y_i} F_X(x - y - \Delta) f_Y(y)dy = \int_{\mathbb{R}} F_X(x - y - \Delta) f_Y(y)dy.$$

De même on montre que :

$$\sum_{i\in\mathbb{Z}} P(X \leq x - y_{i-1})P(Y \in]y_{i-1}, y_i]) \leq \int_{\mathbb{R}} F_X(x - y + \Delta) f_Y(y)dy.$$

Il vient donc :

$$\int_{\mathbb{R}} F_X(x - y - \Delta) f_Y(y)dy \leq P(X + Y \leq x) \leq \int_{\mathbb{R}} F_X(x - y + \Delta) f_Y(y)dy.$$

Puis, en faisant tendre Δ vers 0 (des justifications s'imposent pour le passage à la limite) on obtient :

$$F_{X+Y}(x) = \int_{\mathbb{R}} F_X(x - y) f_Y(y)dy.$$

Enfin, en justifiant la dérivation sous le signe somme, on montre qu'aux points x où F_{X+Y} est dérivable on a :

$$f_{X+Y}(x) = \frac{dF_{X+Y}}{dx}(x) = \int_{\mathbb{R}} \frac{dF_X}{dx}(x - y) f_Y(y)dy = \int_{\mathbb{R}} f_X(x - y) f_Y(y)dy.$$

3.5.4 Démonstration de la proposition 3.5.3

Notons $\mathbb{E}[X] = m$ et f_X la densité de X. On constate que :

$$P(|X - m| > \varepsilon) = \int_{\mathbb{R}} 1_{\{|x-m|>\varepsilon\}} f_X(x)dx.$$

Or il est clair que

$$1_{\{|x-m|>\varepsilon\}} \leq \frac{(x - m)^2}{\varepsilon^2} \quad \forall\ x \in \mathbb{R}.$$

Il vient donc

$$\int_{\mathbb{R}} 1_{\{|x-m|>\varepsilon\}} f_X(x)dx \leq \frac{1}{\varepsilon^2} \int_{\mathbb{R}} (x - m)^2 f_X(x)dx = \frac{\mathrm{Var}[X]}{\varepsilon^2}.$$

d'où le résultat.

3.5.5 Applications mesurables

Une application $\phi : \mathbb{R} \to \mathbb{R}$ est mesurable, si l'image réciproque par ϕ de tout ensemble borélien est un ensemble borélien. Ainsi, si X est une v.a.r. sur $(\Omega, \mathcal{F}, P)$, alors $Y = \phi(X)$ est aussi une v.a.r. sur $(\Omega, \mathcal{F}, P)$. En effet, si $\mathcal{B}$ est la tribu des boréliens de $\mathbb{R}$, on a :

$$\forall B \in \mathcal{B}, \{Y \in B\} = \{\phi(X) \in B\} = \{X \in \phi^{-1}(B)\}$$

or $\phi^{-1}(B) \in \mathcal{B}$ donc $\{X \in \phi^{-1}(B)\} \in \mathcal{F}$ puisque X est une v.a.r.

3.6 Tables

3.6.1 Tables de la loi normale

$\Phi(x) = P(X \leq x)$ où $X \sim N(0,1)$ et $x = x_1 + x_2$										
	x_2									
x_1	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Remarque. Si $x \leq 0$ alors utiliser $\Phi(x) = 1 - \Phi(-x)$.

3.6.2 Quantiles

$u_\alpha = \Phi^{-1}(\alpha)$ où $\alpha = \alpha_1 + \alpha_2$										
	α_2									
α_1	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.50	0.0000	0.0025	0.0050	0.0075	0.0100	0.0125	0.0150	0.0175	0.0201	0.0226
0.51	0.0251	0.0276	0.0301	0.0326	0.0351	0.0376	0.0401	0.0426	0.0451	0.0476
0.52	0.0502	0.0527	0.0552	0.0577	0.0602	0.0627	0.0652	0.0677	0.0702	0.0728
0.53	0.0753	0.0778	0.0803	0.0828	0.0853	0.0878	0.0904	0.0929	0.0954	0.0979
0.54	0.1004	0.1030	0.1055	0.1080	0.1105	0.1130	0.1156	0.1181	0.1206	0.1231
0.55	0.1257	0.1282	0.1307	0.1332	0.1358	0.1383	0.1408	0.1434	0.1459	0.1484
0.56	0.1510	0.1535	0.1560	0.1586	0.1611	0.1637	0.1662	0.1687	0.1713	0.1738
0.57	0.1764	0.1789	0.1815	0.1840	0.1866	0.1891	0.1917	0.1942	0.1968	0.1993
0.58	0.2019	0.2045	0.2070	0.2096	0.2121	0.2147	0.2173	0.2198	0.2224	0.2250
0.59	0.2275	0.2301	0.2327	0.2353	0.2378	0.2404	0.2430	0.2456	0.2482	0.2508
0.60	0.2533	0.2559	0.2585	0.2611	0.2637	0.2663	0.2689	0.2715	0.2741	0.2767
0.61	0.2793	0.2819	0.2845	0.2871	0.2898	0.2924	0.2950	0.2976	0.3002	0.3029
0.62	0.3055	0.3081	0.3107	0.3134	0.3160	0.3186	0.3213	0.3239	0.3266	0.3292
0.63	0.3319	0.3345	0.3372	0.3398	0.3425	0.3451	0.3478	0.3505	0.3531	0.3558
0.64	0.3585	0.3611	0.3638	0.3665	0.3692	0.3719	0.3745	0.3772	0.3799	0.3826
0.65	0.3853	0.3880	0.3907	0.3934	0.3961	0.3989	0.4016	0.4043	0.4070	0.4097
0.66	0.4125	0.4152	0.4179	0.4207	0.4234	0.4261	0.4289	0.4316	0.4344	0.4372
0.67	0.4399	0.4427	0.4454	0.4482	0.4510	0.4538	0.4565	0.4593	0.4621	0.4649
0.68	0.4677	0.4705	0.4733	0.4761	0.4789	0.4817	0.4845	0.4874	0.4902	0.4930
0.69	0.4959	0.4987	0.5015	0.5044	0.5072	0.5101	0.5129	0.5158	0.5187	0.5215
0.70	0.5244	0.5273	0.5302	0.5330	0.5359	0.5388	0.5417	0.5446	0.5476	0.5505
0.71	0.5534	0.5563	0.5592	0.5622	0.5651	0.5681	0.5710	0.5740	0.5769	0.5799
0.72	0.5828	0.5858	0.5888	0.5918	0.5948	0.5978	0.6008	0.6038	0.6068	0.6098
0.73	0.6128	0.6158	0.6189	0.6219	0.6250	0.6280	0.6311	0.6341	0.6372	0.6403
0.74	0.6433	0.6464	0.6495	0.6526	0.6557	0.6588	0.6620	0.6651	0.6682	0.6713
0.75	0.6745	0.6776	0.6808	0.6840	0.6871	0.6903	0.6935	0.6967	0.6999	0.7031
0.76	0.7063	0.7095	0.7128	0.7160	0.7192	0.7225	0.7257	0.7290	0.7323	0.7356
0.77	0.7388	0.7421	0.7454	0.7488	0.7521	0.7554	0.7588	0.7621	0.7655	0.7688
0.78	0.7722	0.7756	0.7790	0.7824	0.7858	0.7892	0.7926	0.7961	0.7995	0.8030
0.79	0.8064	0.8099	0.8134	0.8169	0.8204	0.8239	0.8274	0.8310	0.8345	0.8381
0.80	0.8416	0.8452	0.8488	0.8524	0.8560	0.8596	0.8633	0.8669	0.8705	0.8742
0.81	0.8779	0.8816	0.8853	0.8890	0.8927	0.8965	0.9002	0.9040	0.9078	0.9116
0.82	0.9154	0.9192	0.9230	0.9269	0.9307	0.9346	0.9385	0.9424	0.9463	0.9502
0.83	0.9542	0.9581	0.9621	0.9661	0.9701	0.9741	0.9782	0.9822	0.9863	0.9904
0.84	0.9945	0.9986	1.0027	1.0069	1.0110	1.0152	1.0194	1.0237	1.0279	1.0322
0.85	1.0364	1.0407	1.0450	1.0494	1.0537	1.0581	1.0625	1.0669	1.0714	1.0758
0.86	1.0803	1.0848	1.0893	1.0939	1.0985	1.1031	1.1077	1.1123	1.1170	1.1217
0.87	1.1264	1.1311	1.1359	1.1407	1.1455	1.1503	1.1552	1.1601	1.1650	1.1700
0.88	1.1750	1.1800	1.1850	1.1901	1.1952	1.2004	1.2055	1.2107	1.2160	1.2212
0.89	1.2265	1.2319	1.2372	1.2426	1.2481	1.2536	1.2591	1.2646	1.2702	1.2759
0.90	1.2816	1.2873	1.2930	1.2988	1.3047	1.3106	1.3165	1.3225	1.3285	1.3346
0.91	1.3408	1.3469	1.3532	1.3595	1.3658	1.3722	1.3787	1.3852	1.3917	1.3984
0.92	1.4051	1.4118	1.4187	1.4255	1.4325	1.4395	1.4466	1.4538	1.4611	1.4684
0.93	1.4758	1.4833	1.4909	1.4985	1.5063	1.5141	1.5220	1.5301	1.5382	1.5464
0.94	1.5548	1.5632	1.5718	1.5805	1.5893	1.5982	1.6072	1.6164	1.6258	1.6352
0.95	1.6449	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392
0.96	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663
0.97	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335
0.98	2.0537	2.0749	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904
0.99	2.3263	2.3656	2.4089	2.4573	2.5121	2.5758	2.6521	2.7478	2.8782	3.0902

Remarque. Si $\alpha < 0.5$ alors utiliser $u_\alpha = -u_{1-\alpha}$.

Chapitre 4

Vecteurs Aléatoires

4.1 Vecteurs Aléatoires

4.1.1 Couple de variables aléatoires discrètes

Exemple(couple de v.a.r.)
Pour un échantillon de 1000 naissances on a obtenu des résultats concernant le problème des naissances prématurées. On note X le nombre de mois entiers écoulés entre la conception et la naissance d'un enfant et Y le poids de l'enfant à la naissance (en arrondissant au kg inférieur). Les résultats sont résumés en pourcentage dans le tableau qui suit.

$Y \setminus X$	6	7	8	9
1	8	2	0	0
2	2	5	10	3
3	0	5	25	10
4	0	0	10	20

Le tableau ci-dessus résume les informations de nature "statistique" concernant une naissance prématurée de la façon suivante : $P(X = 7; Y = 3) = 0,05$, $P(X = 8; Y = 4) = 0,1$, etc.

L'observation de ce tableau permet de voir que la simple connaissance des lois de X et Y ne suffit pas à "reconstruire" toute l'information contenue dans le tableau. Il s'avère donc nécessaire d'étudier le couple (X, Y) dans sa globalité.

4.1.2 Loi du couple

Nous n'avons manipulé jusqu'à présent que des v.a.r. Considérons maintenant X et Y deux v.a.r. discrètes à valeurs dans E_X et E_Y et intéressons-nous au couple (X, Y).

Définition 4.1.1 *On appelle loi de probabilité conjointe de* (X, Y) *l'application* $p_{X,Y}$

de $E_X \times E_Y$ *dans* $[0,1]$ *définie par :*

$$\forall (x,y) \in E_X \times E_Y, \qquad p_{X,Y}(x,y) = P(X = x \; ; \; Y = y).$$

Les probabilités p_X *et* p_Y*, définies pour* $(x,y) \in E_X \times E_Y$ *par :*

$$p_X(x) = P(X = x) = \sum_{y \in E_Y} p_{X,Y}(x,y)$$

et

$$p_Y(y) = P(Y = y) = \sum_{x \in E_X} p_{X,Y}(x,y),$$

sont appelées lois marginales de X *et* Y*.*

L'application $F_{X,Y}$ *définie sur* $\mathbb{R}^2$ *par :*

$$\forall (x,y) \in \mathbb{R}^2, \qquad F_{X,Y}(x,y) = P(X \leq x \; ; \; Y \leq y)$$

est appelée fonction de répartition de (X,Y)*.*

Remarque 4.1.1 *L'application* $p_{X,Y}$ *satisfait :*

(i) $p_{X,Y}(x,y) \geq 0$ *pour tout* $(x,y) \in E_X \times E_Y$ *;*

(ii) $\sum_{(x,y) \in E_X \times E_Y} p_{X,Y}(x,y) = 1$ *;*

(iii) Si X *et* Y *sont indépendantes alors* $p_{X,Y} = p_X p_Y$ *et* $F_{X,Y} = F_X F_Y$*.*

4.1.3 Tableau de contingence

Lorsque E_X et E_Y sont finis, il est possible de résumer les informations sur le couple (X,Y) dans un tableau. Supposons que $E_X = \{x_1, x_2, \ldots, x_n\}$ et $E_Y = \{y_1, y_2, \ldots, y_m\}$. Notons $p_{ij} = P(X = x_i; Y = y_j)$, $p_{i\bullet} = \sum_{j=1}^m p_{ij} = P(X = x_i)$ et $p_{\bullet j} = \sum_{i=1}^n p_{ij} = P(Y = y_j)$.

$X \backslash Y$	y_1	y_2	$\ldots$	y_m	$P(X = x_i)$
x_1	p_{11}	p_{12}	$\ldots$	p_{1m}	$p_{1\bullet}$
x_2	p_{21}	p_{22}	$\ldots$	p_{2m}	$p_{2\bullet}$
$\vdots$			$\ldots$		$\vdots$
x_n	p_{n1}	p_{n2}	$\ldots$	p_{nm}	$p_{n\bullet}$
$P(Y = y_j)$	$p_{\bullet 1}$	$p_{\bullet 2}$	$\ldots$	$p_{\bullet m}$	1

où $1 = p_{\bullet\bullet} = \sum_{i=1}^n p_{i\bullet} = \sum_{j=1}^m p_{\bullet j} = \sum_{i=1}^n \sum_{j=1}^m p_{ij}$. Comme on peut le voir sur le tableau les lois marginales de X et Y apparaissent respectivement dans la dernière colonne et la dernière ligne du tableau.

Remarque 4.1.2 *Un tableau comme celui-ci est appelé tableau à double-entrée, ou encore, tableau de contingence.*

4.1.4 Espérance

Proposition 4.1.1 *Soient* $f : \mathbb{R}^2 \to \mathbb{R}$ *et* (X, Y) *un couple de v.a.r. Alors* $Z = f(X, Y)$ *est une v.a.r. discrète à valeurs dans* $E_Z = \{f(x, y); (x, y) \in E_X \times E_Y\}$ *et de loi* p_Z *définie par :*

$$\forall z \in E_Z, \quad p_Z(z) = \sum_{\{(x,y)\in E_X\times E_Y; z=f(x,y)\}} p_{X,Y}(x, y).$$

De plus on a :

$$\mathbb{E}[f(X, Y)] = \sum_{\{(x,y)\in E_X\times E_Y\}} f(x, y)p_{X,Y}(x, y).$$

Démonstration. On suppose X et Y définies sur $(\Omega, \mathcal{F}, P)$.

$$Z = f(X, Y) \in \{z = f(x, y); (x, y) \in E_X \times E_Y\} = E_Z.$$

Soit $z \in E_Z$,

$$\{Z = z\} = \bigcup_{\{(x,y)\in E_X\times E_Y; f(x,y)=z\}} \{X = x; Y = y\} \in \mathcal{F}, \tag{4.1}$$

car il s'agit d'une réunion dénombrable d'événements de $\mathcal{F}$. Donc Z est une v.a.r. discrète, de plus d'après (4.1) on a :

$$P(Z = z) = \sum_{\{(x,y)\in E_X\times E_Y; f(x,y)=z\}} p_{X,Y}(x, y).$$

Enfin,

$$\begin{aligned}
\mathbb{E}(Z) &= \sum_{z\in E_Z} zP(Z = z) \\
&= \sum_{z\in E_Z} \sum_{\{(x,y)\in E_X\times E_Y; f(x,y)=z\}} f(x, y)p_{X,Y}(x, y) \\
&= \sum_{\{(x,y)\in E_X\times E_Y\}} f(x, y)p_{X,Y}(x, y).
\end{aligned}$$

Remarque 4.1.3 *Si* $f : \mathbb{R}^2 \to \mathbb{R}^p$ *avec* $f(x, y) = (f_1(x, y), \ldots, f_p(x, y))$ *alors on notera* $\mathbb{E}(f(X, Y))$ *la quantité* $(\mathbb{E}(f_1(X, Y)), \ldots, \mathbb{E}(f_p(X, Y)))$.

4.1.5 Corrélation

Définition 4.1.2 *Soient* X *et* Y *deux v.a.r. du second ordre (admettant des moments d'ordre 2). On appelle covariance de* X *et* Y *la quantité définie par :*

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

On appelle coefficient de corrélation linéaire le réel $\rho_{X,Y}$ *défini par :*

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

Proposition 4.1.2 *Soit (X, Y) un couple de v.a.r. discrètes.*

(i) $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ *;*

(ii) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$ *;*

(iii) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$ *;*

(iv) $\mathrm{Cov}(aX + bY, Z) = a\mathrm{Cov}(X, Z) + b\mathrm{Cov}(Y, Z)$ *pour tout* $(a, b) \in \mathbb{R}^2$ *;*

(v) Si X et Y sont indépendantes alors $\mathrm{Cov}(X, Y) = 0$*, mais la réciproque est fausse en général ;*

(vi) $|\rho_{X,Y}| \leq 1$ *;*

(vii) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$.

4.2 Variables aléatoires vectorielles

On considère $\mathbb{R}^d$ $(d \geq 1)$ muni de la base canonique. Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et $\boldsymbol{X} : \Omega \to \mathbb{R}^d$ une application. Si les composantes $(X_1, \ldots, X_d)$ de $\boldsymbol{X}$ sont des v.a.r. alors $\boldsymbol{X}$ est appelée v.a.r. vectorielle ou vecteur aléatoire réel.

Toutes les notions que nous avons vues jusqu'ici se généralisent aux vecteurs aléatoires. Bien sûr, le passage du cas unidimensionnel au cas multidimensionnel crée un peu de complexité du point de vue de l'analyse. À ce niveau, le cours de MT22 est prérequis.

Définition 4.2.1 *On appelle $P_{\boldsymbol{X}}$ la loi image de P par $\boldsymbol{X}$. Elle est définie par :*

$$P_{\boldsymbol{X}}(B) = P(\boldsymbol{X} \in B) = P(\{\omega \in \Omega \ ; \ \boldsymbol{X}(\omega) \in B\}),$$

pour tout sous-ensemble B de $\mathbb{R}^d$ (en fait, pour tout borélien de $\mathbb{R}^d$).

S'il existe une fonction f définie sur $\mathbb{R}^d$ telle que pour tout sous-ensemble B de $\mathbb{R}^d$ on ait :

$$P_{\boldsymbol{X}}(B) = \int_B f(x_1, \ldots, x_d) dx_1 \ldots dx_d, \tag{4.2}$$

alors le vecteur aléatoire $\boldsymbol{X}$ est dit à densité.

