



Processus stochastiques et modélisation (Cours et exercices corrigés) L3 MIAGE, 2011-2012

Sylvain Rubenthaler

► **To cite this version:**

Sylvain Rubenthaler. Processus stochastiques et modélisation (Cours et exercices corrigés) L3 MIAGE, 2011-2012. Licence. Processus stochastiques et modélisation (Cours et exercices corrigés) L3 MIAGE, 2011-2012, Université Nice Sophia Antipolis, 2011, pp.95. <cel-00867016>

HAL Id: cel-00867016


<https://cel.archives-ouvertes.fr/cel-00867016>

Submitted on 27 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Processus stochastiques et modélisation
(Cours et exercices corrigés)
L3 MIAGE, 2011-2012



LOGO_UNC_couleurs_web.png

Sylvain Rubenthaler

Table des matières

1	Événements aléatoires et variables aléatoires	1
1.1	Événements et probabilités	1
1.2	Variables aléatoires	3
1.3	Espérance et moments	6
1.3.1	Définitions	6
1.3.2	Propriétés	7
1.4	Fonctions de répartition jointes	8
1.4.1	Définitions générales	8
1.4.2	Indépendance	10
1.5	Sommes et convolutions	12
1.6	Changement de variable	13
1.7	Lois de probabilités usuelles (à connaître par cœur)	15
1.7.1	Lois discrètes	15
1.7.2	Lois continues	15
1.8	Exercices	15
1.8.1	Énoncés des exercices	15
1.8.2	Corrigés des exercices	19
2	Théorèmes limites et méthode de Monte-Carlo	29
2.1	Les différentes notions de convergence	29
2.2	Théorèmes limites	29
2.2.1	Loi des grands nombres	29
2.2.2	Application de la loi des grands nombres	31
2.2.2.1	Dessin de la fonction de répartition	31
2.2.2.2	Dessin de la densité	32
2.2.3	Théorème central-limite	33
2.2.4	Application du TCL	35
2.2.4.1	Sondages	35
2.2.4.2	Planche de Galton	37
2.3	Exercices	38
2.3.1	Énoncés	38
2.3.2	Corrigés	39
3	Probabilités et espérances conditionnelles	43
3.1	Conditionnement dans le cas discret	43
3.2	Sommes aléatoires	46
3.3	Probabilités conditionnelles dans le cas mélangé	47

3.4	Moments et loi d'une somme aléatoire	48
3.5	Conditionnement par une variable continue	51
3.6	Statistiques pour les nuls	53
3.7	Exercices	54
3.7.1	Énoncés	54
3.7.2	Corrigés	55
4	Estimation et test d'hypothèse	65
4.1	Estimation	65
4.1.1	Estimation d'une moyenne par intervalle de confiance	65
4.1.2	Marge d'erreur associée à l'estimation de la moyenne et taille d'échantillon requise pour ne pas excéder la marge d'erreur	66
4.1.3	Estimation d'une moyenne par intervalle de confiance dans le cas d'un petit échantillon ($n < 30$)	67
4.1.4	Estimation d'une proportion par intervalle de confiance	68
4.1.5	Marge d'erreur associée à l'estimation de p et taille d'échantillon requise	68
4.2	Exercices	69
4.2.1	Énoncés des exercices	69
4.2.2	Corrigés des exercices	71
	Liste des symboles	73
	Index	73
A	Table de la loi normale	73
B	Table de la loi de Student	75
C	Fonctions, intégrales et sommes usuelles	77

Préface

Ce cours est une introduction aux probabilités utilisant quelques notions de programmation. Les exemples de programmation seront donnés en scilab¹. Ce cours s'adresse à des étudiants de la filière MIAGE, les notions mathématiques sont simplifiées. Les corrigés des exercices sont volontairement succinct et contiennent involontairement des erreurs. Cela devrait faire réfléchir les étudiants.

Cette version est provisoire. Les chapitres suivants seront ajoutés plus tard.

Informations utiles (examens, corrigés ...) :

<http://www.math.unice.fr/~rubentha/enseignement>

¹<http://www.scilab.org/products/scilab/download>

Chapitre 1

Événements aléatoires et variables aléatoires

1.1 Événements et probabilités

Nous donnons ici des règles de calculs sans rentrer dans le détail des définitions mathématiques.

Définition 1.1.1. Nous notons Ω l'ensemble de toutes les possibilités (un élément quelconque de Ω sera souvent noté ω et s'appellera un aléa). On dira aussi que Ω est «l'ensemble des possibles», l'univers, l'univers des possibles, ...

Un événement (que l'on peut aussi orthographier évènement) est une partie de Ω .

Exemple 1.1.2. Si on jette un dé, $A = \text{«on tire un 6»} = \{\omega \in \Omega, \text{ on tire un 6}\}$ est un événement (dans l'égalité précédente, les trois termes veulent dire la même chose. De même, $B = \text{«le résultat est supérieur ou égal à 3»}$ est aussi un événement.

Définition 1.1.3. Soient A, B deux événements. L'événement «il arrive A ou B » (ce qui veut dire que l'on a au moins l'un des deux) s'appelle la réunion de A et B et se note $A \cup B$. On notera aussi $A \cup B = \{\omega \in \Omega, \omega \in A \text{ ou } \omega \in B\}$.

Exemple 1.1.4. On reprend l'exemple du lancer de dé. Soit $A = \text{«le résultat est pair»}$, $B = \text{«le résultat est supérieur ou égal à 3»}$. Alors $A \cup B = \text{«le résultat est dans } \{2, 3, 4, 5, 6\}$ ».

Définition 1.1.5. Soient A, B deux événements. L'événement «il arrive A et B » (ce qui veut dire que l'on a les deux en même temps) s'appelle l'intersection de A et B et se note $A \cap B$. On notera aussi $A \cap B = \{\omega \in \Omega, \omega \in A \text{ et } \omega \in B\}$.

Exemple 1.1.6. Avec les A, B de l'exemple précédent, $A \cap B = \text{«le résultat est dans } \{4, 6\}$ ».

Définition 1.1.7. Soient une liste au plus dénombrable d'événements A_1, A_2, \dots (au plus dénombrable veut dire que l'on peut numéroter ces événements avec de indices entiers, la liste des indices est finie ou infinie). L'événement «l'un au moins de ces événements a lieu» se note

$$A_1 \cup A_2 \cup \dots = \bigcup_{i=1}^{\infty} A_i.$$

Attention, si on a une liste finie d'événements A_1, \dots, A_n , $\bigcup_{i=1}^{\infty} A_i$ veut dire par convention $A_1 \cup A_2 \cup \dots \cup A_n$. L'événement «tous ces événements ont lieu» se note

$$A_1 \cap A_2 \cap \dots = \bigcap_{i=1}^{\infty} A_i.$$

Définition 1.1.8. La probabilité d'un événement A se note $\mathbb{P}(A)$. Nous avons toujours $\mathbb{P}(\Omega) = 1$. L'événement impossible se note \emptyset et vérifie $\mathbb{P}(\emptyset) = 0$. Pour tout événement A , $0 \leq \mathbb{P}(A) \leq 1$.

Exemple 1.1.9. On reprend l'exemple du lancer de dé ci-dessus. Soit $A = \langle \text{le résultat est } 1 \rangle$. Alors $\mathbb{P}(A) = 1/6$.

Les règles de calcul qui suivent sont plus importantes que les définitions précédentes.

Définition 1.1.10. Deux événements A, B sont dits *disjoints* si $A \cap B = \emptyset$ (on ne peut pas avoir à la fois A et B).

Exemple 1.1.11. Toujours avec le lancer de dé, soit $A = \langle \text{le résultat est pair} \rangle$, $B = \langle \text{le résultat est impair} \rangle$. Alors $A \cap B = \emptyset$, ces deux événements sont disjoints (le résultat ne peut pas être pair et impair).

Proposition 1.1.12. Loi d'addition. Si deux événements A, B sont disjoints alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. Si une liste au plus dénombrable d'événements A_1, A_2, \dots est telle que $\forall i, j \geq 1$, $A_i \cap A_j = \emptyset$, alors $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Exemple 1.1.13. Toujours avec l'exemple du lancer de dé. Soit $A = \langle \text{le résultat est pair} \rangle$, $B = \langle \text{le résultat est égal à } 3 \rangle$. Nous avons $A \cap B = \emptyset$ et donc $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = 1/6 + 3/6 = 4/6 = 2/3$.

Proposition 1.1.14. Loi des probabilités totales. Soit une liste au plus dénombrable d'événements A_1, A_2, \dots telle que $\forall i, j \geq 1$, $A_i \cap A_j = \emptyset$ et $\Omega = \cup_{i=1}^{\infty} A_i$. Soit B un événement. Alors $\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i \cap B)$.

Démonstration. Soient $i, j \geq 1$.

Montrons par l'absurde que $(A_i \cap B) \cap (A_j \cap B) = \emptyset$. Si $\exists \omega \in (A_i \cap B) \cap (A_j \cap B)$, alors $\omega \in A_i \cap A_j$, or $A_i \cap A_j = \emptyset$, nous avons donc là une contradiction.

Montrons que $B = \cup_{i=1}^{\infty} (B \cap A_i)$.

- Soit $\omega \in B$. Nous avons $\omega \in \Omega = \cup_{i=1}^{\infty} A_i$ donc $\exists j$ tel que $\omega \in A_j$. Donc $\omega \in B \cap A_j$. Donc $\omega \in \cup_{i=1}^{\infty} (B \cap A_i)$. Donc $B \subset \cup_{i=1}^{\infty} (B \cap A_i)$.
- Soit $\omega \in \cup_{i=1}^{\infty} (B \cap A_i)$. Il existe j tel que $\omega \in B \cap A_j$, donc $\omega \in B$. Donc $\cup_{i=1}^{\infty} (B \cap A_i) \subset B$.

On déduit de ces deux points que $B = \cup_{i=1}^{\infty} (B \cap A_i)$.

Nous avons par la proposition 1.1.12,

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i).$$

□

Proposition 1.1.15. Propriétés de \mathbb{P} .

Si A, B sont deux événements tels que $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Démonstration. Notons $B \setminus A = \{\omega \in \Omega : \omega \in B, \omega \notin A\}$ (cette définition est valable aussi si $A \not\subset B$). Nous avons $B = A \cup (B \setminus A)$ et $A \cap (B \setminus A) = \emptyset$. Donc, par la proposition 1.1.12, $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$. Ces quantités sont positives donc $\mathbb{P}(A) \leq \mathbb{P}(B)$.

□

Notation 1.1.16. On notera $\mathbb{P}(A, B)$ pour dire $\mathbb{P}(A \cap B)$.

1.2 Variables aléatoires

Définition 1.2.1. Une variable aléatoire à valeurs dans un ensemble E est une application de Ω dans E .

Notation 1.2.2. Nous noterons v.a.r. (variable aléatoire réelle) pour parler d'une variable aléatoire à valeurs réelles.

Exemple 1.2.3. Soit X le résultat d'un lancer de dé. L'ensemble $\{\omega \in \Omega : X(\omega) = 6\}$ est un événement. La notation $\mathbb{P}(X = 6)$ est un raccourci pour dire $\mathbb{P}(\{\omega \in \Omega : X(\omega) = 6\})$. Pour simuler X en scilab, on peut se servir de l'instruction suivante

Algorithme 1.1 Lancer de dé

```
grand(1,1,'uin',1,6)
// grand est le générateur de nombres aléatoires de scilab
// les deux premiers paramètres $(1,1)$ indiquent que l'ordinateur renvoie un
// tableau de taille $1 \times 1$ (donc une seule variable)
// 'uin' indique que le résultat est un entier
// les deux derniers paramètres $(1,6)$ indique que le résultat est entre $1$ et $6$
// 'uin' indique que la variable est uniforme dans $\{1, \dots, 6\}$ ($1, \dots, 6$ ont la même probabilité de //sortir ($1/6$))
```

Voici le résultat de plusieurs appels successifs de cette instruction :

```
->grand(1,1,'uin',1,6)
ans = 4.
->grand(1,1,'uin',1,6)
ans = 5.
->grand(1,1,'uin',1,6)
ans = 2.
->grand(1,1,'uin',1,6)
ans = 5.
```

Définition 1.2.4. Fonction de répartition Soit X une variable aléatoire à valeurs dans \mathbb{R} . La fonction de répartition de X est la fonction $t \in \mathbb{R} \mapsto \mathbb{P}(X \leq t) \in \mathbb{R}$.

Exemple 1.2.5. Soit X le résultat d'un lancer de dé. Nous avons $\forall i \in \{1, \dots, 6\}, \mathbb{P}(X = i) = 1/6$.

- Soit $t < 1$. Nous avons $\{\omega : X(\omega) \leq t\} = \emptyset$ (X n'est jamais $\leq t$) donc $\mathbb{P}(X \leq t) = 0$.
- Soit $t \in [1; 2[$. Nous avons $\{\omega : X(\omega) \leq t\} = \{\omega : X(\omega) = 1\}$ (que l'on peut écrire plus simplement $\{X \leq t\} = \{X = 1\}$). Donc $\mathbb{P}(X \leq t) = \mathbb{P}(X = 1) = 1/6$.
- Soit $t \in [2; 3[$. Nous avons $\{\omega : X(\omega) \leq t\} = \{\omega : X(\omega) \in \{1, 2\}\}$ (que l'on peut écrire plus simplement $\{X \leq t\} = \{X = 1 \text{ ou } 2\}$). Donc $\mathbb{P}(X \leq t) = \mathbb{P}(\{X = 1\} \cup \{X = 2\}) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = 2/6$ (on peut utiliser la proposition 1.1.12 parce que $\{X = 1\} \cap \{X = 2\} = \emptyset$).
- ...
- Soit $t \geq 6$. Nous avons $\{X \leq t\} = \Omega$ donc $\mathbb{P}(X \leq t) = 1$.

FIGURE 1.1 – Fonction de répartition pour le lancer de dé

Nous pouvons maintenant dessiner la fonction de répartition de X (figure 1.1).

Proposition 1.2.6. Propriétés de la fonction répartition Soit X une variables aléatoire à valeurs réelles et soit F sa fonction de répartition. Soient $a, b \in \mathbb{R}$. Nous avons :

1. $\mathbb{P}(X > a) = 1 - F(a)$,
2. $\mathbb{P}(a < X \leq b) = F(b) - F(a)$,
3. $\mathbb{P}(X = x) = F(x) - \lim_{\epsilon \downarrow 0} F(x - \epsilon) = F(x) - F(x-)$ ($F(x-)$ signifie la limite à gauche de F en x).

Démonstration. 1. Nous avons $1 = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(X > a) + \mathbb{P}(X \leq a)$ (le lecteur vérifiera lui-même que nous pouvons bien appliquer la proposition 1.1.12). Donc $\mathbb{P}(X > a) = 1 - \mathbb{P}(X \leq a) = 1 - F(a)$.

2. Nous avons $\mathbb{P}(X \leq b) = \mathbb{P}(X \leq a) + \mathbb{P}(a < X \leq b)$ (même remarque que ci-dessus) donc $\mathbb{P}(a < X \leq b) = F(b) - F(a)$.

3. Ce point est admis. □

Exemple 1.2.7. Reprenons l'exemple précédent. En utilisant la proposition ci-dessus, nous obtenons :

- $\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \leq 2) = 1 - (\mathbb{P}(X = 1) + \mathbb{P}(X = 2)) = 4/6 = 2/3$,
- $\mathbb{P}(X = 2) = F(2) - F(2-) = 2/6 - 1/6 = 1/6$.

Définition 1.2.8. Une variable aléatoire X est dite discrète s'il existe nombre au plus dénombrable de valeurs x_1, x_2, \dots telles que $\forall i, a_i := \mathbb{P}(X = x_i) > 0$. (Notation : nous utilisons ici le symbole «:=» pour dire a_i est défini comme étant égal à $\mathbb{P}(X = x_i)$.)