Remarque 4.2.1 *Bien que peu pratique pour expliciter une densité la relation (4.2) nous dit comment calculer $P(X \in B)$ lorsque f est connue.*

4.2.1 Fonction de répartition

Définition 4.2.2 *Soit $\boldsymbol{X}$ un vecteur aléatoire, on appelle f.d.r. de $\boldsymbol{X}$ la fonction $F_{\boldsymbol{X}} : \mathbb{R}^d \to [0, 1]$ définie par :*

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = P_{\boldsymbol{X}}\left(\prod_{i=1}^{d}]\infty, x_i]\right) = P(X_1 \leq x_1; \ldots; X_d \leq x_d),$$

pour tout $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

Proposition 4.2.1 *(i) La f.d.r. $F_{\boldsymbol{X}}$ d'un vecteur aléatoire $\boldsymbol{X}$ satisfait les propriétés suivantes :*

(a) $F_{\boldsymbol{X}}$ est croissante par rapport à chacun de ses arguments ;

(b) $F_{\boldsymbol{X}}$ est continue à droite par rapport à chacun de ses arguments ;

(c) $F_{\boldsymbol{X}}(\boldsymbol{x}) \to 0$ lorsque $\max_{1 \le i \le d} x_i \to -\infty$ et $F_{\boldsymbol{X}}(\boldsymbol{x}) \to 1$ lorsque $\min_{1 \le i \le d} x_i \to +\infty$;

(d) Si on note pour $1 \le i \le d$ et $a_i < b_i$, $\Delta_{a_i b_i} F_{\boldsymbol{X}}(\boldsymbol{x}) =$

$$F_{\boldsymbol{X}}(x_1, \ldots, x_{i-1}, b_i, x_{i+1}, \ldots, x_d) - F_{\boldsymbol{X}}(x_1, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_d)$$

alors

$$\Delta_{a_1 b_1} \Delta_{a_2 b_2} \ldots \Delta_{a_d b_d} F_{\boldsymbol{X}}(\boldsymbol{x}) \ge 0.$$

(ii) Réciproquement, toute fonction $F : \mathbb{R}^d \to [0,1]$ qui satisfait les conditions ci-dessus est la f.d.r. d'un vecteur aléatoire de dimension d.

Remarque 4.2.2 *Pour $d = 1$ de la condition (a) résulte la condition (d) ; ce qui est faux pour $d > 1$. La condition (d) est une condition dite de compatibilité. Elle correspond au fait que l'on doit toujours avoir :*

$$P(X_1 \in [a_1, b_1], \ldots, X_d \in [a_d, b_d]) \ge 0.$$

Définition 4.2.3 *Les fonctions F_i $(1 \le i \le d)$ définies par*

$$F_i(x) = F(+\infty, \ldots, +\infty, x, +\infty, \ldots, +\infty),$$

où x est le i-ème argument de F, sont appelées f.d.r. marginales de $\boldsymbol{X}$; elles sont les f.d.r. des X_i.

4.2.2 Densité

Proposition 4.2.2 *Soit $\boldsymbol{X}$ un vecteur aléatoire de f.d.r. $F_{\boldsymbol{X}}$.*

(i) S'il existe $f_{\boldsymbol{X}} : \mathbb{R}^d \to \mathbb{R}$ telle que pour tout $(x_1, \ldots, x_d) \in \mathbb{R}^d$:

$$F_{\boldsymbol{X}}(x_1, \ldots, x_d) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f(y_1, \ldots, y_d) dy_1 \ldots dy_d,$$

alors f est la densité de $\boldsymbol{X}$.

(ii) Toute fonction f définie sur $\mathbb{R}^d$, positive, telle que :

$$\int_{\mathbb{R}^d} f(y_1, \ldots, y_d) dy_1 \ldots dy_d = 1,$$

est la densité d'un vecteur aléatoire $\boldsymbol{X}$.

Remarque 4.2.3 *La densité (lorsqu'elle existe) et la f.d.r. d'un vecteur aléatoire sont liées par les relations :*

$$F_{\boldsymbol{X}}(x_1, \dots, x_d) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f_{\boldsymbol{X}}(y_1, \dots, y_d) dy_1 \dots dy_d$$

et

$$f_{\boldsymbol{X}}(x_1, \dots, x_d) = \frac{\partial^d F_{\boldsymbol{X}}}{\partial x_1 \dots \partial x_d}(x_1, \dots, x_d),$$

en tout point de continuité $(x_1, \dots, x_d)$ *de* f.

Définition 4.2.4 *Soit* $\boldsymbol{X}$ *un vecteur aléatoire de densité* $f_{\boldsymbol{X}}$. *Alors, pour* $x \in \mathbb{R}$ *:*

$$f_i(x) = \int_{\mathbb{R}^{d-1}} f_{\boldsymbol{X}}(y_1, \dots, y_{i-1}, x, y_{i+1}, \dots, y_d) dy_1 \dots dy_{i-1} dy_{i+1} \dots \dots dy_d,$$

est la densité marginale de X_i.

Remarque 4.2.4 *Il est clair que pour tout* $x \in \mathbb{R}$ $F_i(x) = \int_{-\infty}^{x} f_i(y) dy$.

4.2.3 Espérance et matrice de covariance

Définition 4.2.5 *Soit* $\boldsymbol{X} = (X_1, \dots, X_d)$ *un vecteur aléatoire à valeurs dans* $\mathbb{R}^d$. *L'espérance de* $\boldsymbol{X}$, *lorsqu'elle existe, est un point de* $\mathbb{R}^d$ *défini par :*

$$\mathbb{E}(\boldsymbol{X}) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d))'.$$

Si $\mathbb{E}(\boldsymbol{X}) = 0$, $\boldsymbol{X}$ *est dit centré. On appelle matrice de variance-covariance (ou de covariance) de* $\boldsymbol{X}$, *la matrice* $\Sigma_{\boldsymbol{X}}$ *qui lorsqu'elle existe est définie par :*

$$(\Sigma_{\boldsymbol{X}})_{ij} = \mathrm{Cov}(X_i, X_j) \quad \textit{pour} \quad 1 \leq i, j \leq d.$$

Remarque 4.2.5 *En fait on peut utiliser des notations vectorielles pour définir* $\Sigma_{\boldsymbol{X}}$ *:*

$$\Sigma_{\boldsymbol{X}} = \mathbb{E}(\boldsymbol{X}\boldsymbol{X}') - \mathbb{E}(\boldsymbol{X})\mathbb{E}(\boldsymbol{X})'.$$

Proposition 4.2.3 *(i)* $\Sigma_{\boldsymbol{X}}$ *est symétrique et positive ;*

(ii) Si les composantes de $\boldsymbol{X}$ *sont indépendantes alors la matrice* $\Sigma_{\boldsymbol{X}}$ *est diagonale.*

Démonstration. (i) D'après la remarque 4.2.5 il est facile de voir que :

$$\boldsymbol{x}'\Sigma_{\boldsymbol{X}}\boldsymbol{x} = \mathbb{E}((\boldsymbol{x}'(\boldsymbol{X} - \mathbb{E}(\boldsymbol{X})))^2) \geq 0.$$

(ii) D'autre part si $i \neq j$ et X_i et X_j sont indépendantes alors $\mathrm{Cov}(X_i, X_j) = 0$ donc $\Sigma_{\boldsymbol{X}}$ est diagonale.

4.3 Transformation d'un vecteur aléatoire

Soit $\boldsymbol{X}$ un vecteur aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{F}, P)$ à valeurs dans $\mathbb{R}^d$ de densité f. Soit g une application de $\mathbb{R}^d$ dans $\mathbb{R}$ telle que $Y = g \circ X$ soit une v.a.r. définie sur $(\Omega, \mathcal{F}, P)$ par :

$$Y(\omega) = g(X_1(\omega), \dots, X_d(\omega)).$$

Proposition 4.3.1 *La f.d.r. de Y est définie par :*

$$\forall y \in \mathbb{R}, \quad F_Y(y) = \int_B f(x_1, \dots, x_d) dx_1, \dots dx_d,$$

où

$$B = \{(x_1, \dots, x_d) \in \mathbb{R}^d ; g(x_1, \dots, x_d) \leq y\}.$$

L'espérance de Y est donnée par :

$$\mathbb{E}[Y] = \int_{\mathbb{R}^d} g(x_1, \dots, x_d) f(x_1, \dots, x_d) dx_1, \dots dx_d.$$

Exemple 4.3.1 *Si X et Y sont indépendantes de densités respectives f_X et f_Y, alors $f_{(X,Y)} = f_X f_Y$. Soit $Z = X + Y$ et $z \in \mathbb{R}$ alors :*

$$\begin{aligned} F_Z(z) &= \int \int_{\{(x,y) \in \mathbb{R}^2 ; x+y \leq z\}} f_{(X,Y)}(x,y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{+\infty} f_X(x) \left(\int_{-\infty}^{z-x} f_Y(y) \right) dx \\ &= \int_{-\infty}^{+\infty} f_X(x) F_Y(z-x) dx. \end{aligned}$$

Résultat que nous avions déjà obtenu au chapitre 3.

Soient E et F deux ouverts de $\mathbb{R}^d$, $\boldsymbol{X} = (X_1, \dots, X_d)$ un vecteur aléatoire à valeurs dans E (c'est-à-dire $P(\boldsymbol{X} \in E) = 1$) et g un C^1-difféomorphisme[1] de E sur F. Alors

$$\boldsymbol{Y} = g(\boldsymbol{X}) = \begin{pmatrix} g_1(X_1, \dots, X_d) \\ \vdots \\ g_d(X_1, \dots, X_d) \end{pmatrix}$$

est un vecteur aléatoire défini sur $(\Omega, \mathcal{F}, P)$ à valeurs dans F.

1. il s'agit d'une fonction C^1 telle que g^{-1} existe et est aussi C^1.

Proposition 4.3.2 *La densité $f_{\boldsymbol{Y}}$ du vecteur aléatoire $\boldsymbol{Y}$ est donnée par :*

$$f_{\boldsymbol{Y}}(y_1,\dots,y_d) = f_{\boldsymbol{X}} \circ g^{-1}(y_1,\dots,y_d)|DJ_{g^{-1}(y_1,\dots,y_d)}|1_F(y_1,\dots,y_d),$$

où $DJ_{g^{-1}}$ (resp. DJ_g) est le jacobien de l'application g^{-1} (resp. g) et nous avons $DJ_{g^{-1}} = (DJ_g)^{-1}$ avec :

$$DJ_g = \text{dét}\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_d}{\partial x_1} & \cdots & \frac{\partial g_d}{\partial x_d} \end{pmatrix}.$$

Exemple 4.3.2 *Soit (X,Y) un vecteur aléatoire à valeurs dans $\mathbb{R}^2$ de densité f. Soit $g:\mathbb{R}^2\to\mathbb{R}^2$ définie par $g_1(x,y)=ax+by$ et $g_2(x,y)=cx+dy$ où $ad-bc\neq 0$. Alors g est un C^1-difféomorphisme de $\mathbb{R}^2$ dans $\mathbb{R}^2$ et la densité h de $(U,V)=g(X,Y)$ est donnée par :*

$$h(u,v) = \frac{1}{|ad-bc|} f\left(\frac{du-bv}{ad-bc}, \frac{-bu+av}{ad-bc}\right), \quad \forall (u,v)\in\mathbb{R}^2.$$

Pour $a=b=c=-d=1$ on a : $h(u,v)=1/2f((u+v)/2,(u-v)/2)$. La densité marginale de $U=X+Y$ est donc $h_U(u)=\int_{\mathbb{R}} h(u,v)dv = \int_{\mathbb{R}} f(y,u-y)dy$. Si X et Y sont indépendantes alors on obtient :

$$h_U(u) = \int_{\mathbb{R}} f_X(y) f_Y(u-y)dy = f_X * f_Y(u),$$

qui est un résultat que nous avions déjà obtenu.

4.4 Vecteur aléatoire Gaussien

Définition 4.4.1 *Un vecteur aléatoire $\boldsymbol{X}$ est dit gaussien si pour tout $\boldsymbol{a}\in\mathbb{R}^d$ fixé, $\boldsymbol{a}'\boldsymbol{X} = \sum_{i=1}^d a_i X_i$ est une v.a.r. gaussienne.*

Proposition 4.4.1 *Soit $\boldsymbol{X}$ un vecteur aléatoire gaussien à valeurs dans $\mathbb{R}^d$.*

(i) chacune des composantes de $\boldsymbol{X}$ est gaussienne ;

(ii) si $u:\mathbb{R}^d\to\mathbb{R}^m$ est une application affine ($u(\boldsymbol{x}) = B\boldsymbol{x}+\boldsymbol{b}$) alors $u\circ\boldsymbol{X}$ est gaussien et si $\boldsymbol{Y} = u\circ\boldsymbol{X} = B\boldsymbol{X}+\boldsymbol{b}$ alors $\Sigma_{\boldsymbol{Y}} = B\Sigma_{\boldsymbol{X}}B'$.

Preuve. (i) Prendre $\boldsymbol{a} = (0,\dots,0,1,0,\dots,0)$ où 1 est la i-ème composante de $\boldsymbol{a}$ puis appliquer la définition.
(ii) $Y_i = \underline{B}_i\boldsymbol{X}+b_i$ où $\underline{B}_i$ est la i-ème ligne de la matrice B. On a alors :

$$(\Sigma_{\boldsymbol{Y}})_{ij} = \text{Cov}(Y_i,Y_j) = \mathbb{E}[Y_iY_j] - \mathbb{E}[Y_i]\mathbb{E}[Y_j]$$

$$= \mathbb{E}\left[\left(\sum_{k=1}^{d} B_{ik}X_k + b_i\right)\left(\sum_{l=1}^{d} B_{jl}X_l + b_j\right)\right]$$
$$-\mathbb{E}\left[\left(\sum_{k=1}^{d} B_{ik}X_k + b_i\right)\right]\mathbb{E}\left[\left(\sum_{l=1}^{d} B_{jl}X_l + b_j\right)\right]$$
$$= \sum_{k=1}^{d}\sum_{l=1}^{d} B_{ik}B_{jl}\mathbb{E}[X_kX_l] + b_i\sum_{k=1}^{d} B_{jk}\mathbb{E}[X_k] + b_j\sum_{l=1}^{d} B_{il}\mathbb{E}[X_l] + b_ib_j,$$

d'où :

$$\mathbb{E}[Y_iY_j] - \mathbb{E}[Y_i]\mathbb{E}[Y_j] = \sum_{k=1}^{d}\sum_{l=1}^{d} B_{ik}B_{jl}\mathrm{Cov}(X_k, X_l)$$
$$= \sum_{k=1}^{d}\sum_{l=1}^{d} B_{ik}B_{jl}(\Sigma_{\boldsymbol{X}})_{kl} = \underline{B}_i\Sigma_{\boldsymbol{X}}\underline{B}_j'.$$

Notation. Soit $\boldsymbol{X}$ un vecteur aléatoire gaussien de moyenne $\boldsymbol{\mu}$ et variance-covariance Σ, alors on note $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$.

Proposition 4.4.2 *(i) Si $X_1, \ldots, X_d$ sont des variables aléatoires indépendantes alors $\boldsymbol{X} = (X_1, \ldots, X_d)$ est gaussien ;*

(ii) Si $\boldsymbol{X}$ est un vecteur aléatoire gaussien, de variance-covariance diagonale, alors les composantes de $\boldsymbol{X}$ sont indépendantes ;

(iii) Si $\boldsymbol{X}$ est un vecteur gaussien à valeurs dans $\mathbb{R}^d$ de moyenne $\boldsymbol{\mu}$ et de variance-covariance Σ, pour que $\boldsymbol{X}$ admette une densité il faut et il suffit que Σ soit inversible. Alors on a :

$$f_{\boldsymbol{X}}(x_1, \ldots, x_d) = \frac{1}{(2\pi)^{d/2}\,(\text{dét}(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{\mu})'\,\Sigma^{-1}\,(\boldsymbol{x} - \boldsymbol{\mu})\right),$$

pour tout $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

4.5 Loi Multinomiale

La loi Binomiale se généralise par la loi Multinomiale. En effet, on considère les événements $A_1, \ldots, A_s$ s'excluant mutuellement et ayant pour probabilités respectives $p_1, \ldots, p_s$.

Théorème 4.5.1 *Pour n épreuves répétées la probabilité que A_1 se produise k_1 fois et A_2 se produise k_2 fois et $\ldots$ et A_s se produise k_s fois est*

$$\frac{n!}{k_1!k_2!\ldots k_s!}p_1^{k_1}p_2^{k_2}\ldots p_s^{k_s}$$

avec $k_1 + \ldots + k_s = n$ et $p_1 + \ldots + p_s = 1$.

Application. Si le vecteur aléatoire X suit une telle loi, X est à valeurs dans $\{(k_1, \ldots, k_s) \in \mathbb{N}^s,\ k_1 + \ldots + k_s = n\}$ et satisfait

$$P\left(X_1 = k_1, \ldots, X_s = k_s\right) = \frac{n!}{k_1!k_2!\ldots k_s!} p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s}.$$

Chacune des variables X_i suit la loi binomiale $B(n, p_i)$ dont la moyenne et la variance sont :

$$E[X_i] = np_i,\ Var(X_i) = np_i(1 - p_i)$$

et les covariances sont données par :

$$Cov(X_i, X_j) = -np_ip_j.$$

Exemple. On lance 5 fois de suite un dé équilibré. Quelle est la probabilité de l'événement “ Obtenir deux fois un nombre pair et une fois le 1, une fois le 3 et une fois le 5” ?
Solution. Soient $A_1 = \{\text{obtenir deux fois un nombre pair}\}$, $A_2 = \{\text{obtenir une fois le 1}\}$, $A_3 = \{\text{obtenir une fois le 3}\}$ et $A_4 = \{\text{obtenir une fois le 5}\}$. On a donc une partition $\Omega = A_1 \cup A_2 \cup A_3 \cup A_4$ et $P(A_1) = 1/2$, $P(A_2) = 1/6$, $P(A_3) = 1/6$ et $P(A_4) = 1/6$.
Soit Y_j le nombre de fois que se réalise l'événement A_j en 5 lancers de dé, $j = 1, 2, 3, 4$.
Alors la variable (Y_1, Y_2, Y_3, Y_4) suit la loi multinomiale et

$$P\left(Y_1 = 2, Y_2 = 1, Y_3 = 1, Y_4 = 1\right) = \frac{5!}{2!}\left(\frac{1}{2}\right)^2\left(\frac{1}{6}\right)^3 = \frac{5}{72}.$$

4.6 Indépendance

Soient X et Y deux v.a.r. définies sur un même espace probabilisé $(\Omega, \mathcal{F}, P)$. Nous avons vu dans le chapitre 1 que deux événements A_1 et A_2 sont indépendants si $P(A_1 \cap A_2) = P(A_1)P(A_2)$. Cette notion “se transporte” sur X et Y en définissant leur indépendance de la manière suivante : X et Y sont indépendantes si $\forall B_1, B_2 \in \mathcal{B}$ les événements $\{X \in B_1\}$ et $\{Y \in B_2\}$ sont indépendants ; bien entendu, cette définition est équivalente à : $\forall B_1, B_2 \in \mathcal{B},\ P(X \in B_1; Y \in B_2) = P(X \in B_1)P(Y \in B_2)$.

D'un point de vue pratique, l'indépendance de X et Y signifie que la réalisation de l'une des deux v.a.r. ne contraint pas la réalisation de l'autre.