La fonction (qui s'applique aux x_i)

$$x_i \mapsto p_X(x_i) = a_i$$

s'appelle la fonction de masse de la variable X .

Proposition 1.2.9. Soit X une variable aléatoire réelle discrète, de fonction de masse p_X et de fonction de répartition F_X . Nous avons la relation ($\forall i$)

$$p_X(x_i) = F_X(x_i) - F_X(x_i-).$$

La fonction F_X est constante par morceaux. Elle ne change de valeurs qu'aux points x_i .

Exemple 1.2.10. Reprenons l'exemple précédent du lancer de dé. La variable X est discrète et nous avons bien $\mathbb{P}(X = 2) = F(2) - F(2-)$.

Définition 1.2.11. Une v.a.r. X est dite continue si sa fonction de répartition F est une fonction continue.

Définition 1.2.12. Soit X une v.a.r. S'il existe une fonction f de \mathbb{R} dans \mathbb{R}^+ telle que $\forall a < b \in \mathbb{R}$,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx,$$

alors cette fonction f s'appelle la densité de probabilité de X (on dit aussi la densité tout court).

Proposition 1.2.13. La définition ci-dessus implique que si X a une densité f alors $\forall a, b \in [-\infty; +\infty]$,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx,$$

et

$$\mathbb{P}(X = a) = 0.$$

Proposition 1.2.14. Soit X une v.a.r. Si X a une densité f alors X est continue et $\forall x \in \mathbb{R}$,

$$F(x) = \int_{-\infty}^x f(t)dt.$$

Proposition 1.2.15. Si X est une v.a.r. de fonction de répartition F telle que F est dérivable, alors X a une densité f qui est égale à $(\forall x)$

$$f(x) = F'(x).$$

Si F est dérivable partout sauf en un nombre fini de points, X est encore continue et elle a pour densité $f = F'$ (que l'on peut calculer partout sauf en un nombre fini de points, on met n'importe quelle valeur pour f aux points où F n'est pas dérivable).

Remarque 1.2.16. S'il y a un nombre fini de points où la dérivée de F est compliquée à calculer, on peut se contenter d'assigner à f des valeurs arbitraires en ces points.

Exemple 1.2.17. Soit X une v.a.r. ayant la fonction de répartition suivante (voir figure 1.2 pour le dessin) (il s'agit de la variable uniforme sur $[0; 1]$)

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } 1 \leq x. \end{cases}$$

FIGURE 1.2 – Fonction de répartition de la variable uniforme sur $[0; 1]$.

Cette fonction F est continue donc X est une variable continue. La fonction F est dérivable partout sauf aux points $0, 1$. Calculons la dérivée $f = F'$, nous obtenons (voir figure 1.3 pour le dessin) :

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{si } 1 < x. \end{cases}$$

Remarquons que les valeurs $f(0)$ et $f(1)$ sont arbitraires.

FIGURE 1.3 – Densité de la variable uniforme sur $[0; 1]$.

Algorithme 1.2 Variable uniforme sur $[0; 1]$

```
grand(1,1,'unf',0,1)
// génère une variable aléatoire uniforme dans  $[0, 1]$ 
// les deux premiers paramètres veulent dire qu'on récupère un tableau  $1 \times 1$ 
// de variables aléatoires, donc une seule variable
```

Voici le résultat de plusieurs appels successifs de cette instruction

```
->grand(1,1,'unf',0,1)
ans = 0.9811097
->grand(1,1,'unf',0,1)
ans = 0.9571669
->grand(1,1,'unf',0,1)
ans = 0.1098618
```

Il existe des v.a.r. qui ne sont ni discrètes ni continues mais nous n'en parlerons pas dans ce cours.

1.3 Espérance et moments

1.3.1 Définitions

Définition 1.3.1. Si X est une v.a.r. discrète (qui prend les valeurs x_1, x_2, \dots) son moment d'ordre m est

$$\mathbb{E}(X^m) = \sum_{i \geq 1} x_i^m \mathbb{P}(X = x_i)$$

si cette série converge absolument (c'est à dire $\lim_{n \rightarrow +\infty} \sum_{i=1}^n |x_i|^m \mathbb{P}(X = x_i) < +\infty$). Dans le cas contraire, on dit que le moment d'ordre m n'existe pas. Le moment d'ordre 1 s'appelle la moyenne.

Définition 1.3.2. Si X est une v.a.r. continue de densité f , son moment d'ordre m est

$$\mathbb{E}(X^m) = \int_{-\infty}^{+\infty} x^m f(x) dx$$

si cette intégrale converge absolument (ce qui est équivalent à :

$\lim_{M \rightarrow +\infty} \int_{-M}^M |x|^m f(x) dx < +\infty$). Dans le cas contraire, on dit que le moment d'ordre m n'existe pas.

Définition 1.3.3. Le moment d'ordre 1 d'une v.a.r. X s'appelle son espérance (on dit aussi «sa moyenne»). Nous avons $\mathbb{E}(1) = 1$ (la moyenne de la variable constante égale à 1 est 1).

Définition 1.3.4. Soit X une v.a.r. de moyenne μ_X . Le moment d'ordre 2 de $X - \mu_X$ s'appelle la variance de X . Ce moment est donc égal à $\mathbb{E}((X - \mu_X)^2) = \mathbb{E}((X - \mathbb{E}(X))^2)$. Nous noterons $\text{Var}(X)$ la variance de X .

Définition 1.3.5. On appelle médiane d'une v.a.r. X toute valeur v telle que

$$\mathbb{P}(X \geq v) \geq 1/2 \text{ et } \mathbb{P}(X \leq v) \geq 1/2.$$

Exemple 1.3.6. Soit X une v.a.r. uniforme sur $[0; 1]$ (voir exemple 1.2.17). Notons f la densité de X . Calculons

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} xf(x)dx \\ &= \int_{-\infty}^0 0dx + \int_0^1 xdx + \int_1^{+\infty} 0dx \\ &= \left[\frac{x^2}{2} \right]_0^1 = \frac{1}{2}. \end{aligned}$$

Calculons maintenant la variance de X

$$\begin{aligned} \mathbb{E}\left(\left(X - \frac{1}{2}\right)^2\right) &= \int_{-\infty}^{+\infty} \left(x - \frac{1}{2}\right)^2 f(x)dx \\ &= \int_{-\infty}^0 0dx + \int_0^1 \left(x - \frac{1}{2}\right)^2 dx + \int_1^{+\infty} 0dx \\ &= \left[\frac{1}{3} \left(x - \frac{1}{2}\right)^3 \right]_0^1 \\ &= \frac{1}{3} \left(\frac{1}{8} - \left(-\frac{1}{8}\right) \right) = \frac{1}{12}. \end{aligned}$$

1.3.2 Propriétés

Si X est une v.a.r. et $g : \mathbb{R} \rightarrow \mathbb{R}$ alors $Y = g(X)$ est encore une v.a.r.

Proposition 1.3.7. Si de plus, X est une variable discrète (qui prend les valeurs x_1, x_2, \dots), alors

$$\mathbb{E}(g(X)) = \sum_{i \geq 1} g(x_i) \mathbb{P}(X = x_i)$$

si cette série converge absolument.

Proposition 1.3.8. Dans le cas où X est une variable continue de densité f , alors

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

si cette intégrale converge absolument.

Exemple 1.3.9. On reprend l'exemple précédent. Calculons

$$\begin{aligned} \mathbb{E}(e^X) &= \int_{-\infty}^{+\infty} e^x f(x)dx \\ &= \int_0^1 e^x dx \\ &= [e^x]_0^1 = e^1 - 1. \end{aligned}$$

Proposition 1.3.10. Linéarité de l'espérance. Soient X, Y deux v.a.r. et $\lambda, \mu \in \mathbb{R}$,

$$\mathbb{E}(\mu X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y).$$

Lemme 1.3.11. Soient X_1, \dots, X_n des v.a.r. et soient h_1, \dots, h_m des fonctions de \mathbb{R}^n dans \mathbb{R} , alors :

$$\mathbb{E}\left(\sum_{j=1}^m h_j(X_1, \dots, X_n)\right) = \sum_{j=1}^m \mathbb{E}(h_j(X_1, \dots, X_n)).$$

Proposition 1.3.12. Croissance de l'espérance. Si X, Y sont deux v.a.r. telles que $\forall \omega, X(\omega) \leq Y(\omega)$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

1.4 Fonctions de répartition jointes

1.4.1 Définitions générales

Définition 1.4.1. Soient X, Y deux v.a.r., leur fonction de distribution jointe est la fonction $\mathbb{R}^2 \rightarrow \mathbb{R}$ (une fonction de deux variables) définie par

$$F_{XY}(x, y) = \mathbb{P}(X \leq x, Y \leq y).$$

Rappelons que $\mathbb{P}(X \leq x, Y \leq y)$ veut dire $\mathbb{P}(\{X \leq x\} \cap \{Y \leq y\})$. Le couple (X, Y) est dit posséder une densité s'il existe une fonction f_{XY} (de deux variables) telle que

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u, v) du dv, \quad \forall x, y. \quad (1.4.1)$$

La fonction $F_X(x) = \lim_{y \rightarrow +\infty} F_{XY}(x, y)$ est égale à la fonction de répartition de la variable X . On l'appelle la fonction de distribution marginale de X . De même, $F_Y(y) = \lim_{x \rightarrow +\infty} F_{XY}(x, y)$ est la fonction de répartition de Y . Si F_{XY} a une densité f_{XY} alors F_X et F_Y ont les densités respectives

$$x \mapsto f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy, \quad y \mapsto f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx.$$

Exemple 1.4.2. Soient X, Y de fonction de répartition jointe

$$F_{XY}(x, y) = \begin{cases} 0 & \text{si } x \text{ ou } y < 0 \\ \min(x, 1) \times \min(y, 1) & \text{sinon.} \end{cases}$$

Algorithme 1.3 Fonction F_{XY}

```
function [z]=FXY(x,y)
if (x<0) then
  r=0;
else
  if (y<0) then
    r=0;
  else
    r=min(1,x)*min(1,y)
  end,
end,
z=r;
endfunction
```

On remarque qu'il existe une densité :

$$f_{XY}(x, y) = \begin{cases} 1 & \text{si } x, y \in [0; 1] \\ 0 & \text{sinon} . \end{cases}$$

Vérifions que nous avons bien la relation (1.4.1) pour $x \in [0; 1]$, $y > 1$:

$$\begin{aligned} \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u, v) du dv &= \int_{-\infty}^0 \int_{-\infty}^y 0 du dv + \int_0^x \int_{-\infty}^0 0 du dv \\ &\quad + \int_0^x \int_0^1 1 dx dy + \int_0^x \int_1^y 0 dx dy \\ &= \int_0^x 1 dx \\ &= x \\ &= \min(1, x) \times \min(1, y) . \end{aligned}$$

Calculons la fonction de répartition marginale de X :

$$F_X(x) = \lim_{y \rightarrow +\infty} F_{XY}(x, y) = \begin{cases} \min(1, x) & \text{si } x \geq 0 \\ 0 & \text{sinon} . \end{cases}$$

Le graphe de F_X est le même que celui de la figure 1.2. Calculons la densité marginale de X :

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \\ &= \begin{cases} 0 & \text{si } x \notin [0; 1] \\ \int_0^1 1 dy & \text{si } x \in [0; 1] \end{cases} \\ &= \begin{cases} 0 & \text{si } x \notin [0; 1] \\ 1 & \text{si } x \in [0; 1] . \end{cases} \end{aligned}$$

Son graphe est le même que celui de la figure 1.3.

Notation 1.4.3. Fonction indicatrice Soit $A \subset E$ des ensemble quelconques. Nous introduisons

$$\mathbf{1}_A : x \in E \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A . \end{cases}$$

Lemme 1.4.4. Si A est un événement, alors

$$\mathbb{E}(\mathbf{1}_A) = \mathbb{P}(A) .$$

Démonstration. La variable aléatoire $\omega \mapsto \mathbf{1}_A(\omega)$ est une variable discrète (à valeurs dans $\{0, 1\}$). Nous avons $\mathbb{P}(\mathbf{1}_A = 1) = \mathbb{P}(A)$, $\mathbb{P}(\mathbf{1}_A = 0) = \mathbb{P}(A^c)$, donc

$$\begin{aligned} \mathbb{E}(\mathbf{1}_A) &= \mathbb{P}(A) \times 1 + \mathbb{P}(A^c) \times 0 \\ &= \mathbb{P}(A) . \end{aligned}$$

□

Exemple 1.4.5. La fonction $1_{[0;1]}$. Son graphe est celui de la figure 1.3.

Algorithme 1.4 Fonction indicatrice de $[0; 1]$

```

function [z]=indicator(x)
if (x<1) then
  if (x>=0) then
    r=1;
  else
    r=0;
  end,
else
  r=0;
end,
z=r;
endfunction

```

Cette fonction est égale à f_X de l'exemple précédent.

1.4.2 Indépendance

Définition 1.4.6. Soient X, Y deux v.a.r. de fonction de répartition jointe F_{XY} et de fonctions de répartition marginales F_X, F_Y (respectivement pour X et Y). Si $F_{XY}(x, y) = F_X(x) \times F_Y(y)$, $\forall x, y$ alors on dit que X et Y sont indépendantes. On notera $X \perp\!\!\!\perp Y$ pour dire que X et Y sont indépendantes.

Remarque 1.4.7. Deux appels successifs de variables aléatoires à l'ordinateur renvoient deux variables indépendantes.

Lemme 1.4.8. Si X, Y sont deux v.a.r. de densité jointe f_{XY} et de densités marginales (respectivement) f_X, f_Y alors $f_{XY} = f_X \times f_Y$ (c'est à dire, $\forall x, y, f_{XY}(x, y) = f_X(x) \times f_Y(y)$).

Proposition 1.4.9. Si X, Y sont deux v.a.r. indépendantes et f, g sont deux fonctions alors $f(X)$ et $g(Y)$ sont deux variables indépendantes et

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X)) \times \mathbb{E}(g(Y)).$$

Proposition 1.4.10. Si X, Y sont deux variables indépendantes et U, V sont deux ensembles, alors les événements $\{X \in U\}$ et $\{Y \in V\}$ vérifient

$$\mathbb{P}(\{X \in U\} \cap \{Y \in V\}) = \mathbb{P}(X \in U) \times \mathbb{P}(Y \in V)$$

et on dit que les événements $\{X \in U\}, \{Y \in V\}$ sont indépendants.

Exemple 1.4.11. Soient X, Y les résultats de deux lancers de dé. On suppose $X \perp\!\!\!\perp Y$. Calculons à l'aide de la proposition précédente

$$\begin{aligned} \mathbb{P}(\{X \in \{1, 2\}\} \cap \{Y \geq 4\}) &= \mathbb{P}(X \in \{1, 2\}) \times \mathbb{P}(Y \geq 4) \\ &= \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

Définition 1.4.12. Soient X, Y deux v.a.r. indépendantes de moyennes μ_X, μ_Y respectivement. La covariance de X et Y est la quantité suivante :

$$\sigma_{XY} := \mathbb{E}((X - \mu_X)(Y - \mu_Y)) = \mathbb{E}(XY) - \mu_X \mu_Y. \quad (1.4.2)$$

Remarque 1.4.13. La deuxième égalité dans (1.4.2) est une conséquence de la proposition 1.3.10 (linéarité de l'espérance). Calculons :

$$\begin{aligned}\mathbb{E}((X - \mu_X)(Y - \mu_Y)) &= \mathbb{E}(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y) \\ (\text{par prop. 1.3.10}) &= \mathbb{E}(XY) - \mu_X \mathbb{E}(Y) - \mu_Y \mathbb{E}(X) + \mu_X \mu_Y.\end{aligned}$$

Remarquons que si $X = Y$, nous avons montré

$$\text{Var}(X) = \mathbb{E}((X - \mu_X)^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2. \quad (1.4.3)$$

Définition 1.4.14. La fonction de répartition jointe de v.a.r. X_1, X_2, \dots, X_n est définie par

$$F(x_1, \dots, x_n) = \mathbb{P}(\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\} \cap \dots \cap \{X_n \leq x_n\}),$$

que l'on peut également noter : $\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$.

Notons F_{X_i} la fonction de répartition de X_i ($\forall i$). Si $F = F_{X_1} F_{X_2} \dots F_{X_n}$ alors les variables X_1, \dots, X_n sont dites indépendantes.

Exemple 1.4.15. Reprenons l'exemple 1.4.2. Nous avons calculé que la fonction de répartition de X est

$$x \mapsto \begin{cases} \min(1, x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} = \mathbf{1}_{[0; +\infty[}(x) \min(1, x),$$

où nous utilisons la notation 1.4.3. On peut montrer de même que la fonction de répartition de Y est $F_Y = F_X$ (ce qui veut dire $F_Y(u) = F_X(u)$, $\forall u$). Nous avons $F_{XY} = F_X F_Y$ donc $X \perp\!\!\!\perp Y$.

Lemme 1.4.16. Soient U_1, U_2, \dots des variables indépendantes, toutes de loi $\mathcal{B}(p)$ ($p \in [0; 1]$ fixé). Alors, $\forall n \in \mathbb{N}^*$,

$$U_1 + \dots + U_n \sim \mathcal{B}(n, p).$$

Démonstration. Soit $n \in \mathbb{N}^*$. La variable $U_1 + \dots + U_n$ est à valeurs dans $\{0, 1, 2, \dots, n\}$. Soit $k \in \{0, 1, \dots, n\}$. Nous avons

$$\{U_1 + \dots + U_n = k\} = \cup_{I \subset \{1, \dots, n\}, \#I=k} (\{U_i = 1, \forall i \in I\} \cap \{U_i = 0, \forall i \in I^c\}),$$

et cette réunion est disjointe. Donc :

$$\begin{aligned}\mathbb{P}(U_1 + \dots + U_n = k) &= \sum_{I \subset \{1, \dots, n\}, \#I=k} \mathbb{P}(\{U_i = 1, \forall i \in I\} \cap \{U_i = 0, \forall i \in I^c\}) \\ (\text{les } U_i \text{ sont indépendants}) &= \sum_{I \subset \{1, \dots, n\}, \#I=k} \prod_{i \in I} \mathbb{P}(U_i = 1) \times \prod_{i \in I^c} \mathbb{P}(U_i = 0) \\ &= \sum_{I \subset \{1, \dots, n\}, \#I=k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} \\ &= C_n^k \left(\frac{1}{2}\right)^n.\end{aligned}$$

□

1.5 Sommes et convolutions

Proposition 1.5.1. Si X, Y sont deux v.a.r. indépendantes de fonctions de densité f_X, f_Y respectivement. Alors, la fonction de densité de leur somme $Z = X + Y$ est la convolution de f_X et f_Y , c'est à dire la fonction :

$$z \mapsto f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-u)f_Y(u)du,$$

remarquons que la dernière quantité est égale à (par changement de variable)

$$\int_{-\infty}^{+\infty} f_Y(z-u)f_X(u)du.$$

Exemple 1.5.2. Soient X, Y deux variables aléatoires uniformes sur $[0; 1]$ et indépendantes (voir exemple 1.2.17 pour la définition de la variable uniforme). Ces variables ont la densité :

$$x \mapsto f(x) = \mathbf{1}_{[0;1]}(x).$$

Calculons la densité de $Z = X + Y$ (nous la notons f_Z). Pour tout $z \in \mathbb{R}$,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f(z-u)f(u)du = \int_{-\infty}^{+\infty} \mathbf{1}_{[0;1]}(z-u)\mathbf{1}_{[0;1]}(u)du \\ &= \int_0^1 \mathbf{1}_{[0;1]}(z-u)du. \end{aligned}$$

Donc si $z \notin [0; 2]$ alors $f_Z(z) = 0$. Si $z \in [-1; 2]$, remarquons que $\mathbf{1}_{[0;1]}(z-u) = 1$ si et seulement si $z-u \in [0; 1]$, c'est à dire $z-1 \leq u \leq z$. Pour $z \in [1; 2]$, $z-1 \in [0; 1]$ et $z \geq 1$ donc

$$f_Z(z) = \int_{z-1}^1 1du = 2-z.$$

Si $z \in [0; 1]$, $z-1 \leq 0$ et $z \in [0; 1]$ donc

$$f_Z(z) = \int_0^z 1du = z.$$

Le dessin de f_Z est dans la figure 1.4.

FIGURE 1.4 – Densité de la somme de deux variables uniformes sur $[0; 1]$ indépendantes.

Proposition 1.5.3. 1. Si X, Y sont deux v.a.r. indépendantes de variances σ_X^2, σ_Y^2 respectivement, alors la variance de $X + Y$ est $\sigma_X^2 + \sigma_Y^2$.

2. Si X_1, \dots, X_n sont des v.a.r. indépendantes telles que $\forall i, X_i$ est de variance σ_i^2 . Alors $X_1 + \dots + X_n$ est de variance $\sigma_1^2 + \dots + \sigma_n^2$.

Démonstration. Nous ne démontrons ici que le premier point. Nous avons :

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

et

$$\begin{aligned} \mathbb{E}((X + Y - \mathbb{E}(X) - \mathbb{E}(Y))^2) &= \mathbb{E}((X - \mathbb{E}(X))^2 + (Y - \mathbb{E}(Y))^2 + 2(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) \\ (\text{car } X \perp\!\!\!\perp Y) &= \sigma_X^2 + \sigma_Y^2 + 2\mathbb{E}(X - \mathbb{E}(X)) \times \mathbb{E}(Y - \mathbb{E}(Y)) \\ &= \sigma_X^2 + \sigma_Y^2 + 0. \end{aligned}$$

□

Proposition 1.5.4. Si X, Y sont deux variables discrètes indépendantes avec X à valeurs dans $\{x_1, x_2, \dots\}$ et Y à valeurs dans $\{y_1, y_2, \dots\}$, alors $Z = X + Y$ est à valeurs dans $\{x_i + y_j : i, j \geq 1\}$ et $\forall i_0, j_0 \geq 1$,

$$\mathbb{P}(Z = x_{i_0} + y_{j_0}) = \sum_{i, j \geq 1 : x_i + y_j = x_{i_0} + y_{j_0}} \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j).$$

Démonstration. Soient $i_0, j_0 \geq 0$. Calculons

$$\begin{aligned} \mathbb{P}(Z = x_{i_0} + y_{j_0}) &= \mathbb{P}(\cup_{i, j \geq 0 : x_i + y_j = x_{i_0} + y_{j_0}} \{X = x_i, Y = y_j\}) \\ (\text{réunion disjointe}) &= \sum_{i, j \geq 0 : x_i + y_j = x_{i_0} + y_{j_0}} \mathbb{P}(X = x_i, Y = y_j) \\ (\text{car } X \perp\!\!\!\perp Y) &= \sum_{i, j \geq 0 : x_i + y_j = x_{i_0} + y_{j_0}} \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j). \end{aligned}$$

□

Exemple 1.5.5. Soient $X \perp\!\!\!\perp Y$ à valeurs dans \mathbb{N} ($\mathbb{N} = \{0, 1, 2, 3, \dots\}$) tels que $\forall i \geq 0$,

$$\mathbb{P}(X = i) = \frac{1}{2^{i+1}}, \quad \mathbb{P}(Y = i) = \frac{2}{3^{i+1}}.$$

Soit $Z = X + Y$. La variable Z est à valeurs dans \mathbb{N} . Soit $n \in \mathbb{N}$, calculons

$$\begin{aligned} \mathbb{P}(Z = n) &= \sum_{j=0}^n \mathbb{P}(X = j) \mathbb{P}(Y = n - j) \\ &= \sum_{j=0}^n \frac{1}{2^{j+1}} \times \frac{2}{3^{n-j+1}} \\ &= \frac{2}{3} \frac{1}{3^n} \sum_{j=0}^n \frac{3^j}{2^j} \\ (\text{somme géométrique}) &= \frac{2}{3} \frac{1}{3^n} \left(\frac{1 - \left(\frac{3}{2}\right)^{n+1}}{1 - \frac{3}{2}} \right) \\ &= \frac{4}{3} \frac{1}{3^n} \left(\left(\frac{3}{2}\right)^{n+1} - 1 \right) \\ &= \frac{1}{2^{n-1}} - \frac{4}{3^{n+1}}. \end{aligned}$$

1.6 Changement de variable

Soit X une v.a.r. à valeurs dans un intervalle $(a; b)$ (que nous notons avec des parenthèses pour dire que l'intervalle peut être ouvert ou fermé en chacune des bornes) et g une fonction bijective $(a; b) \rightarrow (c; d)$ ($(c; d)$ est un autre intervalle) (g est donc soit croissante, soit décroissante). Notons g^{-1} la fonction inverse de g . Posons $Y = g(X)$. La variable Y est à valeurs dans $(c; d)$. Soit F_X, F_Y les fonctions de répartition de X, Y respectivement.

Lemme 1.6.1. *Nous avons les relations :*

1. Si g est croissante : $\forall y \in]c; d[, F_Y(y) = F_X(g^{-1}(y))$.

2. Si g est décroissante : $\forall y \in]c; d[, F_Y(y) = \mathbb{P}(X = g^{-1}(y)) + 1 - F_X(g^{-1}(y))$.

On ne se préoccupe pas de ce qui peut se passer en c et d .

Démonstration. 1. Si g est croissante : $\forall y \in (c; d)$,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X \leq g^{-1}(y)) \\ &= F_X(g^{-1}(y)). \end{aligned}$$

Si g est décroissante : $\forall y \in (c; d)$,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X \geq g^{-1}(y)) \\ &= \mathbb{P}(X = g^{-1}(y)) + \mathbb{P}(X > g^{-1}(y)) \\ &= \mathbb{P}(X = g^{-1}(y)) + 1 - F_X(g^{-1}(y)). \end{aligned}$$

□

Proposition 1.6.2. Formule de changement de variable pour la densité. *Si X a une densité f_X alors Y a une densité f_Y nulle en dehors de $(c; d)$ et telle que :*

1. Si g est croissante : $\forall y \in (c; d)$,

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

2. Si g est décroissante : $\forall y \in (c; d)$,

$$f_Y(y) = -\frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

On ne se préoccupe pas de ce qui peut se passer en c et d .

Exemple 1.6.3. *Soit X variable uniforme sur $[0; 1]$. Soit $Y = X^2$. Soit g la fonction $x \in [0; 1] \mapsto g(x) = x^2 \in [0; 1]$. Cette fonction est bijective et $Y = g(X)$. La densité f_X de X vérifie $f_X(x) = 1$ si $x \in [0; 1]$. La densité f_Y de Y est donc nulle en dehors de $[0; 1]$ et pour $y \in [0; 1]$:*

$$f_Y(y) = \frac{1}{2\sqrt{y}}.$$

Exemple 1.6.4. *Soit $X \sim \mathcal{N}(0, 1)$ (voir section suivante pour la définition). Soit f_X la densité de X . Soient $\sigma > 0, \mu \in \mathbb{R}$. Posons $Y = \sigma X + \mu$. Nous voulons calculer la densité de Y . La fonction*

$$x \in \mathbb{R} \mapsto \sigma x + \mu$$

est bijective croissante de réciproque $g^{-1}(y) = (y - \mu)/\sigma$. Donc la densité de Y est la fonction :

$$y \mapsto \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))} = \frac{\exp(-(y - \mu)^2/2\sigma^2)}{\sqrt{2\pi}} \times \frac{1}{\sigma}.$$

Donc $Y \sim \mathcal{N}(0, 1)$.

1.7 Lois de probabilités usuelles (à connaître par cœur)

Notation 1.7.1. Nous noterons « $X \sim \dots$ » pour dire « X suit la loi ...».

1.7.1 Lois discrètes

- Loi uniforme. Soit E ensemble fini de cardinal n , X est une variable uniforme sur E si $\forall x \in E, \mathbb{P}(X = x) = \frac{1}{n}$.
- Loi de Bernoulli de paramètre p ($\in [0, 1]$), notée $\mathcal{B}(p)$: X à valeurs dans $\{0, 1\}$ telle que $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$.
- Loi binômiale de paramètres n, p ($n \in \mathbb{N}^*, p \in [0, 1]$), notée $\mathcal{B}(n, p)$: X à valeurs dans $\{0, \dots, n\}$ telle que $\forall k \in \{0, \dots, n\}, \mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}$.
- Loi géométrique de paramètre p ($\in [0, 1]$), notée $\mathcal{G}(p)$: X à valeurs dans \mathbb{N}^* telle que $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = (1 - p)^{k-1} p$.
- Loi de Poisson de paramètre λ (> 0), notée $\mathcal{P}(\lambda)$: X à valeurs dans \mathbb{N} telle que $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

1.7.2 Lois continues

- Loi uniforme sur $[a, b]$ ($a < b$), notée $\mathcal{U}([a, b])$: de densité $x \mapsto \frac{1}{b-a} \mathbf{1}_{[a, b]}(x)$.
- Loi exponentielle de paramètre λ ($\lambda > 0$), notée $\mathcal{E}(\lambda)$: de densité $x \mapsto \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}^+}(x)$.
- Loi gaussienne (ou normale) de moyenne m ($\in \mathbb{R}$) et de variance σ^2 ($\in \mathbb{R}^+$), notée $\mathcal{N}(m, \sigma^2)$: de densité $x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$.

1.8 Exercices

1.8.1 Énoncés des exercices

- Soient A, B des événements. Montrer que $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$ (où $B^c = \{\omega \in \Omega, \omega \notin B\}$, c 'est le complémentaire de B).
- Soient A, B des événements. Montrer que $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ (On pourra utiliser le résultat de l'exercice précédent.)
-

(a) Faire un dessin de la fonction de répartition

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x^3 & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geq 1. \end{cases}$$

- Déterminer la fonction de densité correspondante dans les régions : $x \leq 0, 0 < x < 1, x \geq 1$.
- Quelle est la moyenne de cette distribution ? (On demande ici la moyenne de X variable aléatoire de fonction de répartition F .)

- (d) Si X a F pour fonction de répartition, calculer $\mathbb{P}(1/4 \leq X \leq 3/4)$.
4. Soit Z une variable aléatoire discrète pouvant prendre les valeurs 0, 1, 2, 3. La fonction de masse de Z est la suivante

$$p(0) = \frac{1}{4}, p(1) = \frac{1}{2}, p(2) = \frac{1}{8}, p(3) = \frac{1}{8}.$$

- (a) Dessiner la fonction de répartition de Z .
- (b) Calculer $\mathbb{E}(Z)$.
- (c) Calculer la variance de Z .
5. Soient X, Y des v.a.r. de fonctions de répartition F_X, F_Y respectivement.
- (a) Soit $Z = \max(X, Y)$. Montrer que la fonction de répartition de Z est $z \mapsto F_Z(z) = F_X(z)F_Y(z)$.
- (b) Soit $W = \min(X, Y)$. Montre que la fonction de répartition de W est $w \mapsto 1 - (1 - F_X(w))(1 - F_Y(w))$.

6.

Algorithme 1.5 Simulation de variable aléatoire

```
U=grand(1,1,'unf',0,1); // simule une loi uniforme dans [0;1]
X=U^(1/2); // racine carrée de U
disp(X); // affiche le résultat X
```

- (a) L'algorithme ci-dessus simule une variable aléatoire X . Calculer la densité de X (on la notera f_X).
- (b) Calculer la fonction de répartition F_X de X .
- (c) Calculer $\mathbb{E}(X)$.
- (d) Calculer $\text{Var}(X)$.

7.