Exemple 4.6.1 *On lance deux dés honnêtes. Soient X et Z les résultats obtenus. On note $Y = 6 - X$, alors il est facile de voir que :*

$$\begin{cases} P(X = 3; Y = 3) = P(X = 3) = 1/6, \\ P(X = 3)P(Y = 3) = 1/36. \end{cases}$$

Par conséquent X et Y sont dépendantes, ce qui est par ailleurs évident ! En revanche il est clair que X et Z sont indépendantes.

Définition 4.6.1 *Soient $X_1, \ldots, X_d$ des v.a.r. définies sur un même espace probabilisé $(\Omega, \mathcal{F}, P)$. Ces v.a.r. sont dites indépendantes si pour tout $A_1, \ldots, A_d \in \mathcal{B}$ on a :*

$$P(X_1 \in A_1, \ldots, X_d \in A_d) = \prod_{i=1}^{d} P(X_i \in A_i).$$

Indépendance pratique

Proposition 4.6.1 *(i) Soient $X_1, \ldots, X_d$ d v.a.r. discrètes à valeurs dans $E_1, \ldots, E_d$, de lois marginales $p_1, \ldots, p_d$ et de loi conjointe p. Les v.a.r. $X_1, \ldots, X_d$ sont indépendantes si et seulement si :*

$$\forall (x_1, \ldots, x_d) \in E_1 \times \cdots \times E_d, \quad p(x_1, \ldots, x_d) = p_1(x_1) \times \cdots \times p_d(x_d).$$

(ii) Soient $X_1, \ldots, X_d$ d v.a.r. de densités $f_1, \ldots, f_d$. Ces v.a.r. sont indépendantes si et seulement si leur densité conjointe f est égale à : $f_1 \times \cdots \times f_d$ (idem pour la f.d.r.).

(iii) Soient $X_1, \ldots, X_d$ d v.a.r. de densité conjointe f. Si f est à variables séparées, c'est-à-dire si

$$\forall (x_1, \ldots, x_d) \in \mathbb{R}^d, \quad f(x_1, \ldots, x_d) = h_1(x_1) \times \cdots \times h_d(x_d),$$

alors ces v.a.r. sont indépendantes et f_i, la densité marginale de X_i, est égale à $(\int_{\mathbb{R}} h_i(x)dx)^{-1} h_i$.

Proposition 4.6.2 *Les v.a.r. $X_1, \ldots, X_d$ sont indépendantes si et seulement si pour toutes les fonctions réelles h_i pour lesquelles les quantités ci-dessous existent on a :*

$$\mathbb{E}[h_1(X_1) \times \cdots \times h_d(X_d)] = \prod_{i=1}^{d} \mathbb{E}[h_i(X_i)].$$

Remarque 4.6.1 *Pour des fonctions h_i définies par $h_i(x) = 1_{]-\infty, x_i]}(x)$ pour $x \in \mathbb{R}$, on retrouve les f.d.r.*

Proposition 4.6.3 *L'indépendance des v.a.r. $X_1, \ldots, X_n$ entraîne l'indépendance*

(i) de toute sous-suite $(X_{i_1}, \ldots, X_{i_k})$ avec $1 \leq i_1 < \cdots < i_k \leq d$;

(ii) de toute suite de vecteurs aléatoires $(X_1, \ldots, X_{n_1})$, $(X_{n_1+1}, \ldots, X_{n_2})$, …,$(X_{n_k+1}, \ldots, X_n)$ avec $1 \leq n_1 \leq \cdots \leq n_k$ et $1 \leq k \leq n$;

(iii) de toute suite de fonctions $f_1(X_1, \ldots, X_{n_1})$, …, $f_k(X_{n_k+1}, \ldots, X_n)$.

Définition 4.6.2 *Dans une suite infinie de v.a.r. $X_1, X_2, \ldots,$ les v.a.r. sont dites indépendantes si tout sous-ensemble fini de v.a.r. est constitué de v.a.r. indépendantes.*

4.7 Loi de probabilité conditionnelle

4.7.1 v.a. discrète

Soient X et Y deux v.a.r. discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{F}, P)$ à valeurs dans E_X et E_Y. Pour tout $x \in E_X$ tel que $P(X = x) > 0$, la fonction $p_x : E_Y \to [0, 1]$ définie par $p_x(y) = P(Y = y|X = x)$ définit une loi de probabilité sur E_Y. En effet, on a :

(i) $\forall y \in E_Y, \quad p_x(y) \geq 0$,

(ii) $\sum_{y \in E_Y} P(Y = y|X = x) = 1$.

Notons $E_X^* = \{x \in E_X; p_X(x) > 0\}$.

Définition 4.7.1 — *La famille des lois de probabilité $(p_x; x \in E_X^*)$ est appelée famille des lois conditionnelles de Y sachant X.*

— *L'espérance conditionnelle de Y en $\{X = x\}$ $(x \in E_X^*)$ est définie par*

$$\mathbb{E}[Y|X = x] = \sum_{y \in E_Y} yP(Y = y|X = x) = \sum_{y \in E_Y} yp_x(y).$$

— *La variance conditionnelle de Y en $\{X = x\}$ est définie par :*

$$\begin{aligned}\mathrm{Var}(Y|X = x) &= \mathbb{E}[(Y - \mathbb{E}[Y|X = x])^2|X = x] \\ &= \sum_{y \in E_Y} y^2 p_x(y) - (\mathbb{E}[Y|X = x])^2.\end{aligned}$$

Remarque 4.7.1 *Ces définitions sont symétriques en X et Y.*

Définition 4.7.2 *L'espérance conditionnelle de Y sachant X est une v.a.r., notée $\mathbb{E}[Y|X]$, à valeurs dans $E_{Y|X} = \{\mathbb{E}[Y|X = x]; x \in E_X\}$.*

Remarque 4.7.2 *La définition ci-dessus nous permet d'écrire $\mathbb{E}[Y|X] = g(X)$ où g est définie par $g(x) = \mathbb{E}[Y|X = x]$ pour tout $x \in E_X$.*

4.7.2 v.a.r. à densité

Dû au fait que lorsque X est une v.a.r. à densité on a $P(X = x) = 0$ pour tout $x \in \mathbb{R}$, le raisonnement précédent ne peut plus s'appliquer pour des v.a.r. continues. Faisons tout de même le raisonnement suivant. Soit $h > 0$, alors

$$\begin{aligned}P(Y \leq y|X \in [x, x + h]) &= \frac{P(Y \leq y; x \leq X \leq x + h)}{P(x \leq X \leq x + h)} \\ &= \frac{\left(\int_{-\infty}^{y} \int_{x}^{x+h} f(u, v)dudv\right)/h}{(F(x + h) - F(x))/h}.\end{aligned}$$

En supposant la densité de X strictement positive en x et en faisant décroître h vers 0 on obtient la f.d.r. de Y conditionnelle à $\{X = x\}$ que l'on note $F_{Y|X}(\cdot|x)$:

$$\forall y \in \mathbb{R}, \quad F_{Y|X}(y|x) = \frac{1}{f_X(x)} \int_{-\infty}^{y} f(x, v)dv.$$

Ce qui nous permet de définir la densité conditionnelle de Y en $\{X = x\}$.

Définition 4.7.3 *Soit (X, Y) un couple de v.a.r. de densité conjointe f. On note f_X la densité marginale de X définie par : $\forall x \in \mathbb{R}, \quad f_X(x) = \int_{\mathbb{R}} f(x, v)dv$.*

— *Soit $x \in \mathbb{R}$ tel que $f_X(x) > 0$, on appelle densité conditionnelle de Y sachant $\{X = x\}$, la fonction $f_{Y|X}(\cdot|x)$ définie par :*

$$\forall y \in \mathbb{R}, \quad f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f(x, y)}{\int_{\mathbb{R}} f(x, v)dv}.$$

— *On définit l'espérance et la variance conditionnelle de Y sachant $\{X = x\}$, notées respectivement $\mathbb{E}[Y|X = x]$ et $\mathrm{Var}(Y|X = x)$, par :*

$$\mathbb{E}[Y|X = x] = \int_{\mathbb{R}} y f_{Y|X}(y|x)dy$$

et

$$\begin{aligned} \mathrm{Var}(Y|X = x) &= \int_{\mathbb{R}} (y - \mathbb{E}[Y|X = x])^2 f_{Y|X}(y|x)dy \\ &= \int_{\mathbb{R}} y^2 f_{Y|X}(y|x)dy - (\mathbb{E}[Y|X = x])^2. \end{aligned}$$

Remarque 4.7.3 *La définition 4.7.2 reste valable dans ce paragraphe.*

Exemple 4.7.1 *Soit (X, Y) le vecteur aléatoire de loi uniforme sur le disque unité $D = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 < 1\}$. Alors*

$$\forall (x, y) \in \mathbb{R}^2, \quad f_{X,Y}(x, y) = \frac{1}{\pi} 1_D(x, y).$$

Nous avons donc

$$f_X(x) = \int_{\mathbb{R}} f(x, u)du = \frac{2}{\pi}\sqrt{1 - x^2}\, 1_{]-1,1[}(x)$$

et

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{1}{2\sqrt{1 - x^2}} 1_D(x, y), \quad \forall (x, y) \in \mathbb{R}^2, .$$

On calcule alors facilement :

$$\mathbb{E}[Y|X = x] \equiv \int_{\mathbb{R}} y f_{Y|X}(y|x)dy = 1_{]-1,1[}(x)[y^2/2]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} = 0.$$

et

$$\begin{aligned} \mathrm{Var}[Y|X = x] &\equiv \int_{\mathbb{R}} (y - \mathbb{E}[Y|X = x])^2 f_{Y|X}(y|x)dy = \int_{\mathbb{R}} y^2 f_{Y|X}(y|x)dy \\ &= 1_{]-1,1[}(x)[y^3/3]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} = \frac{2}{3}(1 - x^2)^{3/2} 1_{]-1,1[}(x). \end{aligned}$$

4.7.3 Compléments sur l'espérance conditionnelle

Soient X et Y deux v.a.r. conjointes telles que Y admet un moment d'ordre 1.

Proposition 4.7.1 *L'espérance conditionnelle $\mathbb{E}[Y|X]$ est presque sûrement (p.s.) l'unique v.a.r. $g(X)$ (où g est une application borélienne bornée) satisfaisant :*

$$\mathbb{E}[h(X)g(X)] = \mathbb{E}[h(X)Y],$$

pour toute application borélienne bornée h.

Proposition 4.7.2 (espérance totale) $\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$.

Proposition 4.7.3 *(i) Pour tous $a_1, \dots, a_n \in \mathbb{R}$, $\mathbb{E}[\sum_{i=1}^n a_i Y_i | X] = \sum_{i=1}^n a_i \mathbb{E}[Y_i|X]$.*

(ii) Si $Y \geq 0$ p.s. alors $\mathbb{E}[Y|X] \geq 0$ p.s. et si $Y_1 \leq Y_2$ p.s. alors $\mathbb{E}[Y_1|X] \leq \mathbb{E}[Y_2|X]$ p.s.

(iii) Si X et Y sont indépendantes alors $\mathbb{E}[Y|X] = \mathbb{E}[Y]$.

(iv) Si Y et $h(X)Y$ admettent des espérances finies : $\mathbb{E}[h(X)Y|X] = h(X)\mathbb{E}[Y|X]$ p.s.

Proposition 4.7.4 (Lemme de Wald) *Soit $(X_1, X_2, \dots)$ une suite de v.a.r. i.i.d. et N une v.a.r. à valeurs dans $\mathbb{N}$, indépendante de la suite $(X_1, X_2, \dots)$. Soit $S_N = \sum_{i=1}^N X_i$ alors*

$$\mathbb{E}[S_N] = \mathbb{E}[N]\mathbb{E}[X_1].$$

Preuve. $\mathbb{E}[S_N|N = n] = n\mathbb{E}[X_1]$ et $\mathbb{E}[S_N] = \mathbb{E}[\mathbb{E}[S_N|N]] = \mathbb{E}[N\mathbb{E}[X_1]] = \mathbb{E}[N]\mathbb{E}[X_1]$.

4.8 Exercices de cours

EC-4-1 Dans l'exemple de la page 4 :

a- donner E_X et E_Y.

b- calculer les lois marginales p_X et p_Y.

c- a-t-on $p_{X,Y} = p_X p_Y$? Que peut-on en déduire ?

EC-4-2 Vérifier les résultats de la remarque 5.1.1.

EC-4-3 Faire le tableau correspondant à l'exemple de la page 4. En déduire $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathrm{Var}(X)$ et $\mathrm{Var}(Y)$.

EC-4-4 Quelle propriété du tableau caractérise l'indépendance de deux v.a.r. X et Y ?

EC-4-5 Dans l'exemple de la page 4, calculer $\mathbb{E}(XY)$; puis $\mathbb{E}[(X+Y)^2]$ de deux manières différentes.

EC-4-6 Calculer $\mathrm{Cov}(X,Y)$ et $\rho_{X,Y}$ dans l'exemple de la page 4.

EC-4-7 Démontrer la proposition 5.1.2 (utiliser l'inégalité de Cauchy-Scharwz pour le (vi)).

EC-4-8 Soient X et Y deux v.a. telles que $P(X=-1) = P(X=0) = P(X=1) = 1/3$ et $Y = 1_{\{0\}}(X)$. Calculer $\mathrm{Cov}(X,Y)$. X et Y sont-elles indépendantes ?

EC-4-9 Soit (X,Y) un couple de v.a.r. de densité f définie par :

$$\forall (x,y) \in \mathbb{R}^2, \quad f(x,y) = \frac{1}{\pi} 1_D(x,y) \quad \text{où} \quad D = \{(x,y) \in \mathbb{R}^2, x^2+y^2 \leq 1\}.$$

a- $D^+ = (\mathbb{R}^+ \times \mathbb{R}) \cap D$, calculer $P((X,Y) \in D^+)$.

b- $C = [0,1] \times [0,1]$, calculer $P((X,Y) \in C)$.

c- T étant l'intérieur du triangle de sommets $(0,1)$, $(0,-1)$ et $(1,0)$, calculer $P((X,Y) \in T)$.

d- Comment expliquer le $1/\pi$ dans la définition de f ?

EC-4-10 Soient X et Y i.i.d. de loi $U(0,1)$.

a- Donner $F_{X,Y}$.

b- Donner $f_{X,Y}$

c- Comment interpréter la loi du couple (X,Y).

d- Calculer $P((X,Y) \in D)$ où D est le disque de centre O et de rayon 1.

EC-4-11 Soient X et Y i.i.d. de loi $U(0,1)$. Notons $(U,V) = (\min(X,Y), \max(X,Y))$.

a- Donner l'ensemble des valeurs prises par (U,V).

b- Calculer les f.d.r. marginales de U et V.

c- Pourquoi n'a-t-on pas $F_{U,V} = F_U F_V$?

d- U et V sont-elles indépendantes ?

EC-4-12 Vérifier la remarque 4.2.2 pour $d = 2$.

EC-4-13 Soit $T = \{(x, y) \in \mathbb{R}^2; y \geq 1-x\}$, montrer que F définie par $F(x, y) = 1_T(x, y)$ n'est pas une fonction de répartition.

EC-4-14 Calculer la matrice de variance-covariance dans l'exemple de la page 4.

EC-4-15 Calculer la matrice de variance-covariance de (X, Y) et (U, V) de l'exo 11.

EC-4-16 On suppose que $X_1, \dots, X_n$ sont indépendantes.

a- Montrer que $F_{X_1,\dots,X_n} = F_{X_1} \dots F_{X_n}$;

b- Si de plus ces variables aléatoires admettent des densités $f_{X_1}, \dots, f_{X_n}$, montrer que $f_{X_1,\dots,X_n} = f_{X_1} \dots f_{X_n}$;

c- Montrer que si les conditions du b. sont remplies alors :

$$\mathbb{E}(X_1 \dots X_n) = \mathbb{E}(X_1) \dots \mathbb{E}(X_n).$$

EC-4-17 Soit (X, Y) un couple v.a.r. de densité f définie par :

$$f(x, y) = 21_{[0,1]}(x)1_{[0,x]}(y), \quad \forall (x, y) \in \mathbb{R}^2.$$

a- Dans quel ensemble (X, Y) prend ses valeurs ?

b- Calculer $\mathbb{E}(XY)$ puis $\mathrm{Cov}(X, Y)$.

EC-4-18 Soient X, Y et Z i.i.d. de loi $U(0, 1)$. Calculer $P(X \geq YZ)$.

EC-4-19 Soit (U, V, W) le vecteur aléatoire de densité f définie par :

$$f(u, v, w) = \exp(-u)1_F(u, v, w),$$

où $F = \{(u, v, w) \in \mathbb{R}^3; 0 < w < v < u\}$. Montrer que $(U - V, V - W, W)$ est un triplet de v.a.r. i.i.d. de loi $E(1)$.

EC-4-20 Soit $\boldsymbol{X} = (X_1, \dots, X_d)$ un vecteur aléatoire où $X_1, \dots, X_d$ sont i.i.d. de loi $N(\mu, \sigma^2)$. Montrer que $f_{\boldsymbol{X}}$ la densité de $\boldsymbol{X}$ s'écrit :

$$f_{\boldsymbol{X}}(x_1, \dots, x_d) = \frac{1}{(2\pi)^{d/2}\sigma^{2^{1/2}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{d}\left(\frac{x_i - \mu}{\sigma}\right)^2\right).$$

EC-4-21 Soient X et ε deux v.a.r. indépendantes telles que $X \sim N(0, 1)$ et ε satisfait $P(\varepsilon = 1) = P(\varepsilon = -1) = 1/2$. Soit $Y = \varepsilon X$, montrer que :

a- $Y \sim N(0, 1)$;

b- $\mathrm{Cov}(X, Y) = 0$;

c- $P(X + Y = 0) = 1/2$.

En déduire que X et Y ne sont pas indépendantes.

EC-4-22 Dans l'exemple de la page 4 donner :

a- les lois conditionnelles de X sachant Y et de Y sachant X.

b- $\mathbb{E}[Y|X = 8]$ et $\mathrm{Var}[Y|X = 8]$.

EC-4-23 Soient X et Y i.i.d. de loi $P(1/2)$ et $Z = X + Y$. Montrer que la loi de X sachant $Z = k$ $(k > 0)$ est une $B(k, 1/2)$.

EC-4-24 Montrer que si X et Y est un couple de v.a.r. de densité conjointe $f_{X,Y}$ alors :

a- X et Y sont indépendantes si, et seulement si, $f_{X|Y} = f_X$;

b- $f_{X|Y} = f_{Y|X} f_X / f_Y$;

c- à quoi fait penser cette dernière relation ?

EC-4-25 Soit (X, Y) un couple v.a.r. de densité f définie par :

$$f(x, y) = 2 1_{[0,1]}(x) 1_{[0,x]}(y), \quad \forall (x, y) \in \mathbb{R}^2.$$

a- Calculer $f_{X|Y}$ et $f_{Y|X}$.

b- Ces résultats étaient-ils prévisibles ?

c- Calculer $\mathbb{E}(Y|X = x)$ et $\mathrm{Var}(Y|X = x)$.

EC-4-26 Donner la loi de $\mathbb{E}(Y|X)$ dans l'exemple de la page 4.

EC-4-27 Dans l'exercice A.1.25 exprimer la v.a.r. $\mathbb{E}(Y|X)$.