Algorithme 1.6 Simulation de variable aléatoire

```
U=grand(1,1,'unf',0,1);
X=-log(U);
Y=-log(1-U);
disp(X);
disp(Y);
```

- (a) L'algorithme ci-dessus simule des variables aléatoires X, Y . Calculer la densité de X (notée f_X). Calculer la densité de Y (notée f_Y).
- (b) Calculer les fonctions de répartition de X, Y (notées respectivement F_X, F_Y).
- (c) Calculer la fonction de répartition de répartition jointe de X et Y . Les variables X, Y sont-elles indépendantes ?

8. Soit V une v.a.r. de fonction de répartition donnée par la procédure suivante ($A > 0$ est une constante quelconque).

Algorithme 1.7 Fonction de répartition (exercice 8)

```

function [r]=repartition(x)
if (x<0) then
  y=0;
else
  if (x<=1) then
    y=1-(1-x)^A;
  else
    y=1;
  end,
end,
r=y;
endfunction

```

- (a) Donner une expression mathématique de F .
- (b) Calculer la densité de la variable V .
- (c) Faire un dessin de la densité pour $A = 1/2$.
- (d) Calculer $\mathbb{E}(V)$ pour $A = 1/2$.
- (e) Calculer $\text{Var}(V)$ pour $A = 1/2$.
9. Rappelons la définition de $\mathbf{1}_A$ la variable aléatoire fonction indicatrice d'un événement A (voir la notation 1.4.3). Pour tout $\omega \in \Omega$, nous avons

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

En d'autres termes, $\mathbf{1}_A$ est une v.a.r. qui vaut 1 si l'événement A a lieu et 0 sinon. Montrer les points suivants.

- (a) $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$,
- (b) $\mathbf{1}_{A \cap B} = \mathbf{1}_A \times \mathbf{1}_B = \min(\mathbf{1}_A, \mathbf{1}_B)$,
- (c) $\mathbf{1}_{A \cup B} = \max(\mathbf{1}_A, \mathbf{1}_B)$.
- 10.

Algorithme 1.8 Pile ou face

```

b=0;
N=1;
x=grand(1,1,'uin',0,1); // variable aléatoire qui vaut 0 ou 1 (avec proba. 1/2)
while (b==0)
    N=N+1;
    y=grand(1,1,'uin',0,1);
    if (x==y) then
        b=1;
    end,
    x=y;
end,
disp("N=",N);

```

(a) Calculer la fonction de masse de N .

(b) Soit $A = \{\omega : N(\omega) \text{ est pair}\}$ (c'est l'événement « N est pair») et $B = \{\omega : N(\omega) \leq 6\}$. Calculer $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(A \cap B)$.

11. Soient U, W des v.a.r. telles que

$$\mathbb{P}(\{U > u\} \cap \{W > w\}) = \mathbb{P}(U > u)\mathbb{P}(W > w), \forall u, v.$$

Montrer que $U \perp W$.

12. Soient $X \perp Y$ de fonctions de masse

$$\mathbb{P}_X(0) = \frac{1}{2}, \mathbb{P}_X(3) = \frac{1}{2},$$

$$\mathbb{P}_Y(1) = \frac{1}{6}, \mathbb{P}_Y(2) = \frac{1}{3}, \mathbb{P}_Y(3) = \frac{1}{2}.$$

Calculer la fonction de masse de $Z = X + Y$.

13. Soient $U \perp V$ de fonctions de masse

$$\mathbb{P}_U(0) = \frac{1}{3}, \mathbb{P}_U(1) = \frac{1}{3}, \mathbb{P}_U(2) = \frac{1}{3},$$

$$\mathbb{P}_V(1) = \frac{1}{2}, \mathbb{P}_V(2) = \frac{1}{3}.$$

Calculer la fonction de masse de $W = U + V$.

14. Soient $X \perp Y$ de densités avec $X \sim \mathcal{E}(1)$, $Y \sim \mathcal{E}(2)$. Calculer la densité de $X + Y$.

15. On joue à pile ou face 5 fois de suite. Quelle est la probabilité d'obtenir 3 faces en tout ?

16. Soit $X \sim \mathcal{P}(\lambda)$ ($\lambda > 0$). Calculer $\mathbb{P}(X = 2)$, $\mathbb{P}(X \leq 2)$.

17. La durée de vie (en année) d'une ampoule est aléatoire de loi $\mathcal{E}(2)$. Calculer la probabilité que cette durée de vie soit $\geq 1,5$ année. Calculer la probabilité que cette durée soit = 1,5 année.

1.8.2 Corrigés des exercices

1. Nous avons $\Omega = B \cup B^c$ et $B \cap B^c = \emptyset$. Donc, par la loi des probabilités totales, $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$.
2. Nous avons $A \cup B = (A \cap B^c) \cup B$ et $(A \cap B^c) \cap B = \emptyset$ donc $\mathbb{P}(A \cup B) = \mathbb{P}(A \cap B^c) + \mathbb{P}(B)$.
Et par l'exercice précédent :

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

3.

(a)

FIGURE 1.5 – Dessin de la fonction de répartition.

- (b) Aux points où F n'est pas dérivable (c'est à dire en fait, seulement le point 1), on peut prendre la valeur que l'on veut (ci-dessous, $f(1) = 3$). Si vous ne savez pas dériver en 0, vous pouvez mettre n'importe quelle valeur.
 - Pour $x < 0$: $f(x) = F'(x) = 0$.
 - Pour $0 \leq x \leq 1$: $f(x) = F'(x) = 3x^2$.
 - Pour $1 < x$: $f(x) = F'(x) = 0$.

(c) Calculons

$$\begin{aligned} \int_{-\infty}^{+\infty} xf(x)dx &= \int_0^1 3x^3 dx \\ &= \left[\frac{3}{4}x^4 \right]_0^1 = \frac{3}{4}. \end{aligned}$$

(d) Calculons

$$\begin{aligned} \mathbb{P}\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right) &= \mathbb{P}\left(X = \frac{1}{4}\right) + \mathbb{P}\left(\frac{1}{4} < X \leq \frac{3}{4}\right) \\ &= F\left(\frac{1}{4}\right) - F\left(\frac{1}{4}-\right) + F\left(\frac{3}{4}\right) - F\left(\frac{1}{4}\right) \\ &= 0 + \frac{27}{64} - \frac{1}{64} \\ &= \frac{26}{64} = \frac{13}{32}. \end{aligned}$$

4.

- (a) Nous dessinons des petites boules en 0, 1, 2, 3 pour montrer que $F(0) = 1/4$, $F(0-) = 0$, $F(1) = 3/4$, $F(1-) = 1/4$, ...

FIGURE 1.6 – Dessin de la fonction de répartition.

(b) Calculons

$$\begin{aligned}
 \mathbb{E}(Z) &= p(0) \times 0 + p(1) \times 1 + p(2) \times 2 + p(3) \times 3 \\
 &= \frac{1}{2} + \frac{1}{8} \times 2 + \frac{3}{8} \\
 &= \frac{4 + 2 + 3}{8} = \frac{9}{8}.
 \end{aligned}$$

(c) Calculons

$$\begin{aligned}
 \mathbb{E}\left(\left(Z - \frac{7}{8}\right)^2\right) &= p(0) \times \left(0 - \frac{7}{8}\right)^2 + p(1) \times \left(1 - \frac{7}{8}\right)^2 \\
 &\quad + p(2) \times \left(2 - \frac{7}{8}\right)^2 + p(3) \times \left(3 - \frac{7}{8}\right)^2 \\
 &= \frac{1}{4} \times \frac{49}{64} + \frac{1}{2} \times \frac{1}{64} + \frac{1}{8} \times \frac{81}{64} + \frac{1}{8} \times \frac{289}{64} \\
 &= \frac{1}{64} \left(\frac{98}{8} + \frac{4}{8} + \frac{81}{8} + \frac{289}{8}\right) \\
 &= \frac{1}{64} \frac{472}{8} = \frac{59}{64}.
 \end{aligned}$$

5.

(a) Remarquons que $\forall z, \{Z \leq z\} = \{X \leq z\} \cap \{Y \leq z\}$. Nous avons donc

$$\begin{aligned}
 \mathbb{P}(Z \leq z) &= \mathbb{P}(\{X \leq z\} \cap \{Y \leq z\}) \\
 (\text{car } X \perp\!\!\!\perp Y) &= \mathbb{P}(X \leq z)\mathbb{P}(Y \leq z) = F_X(z)F_Y(z).
 \end{aligned}$$

(b) Remarquons que $\forall w, \{W > w\} = \{X > w\} \cap \{Y > w\}$. Nous avons donc

$$\begin{aligned}
 \mathbb{P}(W > w) &= \mathbb{P}(\{X > w\} \cap \{Y > w\}) \\
 (\text{car } X \perp\!\!\!\perp Y) &= \mathbb{P}(X > w)\mathbb{P}(Y > w) \\
 &= (1 - F_X(w))(1 - F_Y(w)).
 \end{aligned}$$

D'où le résultat.

6.

(a) La variable U produite par l'algorithme est une variable uniforme sur $[0; 1]$. Nous avons $X = U^{1/2}$ donc X est à valeurs dans $[0; 1]$. L'application $g : u \in [0; 1] \mapsto u^{1/2} \in [0; 1]$ est bijective croissante de réciproque $x \in [0; 1] \mapsto x^2 \in [0; 1]$. La densité de la variable U est $f_U : u \in \mathbb{R} \mapsto \mathbf{1}_{[0;1]}(u)$. Calculons la densité de X en $x \in [0; 1]$ en utilisant la formule de changement de variable :

$$\begin{aligned}
 f_X(x) &= \frac{f_U(g^{-1}(x))}{g'(g^{-1}(x))} \\
 &= \frac{1}{\left(\frac{1}{2\sqrt{x}}\right)} = 2x.
 \end{aligned}$$

La densité de X est donc $f_X : x \in \mathbb{R} \mapsto \mathbf{1}_{[0;1]}(x) \times 2x$.

- (b) Calculons la fonction de répartition F_X de X . Puisque X est à valeurs dans $[0; 1]$, $\forall t < 0, \mathbb{P}(X < t) = 0$ et $\forall t \geq 1, \mathbb{P}(X \leq t) = 1$. Soit $t \in [0; 1]$,

$$\begin{aligned}\mathbb{P}(X \leq t) &= \int_{-\infty}^t f_X(u) du \\ &= \int_0^t 2u du \\ &= [u^2]_0^t = t^2.\end{aligned}$$

Donc la fonction de répartition de X est $F_X : t \mapsto \begin{cases} 0 & \text{si } t < 0 \\ t^2 & \text{si } t \in [0; 1] \\ 1 & \text{si } t \geq 1. \end{cases}$

- (c) Calculons

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{+\infty} x f_X(x) dx \\ &= \int_0^1 2x^2 dx \\ &= \left[\frac{2x^3}{3} \right]_0^1 = \frac{2}{3}.\end{aligned}$$

- (d) Calculons

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{+\infty} x^2 f_X(x) dx \\ &= \int_0^1 2x^3 dx \\ &= \left[\frac{2x^4}{4} \right]_0^1 = \frac{1}{2}.\end{aligned}$$

Nous avons (par la formule (1.4.3)) $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{1}{2} - \frac{4}{12} = \frac{2}{12} = \frac{1}{6}$.

7.

- (a) La variable U est uniforme sur $]0; 1[$. Soit g la fonction définie sur $]0; 1[$ par $g(t) = -\log(t)$. La fonction g est bijective de $]0; 1[$ sur $]0; +\infty[$ et elle est décroissante. Sa réciproque est $y \in]0; +\infty[\mapsto e^{-y}$. Voir le dessin dans la figure 1.7. Nous avons

FIGURE 1.7 – Dessin de g (exercice 7)

$g'(t) = \frac{1}{t}, \forall t \in]0; 1[$. Calculons, $\forall x \in]0; +\infty[$,

$$\begin{aligned}f_X(x) &= -\frac{\mathbf{1}_{]0; 1[}(g^{-1}(x))}{g'(g^{-1}(x))} \\ &= \frac{1}{\left(\frac{1}{e^{-x}}\right)} = e^{-x}.\end{aligned}$$

Donc, $\forall x \in \mathbb{R}$, $f_X(x) = \mathbf{1}_{]0;+\infty[}(x)e^{-x}$. Soit $h : u \in]0; 1[\mapsto -\log(1-u) \in]0; +\infty[$. La fonction h est bijective croissante de fonction réciproque $y \in]0; +\infty[\mapsto 1 - e^{-y}$. Nous avons $h'(t) = \frac{1}{1-t}$, $\forall t \in]0; 1[$. Calculons, $\forall x \in]0; +\infty[$,

$$\begin{aligned} f_Y(x) &= \frac{\mathbf{1}_{]0;1]}(h^{-1}(x))}{h'(h^{-1}(x))} \\ &= \frac{1}{\left(\frac{1}{1-(1-e^{-x})}\right)} = e^{-x}. \end{aligned}$$

(b) Calculons $\forall t \in]0; +\infty[$,

$$\begin{aligned} \mathbb{P}(X \leq t) &= \mathbb{P}(-\log(U) \leq t) \\ (\text{car } g \text{ est décroissante}) &= \mathbb{P}(U \geq g^{-1}(t)) \\ &= \mathbb{P}(U \geq e^{-t}) \\ &= \mathbb{P}(U > e^{-t}) + \mathbb{P}(U = e^{-t}) \\ &= 1 - \mathbb{P}(U \leq e^{-t}) + 0 \\ &= 1 - e^{-t}. \end{aligned}$$

Et $\forall t \leq 0$, $\mathbb{P}(X \leq t) = 0$. Calculons $\forall t \in]0; 1[$,

$$\begin{aligned} \mathbb{P}(Y \leq t) &= \mathbb{P}(-\log(1-U) \leq t) \\ (\text{car } h \text{ est croissante}) &= \mathbb{P}(U \leq h^{-1}(t)) \\ &= \mathbb{P}(U \leq 1 - e^{-t}) \\ &= 1 - e^{-t}. \end{aligned}$$

Et $\forall t \leq 0$, $\mathbb{P}(Y \leq t) = 0$.

(c) Calculons la fonction de répartition jointe de X et Y , $\forall x, y \in]0; 1[$

$$\begin{aligned} \mathbb{P}(X \leq x, Y \leq y) &= \mathbb{P}(-\log(U) \leq x, -\log(1-U) \leq y) \\ &= \mathbb{P}(U \geq e^{-x}, U \leq 1 - e^{-y}) \\ &= \begin{cases} 0 & \text{si } e^{-x} > 1 - e^{-y} \\ \mathbb{P}(e^{-x} \leq U \leq 1 - e^{-y}) & \text{sinon} \end{cases} \\ &= \begin{cases} 0 & \text{si } e^{-x} > 1 - e^{-y} \\ \mathbb{P}(U = e^{-x}) + 1 - e^{-y} - e^{-x} & \text{sinon} \end{cases} \\ &= \mathbf{1}_{]0;+\infty[}(1 - e^{-y} - e^{-x}) \times (1 - e^{-y} - e^{-x}). \end{aligned}$$

Prenons $x = y = \log(2)$. Nous avons $F_{XY}(x, y) = 1 - \frac{1}{2} - \frac{1}{2} = 0$ et $F_X(x)F_Y(y) = (1 - \frac{1}{2})(1 - \frac{1}{2}) = \frac{1}{4}$. Donc $F_{XY} \neq F_X F_Y$. Donc X et Y ne sont pas indépendantes.

8.

(a) Nous avons

$$F(v) = \begin{cases} 0 & \text{si } v < 0 \\ 1 - (1-v)^A & \text{si } 0 \leq v \leq 1 \\ 1 & \text{si } v > 1. \end{cases}$$

Une autre écriture possible est

$$F(v) = \mathbf{1}_{[0;1]}(v)(1 - (1 - v)^A) + \mathbf{1}_{]1;+\infty]}(v), \forall v$$

(c'est toujours la même fonction).

- (b) Calculons la densité (notée f). La fonction F est dérivable sur $\mathbb{R} \setminus \{0, 1\}$. L'étude en 0 et 1 est un peu plus compliquée et on ne s'en préoccupe pas (cf. remarque 1.2.16). Si $v < 0$, $F'(v) = 0$. Si $v \in]0; 1[$, $F'(v) = A(1 - v)^{A-1}$. Si $v > 1$, $F'(v) = 1$. Nous avons donc

$$\forall v \in \mathbb{R}, f(v) = \begin{cases} 0 & \text{si } v < 0 \\ A(1 - v)^{A-1} & \text{si } v \in [0; 1[\\ 0 & \text{si } v \geq 1. \end{cases}$$

Remarquons que dans l'expression ci-dessus, nous avons pris de manière arbitraire $f(1) = 0$, $f(0) = 0$.

- (c) Le dessin est dans la figure 1.8 .

FIGURE 1.8 – Fonction de densité (exercice 8)

- (d) Calculons

$$\begin{aligned} \mathbb{E}(V) &= \int_{-\infty}^{+\infty} v f(v) dv \\ &= \int_0^1 v \frac{1}{2} (1 - v)^{-1/2} dv \\ (\text{intégration par parties}) &= \left[v \times (-(1 - v)^{1/2}) \right]_0^1 + \int_0^1 (1 - v)^{1/2} dv \\ &= 0 + \left[-\frac{2}{3} (1 - v)^{3/2} \right]_0^1 \\ &= \frac{2}{3}. \end{aligned}$$

- (e) Calculons

$$\begin{aligned} \mathbb{E}(V^2) &= \int_{-\infty}^{+\infty} v^2 f(v) dv \\ &= \int_0^1 v^2 \frac{1}{2} (1 - v)^{-1/2} dv \\ (\text{int. par parties}) &= \left[v^2 \times (-(1 - v)^{1/2}) \right]_0^1 + \int_0^1 2v(1 - v)^{1/2} dv \\ (\text{int. par parties}) &= 0 + \left[2v \times \left(-\frac{2}{3} (1 - v)^{3/2}\right) \right]_0^1 + \int_0^1 \frac{4}{3} (1 - v)^{3/2} dv \\ &= 0 + \left[-\frac{4}{3} \times \frac{2}{5} \times (1 - v)^{5/2} \right]_0^1 \\ &= \frac{8}{15}. \end{aligned}$$

Donc

$$\begin{aligned}\text{Var}(V) &= \mathbb{E}(V^2) - \mathbb{E}(V)^2 \\ &= \frac{8}{15} - \frac{4}{9} \\ &= \frac{24 - 20}{45} = \frac{4}{45}.\end{aligned}$$

9.

(a) Pour tout ω , $1 - \mathbf{1}_{A^c}(\omega) = 0$ si $\omega \in A^c$ (ce qui est équivalent à $\omega \notin A$) et $1 - \mathbf{1}_{A^c}(\omega) = 1$ si $\omega \notin A^c$ (ce qui est équivalent à $\omega \in A$). Donc $1 - \mathbf{1}_{A^c} = \mathbf{1}_A$.

(b) Pour tout ω ,

$$\begin{aligned}\mathbf{1}_A(\omega)\mathbf{1}_B(\omega) &= \begin{cases} 1 & \text{si } \omega \in A \text{ et } \omega \in B \\ 0 & \text{si } \omega \notin A \text{ et } \omega \in B \\ 0 & \text{si } \omega \in A \text{ et } \omega \notin B \\ 0 & \text{si } \omega \notin A \text{ et } \omega \notin B. \end{cases} \\ &= \begin{cases} 1 & \text{si } \omega \in A \cap B \\ 0 & \text{sinon} \end{cases} \\ &= \mathbf{1}_{A \cap B}(\omega),\end{aligned}$$

et

$$\begin{aligned}\min(\mathbf{1}_A(\omega), \mathbf{1}_B(\omega)) &= \begin{cases} 1 & \text{si } \omega \in A \text{ et } \omega \in B \\ 0 & \text{si } \omega \notin A \text{ et } \omega \in B \\ 0 & \text{si } \omega \in A \text{ et } \omega \notin B \\ 0 & \text{si } \omega \notin A \text{ et } \omega \notin B \end{cases} \\ &= \mathbf{1}_{A \cap B}(\omega).\end{aligned}$$

(c) Pour tout ω ,

$$\begin{aligned}\max(\mathbf{1}_A(\omega), \mathbf{1}_B(\omega)) &= \begin{cases} 1 & \text{si } \omega \in A \text{ et } \omega \in B \\ 1 & \text{si } \omega \notin A \text{ et } \omega \in B \\ 1 & \text{si } \omega \in A \text{ et } \omega \notin B \\ 0 & \text{si } \omega \notin A \text{ et } \omega \notin B \end{cases} \\ &= \mathbf{1}_{A \cup B}(\omega).\end{aligned}$$

10. Notons X_1, X_2, \dots les résultats successifs des appels à `grand(1, 1, 'uin', 0, 1)`. Ces variables sont indépendantes de fonction de masse p_X telle que $p_X(0) = 1/2, p_X(1) = 1/2$. La boucle `while` s'arrête dès que deux tirages successifs sont égaux à 0. La variable aléatoire N est à valeurs dans $2, 3, 4, \dots$

(a) Nous avons

$$\begin{aligned}\mathbb{P}(N = 2) &= \mathbb{P}(\{X_1 = 1, X_2 = 1\} \cup \{X_1 = 0, X_2 = 0\}) \\ (\text{événements disjoints}) &= \mathbb{P}(X_1 = 1, X_2 = 1) + \mathbb{P}(X_1 = 0, X_2 = 0) \\ (X_1 \perp\!\!\!\perp X_2) &= \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1) + \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.\end{aligned}$$

De manière générale, soit $k \geq 3$, k pair. Notons

$$C = \{X_1 = 0, X_2 = 1, X_3 = 0, \dots, X_{k-1} = 0, X_k = 1\},$$

$$D = \{X_1 = 1, X_2 = 0, X_3 = 1, \dots, X_k = 0\}.$$

Nous avons

$$\begin{aligned} \mathbb{P}(N = k+1) &= \mathbb{P}(\{N > k\} \cap \{X_k = X_{k+1}\}) \\ &= \mathbb{P}((C \cap \{X_k = X_{k+1}\}) \cup (D \cap \{X_k = X_{k+1}\})) \\ &= \mathbb{P}((C \cap \{X_{k+1} = 1\}) \cup (D \cap \{X_{k+1} = 0\})) \\ \text{(événements disjoints)} &= \mathbb{P}(C \cap \{X_{k+1} = 1\}) + \mathbb{P}(D \cap \{X_{k+1} = 0\}) \\ \text{(} X_1, X_2, \dots \text{ indépendants)} &= \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 1) \dots \mathbb{P}(X_{k+1} = 1) \\ &\quad + \mathbb{P}(X_1 = 1) \dots \mathbb{P}(X_{k+1} = 0) \\ &= \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}} = \frac{1}{2^k}. \end{aligned}$$

Le même résultat et le même raisonnement sont vrais pour k impair.

(b)

$$\begin{aligned} \mathbb{P}(N \text{ pair}) &= \mathbb{P}(\cup_{i \geq 2, i \text{ pair}} \{N = i\}) \\ \text{(événements disjoints)} &= \sum_{i \geq 2, i \text{ pair}} \mathbb{P}(N = i) \\ &= \sum_{i \geq 2, i \text{ pair}} \frac{1}{2^{i-1}} \\ &= \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \dots \\ \text{(série géométrique)} &= \frac{1}{2} \times \frac{1}{1 - \frac{1}{4}} = \frac{4}{6} = \frac{2}{3}. \end{aligned}$$

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(\{N = 2\} \cap \{N = 3\} \cap \dots \cap \{N = 6\}) \\ \text{(événements disjoints)} &= \mathbb{P}(2) + \mathbb{P}(3) + \dots + \mathbb{P}(6) \\ &= \sum_{i=2}^6 \frac{1}{2^{i-1}} \\ \text{(série géométrique)} &= \frac{1}{2} \times \frac{1 - (\frac{1}{2})^5}{1 - (\frac{1}{2})} = \frac{1}{32}. \end{aligned}$$

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(\{N = 2\} \cup \{N = 4\} \cup \{N = 6\}) \\ \text{(événements disjoints)} &= \frac{1}{2} + \frac{1}{8} + \frac{1}{32} \\ &= \frac{16 + 4 + 1}{32} = \frac{21}{32} = \frac{1}{2}. \end{aligned}$$

11. Pour tous événements A, B , nous utilisons la notation $C = A \sqcup B$ pour dire $C = A \cup B$ et $A \cap B = \emptyset$. Notons F_U, F_W les fonctions de répartition de U, W respectivement. Calculons $\forall u, v$

$$(\{U \leq u\} \cap \{W \leq w\})^c = \{U > u\} \cup \{W > w\} \text{ donc :}$$

$$\begin{aligned} \mathbb{P}(U \leq u, W \leq w) &= 1 - \mathbb{P}(\{U > u\} \cup \{W > w\}) \\ &= 1 - \mathbb{P}(\{U > u\} \sqcup (\{U \leq u\} \cap \{W > w\})) \\ &= 1 - [\mathbb{P}(U > u) + \mathbb{P}(U \leq u, W > w)] \\ &= 1 - [1 - \mathbb{P}(U \leq u) + \mathbb{P}(W > w) - \mathbb{P}(U > u, W > w)] \\ &= \mathbb{P}(U \leq u) - (1 - \mathbb{P}(W \leq w)) - (1 - \mathbb{P}(U \leq u))(1 - \mathbb{P}(W \leq w)) \\ &= F_U(u) - (1 - F_W(w)) - (1 - F_U(u))(1 - F_W(w)) \\ &= F_U(u)F_W(w). \end{aligned}$$

12. La variable Z est à valeurs dans $\{1, 2, 3, 4, 5, 6\}$. Calculons

$$\begin{aligned} \mathbb{P}(Z = 1) &= \mathbb{P}(X = 0, Y = 1) \\ (\text{car } X \perp\!\!\!\perp Y) &= p_X(0)p_Y(1) = \frac{1}{12}, \end{aligned}$$

de même :

$$\begin{aligned} \mathbb{P}(Z = 2) &= \mathbb{P}(X = 0, Y = 2) = \frac{1}{6}, \\ \mathbb{P}(Z = 3) &= \mathbb{P}(X = 0, Y = 3) = \frac{1}{4}, \\ \mathbb{P}(Z = 4) &= \mathbb{P}(X = 3, Y = 1) = \frac{1}{12}, \\ \mathbb{P}(Z = 5) &= \mathbb{P}(X = 3, Y = 2) = \frac{1}{6}, \\ \mathbb{P}(Z = 6) &= \mathbb{P}(X = 3, Y = 3) = \frac{1}{4}. \end{aligned}$$

13. La variable W est à valeurs dans $\{1, 2, 3, 4\}$. Calculons

$$\begin{aligned} \mathbb{P}(W = 1) &= \mathbb{P}(U = 0, V = 1) \\ (\text{car } U \perp\!\!\!\perp V) &= \mathbb{P}(U = 0)\mathbb{P}(V = 1) = \frac{1}{6}, \end{aligned}$$

$$\begin{aligned} \mathbb{P}(W = 2) &= \mathbb{P}(\{U = 0, V = 2\} \sqcup \{U = 1, V = 1\}) \\ (\text{réunion disjointe}) &= \mathbb{P}(U = 0, V = 2) + \mathbb{P}(U = 1, V = 1) \\ (\text{car } U \perp\!\!\!\perp V) &= \mathbb{P}(U = 0)\mathbb{P}(V = 2) + \mathbb{P}(U = 1)\mathbb{P}(V = 1) \\ &= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}, \end{aligned}$$

de même :

$$\begin{aligned} \mathbb{P}(W = 3) &= \mathbb{P}(U = 1, V = 2) + \mathbb{P}(U = 2, V = 1) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}, \\ \mathbb{P}(W = 4) &= \mathbb{P}(U = 2, V = 2) = \frac{1}{6}. \end{aligned}$$

14. Notons f la densité de $X + Y$. Notons f_X la densité de X et f_Y la densité de Y . Calculons pour $x < 0$,

$$\begin{aligned} f(x) &= \int_{-\infty}^{+\infty} f_X(t)f_Y(x-t)dt \\ &= \int_{-\infty}^{+\infty} \mathbf{1}_{[0;+\infty[}(t)e^{-t}\mathbf{1}_{[0;+\infty[}2e^{-2(x-t)}dt. \end{aligned}$$

Pour tout $t \in \mathbb{R}$, $\min(t, x-t) < 0$ donc $\mathbf{1}_{[0;+\infty[}(t) \times \mathbf{1}_{[0;+\infty[}(x-t) = 0$. Donc $f(x) = 0$. Prenons maintenant $x \geq 0$,

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} \mathbf{1}_{[0;+\infty[}(t)e^{-t}\mathbf{1}_{[0;+\infty[}2e^{-2(x-t)}dt \\ &= \int_0^{+\infty} e^{-t}\mathbf{1}_{[0;+\infty[}2e^{-2(x-t)}dt \\ &= \int_0^x e^{-t}2e^{-2(x-t)}dt \\ &= \int_0^x 2e^{-2x+t}dt \\ &= 2e^{-2x}[e^t]_0^x \\ &= 2(e^{-x} - e^{-2x}). \end{aligned}$$

Donc la densité cherchée est

$$x \in \mathbb{R} \mapsto \mathbf{1}_{[0;+\infty[}(x) \times 2(e^{-x} - e^{-2x}).$$

15. Notons X_1, \dots, X_5 (à valeurs dans $\{P, F\}$) les résultats successifs. Ces variables sont indépendantes. Notons A l'événement dont on veut calculer la probabilité. Nous avons (ici # veut dire «cardinal»)

$$A = \sqcup_{I \subset \{1,2,\dots,3\}, \#I=3} \{X_i = F, \forall i \in I, X_j = P, \forall j \in I^c\}.$$

Donc

$$\begin{aligned} \mathbb{P}(A) &= \sum_{I \subset \{1,2,\dots,3\}, \#I=3} \mathbb{P}(X_i = F, \forall i \in I, X_j = P, \forall j \in I^c) \\ (\text{les } X_k \text{ sont indépendants}) &= \sum_{I \subset \{1,2,\dots,3\}, \#I=3} \prod_{i \in I} \mathbb{P}(X_i = F) \prod_{j \in I^c} \mathbb{P}(X_j = P) \\ &= \sum_{I \subset \{1,2,\dots,3\}, \#I=3} \frac{1}{2^5} \\ &= C_3^3 \frac{1}{2^5}. \end{aligned}$$

- 16.

$$\mathbb{P}(X = 2) = \frac{\lambda^2}{2!} e^{-2\lambda}$$

$$\begin{aligned} \mathbb{P}(X \leq 2) &= \mathbb{P}(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\}) \\ (\text{réunion disjointe}) &= e^{-2\lambda} + \lambda e^{-2\lambda} + \frac{\lambda^2}{2!} e^{-2\lambda}. \end{aligned}$$

17. Soit X cette durée de vie. Calculons (en utilisant la proposition 1.2.13)

$$\begin{aligned}\mathbb{P}(X \geq 1,5) &= \int_{1,5}^{+\infty} 2e^{-2t} dt \\ &= \left[-e^{-2t}\right]_{3/2}^{+\infty} = e^{-3},\end{aligned}$$

$$\mathbb{P}(X = 1,5) = 0.$$

Chapitre 2

Théorèmes limites et méthode de Monte-Carlo

2.1 Les différentes notions de convergence

On s'intéresse ici à des v.a.r.