EC-4-28 Soient $N, X_1, X_2, \ldots$ des v.a.r. indépendantes. On suppose que $N \sim G(\theta)$ et $X_i \sim P(\lambda)$ pour tout $i \geq 1$. Calculer $\mathbb{E}(X_1 + \cdots + X_N)$.

4.9 Exercices de travaux dirigés

ETD-4-1 Soit (X, Y) un couple de v.a.r. dont la distribution de probabilité conjointe est donnée par le tableau suivant :

$Y \setminus X$	-1	0	1
0	a	2a	a
1	3a/2	3a	b

On pose : $Z = X + 2Y$ et $T = \sup\{X, Y\}$.

a- A quelle condition le tableau ci-dessus définit-il une distribution de probabilité conjointe (dans la suite on supposera cette condition satisfaite) ?

b- Déterminer, en fonction de a seulement, les lois de X, Z et T.

c- Déterminer l'espérance de X, Y et Z.

d- Déterminer $\mathbb{E}[X^2]$ et $\mathrm{Var}(X)$.

e- Est-ce que les v.a.r. X et Y sont indépendantes ?

ETD-4-2 Soient Y_1, Y_2, Y_3, Y_4 des variables aléatoires indépendantes de loi de Bernoulli $B(p)$ ($p \in [0, 1]$). Pour $1 \leq i \leq 3$ on pose $X_i = Y_i Y_{i+1}$.

a- Quelle est la loi de X_i ($i = 1, 2, 3$) ? sa moyenne ? sa variance ?

b- Calculer $\mathbb{E}(X_i X_{i+1})$ puis $Cov(X_i, X_{i+1})$ pour $1 \leq i \leq 2$.

c- Sans calcul, que vaut $Cov(X_1, X_3)$?

d- Calculer $\mathrm{Var}(X_1 + X_2 + X_3)$.

ETD-4-3 Soient X et Y deux v.a.r. indépendantes de même loi $U(0, 1)$. Déterminer la loi de X/Y.

ETD-4-4 Soit (X, Y) un couple de v.a.r. de densité :

$$f(x, y) = \begin{cases} k, & \text{si} \qquad |x| + |y| \leq 1, \\ 0, & \text{sinon,} \end{cases}$$

a- Déterminer k et les lois de X et Y.

b- Déterminer $Cov(X, Y)$ et étudier l'indépendance de X et Y.

ETD-4-5 Soient X et Y deux v.a.r. indépendantes de même loi de densité f définie par $f(t) = te^{-t} 1_{[0,+\infty[}(t)$. On définit les v.a.r. $U = X + Y$ et $V = X/(X + Y)$.

a- Calculer la densité du couple (U, V).

b- Quelle sont les densités de U et de V ? Que peut-on en déduire ?

ETD-4-6 Soient X et Y deux variables aléatoires indépendantes ayant la même densité f définie par $f(x) = \exp(-x) 1_{]0,+\infty[}(x)$.

a- Calculer la densité $X + Y$.

b- Calculer la fonction de répartition G de X/Y. En déduire sa densité g.

c- Calculer la densité jointe de $(X/Y, Y)$. Retrouver la densité g de X/Y.

ETD-4-7 Soient X et Y deux v.a.r. indépendantes de même densité f définie par $f(x) = x^{-2}1_{[1,+\infty[}(x)$. On pose $U = XY$ et $V = X/Y$.

a- Calculer la loi du couple (U, V). U et V sont-elles indépendantes ?

b- Calculer les lois marginales de U et V.

ETD-4-8 Dupont a rendez-vous avec Durand entre 17 et 18 heures, chacun d'eux a promis de ne pas attendre l'autre plus de 10 minutes. On suppose qu'ils arrivent indépendamment à des instants uniformément distribués entre 17 et 18 heures.

a- Déterminer la probabilité d'une rencontre.

b- Dupont fixe son heure d'arrivée à l'instant x. Quelle probabilité a-t-il de rencontrer Durand ?

c- Arrivant à l'heure x, Dupont ne trouve personne. Quelle probabilité a-t-il de rencontrer Durand ?

ETD-4-9 Soit (X, Y) un couple de v.a.r. de densité :

$$f(x,y) = \begin{cases} ((1+x/2)(1+y/2)-1/2)\exp(-3x/2-y), & \text{si} \quad x \geq 0,\ y \geq 0, \\ 0, & \text{sinon}, \end{cases}$$

a- Vérifier que f est une densité de probabilité.

b- Quelles sont les lois de X et Y ?

c- Que valent $\mathbb{E}(X)$ et $\mathbb{E}(Y)$?

d- Déterminer la loi conditionnelle de X sachant $(Y = Z)$, ainsi que $\mathbb{E}[X \mid Y = Z]$?

ETD-4-10 On considère les v.a.r. X, Y et Z, indépendantes et de même loi $N(0, 1)$. On définit les v.a.r. suivantes : $T = X + Y$, $U = X - Y + 2Z$ et $V = -X + Y + Z$. Quelle est la loi du vecteur (T, U, V) ? Déterminer les lois marginales de T, U et V.

ETD-4-11 Soit (X, Y) le vecteur aléatoire de loi uniforme sur $D = \{(x, y) : (x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$.

a- Calculer les densités de probabilités marginales de X et de Y. Les variables aléatoires X et Y sont-elles indépendantes ?

b- On pose $X = R\cos(\Theta)$ et $Y = R\sin(\Theta)$ avec $-\pi < \Theta \leq \pi$ et $R > 0$. Calculer les lois marginales de R et Θ. R et Θ sont-elles indépendantes ?

c- On pose $Z = X/(X^2 + Y^2)$. Calculer l'espérance mathématique de la variable aléatoire Z.

ETD-4-12 Soient X et Y deux v.a.r. et soit f la densité du couple (X, Y) définie par :

$$f(x,y) = e^{(-x/y)-y}/y 1_{]0,+\infty[}(x) 1_{]0,+\infty[}(y)$$

a- Calculer la densité conditionnelle $f_{X|Y}(x \mid y)$.

b- Calculer $\mathbb{E}\left[X \mid Y = y\right]$.

c- Calculer $\mathbb{E}[X \mid Y]$.

ETD-4-13 La densité conditionnelle de X en $\{Y = y\}$ $(y > 1)$ est donnée par

$$f_{X|Y}(x|y) = xy^2 e^{-yx} 1_{]0,+\infty[}(x).$$

La loi de Y étant définie par $f_Y(y) = y^{-2} 1_{]1,+\infty[}(y)$. Pour $x > 0$, calculer la loi conditionnelle de Y sachant $\{X = x\}$, ainsi que $\mathbb{E}[Y \mid X]$.

ETD-4-14 Soient X et Y deux variables aléatoires indépendantes suivant respectivement des lois de Poisson de paramètres $\lambda > 0$ et $\mu > 0$.

a- Montrer que la variable aléatoire $X + Y$ suit une loi de Poisson dont on déterminera le paramètre.

b- Calculer $\mathbb{E}[X \mid X + Y]$ et $\mathrm{Var}(X \mid X + Y)$.

c- Calculer l'espérance mathématique de la variable aléatoire $\mathbb{E}[X \mid X + Y]$.

ETD-4-15 Soient X et Y deux v.a.r. discrètes indépendantes, de lois respectives $B(n_1, p)$ et $B(n_2, p)$.

a- Déterminer la loi de X en $\{X + Y = n\}$ en indiquant les valeurs possibles de n.

b- Déterminer $\mathbb{E}[X \mid (X + Y)]$.

ETD-4-16 On a une machine comprenant deux composants de durée de vie T_1 et T_2, respectivement. La densité conjointe est donnée par :

$$f(x, y) = c \exp\left(-(x + 2y)\right), \ x \geq 0, \ y \geq 0$$

a- Calculer la constante c.

b- i- La fonction de répartition conjointe.

ii- Les fonctions de répartitions marginales.

iii- Les densités marginales ; est-ce que les v.a. sont indépendantes ?

c- Supposons que les deux composants sont connectés en série. Alors la durée de vie du système, notée T_0 est égale à $T_0 = \min(T_1, T_2)$. Trouver la fonction de répartition et la densité de T_0.

d- Même question lorsque les composants sont connectés en parallèle, c.a.d. $T_0 = \max(T_1, T_2)$.

ETD-4-17 Un dé est jeté 12 fois.

a- Trouver la probabilité que chaque face apparaisse deux fois.

b- Soit X le nombre d'apparitions de 6 et Y celui de 1. Trouver la distribution jointe de (X, Y).

c- Déterminer $\mathrm{Cov}(X, Y)$.

ETD-4-18 On considère un élément d'un système physique donné. On suppose que la durée de vie de cet élément (durée entre la mise en service et la première panne) est une variable aléatoire de loi exponentielle (loi de densité $\lambda e^{-x\lambda} 1_{R^+}(x)$ où $\lambda > 0$). Chaque fois que l'élément tombe en panne il est immédiatement remplacé par un autre dont la durée de vie est indépendante de celle des

éléments précédents et de même loi exponentielle.
Soient $X_1, X_2, ...$ les durées de vie des éléments succesifs. On pose $S_n = X_1 + \cdots + X_n$. Calculer, pour $n \geq 1$ fixé,

a- La loi de S_n (indication : procéder par récurrence sur $n \geq 1$).

b- La loi du couple (S_n, X_{n+1}).

c- La probabilité $P(S_n \leq t < S_n + X_{n+1})$ où $t > 0$.

d- La loi et l'espérance mathématique de la v.a. N représentant le nombre d'éléments utilisées en remplacement avant et jusqu'à l'instant t. (un calcul explicite est demandé, vous pouvez utiliser le résultat obtenu en 3.)

ETD-4-19 a- Soient X et Y deux variables aléatoires de lois respectives $\mathbb{E}(\lambda)$ et $\mathbb{E}(\mu)$. Calculer $P(X < Y)$.

b- Soit $D = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 \leq R\}$ et f l'application définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ telle que :

$$f(x, y) = c(x^2 + y^2)1_D(x, y).$$

i- Pour quelle valeur de c, f est-elle une densité ?

ii- Calculer les densités marginales de X et Y.

c- Soit X une variable aléatoire de loi $E(\lambda)$. Soient a et b deux réels tels que $0 \leq a < b$. Calculer $\mathbb{E}[X|a \leq X \leq b]$.

d- Soit X une variable aléatoire suivant la loi de Poisson $\mathcal{P}(\lambda)$ $(\lambda > 0)$. Soit Y une variable aléatoire qui conditionnellement à X, suit une loi uniforme sur $\{0, 1, \ldots, X\}$. Calculer l'espérance mathématique de Y.

ETD-4-20 On considère le vecteur gaussien centré (X, Y), dont la matrice de covariance est donnée par : $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ avec $|\rho| < 1$. Calculer $\mathbb{E}[\sup(X, Y)]$.

ETD-4-21 Soit $\boldsymbol{X} = (X_1, \ldots, X_n)$ où les X_i sont indépendantes et de loi $N(m_i, 1)$ et $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ définie par : $Y_1 = \frac{1}{n}\sum_{i=1}^n X_i$ et $Y_k = X_k - Y_1$ pour $k \in \{2, \ldots, n\}$.

a- Quelle est la loi de $\boldsymbol{Y}$?

b- Montrer que Y_1 et $(Y_2, \ldots, Y_n)$ sont indépendants.

ETD-4-22 Soit (X, Y) un couple de variables aléatoires de densité $f_{(X,Y)}$ définie par

$$\forall (x, y) \in \mathbb{R}^2 \quad f_{(X,Y)}(x, y) = \frac{1}{8}(x^2 - y^2)\exp(-x)1_{]0,+\infty[}(x)1_{]-x,x[}(y).$$

a- Les variables aléatoires X et Y sont-elles indépendantes ?

b- Quelle est la loi marginale de X ? Appartient-elle à une famille de lois connue ?

c- Calculer $\mathbb{E}[Y|X = x]$ et $\mathrm{Var}[Y|X = x]$. A quoi $\mathrm{Var}[Y|X]$ est-elle presque sûrement égale ?

ETD-4-23 Dans le plan xOy on note C le demi-cercle $\{(x,y) \in \mathbb{R}^2; x^2+y^2=1 \text{ et } x>0\}$ et D la droite d'équation $x=1$. On choisit au hasard un point P de C de manière uniforme et on note $\Theta \in]-\pi/2, \pi/2[$ l'angle orienté (Ox, OP).

a- Donner la densité f_Θ de Θ (vérifier que la fonction donnée est bien une densité).

b- On note C' le demi-disque $\{(x,y) \in \mathbb{R}^2; x^2+y^2 \leq 1 \text{ et } x>0\}$. On choisit un point P' de manière uniforme dans C' et on note Θ' l'angle orienté (Ox, OP'). Calculer la fonction de répartition de Θ' (on pourra s'aider d'un dessin) ; en déduire que Θ et Θ' suivent la même loi de probabilité.

c- On note T l'intersection de D et de la droite (OP) et Y l'ordonnée de T. Illustrer la situation par un dessin. Quelle relation lie Θ et Y ? En déduire la loi de Y en donnant sa densité (on rappelle que $\arctan(x)' = 1/(1+x^2)$).

d- La variable aléatoire $|Y|$ admet-elle une espérance mathématique ? une variance ?

ETD-4-24 Soit Z_1, Z_2 et Z_3 trois variables aléatoires indépendantes suivant des lois exponentielles de paramètres respectifs $\lambda_1 > 0$, $\lambda_2 > 0$ et $\lambda_3 > 0$. On pose $T_1 = \min(Z_1, Z_3)$ et $T_2 = \min(Z_2, Z_3)$.

a- Montrer que pour tous $s, t \in \mathbb{R}^+$ on a

$$\{T_1 > t, T_2 > s\} = \{Z_1 > t; Z_2 > s; Z_3 > \max(s,t)\},$$

en déduire que $P(T_1 > t, T_2 > s) = \exp(-\lambda_1 t - \lambda_2 s - \lambda_3 \max(t,s))$.

b- En déduire les lois marginales de T_1 et T_2. A quelle famille de lois usuelles appartiennent-elles ?

c- On note $F_{1,2}$, F_1 et F_2 les fonctions de répartition respectives de (T_1, T_2), T_1 et T_2. Montrer que pour tous $t, s \in \mathbb{R}^+$ on a

$$F_{1,2}(t,s) = P(T_1 > t, T_2 > s) + F_1(t) + F_2(s) - 1.$$

d- Que vaut $P(Z_i = Z_j)$ pour $1 \leq i \neq j \leq 3$? En déduire que

$$P(T_1 = T_2) = P(Z_3 \leq \min(Z_1, Z_2)).$$

Calculer $P(T_1 = T_2)$ à l'aide d'une intégrale triple. T_1 et T_2 sont-elles indépendantes ?

ETD-4-25 Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi normale de moyenne finie $\mu \in \mathbb{R}$ et de variance $\sigma^2 \in]0, +\infty[$.

a- (i) Calculer $\mathbb{E}(Y_1^3)$ où $Y_1 = (X_1 - \mu)/\sigma$.

(ii) En déduire $\mathbb{E}(X_1^3)$.

(iii) Calculer $\mathrm{Cov}(X_1, X_1^2)$.

b- On suppose désormais que les variables aléatoires X_i sont centrées et réduites. On admet que $\mathrm{Var}(X_1^3) = 15$.

(i) Quelle est la limite presque sûre de $n^{-1}\sum_{i=1}^{n} X_i^3$?

(ii) Donner pour n grand, une approximation de la loi de $\sum_{i=1}^{n} X_i^3/\sqrt{n}$.

ETD-4-26 Soit (X, Y) un couple de v.a. dont la densité est :

$$f_{X,Y}(x, y) = 4xe^{-(x+y)}1_{]0,y[}(x)1_{]0,+\infty[}(y).$$

(a) X et Y sont-elles indépendantes ? Déterminer les lois marginales de X et Y.

(b) Calculer $Cov(X, Y)$. (Conseil : balayer le domaine de sorte que x soit la variable "d'intégration" indépendante.)

(c) On pose : $U = \dfrac{X}{X+Y}$ et $S = X + Y$.

i. Déterminer la loi du couple (U, S).

ii. U et S sont-elles indépendantes ?

On rappelle que : $\forall n \in \mathbb{N}^*,\ \forall \alpha > 0,\ \Gamma(n) = \displaystyle\int_0^\infty t^{n-1}e^{-\alpha t}dt = \frac{(n-1)!}{\alpha^n}$.

ETD-4-27 Soit (X, Y) un couple de variables aléatoires à valeurs dans $E_1 \times E_2 \subset \mathbb{N} \times \mathbb{N}$. On note $p_{X,Y}$ la loi jointe du couple (X, Y) définie par :

$$p_{X,Y}(n, m) = P(X = n; Y = m) = p_{n,m}, \quad \forall (n, m) \in E_1 \times E_2.$$

On notera $p_{\bullet,m} = \sum_{n\in E_1} p_{n,m}$ et $p_{n,\bullet} = \sum_{m\in E_2} p_{n,m}$.
On appelle fonction génératrice jointe du couple (X, Y) l'application $g_{X,Y}$ définie [2] par :

$$g_{X,Y}(u, v) = \sum_{n\in E_1}\sum_{m\in E_2} p_{n,m}u^n v^m.$$

On notera g_X et g_Y les fonctions génératrices respectives de X et Y.

I) a- Montrer que $g_{X,Y}(1, 1) = 1$ et que $g_{X,Y}(u, 1) = g_X(u)$.

b- Démontrer que si X et Y sont indépendantes alors $g_{X,Y} = g_X \times g_Y$.

c- Démontrer que :

— $\dfrac{\partial g_{X,Y}}{\partial u}(1, 1) = \mathbb{E}(X)$;

— $\dfrac{\partial^2 g_{X,Y}}{\partial u^2}(1, 1) = \mathbb{E}(X^2) - \mathbb{E}(X)$;

— $\dfrac{\partial^2 g_{X,Y}}{\partial u \partial v}(1, 1) = \mathbb{E}(XY)$.

d- En déduire des expressions de $\mathrm{Cov}(X, Y)$ et $\mathrm{Var}(X)$ en fonction de $g_{X,Y}$.

2. Tout au long de l'exercice on supposera que les fonctions définies par des sommes sont bien définies et qu'il est permis d'intervertir dérivation et sommation.

II) On suppose maintenant que la fonction génératrice jointe de (X, Y) est définie par :

$$g_{X,Y}(u, v) = \frac{1}{6} + \frac{u}{3} + \frac{v}{3} + \frac{uv}{6}.$$

a- Déterminer g_X, en déduire la loi marginale de X, son espérance, sa variance.

b- Déterminer la loi marginale de Y.

c- Calculer la covariance de X et Y.

d- X et Y sont-elles indépendantes ?

III) On suppose maintenant (X, Y) est à valeurs dans $\{0, 1\} \times \{0, 1\}$ et que :

$$p_{X,Y}(0, 0) = p_{X,Y}(1, 1) = 2p_{X,Y}(0, 1) = 2p_{X,Y}(1, 0).$$

Calculer la fonction génératrice jointe $g_{X,Y}$ de (X, Y).

Chapitre 5

Convergences

5.1 Convergences

Le sens que l'on donne au mot convergence dépend étroitement de l'objet mathématique manipulé. Par exemple, on dit d'une suite numérique $(x_n)_{n\geq 1}$ qu'elle converge vers x si $\lim_{n\to+\infty} |x_n - x| = 0$.