Définition 2.1.1. On dit que X_n converge presque sûrement vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$ si

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = \lim_{n \rightarrow +\infty} X_n(\omega) = X\}) = 1 .$$

Définition 2.1.2. Soit $p > 0$, on dit que X_n converge dans L^p vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{L^p} X$ si $\mathbb{E}(|X - X_n|^p) \xrightarrow[n \rightarrow +\infty]{} 0$

Définition 2.1.3. On dit que X_n converge en probabilité vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{proba.} X$ si $\forall \epsilon > 0$, $\mathbb{P}(|X - X_n| \geq \epsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.

Définition 2.1.4. On dit que X_n converge en loi vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{loi} X$ si $\forall \phi$ continue bornée de \mathbb{R} dans \mathbb{R} , $\mathbb{E}(\phi(X_n)) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(\phi(X))$.

Attention, ces notions ne sont pas équivalentes.

2.2 Théorèmes limites

2.2.1 Loi des grands nombres

Théorème 2.2.1. Loi des grands nombres. Soient X_1, X_2, X_3, \dots des v.a.r. indépendantes et de même loi (nous dirons aussi indépendantes et identiquement distribuées, ou i.i.d.) telles que $\mathbb{E}(X_1^2) < \infty$. Alors

$$S_n := \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1). \quad (2.2.1)$$

La quantité S_n ci-dessus s'appelle la moyenne empirique de X_1, \dots, X_n .

D'après la définition 2.1.1, (2.2.1) veut dire que $\mathbb{P}(\{\omega : S_n(\omega) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(X)\}) = 1$. La convergence n'a pas lieu forcément pour tous les ω .

Exemple 2.2.2. Soient des variables i.i.d. X_1, X_2, \dots de loi $\mathcal{B}(1/2)$. Nous avons

$$\begin{aligned} \mathbb{E}(X_1^2) &= \mathbb{P}(X_1 = 0) \times 0^2 + \mathbb{P}(X_1 = 1) \times 1^2 \\ &= \frac{1}{2} < \infty. \end{aligned}$$

Nous pouvons donc appliquer la loi des grands nombres :

$$S_n := \frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1) = \frac{1}{2}.$$

La convergence n'a pas lieu pour tous les ω . Il existe en effet $\omega_0 \in \Omega$ tel que $X_i(\omega_0) = 1, \forall i$. Dans ce cas

$$S_n(\omega_0) \xrightarrow[n \rightarrow +\infty]{} 1.$$

Le théorème nous dit que l'ensemble $\{\omega \in \Omega : S_n(\omega) \text{ ne converge pas vers } \mathbb{E}(X_1)\}$ est de probabilité nulle et que l'ensemble $\{\omega \in \Omega : S_n(\omega) \xrightarrow[n \rightarrow +\infty]{} \frac{1}{2}\}$ est de probabilité 1.

Algorithme 2.1 Illustration de la loi des grands nombres

```

N=100 ;
X=[] ;
S=0 ;
for i=1 :N
    U=grand(1,1,'bin',1,1/2) ; // on tire une variable de Bernoulli de paramètre 1/2
    S=S+U ; // S va contenir les sommes partielles des variables de Bernoulli
    X=[X,S/i] ; // on range dans un vecteur X les quantités S/i (correspond au  $S_i$  de l'exemple
ci-dessus)
end,
xbase() ; // efface la fenêtre graphique
plot2d([1 :N],X,2) ; // dessine un graphe des valeurs dans X (les points sont reliés par
// des traits bleus)
plot2d([0 :N],0.5*ones(1,N+1),0) ; // dessine une ligne horizontale  $y = 0,5$  en
// pointillés
plot2d([0 :N],0*ones(1,N+1),rect=[0,-0.2,N,1]) ; // dessine une ligne horizontale
//  $y = 0$  en noir en fixant les bornes (rect=[xmin,ymin,xmax,ymax])

```

Une première exécution de l'algorithme affiche le graphique suivant :

FIGURE 2.1 – Essai n° 1

On voit une suite $S_i(\omega)$ qui converge vers $1/2$. Lançons une deuxième exécution, nous obtenons :

FIGURE 2.2 – Essai n° 2

Le dessin est différent mais nous avons toujours convergence vers $1/2$. Voir <http://www-sop.inria.fr/mefisto/java/tutorial1/tutorial1.html> (Applet : Pile ou face) pour une version dynamique des illustrations ci-dessus.

2.2.2 Application de la loi des grands nombres

2.2.2.1 Dessin de la fonction de répartition

Supposons que nous disposions de réalisations indépendantes $X_1(\omega), X_2(\omega), \dots$ (dans \mathbb{R}) d'une même loi (c'est à dire que les variables X_1, X_2, \dots sont i.i.d. et ont toute la même fonction de répartition, que nous noterons F). Nous aimerions faire un dessin approximatif de F . Soit $t \in \mathbb{R}$. Pour tout i , nous introduisons la variable $U_i = \mathbf{1}_{1-\infty:t]}(X_i)$. Nous avons, en utilisant la propriété de croissance de l'espérance (proposition 1.3.12) $\mathbb{E}(U_i^2) \leq \mathbb{E}(U_i) = 1 < \infty$. Donc, par la loi des grands nombres

$$S_n := \frac{U_1 + \dots + U_n}{n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbb{E}(U_1).$$

Or $\mathbb{E}(U_1) = \mathbb{P}(X \leq t) \times 1 + \mathbb{P}(X > t) \times 0 = \mathbb{P}(X \leq t)$. Cette remarque nous donne un moyen de calculer $F(t)$. Ce genre de méthode s'appelle une méthode de Monte-Carlo. Commençons par écrire une fonction qui simule une variable aléatoire.

Algorithme 2.2 Fonction «variable»

```

fonction [z]=variable(u) // le paramètre u ne joue aucun rôle dans le résultat
  z=grand(1,1,'unf',0,1); // ici nous simulons en fait une variable connue (uniforme
// sur [0; 1])
endfonction

```

Calculons ensuite S_N pour un certain N (et un certain t définissant les variables U_i).

Algorithme 2.3 Calcul de S_N

```

N=10000; // nous prenons N «très grand»
t=0.5; // nous fixons t
s=0;
for i=1 :N
  X=variable(1); // nous tirons une variable X avec la fonction «variable»
  if (X<t) then
    U=1;
  else
    U=0;
  end, // U vaut 1 si X ≤ t et 0 sinon
  s=s+U;
end,
// s contient la somme de toutes les variables U tirées dans la boucle
disp(s/N); // affiche s/N

```

Voici un exemple de ce que l'on obtient en plusieurs exécutions de ce programme

```

-> 0.57
-> 0.47
-> 0.4911
-> 0.4996
-> 0.5039

```

Nous n'obtenons pas deux fois la même chose mais à chaque fois, le résultat est proche de $\mathbb{E}(U_1) = 1/2$.

Dans le même esprit, l'algorithme 2.4 dessine une approximation numérique de la fonction de répartition. Le résultat est dans la figure 2.3.

Algorithme 2.4 Approximation numérique de la fonction de répartition

```

N=100 ; % on choisit N «très grand»
a=-1 ;
b=2 ; % on va dessiner F entre a et b
h=0.1 ; % pas d'approximation de la fonction F
X=[] ;
for i=1 :N
    x=variable(1) ;
    X=[X,x] ;
end, // on remplit un vecteur X avec des tirages de la variable aléatoire (X1, X2, ...)
t=a-h ;
T=[] ; // T va contenir les valeurs de t pour lesquelles on calcule F(t)
F=[] ; // F va contenir les approximations de F(t)
while(t<b) // t parcourt a,a+h,a+2h,...
    t=t+h ;
    T=[T,t] ;
    C=find(X<t) ; // C contient les indices i pour lesquels Xi ≤ t
    F=[F,length(C)/N] ; // length(C) est la longueur de C
// length(C)/N est proche de F(t) par la loi des grands nombres
end,
xbasc() ;
plot2d(T,F,5,rect=[a,-0.2,b,1.2]) ;

```

FIGURE 2.3 – Approximation numérique de la fonction de répartition

Cet algorithme tire des variables aléatoires i.i.d. X_1, X_2, \dots, X_N de même fonction de répartition F . Pour chaque $t \in \{a + kh : k \in \mathbb{N}, a + kh < b\}$, l'algorithme calcule $\#\{i : X_i \leq t\}$. fixons un tel t . Notons, $\forall i, U_i = \mathbf{1}_{]-\infty, t]}(X_i)$. Nous remarquons que

$$\#\{i : X_i \leq t\} = \frac{U_1 + \dots + U_N}{N}.$$

Donc, par la loi des grands nombres, $\#\{i : X_i \leq t\}$ devrait être proche de

$$\begin{aligned}
 \mathbb{E}(U_1) &= \mathbb{P}(U_1 = 0) \times 0 + \mathbb{P}(U_1 = 1) \times 1 \\
 &= \mathbb{P}(U_1 = 1) \\
 &= \mathbb{P}(X_1 \leq t) \\
 &= F(t).
 \end{aligned}$$

2.2.2.2 Dessin de la densité

Supposons que nous disposions de réalisations indépendantes $X_1(\omega), X_2(\omega), \dots$ (dans \mathbb{R}) d'une même loi continue (c'est à dire que les variables X_1, X_2, \dots sont i.i.d. et ont toute la même

densité, que nous noterons f). Nous supposons ici que f est continue. Nous aimerions faire un dessin approximatif de f . Soit $a < b \in \mathbb{R}$. Pour tout i , nous introduisons la variable $U_i = \mathbf{1}_{[a,b]}(X_i)$. Nous avons, en utilisant la propriété de croissance de l'espérance (proposition 1.3.12) $\mathbb{E}(U_i^2) \leq \mathbb{E}(U_i) = 1 < \infty$. Donc, par la loi des grands nombres

$$S_n := \frac{U_1 + \cdots + U_n}{n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbb{E}(U_1).$$

Or

$$\begin{aligned} \mathbb{E}(U_1) &= \mathbb{P}(U_1 = 0) \times 0 + \mathbb{P}(U_1 = 1) \times 1 \\ &= \mathbb{P}(X_1 \in [a; b]) \\ &= \int_a^b f(x) dx. \end{aligned}$$

Puisque f' est bornée,

$$\left| \frac{1}{b-a} \int_a^b f(x) dx - f(a) \right| \leq (b-a) \times \sup_{t \in \mathbb{R}} |f'(t)|$$

(nous ne démontrons pas ce point). Cette remarque nous donne un moyen de calculer $f(t)$. Nous prenons $b-a$ «petit», et n «grand» (voir ... pour une discussion sur le choix de ces paramètres). Nous allons maintenant écrire un algorithme basé sur ces idées. Commençons par écrire une fonction qui simule une variable aléatoire.

Algorithme 2.5 Simulation de variable aléatoire

```

fonction [z]=mystere(x)
    u=grand(1,2,'unf',0,1); // on remplit un vecteur 1 x 2 avec deux variables aléatoires
    // ~ U([0; 1]) indépendantes
    z=cos(2*pi*u(1,1))*sqrt(-log(u(1,2)));
endfonction

```

L'algorithme 2.6 dessine une approximation de la densité. Un dessin produit par ce programme se trouve dans la figure 2.4.

FIGURE 2.4 – Dessin approximatif de la densité

2.2.3 Théorème central-limite

Théorème 2.2.3. Théorème central-limite (ou TCL). Soient X_1, X_2, \dots des v.a.r. i.i.d. telles que $\text{Var}(X_1) = \sigma^2 < \infty$. Alors

$$\sqrt{n} \left(\frac{X_1 + \cdots + X_n}{n} - \mathbb{E}(X_1) \right) \xrightarrow[n \rightarrow +\infty]{\text{loi}} Z$$

avec $Z \sim \mathcal{N}(0, \sigma^2)$.

Exemple 2.2.4. Soient U_1, U_2, \dots i.i.d. de loi $\mathcal{B}(1/2)$. Posons $\forall i, V_i = 2(U_i - 1/2)$. Par la loi des grands nombres,

$$\frac{V_1 + \cdots + V_n}{n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbb{E}(V_1).$$

Algorithme 2.6 Programme pour dessiner la densité (approximativement)

```

a=-3 ;
b=3 ; // on va dessiner la fonction entre a et b
h=0.2 ; // on choisit un pas d'approximation
N=10000 ; // on choisit N «grand»
X=[] // X va contenir les variables aléatoires simulées avec la fonction «mystere»
for i=1 :N
    x=mystere(1) ;
    X=[X x] ; // on remplit X
end,
t=a-h ;
T=[] ;
approx=[] ; // approx va contenir les valeurs successives de l'approximation de f
while (t<b) // t parcourt a,a+h,... on s'arrête quand on dépasse b
    t=t+h/2 ;
    T=[T t] ; // T contient les valeurs successives de t
    C=find(X<t+h) ; // C contient les indices de X pour lesquels la valeur est <t+h
    D=find(X>t) ; // D contient les indices de X pour lesquels la valeur est >t
    E=intersect(C,D) ; // E contient les valeurs communes à C et D
    approx=[approx length(E)/(h*N)] ; // on remplit le vecteur approx
end,
clf() ;
plot2d(T,approx,5,rect=[a-0.5,-0.5,b+0.5,max(approx)+0.5]) ;
drawaxis(x=[a-0.5,1,b+0.5],y=0,dir='d',tics='v') ; // dessin d'un axe horizontal
drawaxis(x=0,y=[-0.5,1,max(approx)+0.5],dir='r',tics='v') ; // dessin d'un axe vertical

```

Et

$$\begin{aligned}
 \mathbb{E}(V_1) &= \mathbb{P}(V_1 = -1) \times (-1) + \mathbb{P}(V_1 = 1) \times 1 \\
 &= \mathbb{P}(U_1 = 0) \times (-1) + \mathbb{P}(U_1 = 1) \times 1 \\
 &= 0.
 \end{aligned}$$

Le théorème central-limite nous dit que pour toute fonction bornée ϕ ,

$$\mathbb{E}\left(\phi\left(\sqrt{n}\left(\frac{V_1 + \dots + V_n}{n}\right)\right)\right) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(\phi(Z))$$

avec une certaine variable $Z \sim \mathcal{N}(0, \sigma^2)$ où

$$\begin{aligned}
 \sigma^2 &= \text{Var}(V_1) \\
 &= \mathbb{E}(V_1^2) - \mathbb{E}(V_1)^2 \\
 &= \mathbb{P}(V_1 = -1) \times (-1)^2 + \mathbb{P}(V_1 = 1) \times 1^2 \\
 &= 1.
 \end{aligned}$$

Le théorème central-limite peut aussi s'interpréter de la manière suivante : pour n «grand», $\sqrt{n}\left(\frac{V_1 + \dots + V_n}{n}\right)$ est (presque) de loi $\mathcal{N}(0, \sigma^2)$. Le programme suivant simule cette variable $\sqrt{n}\left(\frac{V_1 + \dots + V_n}{n}\right)$.

Algorithme 2.7 Simulation de variable aléatoire

```

N=1000;
U=grand(1,N,'bin',1,1/2); // au lieu de faire une boucle, on simule N variables de
// Bernoulli indépendantes que l'on range dans une matrice de taille 1×N
V=2*(U-0.5*ones(1,N)); // sur chaque composante des vecteurs,
// nous avons  $V_i = 2(U_i - 1/2)$ 
disp(sum(V)/sqrt(N));

```

Des appels successifs à ce programme renvoient par exemple :

```

- 0.4427189
0.1897367
3.2887688
- 0.1264911
0.1264911
0.3794733 ...

```

Nous obtenons un résultat différent à chaque fois. Ces variables sont indépendantes et de même loi (approximativement $\mathcal{N}(0,1)$). Nous appliquons les idées de 2.2.2.2 pour faire un dessin approximatif de la densité (voir algorithme 2.8). Nous

Algorithme 2.8 «Vérification» du théorème central-limite

```

N1=10000;
N2=10000;
h=0.2;
a=-3;
b=3;
X=[];
for i=1 :N2
    U=grand(1,N1,'bin',1,1/2);
    V=2*(U-0.5*ones(1,N1));
    X=[X sum(V)/sqrt(N1)];
end,
clf();
// à partir de là, le code pour faire le dessin est identique à celui de l'algorithme 2.6

```

obtenons le dessin de la figure 2.5 dans lequel on reconnaît la densité de la loi $\mathcal{N}(0,1)$.