Exemple 5.1.1 *$x_n = 1 + 1/n$ pour $n \geq 1$ et $x = 1$, alors $|x_n - x| = 1/n \to 0$ lorsque $n \to +\infty$.*

Si l'on considère maintenant $(f_n)_{n\geq 1}$ une suite de fonctions définies sur $[0,1]$ nous allons pouvoir donner plusieurs sens à la convergence de la suite $(f_n)_{n\geq 1}$ vers une fonction f.

- *Convergence simple* : on dit que $(f_n)_{n\geq 1}$converge simplement vers f si pour tout $t \in [0,1]$, la suite numérique $(f_n(t))_{n\geq 1}$ converge vers le réel $f(t)$.
- *Convergence uniforme* : on dit que $(f_n)_{n\geq 1}$ converge uniformément vers f si

$$\lim_{n\to+\infty} \sup_{0\leq t\leq 1} |f_n(t) - f(t)| = 0.$$

Comme pour tout $t \in [0,1]$,

$$|f_n(t) - f(t)| \leq \sup_{0\leq t\leq 1} |f_n(t) - f(t)|$$

la convergence uniforme entraîne la convergence simple, la réciproque étant fausse.

Exemple 5.1.2 *$f_n(t) = t(t + 1/n)$ et $f(t) = t^2$. $(f_n)_{n\geq 1}$ converge uniformément vers f sur $[0,1]$ (donc simplement).*

Exemple 5.1.3 *On considère la suite de fonctions :*

$$f_n(t) = \begin{cases} 0 & si \quad t \in [0, 1 - 1/n], \\ 2n(t-1) + 2 & si \quad t \in [1 - 1/n, 1 - 1/(2n)], \\ 2n(1-t) & si \quad t \in [1 - 1/(2n), 1], \end{cases}$$

et $f(t) = 0$. $(f_n)_{n \geq 1}$ converge simplement vers f mais pas uniformément.

Pour les suites de variables aléatoires réelles il y a aussi de nombreuses façons de donner un sens à la convergence vers une autre variable aléatoire. Nous nous limiterons ici à trois types de convergence :

— la convergence presque sûre ;
— la convergence en probabilité ;
— la convergence en distribution ou en loi.

Dans la suite de ce chapitre, X est une variable aléatoire réelle et $(X_n)_{n \geq 1}$ désigne une suite de variables aléatoires réelles. On note F_n la fonction de répartition de X_n pour tout $n \geq 1$ et F celle de X. On suppose les variables X et X_n définies sur un même espace probabilisé $(\Omega, \mathcal{F}, P)$.

5.1.1 Convergence presque sûre

Définition 5.1.1 *On dit que $(X_n)_{n \geq 1}$ converge presque sûrement vers X et on note $X_n \xrightarrow{p.s.} X$, $n \to +\infty$ si :*

$$P\left(\omega \in \Omega \ : \ \lim_{n \to +\infty} X_n(\omega) = X(\omega)\right) = 1.$$

Théorème 5.1.1 (loi forte des grands nombres) *On suppose que les v.a.r. X_n sont i.i.d. et qu'elles admettent une moyenne $\mu = \mathbb{E}(X)$. Soit $S_n = \sum_{i=1}^n X_i/n$, $n \geq 1$. Alors $S_n \xrightarrow{p.s.} \mu$, $n \to +\infty$.*

Exemple 5.1.4 *On lance indéfiniment un dé et on note X_i le résultat obtenu au i-ème lancer. Les X_i sont i.i.d. et $\mathbb{E}(X_i) = (1+2+3+4+5+6)/6 = 7/2$. D'après le théorème 5.1.1, $S_n = (X_1 + \cdots + X_n)/n \xrightarrow{p.s.} 7/2$, $n \to +\infty$.*

Remarque. La v.a.r. S_n est appelée **moyenne empirique** et est souvent notée $\bar{X}_n$ en statistique. L'origine de son nom est la suivante. On définit Z_n une variable aléatoire discrète uniformément distribuée sur $(X_1, \ldots, X_n)$. Alors $\mathbb{E}(Z_n) = \sum_{i=1}^n X_i/n = S_n$, donc S_n peut être vue comme une moyenne.

Exemple 5.1.5 *On répète indéfiniment une expérience en s'intéressant à la survenue d'un événement A. On note X_i la variable aléatoire de Bernoulli égale à 1 si A est réalisé à la i-ème expérience et 0 sinon.* $\mathbb{E}(X_i) = 0 \times P(\bar{A}) + 1 \times P(A) = P(A)$ *et* $S_n = \sum_{i=1}^{n} X_i/n$ *est donc la fréquence de réalisation de l'événement A au cours des n expériences. D'après le théorème 5.1.1 on a :* $S_n \xrightarrow{p.s.} P(A)$, $n \to +\infty$.

Remarque. Ce résultat est la justification d'une intuition bien naturelle : “une probabilité est une fréquence limite”.

Exemple 5.1.6 $X_1, X_2, \ldots$ *sont des v.a.r. i.i.d. de loi* $\mathcal{L}$ *quelconque et le moment d'ordre k* ($\mu_k = \mathbb{E}(X_1^k)$) *existe. Alors* $X_1^k, X_2^k, \ldots$ *est une suite i.i.d. de v.a.r. admettant une moyenne* μ_k *donc, d'après le théorème 5.1.1,* $\sum_{i=1}^{n} X_i^k/n \xrightarrow{p.s.} \mu_k$, $n \to +\infty$.

Proposition 5.1.1 *On suppose la suite* $(X_n)_{n\geq 1}$ *i.i.d. Soit* $g : \mathbb{R} \to \mathbb{R}$ *une application telle que :*

(i) $g(X_n)$ *est une v.a.r. ;*

(ii) $\mathbb{E}(g(X_n))$ *existe.*

Alors $\sum_{i=1}^{n} g(X_i)/n \xrightarrow{p.s.} \mathbb{E}(g(X_1))$, $n \to +\infty$.

Preuve. D'après (i) $(g(X_n))_{n\geq 1}$ est une suite i.i.d. de v.a.r. D'après (ii) $g(X_n)$ admet un moment d'ordre 1. Les hypothèses du théorème 5.1.1 étant satisfaites la conclusion est immédiate.

Lemme 5.1.1 (de Borel-Cantelli) *Soit* $(\Omega, \mathcal{F}, P)$ *un espace probabilisé et* $(A_n)_{n\geq 1}$ *une suite d'événements de* $\mathcal{F}$. *On définit :*

$$A = \limsup_{n\to+\infty} A_n = \bigcap_{n=1}^{+\infty} \bigcup_{m=n}^{+\infty} A_m,$$

alors :

- *si* $\sum_{n=1}^{+\infty} P(A_n) < \infty$ *alors* $P(A) = 0$ *;*
- *si* $\sum_{n=1}^{+\infty} P(A_n)$ *diverge et si les événements* $(A_n)_{n\geq 1}$ *sont indépendants alors* $P(A) = 1$.

Preuve. Voir le document attaché 5.7.

5.1.2 Convergence en probabilité

Définition 5.1.2 *On dit que la suite de v.a.r.* $(X_n)_{n\geq 1}$ *converge en probabilité vers la v.a.r. X et on note* $X_n \xrightarrow{P} X$ *si :*

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P\left(|X_n - X| \geq \varepsilon\right) = 0.$$

Proposition 5.1.2 *La convergence presque sûre entraîne la convergence en probabilité.*

Preuve (facultative).

Théorème 5.1.2 (loi faible des grands nombres) *Si $(X_n)_{n\geq 1}$ est une suite i.i.d. de v.a.r. admettant un moment d'ordre 1 alors $S_n = (X_1 + \cdots + X_n)/n$ converge en probabilité vers $\mathbb{E}(X_1)$.*

Preuve. D'après la proposition 5.1.2, le théorème 5.1.2 est une conséquence immédiate du théorème 5.1.1. Toutefois lorsque les variables X_n admettent une variance finie σ^2 il existe une démonstration simple de ce résultat. Notons tout d'abord que : $\mathbb{E}(S_n) = \mathbb{E}(X_1) = \mu$ et $\mathrm{Var}(S_n) = \sigma^2/n$, donc d'après l'inégalité de Bienaymé-Tchebycheff on a :

$$P(|S_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \to 0, \quad n \to +\infty.$$

D'où le résultat.

Exemple 5.1.7 *Les exemples 5.1.4-5.1.6 restent vrais pour la convergence en probabilité.*

Exemple 5.1.8 *Soit $n \geq 1$ un entier. Il existe deux entiers p et k uniques tels que : $n = 2^k + p$ et $0 \leq p < 2^k$. Soit X_n la variable aléatoire de Bernoulli égale à 1 avec probabilité $1/2^k$. Si $(X_n)_{n\geq 1}$ est une suite de v.a.r. indépendantes, elle converge vers 0 en probabilité mais elle ne converge pas vers 0 presque sûrement.*

5.1.3 Convergence en loi

Définition 5.1.3 *On dit que $(X_n)_{n\geq 1}$ converge en loi ou en distribution vers la v.a.r. X si pour tout x où F est continue on a :*

$$\lim_{n\to+\infty} F_n(x) = \lim_{n\to+\infty} P(X_n \leq x) = F(x).$$

Cette convergence est notée : $X_n \xrightarrow{\mathcal{L}} X$ ou $X_n \xrightarrow{\mathcal{D}} X$, $n \to +\infty$.

Théorème 5.1.3 (de la limite centrale) *On suppose les v.a.r. $X_1, X_2, \ldots$ i.i.d. admettant une moyenne μ et une variance σ^2. Alors si $S_n = \sum_{i=1}^n X_i/n$, on a :*

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathrm{Var}(S_n)}} = \sqrt{n}\left(\frac{S_n - \mu}{\sigma}\right) \xrightarrow{\mathcal{L}} X, \quad n \to +\infty,$$

où X suit une loi normale centrée réduite. Ce résultat peut encore s'écrire :

$$\lim_{n\to+\infty} P\left(\sqrt{n}\left(\frac{S_n-\mu}{\sigma}\right)\leq x\right)=\Phi(x),$$

où Φ est la f.d.r. d'une loi normale centrée réduite.

Remarque. Sous les hypothèses du théorème 5.1.3, le théorème de la limite centrale conduit souvent à faire, pour n assez grand (en pratique on prend souvent $n \geq 30$), l'approximation suivante :

$$P(X_1+\cdots+X_n\leq x)=P\left(\sqrt{n}\left(\frac{S_n-\mu}{\sigma}\right)\leq\frac{x-n\mu}{\sqrt{n}\sigma}\right)\approx\Phi\left(\frac{x-n\mu}{\sqrt{n}\sigma}\right). \quad (5.1)$$

Remarque. Si les v.a.r. X_i sont i.i.d. de loi $N(\mu,\sigma^2)$ alors il y a égalité dans (5.1).

Exemple 5.1.9 *Approximation de la loi binomiale. Si $X\sim B(n,p)$ alors $X = X_1+\cdots+X_n$ où les v.a.r. sont i.i.d. de loi $B(p)$. Pour n "grand" d'après (5.1) on a :*

$$P(X\leq x)\approx\Phi\left(\frac{x-np}{\sqrt{np(1-p)}}\right).$$

Remarque. Dans l'exemple 5.1.9 on approche une loi discrète par une loi continue. Comme $X\in\{0,\ldots,n\}$ on a pour $x\in\{0,\ldots,n\}$: $P(X\leq x)=P(X<x+1)$ mais pour le terme de droite l'approximation (5.1) conduit à : $P(X<x+1)\approx \Phi\left(\frac{x+1-np}{\sqrt{np(1-p)}}\right)$. Pour compenser ce "défaut" on introduit parfois un terme correctif, appelé *correction de continuité* :

$$\forall x\in\{0,1,\ldots,n\}\quad P(X\leq x)\approx\Phi\left(\frac{x+0,5-np}{\sqrt{np(1-p)}}\right).$$

Exemple 5.1.10 *Soit $X_1, X_2, \ldots$ une suite i.i.d. de v.a.r. de même loi que la variable aléatoire X. Dans l'exemple 5.1.5 si $A=\{X\leq x\}$ alors :*

$$F_n(x)=\frac{1}{n}\sum_{i=1}^{n}1(X_i\leq x)\xrightarrow{p.s.}F(x)=P(X\leq x),\quad n\to+\infty.$$

La fonction (aléatoire) F_n s'appelle fonction de répartition empirique. Par le théorème de la limite centrale on obtient de plus que pour tout $x\in\mathbb{R}$ tel que $0<F(x)<1$:

$$\sqrt{n}(F_n(x)-F(x))\xrightarrow{\mathcal{L}}N(0,F(x)(1-F(x))),\quad n\to+\infty.$$

5.2 Fonction caractéristique

Fonction caractéristique

Définition 5.2.1 *Soit X une v.a.r. absolument continue à valeurs dans $\mathbb{R}$ (resp. discrète à valeurs dans $\{x_i; i \in I \subset \mathbb{N}\}$) de densité f_X (resp. de loi $\{p_i; i \in I\}$) ; on appelle fonction caractéristique de X la fonction $\varphi_X : \mathbb{R} \to \mathbb{C}$ définie par :*

$$\varphi_X(t) = \mathbb{E}(\exp(itX)) = \int_{\mathbb{R}} \exp(itx) f(x) dx$$

$$\left(\textit{resp.} \sum_{i \in I} \exp(itx_i) p_i \right).$$

Proposition 5.2.1 *Si φ_X est la fonction caractéristique de la v.a.r. X, la fonction caractéristique φ_Y de $Y = aX + b$ est donnée par :*

$$\varphi_Y(t) = \exp(itb)\varphi_X(at).$$

Proposition 5.2.2 *Soit φ_X la fonction caractéristique d'une v.a.r. X ; on a :*

(i) $\varphi_X(0) = 1$.

(ii) *Pour tout $t \in \mathbb{R}$, $|\varphi_X(t)| \leq 1$.*

(iii) *φ_X est uniformément continue sur $\mathbb{R}$.*

Exemple 5.2.1 *Soit X de loi $E(\lambda)$. Alors φ_X est définie par :*

$$\begin{aligned}\varphi_X(t) = \mathbb{E}(\exp(itX)) &= \int_{\mathbb{R}} \lambda \exp(itx) \exp(-\lambda x) 1_{[0,+\infty[}(x) dx \\ &= \int_0^{+\infty} \lambda \exp((it - \lambda)x) dx = \frac{\lambda}{\lambda - it},\end{aligned}$$

après quelques calculs.

Théorème 5.2.1 *Soit $X_1, X_2, \dots$ une suite de v.a.r. ; cette suite converge vers la v.a.r. X si, et seulement si, pour tout réel t, la suite $(\varphi_{X_n}(t))_{n\geq 1}$ converge vers $\varphi_X(t)$.*

Exemple 5.2.2 *A l'aide du théorème 5.2.1 on peut montrer que si $(X_n)_{n\geq 1}$ est une suite de v.a.r. Binomiale telle que $X_n \sim B(n, \lambda_n)$ et $\lim_{n\to+\infty} n\lambda_n = \lambda \in]0, +\infty[$ alors $(X_n)_{n\geq 1}$ converge en loi vers une variable aléatoire $X \sim P(\lambda)$.*

Pour les variables aléatoires discrètes, il est aussi possible d'utiliser un théorème de continuité des fonction génératrices.

Théorème 5.2.2 *Soit $X_1, X_2, \dots$ une suite de v.a.r. discrètes de fonctions génératrices g_n ; cette suite converge vers la v.a.r. X de fonction génératrice g si, et seulement si, pour tout réel $z \in]0, 1[$, la suite $(g_n(z))_{n\geq 1}$ converge vers $g(z)$.*

Exemple 5.2.3 *A l'aide du théorème 5.2.2 on peut aussi démontrer le résultat de l'exemple 5.2.2.*

TABLE 5.1 – Fonctions caractéristiques des lois usuelles.

Lois	Fonction caractéristique
Bernoulli $B(p)$	$pe^{it} + (1-p)$
Binomiale $B(n,p)$	$(pe^{it} + (1-p))^n$
Poisson $P(\lambda)$	$\exp(-\lambda(1-e^{it}))$
Uniforme sur $[0,1]$	$\frac{1}{it}(\exp(it)-1)$
Normale $N(\mu,\sigma^2)$	$\exp(it\mu - \sigma^2 t^2/2)$
Exponentielle $E(\lambda)$	$\frac{\lambda}{\lambda - it}$

5.3 Méthode de Monte-Carlo

Dans l'Annexe (Chapitre 6,) nous complétons cette section.

5.3.1 Simulation

Proposition 5.3.1 *Soit X une variable aléatoire absolument continue de fonction de répartition F. La variable aléatoire $Y = F(X)$ suit une loi uniforme sur $[0,1]$.*

Démonstration. F est une fonction croissante (voir dans l'Annexe le cas général) à valeurs dans $[0,1]$ et F est continue car X est une v.a.r. absolument continue. La variable Y est donc à valeurs dans $[0,1]$. Soit alors $y \in]0,1[$ et évaluons $P(Y \leq y)$.

L'événement $\{Y \leq y\}$ signifie que $\{F(X) \leq y\}$.

Soit $A = \{x \in \mathbb{R}; F(x) \leq y\}$, A est une partie non vide (car $\lim_{x\to-\infty} F(x) = 0$) et majorée (car $\lim_{x\to+\infty} F(x) = 1$) de $\mathbb{R}$. Notons u sa borne supérieure, comme F est continue, on a par conséquent $F(u) = y$. D'où :

$$P(Y \leq y) = P(F(X) \leq y) = P(X \leq u) = F(u) = y.$$

La variable aléatoire Y suit donc une loi uniforme sur l'intervalle $[0,1]$.

Remarque. La proposition 5.3.1 sert de base à la simulation de variables aléatoires. En effet, les langages de programmation courants (comme le C, le fortran, etc.) possèdent des générateurs aléatoires. Par exemple, en C, chaque appel de la fonction `drand48()` renvoie un nombre aléatoire compris entre 0 et 1 et issu d'une loi uniforme sur $[0,1]$. De plus les appels successifs de cette fonction correspondent à des valeurs indépendantes (donc à des v.a.r. indépendantes). Il est possible d'initialiser le générateur (en utilisant la fonction `srand48()`) soit sur un nombre fixe soit sur une variable dépendant du temps. Dans le premier cas, à chaque appel du programme on génère le même échantillon alors que dans le second cas, à chaque appel du programme, on génère un échantillon différent et indépendant des échantillons obtenus lors des autres appels.

Exemple 5.3.1 *Pour générer un échantillon de 100 valeurs issues d'une loi exponentielle $E(1)$ on procède comme suit :*

$$\forall x \geq 0, \quad F(x) = 1 - \exp(-x) \Leftrightarrow g(y) = F^{-1}(y) = \log(1/(1-y)), \quad \forall y \in [0,1[.$$

Ensuite il suffit d'écrire le petit programme qui suit :

```
#include <stdio.h>
#include <stdlib.h>
#include <stddef.h>
#include <math.h>

double g(double y) /* inverse de la fdr d'une E(1) */
{return -log(1-y);}

void main(int argc, char *argv[])
{
  int i;
  srand48(time(NULL)); /* initialisation al\'eatoire */
  /* ou remplacer par ``srand48(100001);'' */
  for(i=1;i<=100;i++)
  {
    printf("%d-\`eme valeur g\'en\'er\'ee %f\n",i,g(drand48()));
  }
}
```

5.3.2 Approximation stochastique

Ce paragraphe est purement méthodologique. On y trouve, via des exemples, la base de l'utilisation de la méthode de Monte-Carlo.

Exemple 5.3.2 *Approximation d'une intégrale $I = \int_0^1 g(x)dx$.*
Soit $U \sim \mathcal{U}(0,1)$ alors on a $I = \mathbb{E}(g(U))$. Soit $U_1, U_2, \ldots$ une suite i.i.d. de v.a.r. de loi $\mathcal{U}(0,1)$, d'après le théorème 5.1.1, si I est finie on a :

$$\hat{I}_n = \sum_{i=1}^{n} \frac{g(U_i)}{n} \xrightarrow{p.s.} I, \quad n \to +\infty.$$

Programme C :

```
#include <stdio.h>
#include <stdlib.h>
#include <stddef.h>
```

```
#include <math.h>

double g(double x)
{
  return log(1+x);
}

void main(int argc, char *argv[])
{
  int i;
  int n = atoi(argv[1]);
  double Estimateur =0.0;

  for(i=1;i<=n;i++)
    {
      Estimateur += g(drand48());
    }
  Estimateur = Estimateur/(double)n;
  printf("Vraie valeur = %f, Estimation = %f,
          Erreur relative = %f\n", log(4)-1,
          Estimateur, (log(4)-1-Estimateur)/Estimateur);
}
```

Dans la procédure principale ci-dessus, les appels successifs de la fonction `drand48` *correspondent aux réalisations de* $U_1, U_2, \ldots$

Résultat : Pour $g(x) = \log(1+x)$ *on a* $I = [(x+1)\log(x+1) - (x+1)]_0^1 = \log(4) - 1 \approx 0,3863$.

n	*10*	*50*	*100*	*500*	*1000*	*100000*
$\hat{I}_n$	*0,1886*	*0,3444*	*0,3659*	*0,3734*	*0,3796*	*0,3860*
$\left\lvert \frac{\hat{I}_n - I}{I} \right\rvert$	*1,0486*	*0,1216*	*0,0557*	*0,0347*	*0,0176*	*0,0009*

Exemple 5.3.3 *Approximation de l'intégrale* $I = \int_a^b \exp(-x^2)dx$.
Il suffit de remarquer que $\frac{1}{b-a} I = \mathbb{E}(\exp(-X^2))$ *où* $X \sim \mathcal{U}(a,b)$. *Rappelons que si* F *est une f.d.r. continue alors* $X = F^{-1}(U)$ *a pour f.d.r.* F *si* $U \sim \mathcal{U}(0,1)$. *Or pour* $x \in [a,b]$ *on a* $F(x) = \frac{x-a}{b-a}$, *donc pour* $y \in [0,1]$ *on a* $F^{-1}(y) = a + (b-a)y$. *Soit* $U_1, U_2, \ldots$ *une suite i.i.d. de v.a.r. de loi* $\mathcal{U}(0,1)$, *alors :*

$$\hat{I}_n = \frac{b-a}{n} \sum_{i=1}^{n} \exp(-(a + (b-a)U_i)^2) \xrightarrow{p.s.} I, \quad n \to +\infty.$$

Exemple 5.3.4 *Approximation de* $I = \int_0^{+\infty} g(x)dx$.
On a par exemple : $I = \int_0^{+\infty} h(x)\exp(-x)dx$ *où* $h(x) = g(x)\exp(x)$, *donc* $I = \mathbb{E}(h(X))$ *avec* $X \sim E(1)$ *si cette dernière espérance est bien définie.*
pour $x \in [0, +\infty[$ *on a* $F(x) = P(X \leq x) = 1 - \exp(-x)$ *donc pour* $y \in [0, 1[$ *on a* $F^{-1}(y) = \log(1/(1-y))$. *Par conséquent, si* $U \sim \mathcal{U}(0,1)$ *la v.a.r.* $X = F^{-1}(U)$ *a pour f.d.r.* F. *On en déduit, comme dans l'exemple précédent que :*

$$\hat{I}_n = \frac{1}{n}\sum_{i=1}^{n} h\left(\log\left(\frac{1}{1-U_i}\right)\right) \xrightarrow{p.s.} I, \quad n \to +\infty,$$

si $U_1, U_2, \ldots$ *est une suite i.i.d. de v.a.r. de loi* $\mathcal{U}(0,1)$.

Remarque. Dans l'exemple précédent, pour que la qualité de la convergence soit bonne il faut que la densité de X soit élevée sur la partie de $[0, +\infty[$ où la fonction h apporte sa contribution principale à I.

Exemple 5.3.5 *Dans l'exemple 5.3.2 on a :*

$$\mathbb{E}(\hat{I}_n) = I \quad \textit{et} \quad \mathrm{Var}(\hat{I}_n) = \frac{1}{n}\left(\int_0^1 g^2(x)dx - \left(\int_0^1 g(x)dx\right)^2\right) = \frac{\sigma^2}{n}$$

et d'après le théorème de la limite centrale on a :

$$\sqrt{n}\left(\frac{\hat{I}_n - I}{\sigma}\right) \xrightarrow{\mathcal{L}} N(0,1).$$

On en déduit que :

$$\lim_{n\to+\infty} \frac{n}{\sigma^2}\mathbb{E}\left[(\hat{I}_n - I)^2\right] = 1.$$

Par conséquent on a l'équivalence :

$$\mathbb{E}\left[(\hat{I}_n - I)^2\right] \overset{+\infty}{\sim} \frac{1}{n}\left(\int_0^1 g^2(x)dx - \left(\int_0^1 g(x)dx\right)^2\right).$$

Le théorème de la limite centrale nous permet donc d'évaluer la vitesse à laquelle l'erreur quadratique moyenne $\mathbb{E}\left[(\hat{I}_n - I)^2\right]$ tend vers 0 lorsque n tend vers l'infini.

Remarque. Attention dans l'exemple ci-dessus il s'agit d'une erreur en moyenne et non d'une erreur absolue.

5.4 Autres résultats

Les résultats suivants sont souvent très utiles.

Proposition 5.4.1 (Slutsky) *Soient* $(X_n)_{n\geq 1}$*,* $(Y_n)_{n\geq 1}$ *et* $(Z_n)_{n\geq 1}$ *trois suites de v.a.r. telles que* $X_n \xrightarrow{P} x$*,* $Y_n \xrightarrow{P} y$ *et* $Z_n \xrightarrow{\mathcal{L}} Z$ *où* x *et* y *sont deux nombres réels et* Z *une v.a.r. Alors*

$$X_n Z_n + Y_n \xrightarrow{\mathcal{L}} xZ + y, \quad n \to +\infty.$$

Proposition 5.4.2 *Soit* $X_1, X_2, \ldots$ *une suite de v.a.r. i.i.d. admettant une moyenne* μ *et une variance* σ^2*. Soit* $f \in C^1(\mathbb{R})$*, alors si* $S_n = (X_1 + \cdots + X_n)/n$ *on a :*

$$\sqrt{n}(f(S_n) - f(\mu)) \xrightarrow{\mathcal{L}} X, \quad n \to +\infty,$$

où $X \sim N(0, (f'(\mu)\sigma)^2)$*.*

Preuve. On a :

$$\sqrt{n}(f(S_n) - f(\mu)) = f'(S_n^*)\sqrt{n}(S_n - \mu),$$

or $|\mu - S_n^*| \leq |\mu - S_n|$ et d'après le théorème 5.1.2 pour tout $\varepsilon > 0$ on a $\lim_{n\to+\infty} P(|S_n - \mu| > \varepsilon) = 0$ donc $\lim_{n\to+\infty} P(|S_n^* - \mu| > \varepsilon) = 0$ soit $S_n^* \xrightarrow{P} \mu$ lorsque $n \to +\infty$. Comme f' est continue, d'après la proposition 5.1.1 (valable pour la convergence en probabilité) on a $f'(S_n^*) \xrightarrow{P} f'(\mu)$. D'après le théorème 5.1.3 on a $\sqrt{n}(S_n - \mu) \xrightarrow{\mathcal{L}} N(0, \sigma^2)$, $n \to +\infty$. Enfin d'après la proposition de Slutsky on a :

$$f'(S_n^*)\sqrt{n}(S_n - \mu) \xrightarrow{\mathcal{L}} N(0, (\sigma f'(\mu))^2), \quad n \to +\infty,$$

d'où le résultat.

5.5 Exercices de cours

EC-5-1 Vérifier le résultat de l'exemple 5.1.2. Est-ce toujours vrai sur $\mathbb{R}$?

EC-5-2 Vérifier le résultat de l'exemple 5.1.3 (aidez-vous d'un dessin).

EC-5-3 On a observé que sur 100 jours de présence à l'UTC un étudiant a mangé 75 fois au R.U. Qu'est-ce qui justifie de dire que l'étudiant, un jour donné, a une probabilité égale à 0,375 de manger au R.U. ?

EC-5-4 Soit $X_1, X_2, \ldots$ une suite de v.a.r. identiquement distribuées mais non indépendantes. Trouver des contre-exemples aux résultats des théorèmes 5.1.1, 5.1.2 et 5.1.3.

EC-5-5 Soit $X_1, X_2, \ldots$ une suite de v.a.r. i.i.d. de loi $E(1)$. Vers quoi et en quel sens ? converge la suite :

$$\sum_{i=1}^{n} X_i^2/n, \quad n \geq 1 \text{ ?}$$

EC-5-6 Soit $X_1, X_2, \ldots$ une suite de v.a.r. telle que $X_n \sim \delta_{1/n}$ (c'est-à-dire $P(X_n = 1/n) = 1$). Montrer que $(X_n)_{n \geq 1}$ converge en loi vers la variable aléatoire $X \sim \delta_0$.

EC-5-7 Soit $X \sim B(100; 0,3)$. Donner une approximation de $P(X \leq 50)$.

EC-5-8 Soit $X \sim B(1000; 0,002)$. Donner une approximation de $P(X \leq 3)$.

EC-5-9 Calculer la fonction caractéristique d'une loi $B(p)$; d'une loi $B(n; p)$; d'une loi $U(0,1)$.

EC-5-10 Démontrer le résultat de l'exemple 5.2.2.

EC-5-11 Comment simuler une loi $B(p)$? une loi $B(n,p)$?

EC-5-12 On appelle loi de Weibull la loi d'une variable aléatoire ayant pour fonction de répartition F, la fonction définie par :

$$F(x) = \left(1 - \exp\left(-\left(\frac{x}{\sigma}\right)^{\theta}\right)\right) 1_{[0,+\infty[}(x),$$

avec $(\theta, \sigma) \in]0, +\infty[^2$. Comment simuler des variables aléatoires suivant une telle loi ?

EC-5-13 Donner un algorithme permettant d'approcher $I = \int_0^{+\infty} \exp(-x^2)dx$ en simulant des lois $E(1)$.

EC-5-14 Soit $X \sim B(n,p)$ où $p \in]0,1[$. Montrer que pour n assez grand on a approximativement :

$$\sqrt{n}\left(\left(\frac{X}{n}\right)^2 - p^2\right) \sim N(0; 4p^3(1-p)).$$

EC-5-15 On lance indéfiniment un dé équilibré et on note $X_1, X_2, \ldots$ les résultats obtenus. Soit $M_n = \max\{X_1, \ldots, X_n\}$, montrer que M_n converge presque surement vers 6.

5.6 Exercices de travaux dirigés

ETD-5-1 Montrer que dans l'exemple 5.1.8 il y a bien convergence en probabilité, puis montrer, à l'aide du lemme 5.1.1, qu'il n'y a pas convergence presque sûre.

ETD-5-2 Soit Y une v.a. de densité : $be^{-x}1_{[2,\infty[}(x)$ où b est une constante positive.

a- Déterminer la valeur de b.

b- On pose $X = Y - [Y]$, où $[Y]$ est la partie entière de Y. Calculer $E(X)$.

c- Soient $X_1, Y_1, X_2, Y_2, \cdots$ des v.a. indépendantes. On suppose que pour $n \geq 0$, X_n suit la loi de X, et Y_n celle de Y. Etudier la convergence presque sûre de

$$\frac{X_1 - Y_1 + X_2 - Y_2 + \cdots + X_n - Y_n}{n}$$

lorsque $\quad n \to +\infty$.

ETD-5-3 Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes, de même loi, admettant une moyenne μ et une variance $\sigma^2 < +\infty$. On note $S_n = (X_1 + \cdots + X_n)/n$.

a- Calculer $\mathbb{E}(S_n)$ et $\mathrm{Var}(S_n)$.

b- Lorsque n tend vers l'infini, vers quoi converge la suite S_n et en quel sens (justifier la réponse) ?

c- Pour n “grand”, par quelle loi peut-on approcher $\dfrac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathrm{Var}(S_n)}}$ (justifier la réponse) ?

d- En déduire une approximation de $P(S_n \leq x)$.

e- Soit $X \sim B(n,p)$. Donner une approximation de $P(X/n \leq x)$ pour n grand. Par quelle loi préfère-t-on approcher la loi de X lorsque $p \ll np \ll n$ (le symbôle $\ll$ signifie “très inférieur à”) ?

ETD-5-4 Soit $X_1, X_2, \ldots$ une suite de variables aléatoires telle que :

a- $\lim_{n\to+\infty} \mathbb{E}(X_n) = \mu \in \mathbb{R}$;

b- $\lim_{n\to+\infty} \mathrm{Var}(X_n) = 0$.

Montrer que la suite $\{X_n; n \in \mathbb{N}^*\}$ converge en probabilité vers la variable aléatoire certaine égale à μ.

ETD-5-5 Soit $Y_1, Y_2, \ldots$ une suite de variables aléatoires réelles i.i.d. telle que $\mathbb{E}(Y_1) = -1$. On pose $S_n = Y_1 + \cdots + Y_n$, pour $n \geq 1$.

a- Soit α un réel. Etudier la convergence presque sûre de $\frac{S_n}{n^\alpha}$.

b- Etudier la convergence presque sûre de $\frac{1}{n}\sum_{j=1}^{n}(-1)^j Y_j$.

c- Soit $T = \sum_{j=1}^{\infty} 1_{\{S_j \geq 0\}}$. Calculer $P(T = \infty)$.

ETD-5-6 Soit $X_1, X_2, \ldots$ une suite de variables aléatoires de Bernoulli indépendantes de paramètre $1/2$. On pose :

$$S_n = \frac{X_1}{2} + \frac{X_2}{2^2} + \cdots + \frac{X_n}{2^n}.$$

a- Déterminer la loi de S_n.

b- Montrer que la suite $\{S_n ; n \in \mathbb{N}^*\}$ converge en loi vers une variable aléatoire de loi uniforme sur $[0,1]$.

ETD-5-7 Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes et identiquement distribuées de lois normales de moyennes $\mu \in \mathbb{R}$ et de variances $\sigma^2 \in]0,+\infty[$. Soit X une variable aléatoire de loi normale centrée réduite.

a- Montrer que pour tout $\varepsilon > 0$ on a :

$$P(|X| \geq \varepsilon) \leq \frac{1}{\varepsilon}\sqrt{\frac{2}{\pi}}\exp(-\varepsilon^2/2).$$

b- On note $S_n = (X_1 + \cdots + X_n)/n$. Calculer l'espérance mathématique et la variance de S_n. Quelle est la loi de la variable aléatoire $U_n = \sqrt{n}(S_n - \mu)/\sigma$?

c- Soit $a > 0$. Déduire des deux questions précédentes que :

$$\lim_{n\to+\infty} P\left(|S_n - \mu| > a\right) = 0.$$

d- On suppose toujours les variables aléatoires $X_1, X_2, \ldots$ indépendantes et identiquement distribuées, de moyennes $\mu \in \mathbb{R}$, de variances $\sigma^2 \in]0,+\infty[$ mais de loi quelconque. En utilisant l'inégalité de Bienaymé-Tchebyshev, montrer que le résultat de la question précédente reste vrai.

e- Que se passe-t-il si $X_i = X$ pour tout $i \in \mathbb{N}^*$? Quelle hypothèse fondamentale n'est pas satisfaite dans ce cas ?

ETD-5-8 Dans une usine on produit des résistances dont la valeur en Ohms suit une loi normale de moyenne 100 et de variance 0,26. On considère qu'une résistance est commercialisable si sa valeur est de 100 Ohms à 1% près.

a- Quelle est la probabilité p qu'une résistance soit commercialisable ?

b- Les résistances sont fabriquées indépendamment les unes des autres par lots de taille n. On note S_n le nombre de résistances commercialisables dans un lot de taille n. Quelle est la loi de S_n ? sa moyenne ? sa variance ?

c- Soit X une variable aléatoire de loi binomiale de paramètres n et p. Donner, pour n grand, une approximation de $P(X \leq x)$ en fonction de n,p, x et Φ (Φ est la fonction de répartition d'une loi normale centrée réduite).

d- Quelle est la taille minimale des lots qui assure qu'au moins 99% des lots contiennent plus de 90% de resistances commercialisables ?

Calculs numériques : $\Phi(1,961) \approx 0,975$ et $\Phi(-2,232) \approx 0,01$.

ETD-5-9 Soit Y une v.a. réelle uniformément distribuée sur l'intervalle $[3,6]$. Pour tout $n > 0$, on pose $X_n = 5n^2$, si $3 \leq Y \leq 3 + (4/n^2)$ et $X_n = 0$, sinon.

a- Déterminer $E(X_n)$ et $E(X_n^2)$.

b- Calculer $E(X_{n+1}X_{n+2})$.

c- La suite $(X_n)_{n\geq 1}$ converge t-elle p.s. vers une limite ?

d- La suite $(X_n)_{n\geq 1}$ converge t-elle en probabilité vers une limite ?

ETD-5-10 Un appareil électronique fait des mesures dont l'erreur sur la seconde décimale est supposée uniformément distribuée sur $]-0,05,+0,05[$. Quelle est la probabilité que la valeur absolue de l'erreur commise sur la somme de 1000 mesures soit inférieure à 2 ?

Solution. On note $X_1, \ldots, X_{1000}$ les erreurs commises sur chacune des 1000 mesures. Ces variables aléatoires sont indépendantes et identiquement distribuées, de lois $U(-0,05; 0,05)$. L'erreur E commise sur la somme est donc la somme des erreurs commises sur chacune des mesures, soit $E = X_1 + \cdots + X_{1000}$. On a donc

$$\mathbb{E}[E] = 1000 \times \mathbb{E}[X_1] = 1000 \times \int_{\mathbb{R}} 10x 1_{[-0,05;+0,05]}(x)dx$$

$$= 1000 \times 10[x^2/2]_{-0,05}^{0,05} = 10000 \times 0 = 0$$

et

$$\mathrm{Var}[E] = 1000 \times \mathrm{Var}[X_1] = 1000 \times \int_{\mathbb{R}} 10x^2 1_{[-0,05;+0,05]}(x)dx$$

$$= 1000 \times 10[x^3/3]_{-0,05}^{0,05} = 10000 \times 0,000250/3 = 5/6.$$

Or d'après le théorème de la limite centrale on a, en notant $E_n = X_1 + \cdots + X_n$:

$$\frac{E_n - \mathbb{E}[E_n]}{\sqrt{\mathrm{Var}(E_n)}} \xrightarrow{\mathcal{L}} X, \quad n \to +\infty,$$

où X est une variable aléatoire normale centrée-réduite. En considérant que $n = 1000$ est suffisamment grand pour utiliser l'approximation :

$$P((E - \mathbb{E}[E])/\sqrt{\mathrm{Var}(E)} \leq x) = P(E \leq x\sqrt{5/6}) \approx \Phi(x),$$

ou encore

$$P(E \leq x) \approx \Phi(x\sqrt{6/5}),$$

on obtient :

$$P(|E| \leq 2) \approx P(|X| \leq 2\sqrt{6/5}) = \Phi(2\sqrt{6/5}) - \Phi(-2\sqrt{6/5}) = 2\Phi(\sqrt{24/5}) - 1$$

$$\approx 2\Phi(2,19) - 1 \approx 2 \times 0,9857 - 1 = 0,9714..$$

On estime donc que l'on a moins de 3% de chance d'avoir une erreur totale plus grande que 2 en valeur absolue.