FIGURE 2.5 – Dessin approximatif de la densité de $\mathcal{N}(0,1)$

2.2.4 Application du TCL

2.2.4.1 Sondages

On s'intéresse au nombre de gens qui achètent de la lessive Ariel en France. On ne peut pas interroger toute la population et on se contente donc d'un échantillon de personnes. Introduisons

la variable

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ème personne interrogée achète Ariel} \\ 0 & \text{si la } i\text{-ème personne interrogée n'achète pas Ariel.} \end{cases}$$

Les variables X_i sont supposées i.i.d. avec $\mathbb{P}(X_i = 1) = p$ (ce sont nos hypothèses de modélisation). La quantité p est celle que nous cherchons à déterminer. Remarquons que $\mathbb{E}(X_1) = p \times 1 + (1 - p) \times 0 = p$. Par la loi (forte) des grands nombres

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbb{E}(X_1) = p.$$

Quelle taille n d'échantillon sélectionner pour que $\frac{X_1 + \dots + X_n}{n}$ soit proche de p ? Supposons que l'on veuille n tel que la probabilité de se tromper de plus de 0,01 dans notre estimée de p soit plus petite que 0,1, c'est à dire

$$\mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - p\right| \geq 0,01\right) \leq 0,1. \quad (2.2.2)$$

Notons $\sigma^2 = \text{Var}(X_1)$. Nous avons

$$\begin{aligned} \mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - p\right| \geq 0,01\right) &= \mathbb{P}\left(\left|\frac{(X_1 - p) + \dots + (X_n - p)}{\sigma \sqrt{n}}\right| \geq \frac{\sqrt{n} \times 0,01}{\sigma}\right) \\ (\text{par TCL}) &\approx \mathbb{P}\left(Z \geq \frac{\sqrt{n} \times 0,01}{\sigma}\right) \text{ avec } Z \sim \mathcal{N}(0,1) \\ &= 2 \int_{\frac{\sqrt{n} \times 0,01}{\sigma}}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ &= 2 \left(1 - \int_{-\infty}^{\frac{\sqrt{n} \times 0,01}{\sigma}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx\right). \end{aligned} \quad (2.2.3)$$

Nous voyons sur une table (cf. annexe A) qu'il suffit de prendre n tel que $\sqrt{n} \times 0,01/\sigma \geq 1.65$. Calculons

$$\begin{aligned} \text{Var}(X_1) &= \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 \\ &= p \times 1^2 + (1 - p) \times 0^2 - p^2 \\ &= p - p^2 = p(1 - p). \end{aligned}$$

Nous avons alors que

$$n \geq \left(\frac{1,65 \times \sqrt{p(1-p)}}{0,01}\right)^2$$

réalise (??). Mais justement, nous ne connaissons pas p . Nous étudions la fonction

$$p \in [0,1] \mapsto p(1-p).$$

FIGURE 2.6 – Parabole $p \mapsto p(1-p)$

C'est une parabole qui atteint son max. en 1/2. Donc, $\forall p \in [0,1]$,

$$\left(\frac{1,65 \times \sqrt{p(1-p)}}{0,01}\right)^2 \leq \left(\frac{1,65 \times \sqrt{0,5 \times 0,5}}{0,01}\right)^2.$$

Donc il suffit de prendre $n = \left(\frac{1,65 \times \sqrt{0,5 \times 0,5}}{0,01}\right)^2$.

2.2.4.2 Planche de Galton

Voir l'animation : <http://www-sop.inria.fr/sysdys/java/tutorial1/node11.html>. On appelle ce dispositif une planche de Galton. Voir la figure 2.7 pour une capture d'écran de cette expérience. Notons n le nombre de niveau sur la planche de Galton. On va supposer n pair.

FIGURE 2.7 – Planche de Galton

Chaque bille rebondit aléatoirement à chaque niveau. Par convention, elle se déplace à chaque niveau de $+1$ ou -1 en abscisse (un cran vers la gauche ou vers la droite). On obtient $+1$ (ou -1) avec probabilité $1/2$. Nous avons donc l'abscisse finale d'une bille qui a la loi de

$$Z = X_1 + X_2 + \cdots + X_n,$$

avec X_1, \dots, X_n i.i.d. de loi suivante

$$X_1 = \begin{cases} +1 & \text{avec proba. } 1/2 \\ -1 & \text{avec proba. } 1/2. \end{cases}$$

Une telle variable Z est à valeurs dans $\{-n, -n+2, \dots, 0, 2, \dots, n\}$, c'est à dire $\{-n+2k : k \in \{1, 2, \dots, n\}\}$. Calculons, $\forall k \in \{0, 1, \dots, n\}$, la quantité $\mathbb{P}(Z = -n+2k)$. Nous remarquons que

$$\begin{aligned} \{X_1 + \cdots + X_n = -n+2k\} \\ = \sqcup_{\{j_1, \dots, j_k\} \subset \{1, \dots, n\}} \{X_i = 1, \forall i \in \{X_{j_1}, \dots, X_{j_k}\} \text{ et } X_j = -1, \forall j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_k\}\}. \end{aligned}$$

Donc

$$\begin{aligned} \mathbb{P}(Z = -n+2k) \\ &= \sum_{\{j_1, \dots, j_k\} \subset \{1, \dots, n\}} \mathbb{P}(X_i = 1, \forall i \in \{X_{j_1}, \dots, X_{j_k}\} \text{ et } X_j = -1, \forall j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_k\}) \\ &\quad \text{(car les } X_1, \dots, X_n \text{ sont indépendants)} \\ &= \sum_{\{j_1, \dots, j_k\} \subset \{1, \dots, n\}} \left(\frac{1}{2}\right)^n \\ &= C_n^k \left(\frac{1}{2}\right)^n. \end{aligned}$$

Rappelons que C_n^k est le nombre de parties à k éléments dans un ensemble à n éléments. L'animation référencée ci-dessus n'est pas facile à comprendre car le programme renormalise le graphe au cours de l'exécution. Plusieurs billes sont jetées donc on obtient Z_1, \dots, Z_m variables i.i.d. de même loi que Z . Pour tout k impair, la k -ème barre de l'histogramme en partant de l'origine a la hauteur

$$\frac{\sqrt{n}}{2} \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{(k)}(Z_i),$$

(la renormalisation compliquée $\frac{\sqrt{n}}{2}$ vient de la renormalisation automatique du graphe) cette dernière quantité

$$\xrightarrow[m \rightarrow +\infty]{\text{p.s.}} \frac{\sqrt{n}}{2} \mathbb{P}(Z = k) \quad (2.2.4)$$

par la loi des grands nombres. Fixons k impair et notons $c = k/\sqrt{n}$. La k -ème barre de l'histogramme est à l'abscisse c . Nous avons

$$\begin{aligned} \frac{\sqrt{n}}{2} \mathbb{P}(X_1 + \dots + X_n = k) &= \frac{\sqrt{n}}{2} \mathbb{P}(X_1 + \dots + X_n \in [k-1; k+1]) \\ &= \frac{\sqrt{n}}{2} \mathbb{P}\left(\frac{X_1 + \dots + X_n}{\sqrt{n}} \in [c-1/\sqrt{n}; c+1/\sqrt{n}]\right). \end{aligned}$$

Le théorème de Dini (que nous n'énonçons pas dans ce polycopié) et le TCL nous donnent ensuite

$$\frac{\sqrt{n}}{2} \mathbb{P}(X_1 + \dots + X_n = k) \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \quad (2.2.5)$$

avec

$$\sigma^2 = \text{Var}(X_1) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1) = 2.$$

Les équations (2.2.4) et (2.2.5) expliquent la forme en cloche de la simulation ci-dessus.

2.3 Exercices

2.3.1 Énoncés

- Soient U_1, U_2, \dots indépendantes et identiquement distribuées de loi $\mathcal{E}(1)$ (loi exponentielle de paramètre 1).
 - Calculer $\mathbb{E}(U_1)$, $\text{Var}(U_1)$.
 - Estimer $\mathbb{P}(U_1 + \dots + U_n \geq n(1 + \alpha))$ pour $n = 100$, $\alpha = 1/10$.
- Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\forall x, y, |f(x) - f(y)| \leq C \inf(1, |x - y|)$ pour une certaine constante C .
 - Si $X_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} X$ (rappel : pour p.t. ω , $X_n(\omega) \xrightarrow[n \rightarrow +\infty]{} X(\omega)$), montrer que $\mathbb{E}(f(X)) - \mathbb{E}(f(X_n)) \xrightarrow[n \rightarrow +\infty]{} 0$.
 - Soit $\epsilon > 0$, toujours sous l'hypothèse $X_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} X$, montrer que $\mathbb{P}(|f(X_n) - f(X)| \geq \epsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.
- On achète un stock d'ampoules pour un lampadaire. Les ampoules ont une durée de vie de loi $\mathcal{E}(\lambda)$. La première ampoule dure un temps X_1 , on la remplace immédiatement et la deuxième qui dure un temps $X_2 \dots$. Soit $T > 0$. On admet que le nombre d'ampoules N grillées pendant le temps T est tel que N est de loi $\mathcal{P}(\lambda T)$. On suppose que $\lambda T \in \mathbb{N}$.
 - Calculer $m = \mathbb{E}(N)$.
 - Soit $p \in \mathbb{N}^*$. Montrer que $\mathbb{P}(N \geq m + p) = \mathbb{P}(X_1 + \dots + X_{m+p} \leq T)$.
 - On suppose maintenant que $\lambda = 1$, $T = 20$, $p = 5$. Donner une valeur numérique approchée de $\mathbb{P}(N \geq m + p)$ à l'aide de la table jointe.
 - Avec les mêmes valeurs numériques que ci-dessus, combien d'ampoules faut-il acheter au minimum pour que

$$\mathbb{P}(\text{se retrouver à court d'ampoules avant le temps } T) < 0.05 ?$$

4. Pour sa migration annuelle, une grenouille part d'une mare située sur un plan au point de coordonnées $(-25, 0)$ dans le repère orthonormé xOy . Elle est repérée par sa position Z_n au temps n . On suppose que :

au temps 0, sa position est $Z_0 = (-25, 0)$

$$\text{et } \forall n \geq 0, Z_{n+1} = Z_n + (1, 0) + U_n,$$

où les variables U_n sont i.i.d. avec $\mathbb{P}(U_n = (0, 1/\sqrt{2})) = 1/2$, $\mathbb{P}(U_n = (0, -1/\sqrt{2})) = 1/2$. Ainsi à chaque étape de sa progression, la grenouille avance de +1 dans la direction Ox et se déporte en même temps de $\pm 1/\sqrt{2}$ dans la direction perpendiculaire Oy . Sur l'axe des ordonnées se trouve cette année une autoroute neuve. On décide de creuser des tunnels sous l'autoroute le long d'une certaine zone pour permettre le passage de cette grenouille. La zone à tunnels se situe entre des points d'ordonnées a et b . Si la grenouille arrive dans cette zone, elle passe dans un tunnel et sinon elle se fait écraser. Voir figure 2.8 pour un dessin.

FIGURE 2.8 – Plan

- (a) À quel instant passe-t-elle par l'autoroute ?
- (b) Supposons que l'on construise une zone de tunnels entre les points d'ordonnées -5 et 5 (compris). Donner une approximation de la probabilité qu'a la grenouille de passer par un tunnel. (Dans les calculs, on arrondira au deuxième chiffre après la virgule pour simplifier.)
- (c) On décide de construire une zone de tunnels entre des points d'ordonnées $-x$ et $+x$ ($x > 0$). Donner une valeur approximative de x telle que la probabilité de survie de la grenouille soit 0.9. (Dans les calculs, on arrondira au deuxième chiffre après la virgule pour simplifier.)

2.3.2 Corrigés

1. (a)

$$\begin{aligned} \mathbb{E}(U_1) &= \int_0^{+\infty} x e^{-x} dx \\ &= [-x e^{-x}]_0^{+\infty} + \int_0^{+\infty} e^{-x} dx \\ &= 0 + [-e^{-x}]_0^{+\infty} \\ &= 1. \end{aligned}$$

- (b)

$$\begin{aligned} \mathbb{E}(U_1^2) &= \int_0^{+\infty} x^2 e^{-x} dx \\ &= [-x^2 e^{-x}]_0^{+\infty} + \int_0^{+\infty} 2x e^{-x} dx \\ &= [-2x e^{-x}]_0^{+\infty} + \int_0^{+\infty} 2e^{-x} dx \\ &= 2. \end{aligned}$$

Donc $\text{Var}(U_1) = 1$.

(c) Les variables U_1, U_2, \dots sont L^2 , on peut donc appliquer le théorème central-limite.

$$\begin{aligned} \mathbb{P}(U_1 + \dots + U_n \geq n(1 + \alpha)) &= \mathbb{P}\left(\frac{U_1 - 1 + \dots + U_n - 1}{\sqrt{n}} \geq \sqrt{n}\alpha\right) \\ \text{(TCL)} &\approx \mathbb{P}(Z \geq 1) \\ &\text{avec } Z \sim \mathcal{N}(0, 1). \end{aligned}$$

Et on lit sur la table que cette dernière valeur vaut (à peu près) $1 - 0.8413 = 0,1587$.

2. (a)

$$\begin{aligned} \mathbb{E}(N) &= \sum_{n \geq 0} n \frac{(\lambda T)^n e^{-\lambda T}}{n!} \\ &= \sum_{n \geq 1} n \frac{(\lambda T)^n e^{-\lambda T}}{n!} \\ &= (\lambda T) e^{-\lambda T} \sum_{k \geq 0} \frac{(\lambda T)^k}{k!} \\ &= \lambda T \end{aligned}$$

(b)

$$\begin{aligned} \mathbb{P}(N \geq m + p) &= \mathbb{P}(\text{on a grillé plus de } m + p \text{ ampoules dans } [0, T]) \\ &= \mathbb{P}(\text{les } m + p \text{ premières ampoules ont déjà grillé} \\ &\quad \text{quand on arrive en } T) \\ &= \mathbb{P}(X_1 + \dots + X_{m+p} < T) \end{aligned}$$

(c) On remarque que $\text{Var}(X_1) = 1/\lambda^2$, $\mathbb{E}(X_1) = 1/\lambda$.

$$\begin{aligned} \mathbb{P}(N \geq m + p) &= \mathbb{P}(X_1 + \dots + X_{m+p} \leq T) \\ &= \mathbb{P}\left(\frac{X_1 - \mathbb{E}(X_1) + \dots + X_{m+p} - \mathbb{E}(X_{m+p})}{(1/\lambda)\sqrt{m+p}}\right. \\ &\quad \left.< \frac{T - (m+p)/\lambda}{(1/\lambda)\sqrt{m+p}}\right) \\ \text{(TCL)} &\approx \int_{-\infty}^{\frac{T - (m+p)/\lambda}{(1/\lambda)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt. \end{aligned}$$

On calcule $\frac{T - (m-1+p)/\lambda}{(1/\lambda)\sqrt{m-1+p}} = -1$. On a par parité :

$$\begin{aligned} \int_{-\infty}^{-1} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt &= \int_1^{+\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= 1 - \int_{-\infty}^1 \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ \text{(d'après la table)} &= 1 - 0,8413 = 0.1587. \end{aligned}$$

(d) Ici, on cherche p pour que $\mathbb{P}(N \geq m + p) \leq 0.05$. Comme avant :

$$\begin{aligned} \mathbb{P}(N \geq m + p) &\stackrel{\text{TCL}}{\approx} \int_{-\infty}^{\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= 1 - \int_{-\infty}^{\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt. \end{aligned}$$

On regarde la table et on voit qu'il faut prendre $-\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}} \geq 1.65$. Une rapide étude de fonction montre qu'il faut prendre $m + p \geq 29$.