ETD-5-11 Dans une population une personne sur 1000 risque un accident d'un certain type chaque année. Une compagnie d'assurance assure 5000 personnes de cette population. Quelle est la probabilité qu'au plus deux des assurés soient victimes dudit accident ?

Solution. On considère que les personnes assurées par cette compagnie d'assurance sont indépendantes et l'on considère qu'elles ont toutes la même probabilité 1/1000 d'être victime de l'accident considéré. Par conséquent, le nombre N de personnes assurées qui auront l'accident est une variable aléatoire discrète de loi binomiale $B(n,p)$ où $n = 5000$ est "grand", $p = 0,001$ est "petit" alors que le produit $np = 5$ est "moyen". Nous sommes donc dans une situation où l'on considère comme acceptable d'approcher la loi binomiale $B(n,p)$ par une loi de poisson $\mathcal{P}(np)$.
On suppose donc que la loi de N est bien approchée par la loi de poisson $\mathcal{P}(5)$ et donc la probabilité qu'il y ait au plus deux assurés victimes dudit accident est :

$$P(N \leq 2) \approx P(X \leq 2)$$

où $X \sim \mathcal{P}(5)$ soit :

$$P(N \leq 2) \approx \exp(-5)(1 + \frac{5}{1!} + \frac{5^2}{2!}) \approx 0,125.$$

ETD-5-12 Utiliser le théorème de la limite centrale pour donner une approximation de la loi Gamma $\gamma(n,\lambda)$ (loi d'Erlang) lorsque n tend vers l'infini. Avec quelles lois classiques peut-on faire un raisonnement identique ?

Solution. Soit Y_n une variable aléatoire de loi $\gamma(n,\lambda)$ où $\lambda > 0$ et $n \in \mathbb{N}^*$. D'après le cours on sait que Y_n est la somme de n variables aléatoires $X_1, \ldots, X_n$, indépendantes et identiquement distribuées, de lois exponentielles $E(\lambda)$. On a donc :

$$Y_n = X_1 + \cdots + X_n.$$

Or les variables aléatoires $X_1, \ldots, X_n$ admettent une moyenne $1/\lambda$ et une variance $1/\lambda^2$. Le théorème de la limite centrale s'applique donc à Y_n, c'est-à-dire que lorsque n tend vers l'infini on a :

$$\frac{Y_n - \mathbb{E}[Y_n]}{\sqrt{\mathrm{Var}(Y_n)}} \xrightarrow{\mathcal{L}} X, \tag{5.2}$$

où X est normale centrée-réduite. Or il est facile de voir que :

$$\mathbb{E}[Y_n] = n/\lambda \quad \text{et} \quad \mathrm{Var}[Y_n] = n/\lambda^2.$$

Donc d'après (5.2), et par définition de la convergence en loi, puisque la fonction de répartition de X est continue en tout point $x \in \mathbb{R}$, on a :

$$\lim_{n \to +\infty} P\left(\frac{Y_n - n/\lambda}{\sqrt{n}/\lambda} \leq x\right) = \Phi(x),$$

où Φ est la fonction de répartition de X. Donc lorsque n tend vers l'infini on fait l'approximation :

$$P\left(\frac{Y_n - n/\lambda}{\sqrt{n}/\lambda} \leq x\right) \approx \Phi(x),$$

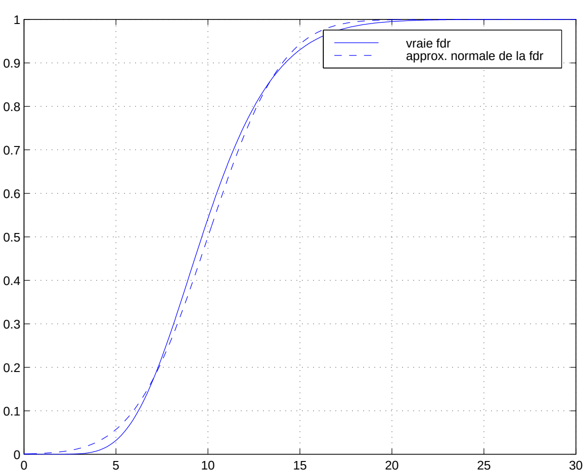
ou encore :

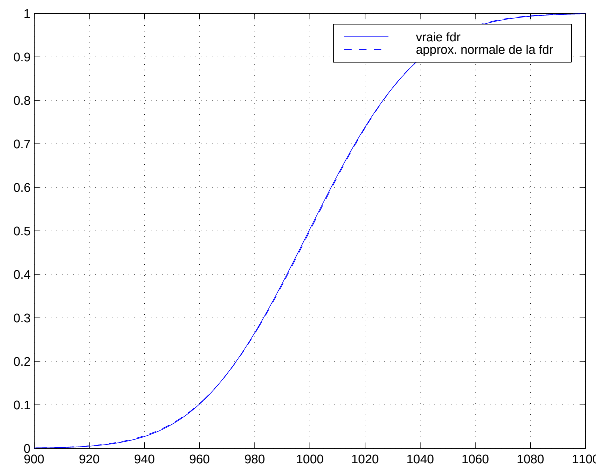
$$P\left(Y_n \leq \frac{x\sqrt{n}+n}{\lambda}\right) \approx \Phi(x).$$

En posant $y = (x\sqrt{n}+n)/\lambda$ on obtient :

$$P\left(Y_n \leq y\right) \approx \Phi\left(\frac{\lambda y}{\sqrt{n}} - \sqrt{n}\right).$$

Remarque. Ci-après, on approche la fonction de répartition d'une loi Gamma $\gamma(n,\lambda)$ en utilisant l'approximation normale. Pour ces figures on a $\lambda = 1$ et n vaut successivement 10 et 1000.





Le raisonnement précédent est valable pour toutes les lois obtenues par convolution de lois identiques (c'est-à-dire pour des sommes de variables aléatoires indépendantes et identiquement distribuées) admettant un moment d'ordre 2. Le procédé d'approximation peut donc être utilisé pour les lois Binomiale, Gamma, Chi-deux, etc.

ETD-5-13 Soit f la fonction définie par

$$f(x) = \alpha(1-x)^{\beta} 1_{[0,1]}(x),$$

où $\beta > 0$.

a- Calculer α pour que f soit une densité.

b- Donner un algorithme permettant de simuler des v.a. de densité f.

c- Soit $\lambda > \beta$ fixé. Donner un algorithme permettant d'approcher l'intégrale I définie par

$$I = \int_0^1 (1-u)^{\lambda} du.$$

ETD-5-14 Trouver un algorithme permettant d'approcher π via la loi forte de grands nombres.

ETD-5-15 On jette indéfiniment un dé équilibré. On note $X_1, X_2, \ldots$ les résultats successifs obtenus et $M_n = \min\{X_1, X_2, \ldots, X_n\}$ pour $n \geq 1$.

a- Soit $\varepsilon \in]0,1[$, exprimer l'événement $\{|M_n - 1| > \varepsilon\}$ à l'aide de $X_1, \ldots, X_n$ (traiter $n = 1$ et $n = 2$ puis généraliser).

b- Montrer que $(M_n)_{n\geq 1} \xrightarrow{P} 1$. Peut-on en déduire la convergence presque-sûre de $(M_n)_{n\geq 1}$ vers 1 (justifiez votre réponse sans chercher à justifier une éventuelle convergence presque-sûre) ?

ETD-5-16 Dans un atelier on assemble 100 ordinateurs par jour. En général, 99% des ordinateurs assemblés fonctionnent lorsqu'on les teste. Soit X le nombre d'ordinateurs en état de marche parmi les ordinateurs assemblés un jour donné.

a- Quelle est la loi de X ?

b- Quelle relation existe-t-il entre les événements $\{X \leq 95\}$ et $\{|X - 99| \geq 4\}$?

c- Donner une majoration de $P(|X-99| \geq 4)$, puis en déduire une majoration de $P(X < 95)$.

ETD-5-17 Un élevage intensif produit des animaux en continu pour la filière alimentaire. Chaque animal produit a une chance sur mille d'être impropre à la consommation.

a- Quelle est la probabilité que le premier animal impropre à la consommation apparaisse après les 100 premiers animaux produits ? En donner une approximation en développant $\varepsilon \to (1-\varepsilon)^\alpha$ au voisinage de 0.

b- Quelle est la probabilité que le premier animal soit impropre à la consommation mais qu'ensuite tous les animaux soient sains ?

c- Un lot de cent mille animaux est mis sur le marché. Soit X le nombre d'animaux du lot impropres à la consommation. Quelle est la loi de X ? Par quelle autre loi peut-on approcher la loi de X ? Quelle est la probabilité approchée,que le lot contienne au moins un animal impropre à la consommation ?

ETD-5-18 On lance n fois un dé et soit F_n la v.a.r. égale à la fréquence d'apparition du six, définie par : $F_n = \frac{X_n}{n}$, où X_n est la v.a.r. égale au nombre d'apparitions du six.
Trouver n tel que : $P(|F_n - \frac{1}{6}| < 0.99) \geq 0.99$.

ETD-5-19 Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de v.a. i.i.d. de loi de Bernoulli, de même paramètre p.

(a) Soit $Y_n = X_n.X_{n+1}$ pour tout $n \in \mathbb{N}^*$; déterminer la loi de Y_n, son espérance et sa variance.

(b) Soit $S_n = \frac{Y_1+\cdots+Y_n}{n}$: déterminer $E(S_n)$ et $Var(S_n)$.

(c) Montrer que pour tout $\epsilon > 0$, on a $\lim_{n\to+\infty} P(|S_n - p^2| \geq \epsilon) = 0$.

ETD-5-20 Soit $X_1, X_2, \dots$ une suite de variables aléatoires indépendantes et identiquement distribuées suivant une loi de Poisson de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$. On définit $S_n = X_1 + \cdots + X_n$ pour $n \geq 1$.

(a) i. Montrer que la fonction génératrice g d'une variable aléatoire X_i est définie par $g(u) = \exp(-\lambda + \lambda u)$. En déduire $\mathbb{E}(X_i)$ et $\operatorname{Var}(X_i)$.

ii. Montrer que $\mathcal{P}(n\lambda)$ est la loi de S_n pour $n \geq 1$.

iii. A l'aide du théorème de la limite centrale, donner une approximation de la fonction de répartition de la loi de S_n pour n grand (préciser les valeurs de n qui rendent cette approximation licite).

iv. Le nombre de clients se présentant au guichet d'une banque un jour j est une variable aléatoire qui suit une loi de Poisson de paramètre 12. On admet que les variables correspondant à des jours différents sont indépendantes. Quelle est la probabilité d'avoir au moins 250 clients durant un mois de 22 jours ouvrables ?

Indication. On a $\Phi(0,862) \approx 0,8$ où Φ est la f.d.r. d'une loi normale centrée réduite.

(b) Calculer $\mathbb{E}[X_1|X_2 = k]$, pour $k \in \mathbb{N}$ et en déduire $\mathbb{E}[X_1]$.

5.7 Démonstrations

Démonstration de la proposition 4.2.2, Convergence en probabilité
On a :

$$P(\omega \in \Omega : \lim_{n \to +\infty} X_n(\omega) = X(\omega)) = 1$$

donc

$$P(\omega \in \Omega : \overline{\lim}_{n \to +\infty} |X_n(\omega) - X(\omega)| \geq \varepsilon) = 0.$$

Il vient donc :

$$\begin{aligned} 0 &= \lim_{n \to +\infty} \searrow P(\cup_{p \geq n}\{|X_p(\omega) - X(\omega)| \geq \varepsilon\}) \\ &\geq \limsup_{n \to +\infty} P(\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \varepsilon), \end{aligned}$$

d'où $\lim_{n \to +\infty} P(|X_n - X| \geq \varepsilon) = 0$.

Démonstration du lemme 4.2.1 Convergence presque-sûre Il s'agit en fait de résoudre l'exercice A.3.2 du chapitre 1.

I. a. Les événements élémentaires qui constituent A sont les $\omega \in \Omega$ réalisant une infinité d'événements A_n. Par conséquent on a $A \subset \cup_{m=n}^{+\infty} A_m$ pour tout $n \geq 1$.

b.

$$P(A) \leq P\left(\bigcup_{m=n}^{+\infty} A_m\right) \leq \sum_{m=n}^{+\infty} P(A_m).$$

Donc $P(A)$ est majorée par le "terme reste" d'une série convergente. Ce "terme reste" tend nécessairement vers 0 lorsque $n \to +\infty$.

II. a.

$$\overline{A} = \overline{\left(\bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty} A_m\right)} = \bigcup_{n=1}^{+\infty}\overline{\left(\bigcup_{m=n}^{+\infty} A_m\right)} = \bigcup_{n=1}^{+\infty}\bigcap_{m=n}^{+\infty}\overline{A}_m.$$

b. Soit f la fonction C^1 définie sur $\mathbb{R}$ par $f(x) = 1 - x - \exp(-x)$. On a $f'(x) = -1 + \exp(-x)$; donc f' est positive sur $\mathbb{R}^-$ et négative sur $\mathbb{R}^+$. f admet donc un maximum en 0 or $f(0) = 0$ ce qui démontre le résultat.

c. Pour $n \geq 1$ on a :

$$\begin{aligned} P\left(\bigcap_{m=n}^{+\infty}\overline{A}_m\right) &= \prod_{m=n}^{+\infty} P(\overline{A}_m) \qquad \text{(par l'indépendance des } A_m\text{)} \\ &= \prod_{m=n}^{+\infty}(1 - P(A_m)) \\ &\leq \prod_{m=n}^{+\infty}\exp(-P(A_m)) \quad \text{(d'après } \mathbf{b}\text{)} \\ &\leq \exp\left(-\sum_{m=n}^{+\infty} P(A_m)\right), \end{aligned}$$

or, comme $\sum_{m=1}^{+\infty} P(A_m) = +\infty$, on a pour tout $n \geq 1$: $\sum_{m=n}^{+\infty} P(A_m) = +\infty$ et donc

$$\exp\left(-\sum_{m=n}^{+\infty} P(A_m)\right) = 0.$$

d. D'après **a** on a :

$$P(\overline{A}) = P\left(\bigcup_{n=1}^{+\infty}\left(\bigcap_{m=n}^{+\infty}\overline{A}_m\right)\right) \leq \sum_{n=1}^{+\infty} P\left(\bigcap_{m=n}^{+\infty}\overline{A}_m\right) = 0$$

d'après le **c**. Il vient donc $P(A) = 1$.

Chapitre 6

Annexe

*

6.1 Génération de nombres aléatoires.

Au cours des dernières décennies nombre de problèmes de nature statistique, mais aussi de nature déterministe, ont vu le développement de méthodes de traitement nécessitant de générer, via un ordinateur, des suites de nombres pouvant être considérées comme étant les réalisations de variables aléatoires indépendantes, voire identiquement distribuées, suivant des lois fixées a priori par l'utilisateur de la méthode. Le nom générique de ces méthodes est *méthodes de Monte-Carlo.*
Le codage machine d'un nombre dans un ordinateur rend a priori vein l'espoir de générer des suites de nombres qui soient réellement les réalisations de variables aléatoires i.i.d. D'autant plus que les suites produites sont généralement récurrentes et périodiques. Mais, nous ne déveloperons pas davantage cet aspect.
La plupart des languages de programmation fournissent deux fonctions :

— `SEED` qui définit la graîne ;

— `RANDOM` dont les appels successifs produisent la suite $(u_n; n \geq 1)$.

Ainsi, l'algorithme :

```
SEED(TIME) ;
for i=1:100,
  U(i)=RANDOM;
end;
```

produira un vecteur $(U_1, \ldots, U_{100})$ qui pourra être considéré comme les réalisations de 100 variables aléatoires i.i.d. de loi $U(0,1)$.
Pour évaluer le caractère i.i.d. suivant un loi $U(0,1)$ des "sorties" $U_1, U_2, dots$ il existe des méthodes statistiques.
Enfin, les problèmes abordés nécessitent rarement que les nombres aléatoires produits soient approximativement distribués suivant des lois $U(0,1)$; il peut être nécessaire de générer des suites de nombres i.i.d. suivant des lois $\mathcal{P}(\lambda)$, $E(\lambda)$, $N(\mu, \sigma^2)$,

etc. Nous verrons alors des méthodes permettant d'obtenir une variable aléatoire de loi connue à partir d'une ou plusieurs variables aléatoires i.i.d. de loi $U(0,1)$.
La nécessité de produire des nombres pseudo-aléatoires de loi connue, est illustrée par l'exemple suivant.

Exemple 6.1.1 Approximations d'intégrales
On souhaite estimer par un calcul approché la valeur d'une intégrale du type :

$$I = \int_{\mathbb{R}^d} f(x_1, \ldots, x_d) dx_1 \ldots dx_d.$$

Il existe des méthodes déterministes pour résoudre un tel problème, comme par exemple la méthode des rectangles ou des trapèzes lorsque $d = 1$. L'efficacité des méthodes déterministes chute parfois lorsque la dimension d de l'espace sur lequel on intégre augmente. On peut alors utiliser des méthodes stochastiques comme méthodes alternatives. Voici un exemple de programme qu'il est parfois possible de mettre en œuvre.

1. `Choisir` M `''assez grand'' et écrire` I `comme l'espérance d'une fonction` h `d'un vecteur aléatoire` $(X_1, \ldots, X_d)$ `admettant la densité` g`, alors` $I = \mathbb{E}(h(X_1, \ldots, X_d))$ ` ;`
2. `Générer un vecteur aléatoire de loi de densité` g ` ;`
3. `Calculer l'image par` h `du vecteur obtenu en 2. ;`
4. `Répéter 2. et 3.` M `fois ;`
5. `Faire la moyenne des résultats obtenus en 4.`

D'après la loi forte des grands nombres le résultat obtenu en 5. doit être d'autant plus proche de I que M est grand.
L'étape 1. est réussie dès lors que nous sommes capables de trouver un fonction g définie sur $\mathbb{R}^d$, positive, d'intégrale égale à 1 et telle que $f(x_1, \ldots, x_d) > 0$ entraîne $g(x_1, \ldots, x_d) > 0$ pour tout $(x_1, \ldots, x_d) \in \mathbb{R}^d$. Alors on a $h = f/g$ avec la convention $0/0 = 0$.

Remarque. Cet exemple, nous montre que les méthodes de Monte-Carlo, requièrent de :
- simuler des variables ou vecteurs aléatoires de lois connues ;
- simuler des variables ou vecteurs aléatoires indépendants ;
- simuler des variables ou vecteurs aléatoires en grande quantié.

6.2 Générer des lois uniformes-Générateurs de nombres pseudo-aléatoires

Rappelons qu'un générateur pseudo-aléatoire est une fonction `RANDOM` (au sens informatique du terme) dont les appels successifs renvoient les termes d'une suite définie par récurrence comme suit :

$$X_{i+1} = h(X_i, \ldots, X_{i-s+1}), \quad \text{pour} \quad i \geq s.$$

Une telle suite est initialisée par la fonction `SEED` qui est définie par la graîne qui initialise les premiers termes $X_1, \ldots, X_s$ de la suite, pour un entier $s \geq 1$ fixé.
On cherche à produire des suites qui satisfassent les critères i.i.d. et loi uniforme sur $[0, 1]$ d'un point de vue statistique. C'est-à-dire que l'on met en œuvre des tests statistiques permettant de vérifier que la suite $(U_i; i \geq 1)$ satisfait :

(i) Pour tout $b \in [0, 1]$:

$$\frac{\text{Card}\{1 \leq i \leq n; 0 \leq U_i \leq b\}}{n} \to b, \quad \text{lorsque} \quad n \to +\infty;$$

(ii) Pour tout $k \geq 2$, les k-uplets $(U_i, \ldots, U_{i+k-1})$ sont uniformément distribués sur $[0, 1]^k$.

Remarque. La propriété (i) est satisfaite par la loi forte des grands nombres si les variables sont i.i.d. de loi $U(0, 1)$. Notons que la propriété (ii) implique l'indépendance des k variables aléatoires composant le vecteur aléatoire.

6.3 Générer des lois non-uniformes

6.3.1 Génération par inversion

On définit F^{-1} l'inverse d'une f.d.r. F par :

$$F^{-1}(x) = \inf\{y \in \mathbb{R}; F(y) \geq x\}.$$

La méthode d'inversion est alors basée sur le résultat suivant.

Théorème 6.3.1 *Soit U une variable aléatoire de loi $U(0, 1)$ et F la f.d.r. d'une variable aléatoire réelle. Alors $X = F^{-1}(U)$ admet F pour f.d.r.*

Preuve. Le théorème est démontré si pour tout $x \in \mathbb{R}$ on a :

$$\{F^{-1}(U) \leq x\} = \{U \leq F(x)\} \tag{6.1}$$

car alors

$$P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

Montrons donc (6.1). Si $U \leq F(x)$ alors

$$\{y \in \mathbb{R}; F(y) \geq U\} \supset \{y \in \mathbb{R}; F(y) \geq F(x)\} \supset [x, +\infty[$$

car F est c.à.d. et croissante. Donc

$$F^{-1}(U) = \inf\{y \in \mathbb{R}; F(y) \geq U\} \leq \inf[x, +\infty[= x.$$

Si $F^{-1}(U) \leq x$ alors $x \in \{y \in \mathbb{R}; F(y) \geq U\}$ car F est croissante sur $\mathbb{R}$. On a donc $F(x) \geq U$. $\square$

Exemple 6.3.1 *Dans un programme est définie la fonction* `BERNOULLI` *par :*

```
fonction BERNOULLI(p)
  if (1-p <= RANDOM ) then
    BERNOULLI=1;
  else
    BERNOULLI=0;
  end;
```

Le programme principal est le suivant :

```
p=0.7;
n=50;
X=0;
for (i=1:n)
  X=X+BERNOULLI(p);
end;
```

On montre que la variable X obtenue en fin d'exécution de ce programme suit une loi $B(50;0.7)$.

Exemple 6.3.2 *Dans un programme est définie la fonction* `EXPO` *par :*

```
fonction EXPO(lambda)
  EXPO=ln(1/(1-RANDOM))/lambda;
```

Le programme principal est le suivant :

```
lambda=3;
n=50;
X=0;
for (i=1:n)
  X=X+EXPO(lambda);
end;
```

On montre que la variable X obtenue en fin d'exécution de ce programme suit une loi (d'Erlang) $\Gamma(50;3)$.

6.3.2 Limites de la génération par inversion

Une des limites évidente de la méthode de génération par inversion est que il est parfois difficile, voire impossible, de calculer explicitement F^{-1}. Voici des pistes qui permettent de contourner cette difficulté. On considère dans la suite U de loi $U(0,1)$.

Inversion par approximation de F^{-1}. Une première méthode consiste à déterminer une fonction H telle que :

$$|F^{-1}(y) - H(y)| \leq 10^{-k}, \quad \forall y \in]0,1[.$$

Alors pour $U \sim U(0,1)$ on aura $H(U) \sim F$ approximativement car $H(U)$ ne differera de $F^{-1}(U)$ qu'après la k^{e} décimale.

Inversion par recherche de zéros. On pourra chercher à résoudre numériquement l'équation $F(X) = U$ à U fixé. La solution X trouvée suivra alors la loi de f.d.r. F. Parmi les méthodes classiques on peut citer la dichotomie ou encore la méthode de Newton.

Inversion par découpage de F. Cette méthode consiste à choisir $k+1$ réels $z_0 < z_1 < \cdots < z_k$ tels que les quantités $F(z_l) - F(z_{l-1})$ soient petites pour tout $1 \leq l \leq k$, avec de plus $F(z_0) = 0$ et $F(z_k) = 1$. Alors on génère X de la manière qui suit :

(a) trouver l'indice l tel que $F(z_l) \leq U < F(z_{l+1})$;

(b) $X = z_l + (z_{l+1} - z_l)\dfrac{U - F(z_l)}{F(z_{l+1}) - F(z_l)}$.

Alors X suit approximativement une loi de f.d.r. F ;

6.3.3 Génération par rejet : esprit de la méthode

Le but est de générer une variable aléatoire X dont la densité f satisfait :

— $\{x \in \mathbb{R}; f(x) > 0\} = [a,b]$, où $a < b$ sont deux réels ;

— $\sup\{f(x); x \in [a,b]\} = c < +\infty$.

Ces propriétés sont illustrées sur la figure 6.1 qui suit. L'algorithme qui permet de

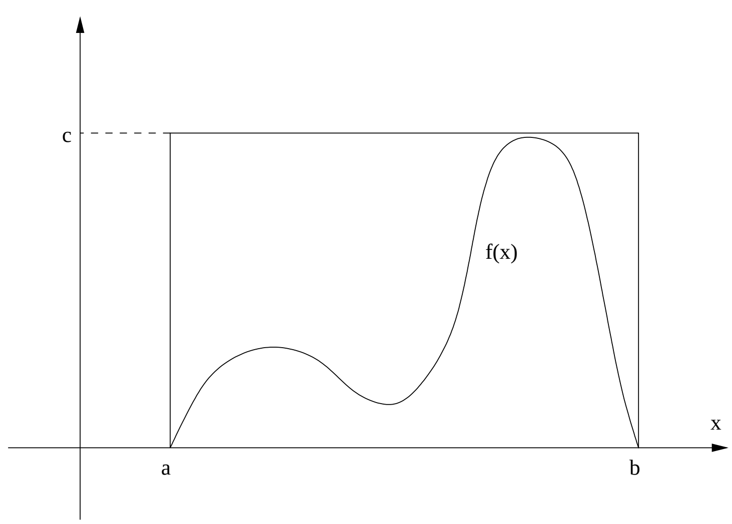


Figure 6.1 – Allure de la densité

générer X est le suivant.

1. `Générer` $R \sim U(a,b)$ `;`

```
2.  Générer S ~ U(0,c) ;
3.  Si S ≤ f(R) alors
           X = R (acceptation)
    Sinon
           revenir en 1. (rejet).
```

Théorème 6.3.2 *La variable aléatoire X obtenue via l'algorithme ci-dessus admet f pour densité.*

Preuve. Remarquons que $(R,S) \sim U([a,b] \times [0,c])$ donc sa densité jointe g est définie par :

$$g(r,s) = \frac{1}{c(b-a)} 1((r,s) \in [a,b] \times [0,c]).$$

Calculons maintenant, pour $x \in \mathbb{R}$ la probabilité $P(X \leq x)$. Tout d'abord on a :

$$P(X \leq x) = \begin{cases} 0 & \text{si} \quad x < a, \\ 1 & \text{si} \quad x > b. \end{cases}$$

Maintenant, pour $x \in [a,b]$, on a :

$$\begin{aligned} P(X \leq x) &= P(R \leq x | S \leq f(R)) \\ &= \frac{P(R \leq x; S \leq f(R))}{P(S \leq f(R))} \\ &= \frac{\int_a^x \int_0^{f(r)} g(r,s) dr ds}{\int_a^b \int_0^{f(r)} g(r,s) dr ds} \\ &= \int_a^x f(r) dr = F(x). \end{aligned}$$

□

6.3.4 Génération par rejet pour un vecteur aléatoire à densité

Soit $(X_1, \ldots, X_d)$ un vecteur aléatoire de densité f définie sur $\mathbb{R}^d$. Soit g une autre densité sur $\mathbb{R}^d$ telle que :

(i) On sache générer $V = (V_1, \ldots, V_d) \sim g$;

(ii) Il existe $\alpha \geq 1$ tel que $f \leq \alpha g$.

Alors la méthode générale de rejection consiste à choisir $X = (X_1, \ldots, X_d)$ de la manière qui suit :

```
1.  Générer V ~ g ;
2.  Générer Y ~ U(0, αg(V_1,...,V_d)) ;
```

```
3.  Si Y ≤ f(V_1, ..., V_d), alors
         X = V (acceptation-fin) ;
    Sinon
         revenir en 1. (rejet) ;
```

Théorème 6.3.3 *Le vecteur aléatoire X simulé par l'algorithme ci-dessus, admet f pour densité.*

Dans la méthode de rejet, la loi de densité g est dite *loi a priori* et g est dite *densité a posteriori*.

Un autre algorithme couramment associé à la méthode de rejet est

```
Répéter
     Simuler X de densité g
    U ← Random
Jusqu'à (cUg(X)) < f(X))
```

Cet algorithme est une conséquence du théorème suivant :

Théorème 6.3.4 *Soient f et g deux densités de probabilités sur $(\mathbb{R}^d, P)$ telles qu'il existe une constante c vérifiant :*

$$\forall x \in \mathbb{R}^d,\ cg(x) \geq f(x).$$

Soit X une variable aléatoire de densité g et U une variable aléatoire de loi uniforme sur $[0,1]$, indépendante de X. Alors la loi conditionnelle de X sachant l'événement "$'cUg(X) < f(X)$"', a pour densité de probabilité f.

L'idée sous-jacente étant celle de partir d'une densité g facile à simuler, pour simuler une loi de densité f quelconque.

Exemple 6.3.3 *Soit n un paramètre positif (strictement)*

1. *Montrer que $g(x,y) = ny^{-n}e^{-xy}1_{[0,\ +\infty)}(x)1_{[1,\ +\infty)}(y)$ est une densité de probabilité.*
   Corrigé. $g(x,y) \geq 0$ *et* $\int_1^{+\infty}\int_0^{+\infty} g(x,y)dxdy = \int_1^{+\infty} ny^{-n}\left(\int_0^{+\infty} e^{-xy}dx\right)dy$

$$= \int_1^{+\infty} ny^{-n-1}dy = n\left[-\frac{y^{-n}}{n}\right]_1^{+\infty} = 1.$$

   donc la fonction $g(x,y)$ est bien une densité de probabilité.
2. *Desormais on considère que $g(x,y)$ de la question précédente est la densité jointe de (X,Y). Déterminer la densité de Y.*
   Corrigé. *La densité de Y est donnée par : pour $y \geq 1$,*
   $g_Y(y) = \int_0^{+\infty} g(x,y)dx = ny^{-n}\left(\int_0^{+\infty} e^{-xy}dx\right) = ny^{-n}\left[\frac{e^{-xy}}{y}\right]_0^{+\infty} = ny^{-n-1}.$
   Soit $g_Y(y) = ny^{-n-1}1_{[1,+\infty)}(y)$.

3. *Pour $y \geq 1$, donner la loi (densité) conditionnelle de X sachant $\{Y = y\}$. De quelle loi usuelle s'agit-il ?*
Corrigé. *Pour $y \geq 1$, $g(x|y) = \frac{g(x,y)}{g_Y(y)} = ye^{-xy}1_{[0,+\infty)}(x)$, c'est à dire que la loi conditionnelle de X sachant $\{Y = y\}$ est la loi exponentielle de paramètre y.*

4. *En justifiant votre réponse dire si l'algorithme de génération de X suivant est correct*

$$Y \longleftarrow \exp(-\ln(Random)/n)$$

$$X \longleftarrow -\ln(Random)/Y$$

$$\textit{retourner } X$$

Corrigé. *La fonction de répartition de Y est donnée par $F_Y(y) = (1 - y^{-n})1_{[1,+\infty)}(y)$ et $F_Y^{-1}(u) = (1-u)^{-1/n}1_{[0,1]}(u)$, de plus en utilisant le fait que si U suit une loi $U(0,1)$, la v.a. $1-U$ aussi. Ces deux arguments permetent de justifier la première ligne de l'algorithme.*
La deuxième ligne est une application directe de la méthode d'inversion dans le cas d'une v.a. de loi exponentielle de paramètre y.
La dernière ligne provient de : La relation $g(x|y) = \frac{g(x,y)}{g_y(y)}$ permet de justifier que si les densités g_Y et $g(.|y)$ sont facilement simulables alors en simulant d'abord Y de densité g_Y, puis ayant obtenu y on simule la densité conditionnelle $g(.|y)$ ceci fournit une simulation de X (voir aussi la preuve du Théorème 2 du Chapitre 2).

6.3.5 Génération par rejet pour un vecteur aléatoire discret

On veut générer X une variable aléatoire discrète à valeurs dans E un ensemble fini ou infini dénombrable. Sa loi p est donnée par $(p(x); x \in E)$. Si l'on sait générer la variable aléatoire discrète V de loi q définie sur E et s'il existe une constante α telle que $p \leq \alpha q$ on procède ainsi :

1. `Générer` V `de loi` q `;`
2. `Générer` Y `de loi` $U(0, \alpha q(V))$ `;`
3. `Si` $Y \leq p(V)$`, alors`
      $X = V$ `(acceptation-fin) ;`
   `Sinon`
      `revenir en 1. (rejet) ;`

Théorème 6.3.5 *La variable aléatoire X résultant de l'algorithme ci-dessus admet la loi p.*

Preuve. Puisque $X = V$ en fin d'algorithme et que $V \in E$ on a $X \in E$. Soit $x \in E$, on a :

$$\begin{aligned} P(X = x) &= P(V = x | Y \leq p(V)) \\ &= \frac{P(V = x; Y \leq p(V))}{P(Y \leq p(V))} \\ &= \frac{P(Y \leq p(x) | V = x) P(V = x)}{P(Y \leq p(V))} \\ &= \frac{P(Y \leq p(x) | V = x) P(V = x)}{\sum_{y \in E} P(Y \leq p(y) | V = y) P(V = y)} \end{aligned}$$

or on montre facilement que

$$P(Y \leq p(y) | V = y) = \frac{p(y)}{\alpha q(y)},$$

d'où il vient :

$$P(X = x) = p(x), \quad \forall x \in E.$$

$\square$

6.4 Méthode de Box-Müller pour la loi normale

On génère un couple de variables aléatoires i.i.d. de loi $N(0,1)$ par l'algorithme que suit.

1. `Générer` U_1 `et` U_2 `i.i.d. de loi` $U(0,1)$ `;`
2. $R = \sqrt{-2\ln(U_1)}$ `et` $\Theta = 2\pi U_2$ `;`
3. $X_1 = R\cos(\Theta)$ `et` $X_1 = R\sin(\Theta)$`.`

Théorème 6.4.1 *Les variables aléatoires X_1 et X_2 sont i.i.d. de loi $N(0,1)$.*

Preuve. Tout d'abord un calcul rapide que les densités g_R et g_Θ de R et Θ sont définies par :

$$g_R(r) = re^{-r^2/2} 1(r \geq 0) \quad \text{et} \quad g_\Theta(\theta) = \frac{1}{2\pi} 1(0 \leq \theta < 2\pi).$$

Soit h la transformation de $(R, \Theta) \mapsto (X_1, X_2)$, *i.e.* :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{lcl} h_1(r,\theta) &=& r\cos(\theta) = x_1, \\ h_2(r,\theta) &=& r\sin(\theta) = x_2. \end{array} \right.$$

On a alors la densité f de (X_1, X_2) qui est donnée par :

$$f(x_1, x_2) = (g \circ h^{-1})(x_1, x_2) |\text{dét}(J_{h^{-1}})(x_1, x_2)|$$

or comme $\text{dét}(J_{h^{-1}}) = 1/\text{dét}(J_h)$ on obtient :

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi}\sqrt{x_1^2 + x_2^2}\exp(-(\sqrt{x_1^2 + x_2^2})^2/2)\frac{1}{\sqrt{x_1^2 + x_2^2}} \\ &= \frac{1}{\sqrt{2\pi}}\exp(-x_1^2/2) \times \frac{1}{\sqrt{2\pi}}\exp(-x_2^2/2), \end{aligned}$$

d'où le résultat. □

6.5 Générer des vecteurs aléatoire avec SCILAB

SCILAB est un logiciel de calcul scientifique gratuit distribué par l'INRIA. Ce logiciel dispose d'une librairie contenant des fonctions qui permettent de générer des tableaux de variables aléatoires suivant de nombreuses lois. Parmi elles on peut citer les lois

— Binomiales, Poisson, géométriques, etc.

— Uniformes, exponentielles, Gaussiennes, Gamma, χ^2, etc.

Fonction `rand`. Cette fonction génère des matrices aléatoires dont les composantes sont i.i.d. de loi $U(0, 1)$ où $N(0, 1)$. Les lignes de programme qui suivent génèrent une matrice réelle `M` de taille 100×50 dont les composantes sont i.i.d. de loi $U(0, 1)$:

```
n=100;
m=50;
M=rand(n,m,'uniform');
```

en changeant `uniform` en `normal` on obtient une matrice de même taille, dont les composantes sont i.i.d. de loi $N(0, 1)$.

Fonction `grand`. La fonction `grand` de SCILAB est très complète. Elle permet entre autres de générer des tableau de variables aléatoires de lois connues. Pour comprendre toutes les fonctionnalités de `grand` nous renvoyons à l'aide de SCILAB. Nous donnons dans les lignes de programme qui suivent quelques éléments de l'utilisation de cette fonction.

```
n=100;
m=50;
moy=10;
e_type=3;
M=grand(n,m,'nor',moy,e_type);
```

Le programme ci-dessus génère une matrice `M` de taille 100×10 dont les composantes sont i.i.d. de loi $N(10, 9)$.

```
n=2;
moy=[1;2];
mat_var_cov=[2,.25;0.25,3];
M=grand(n,'mn',moy,mat_var_cov);
```

Le programme ci-dessus génère une matrice `M` de taille 2×2 dont les colones sont i.i.d. de loi $N(\mu, \Sigma)$, où :

$$\mu = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} 2 & 0,25 \\ 0,25 & 3 \end{pmatrix}.$$

Les fonctions `cdfxxx` Ces fonctions `cdf` (pour *cumulative distribution function* soit f.d.r. en français) renvoient des caractéristiques des lois, telles que les f.d.r. et les quantiles. Par exemple, `cdfnor` renvoie des caractéristiques de la loi normale, `cdfbin` renvoie des caractéristiques de la loi binomiale, etc. Nous renvoyons à l'aide SCILAB pour l'utilisation de ces fonctions.