3. (a)

$$\begin{aligned} \mathbb{P}(N \geq m + p) &= \mathbb{P}(\text{on a grillé plus de } m + p \text{ ampoules dans } [0, T]) \\ &= \mathbb{P}(\text{les } m + p \text{ premières ampoules ont déjà grillé} \\ &\quad \text{quand on arrive en } T) \\ &= \mathbb{P}(X_1 + \dots + X_{m+p} < T) \end{aligned}$$

(b) On remarque que $\text{Var}(X_1) = 1/\lambda^2$, $\mathbb{E}(X_1) = 1/\lambda$.

$$\begin{aligned} \mathbb{P}(N \geq m + p) &= \mathbb{P}(X_1 + \dots + X_{m+p} \leq T) \\ &= \mathbb{P}\left(\frac{X_1 - \mathbb{E}(X_1) + \dots + X_{m+p} - \mathbb{E}(X_{m+p})}{(1/\lambda)\sqrt{m+p}}\right. \\ &\quad \left.< \frac{T - (m+p)/\lambda}{(1/\lambda)\sqrt{m+p}}\right) \\ (\text{TCL}) &\approx \int_{-\infty}^{\frac{T-(m+p)/\lambda}{(1/\lambda)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt. \end{aligned}$$

On calcule $\frac{T-(m-1+p)/\lambda}{(1/\lambda)\sqrt{m-1+p}} = -1$. On a par parité :

$$\begin{aligned} \int_{-\infty}^{-1} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt &= \int_1^{+\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= 1 - \int_{-\infty}^1 \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ (\text{d'après la table}) &= 1 - 0,8413 = 0.1587. \end{aligned}$$

(c) Ici, on cherche p pour que $\mathbb{P}(N \geq m + p) \leq 0.05$. Comme avant :

$$\begin{aligned} \mathbb{P}(N \geq m + p) &\stackrel{\text{TCL}}{\approx} \int_{-\infty}^{\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= 1 - \int_{-\infty}^{\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt. \end{aligned}$$

On regarde la table et on voit qu'il faut prendre $-\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}} \geq 1.65$. Une rapide étude de fonction montre qu'il faut prendre $m + p \geq 29$.

4. (a) À chaque pas de temps, la grenouille se déplace de 1 vers la droite (et de manière aléatoire vers le haut ou le bas) donc elle passe par l'axe des ordonnées (c'est à dire l'autoroute) au temps 25.
- (b) L'ordonnée de la grenouille au temps n peut s'écrire $V_1 + \dots + V_n$ où $V_n = 1/\sqrt{2}$ avec probabilité $1/2$ et $V_n = -1/\sqrt{2}$ avec probabilité $1/2$ (pour tout k , V_k est la composante verticale du vecteur U_k). Les variables V_k sont d'espérance $m = 0$ et de variance $\sigma^2 = 1/2$. La probabilité de passer par un tunnel est :

$$\begin{aligned} \mathbb{P}(\text{ordonnée de } Z_{25} \in [-5, 5]) &= \mathbb{P}(|V_1 + \dots + V_{25}| \leq 5) \\ &= \mathbb{P}\left(\left|\frac{V_1 + \dots + V_{25} - 25m}{\sigma \sqrt{25}}\right| \leq \sqrt{2}\right). \end{aligned}$$

Les variables V_i sont i.i.d., intégrables et de variance finie donc par le théorème central-limite :

$$\mathbb{P}(\text{ordonnée de } Z_{25} \in [-5, 5]) \approx \int_{-\sqrt{2}}^{+\sqrt{2}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt = -1 + 2 \int_{-\infty}^{\sqrt{2}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt.$$

On trouve sur la table jointe au sujet que $\mathbb{P}(\text{ordonnée de } Z_{25} \in [-5, 5]) \approx 0.84$.

- (c) On veut trouver x tel que $\mathbb{P}(\text{ordonnée de } Z_{25} \in [-x, x]) \approx 0.9$. On a par le théorème central-limite :

$$\begin{aligned} \mathbb{P}(\text{ordonnée de } Z_{25} \in [-x, x]) &= \mathbb{P}(|V_1 + \dots + V_{25}| \leq x) \\ &= \mathbb{P}\left(\left|\frac{V_1 + \dots + V_{25} - 25m}{\sigma \sqrt{25}}\right| \leq \frac{x}{5}\right) \\ &\approx \int_{-x/\sqrt{2/5}}^{x/\sqrt{2/5}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= -1 + 2 \int_{-\infty}^{x/\sqrt{2/5}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt. \end{aligned}$$

D'après la table, il faut $x/\sqrt{2/5} \approx 1.65$ donc $x \approx 5.83$. La grenouille se trouve toujours sur des points de coordonnées entières donc il suffit de prendre $x = 5$.

Chapitre 3

Probabilités et espérances conditionnelles

FIGURE 3.1 – Xkcd : <http://xkcd.com>

3.1 Conditionnement dans le cas discret

Définition 3.1.1. La probabilité conditionnelle $\mathbb{P}(A|B)$ d'un événement A sachant un événement B est définie par

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ si } \mathbb{P}(B) > 0. \quad (3.1.1)$$

On attribue une valeur arbitraire à $\mathbb{P}(A|B)$ si $\mathbb{P}(B) = 0$.

Définition 3.1.2. Soient X, Y des v.a. discrètes à valeurs dans un ensemble I . La fonction de masse conditionnelle $p_{X|Y}(\cdot|y)$ de X sachant $Y = y$ est définie par

$$\forall x, p_{X|Y}(x|y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \text{ si } \mathbb{P}(Y = y) > 0. \quad (3.1.2)$$

On attribue une valeur arbitraire à $p_{X|Y}(x|y)$ si $\mathbb{P}(Y = y) = 0$. On peut réécrire (3.1.2) en terme des fonctions de masse p_{XY}, p_Y :

$$p_{X|Y}(x|y) = \frac{p_{XY}(x, y)}{p_Y(y)}, \text{ si } p_Y(y) > 0.$$

Remarque 3.1.3. Observons que $x \mapsto p_{X|Y}(x|y)$ est une fonction de masse (à y fixé), c'est à dire : $p_{X|Y}(x|y) \geq 0$ et $\sum_{\xi \in I} p_{X|Y}(\xi|y) = 1, \forall x, y$. La loi des probabilités totales nous donne

$$\mathbb{P}(X = x) = \sum_{y \in I} p_{X|Y}(x|y)p_Y(y).$$

Dans les cas où $p_Y(y) = 0$, $p_{X|Y}(x|y)$ est défini par une valeur quelconque mais cela n'influe pas sur la somme ci-dessus. De manière générale, les valeurs choisies pour $p_{X|Y}(x|y)$ dans les cas $p_Y(y) = 0$ n'influent pas sur les calculs ci-dessous.

Exemple 3.1.4. Soit $N \sim \mathcal{B}(q, M)$ (pour des constantes $q \in]0; 1[$, $M \in \mathbb{N}^* = \mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\}$). On tire N , puis on tire $X \sim \mathcal{B}(p, N)$

(p est une constante $\in]0; 1[$). Calculons la loi de X .

La variable X est à valeurs dans $\{1, 2, \dots, M\}$. Les données du problème sont :

$$p_{X|N}(k|n) = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

et

$$p_N(n) = C_M^n q^n (1-q)^{M-n}.$$

Nous appliquons la loi des probabilités totales pour obtenir ($\forall k \in \{1, 2, \dots, M\}$)

$$\begin{aligned} \mathbb{P}(X = k) &= \sum_{n=0}^M p_{X|N}(k|n) p_N(n) \\ &= \sum_{n=0}^M \mathbf{1}_{[0;n]}(k) C_n^k p^k (1-p)^{n-k} C_M^n q^n (1-q)^{M-n} \\ &= \sum_{n=k}^M C_n^k p^k (1-p)^{n-k} C_M^n q^n (1-q)^{M-n} \\ &= \sum_{n=k}^M \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \frac{M!}{n!(M-n)!} q^n (1-q)^{M-n} \\ &= \frac{M!}{k!} p^k (1-q)^M \left(\frac{q}{1-q}\right)^k \sum_{n=k}^M \frac{1}{(n-k)!(M-n)!} (1-p)^{n-k} \left(\frac{q}{1-q}\right)^{n-k} \\ &= \frac{M!}{k!(M-k)!} p^k (1-q)^M \left(\frac{q}{1-q}\right)^k \sum_{n=k}^M \frac{(M-k)!}{(n-k)!(M-n)!} (1-p)^{n-k} \left(\frac{q}{1-q}\right)^{n-k} \\ &\quad \text{(changement d'indice } j = n - k) \\ &= \frac{M!}{k!(M-k)!} p^k (1-q)^M \left(\frac{q}{1-q}\right)^k \sum_{j=0}^{M-k} C_{M-k}^j (1-p)^j \left(\frac{q}{1-q}\right)^j \\ &\quad \text{(formule du binôme)} \\ &= \frac{M!}{k!(M-k)!} p^k (1-q)^{M-k} \left(1 + \frac{(1-p)q}{1-q}\right)^{M-k} \\ &= C_M^k (pq)^k (1-pq)^{M-k}. \end{aligned}$$

Donc X est de loi $\mathcal{B}(pq, M)$.

Exemple 3.1.5. Problème de Monty Hall (on pourra consulter http://fr.wikipedia.org/wiki/Probl%C3%A8me_de_Monty_Hall). Il s'agit d'un jeu télévisé. Le candidat se trouve devant trois portes. Derrière l'une d'entre elles se trouve une voiture et derrière les deux autres se trouvent des chèvres. La position de la voiture est une variable aléatoire uniforme à valeurs dans $\{A, B, C\}$ (A, B, C sont les noms des trois portes, la voiture a la probabilité $1/3$ de se trouver derrière chacune d'entre elles). Les trois portes A, B, C sont alignées de gauche à droite. Le jeu se déroule suivant les phases suivantes.

- Le candidat sélectionne une porte.

- Le présentateur de l'émission ouvre, parmi les deux portes restantes, une porte derrière laquelle se trouve une chèvre. Si le candidat a sélectionné une porte avec une chèvre, le présentateur n'a pas le choix de la porte qu'il va ouvrir. Dans le cas contraire, le présentateur a deux choix possibles et tire à pile ou face pour décider quel porte il ouvre.
- Le candidat ouvre ensuite la porte de son choix et gagne ce qu'il trouve derrière celle-ci (une chèvre ou une voiture).

On part du principe que le candidat préfère gagner la voiture. Nous allons examiner deux stratégies :

- Le candidat ouvre la porte qu'il a sélectionnée, indépendamment de la porte ouverte par le présentateur.
- Le candidat change de porte, c'est à dire qu'il ouvre la porte qui n'est ni sélectionnée par lui, ni ouverte par le présentateur.

Question : quelle stratégie offre le plus de chances de gagner la voiture ? La réponse est qu'il vaut mieux changer de porte. Nous allons ici en faire une démonstration. Voir l'exercice 1 pour une simulation numérique.

Soit U à valeurs dans $\{P, F\}$ la variable «pile ou face» utilisée par le présentateur (s'il en a besoin) (cette variables est indépendante des autres variables). On peut supposer, par exemple, que si le candidat a sélectionné une porte avec une voiture, le présentateur choisit la porte avec une chèvre la plus à gauche si $U = P$ et l'autre porte avec une chèvre si $U = F$. Supposons que le candidat ait choisi la porte A (pour des raisons de symétrie, les autres cas sont équivalents). Notons $A = c$ l'événement «une chèvre est derrière la porte A », $A = v$ l'événement «la voiture est derrière la porte A », ... etc ..., $B = p$ pour l'événement «le présentateur ouvre la porte B ». Nous voulons calculer :

$$\mathbb{P}(A = c|B = p) = \frac{\mathbb{P}(\{A = c\} \cap \{B = p\})}{\mathbb{P}(B = p)} = \frac{\mathbb{P}(C = v)}{\mathbb{P}(B = p)}.$$

Or

$$\begin{aligned} \mathbb{P}(B = p) &= \mathbb{P}(\{C = v\}) + \mathbb{P}(\{A = v\} \cap \{U = P\}) \\ &= \frac{1}{3} + \frac{1}{3} \times \frac{1}{2} \\ &= \frac{1}{2}. \end{aligned}$$

Donc

$$\mathbb{P}(A = c|B = p) = \frac{(1/3)}{(1/2)} = \frac{2}{3}.$$

On en conclut que le candidat a intérêt à changer de porte.

Définition 3.1.6. Soient X, Y des variables discrètes à valeurs dans un ensemble I . Soit g une fonction telle que $\mathbb{E}(|g(X)|) < \infty$. Nous définissons l'espérance conditionnelle de $g(X)$ sachant $Y = y$ par

$$\mathbb{E}(g(X)|Y = y) = \sum_{x \in I} g(x)p_{X|Y}(x|y), \text{ si } p_Y(y) > 0,$$

et l'espérance conditionnelle n'est pas définie en y si $p_Y(y) = 0$. La loi des probabilités totales nous donne

$$\mathbb{E}(g(X)) = \sum_{y \in I} \mathbb{E}(g(X)|Y = y)p_Y(y).$$

La quantité $\mathbb{E}(g(X)|Y = y)$ est une fonction de y . Nous notons $\mathbb{E}(g(X)|Y)$ la variable aléatoire telle $\mathbb{E}(g(X)|Y)(\omega) = \mathbb{E}(g(X)|Y = y)$ pour tout ω tel que $Y(\omega) = y$. La loi des probabilités totales nous donne

$$\mathbb{E}(g(X)) = \mathbb{E}(\mathbb{E}(g(X)|Y)).$$

Proposition 3.1.7. Soient X, X_1, X_2, Y des variables discrètes à valeurs dans un ensemble I (définies conjointement). Soient g, g_1, g_2 des fonctions telles que $\mathbb{E}(|g(X)|) < \infty$ (pareil pour g_1, g_2, X_1, X_2). Soit h une fonction bornée. Soit $v : \mathbb{R}^2 \rightarrow \mathbb{R}$ telle que $\mathbb{E}(|v(X, Y)|) < \infty$. Nous prenons $c_1, c_2 \in \mathbb{R}$. Nous avons les propriétés suivantes.

1. $\mathbb{E}(c_1 g_1(X_1) + c_2 g_2(X_2)|Y = y) = c_1 \mathbb{E}(g_1(X_1)|Y = y) + c_2 \mathbb{E}(g_2(X_2)|Y = y)$
2. Si $g(x) \geq 0, \forall x$, alors $\mathbb{E}(g(X)|Y = y) \geq 0, \forall y$.
3. $\mathbb{E}(v(X, Y)|Y = y) = \mathbb{E}(v(X, y)|Y = y)$
4. Si $X \perp\!\!\!\perp Y$, alors $\mathbb{E}(g(X)|Y = y) = \mathbb{E}(g(X))$.
5. $\mathbb{E}(g(X)h(Y)|Y = y) = h(y)\mathbb{E}(g(X)|Y = y)$
6. $\mathbb{E}(g(X)h(Y)) = \sum_{y \in I} h(y)\mathbb{E}(g(X)|Y = y)p_Y(y) = \mathbb{E}(h(Y)\mathbb{E}(g(X)|Y))$

Si nous appliquons ces formules avec $g \equiv 1$ ou $h \equiv 1$, nous obtenons

$$\mathbb{E}(c|Y = y) = c,$$

$$\mathbb{E}(h(Y)|Y = y) = h(y),$$

$$\mathbb{E}(g(X)) = \sum_{y \in I} \mathbb{E}(g(X)|Y = y)p_Y(y) = \mathbb{E}(\mathbb{E}(g(X)|Y)). \quad (3.1.3)$$

Voir la démonstration de la proposition 3.4.1 pour des exemples d'utilisation de ces propriétés.

3.2 Sommes aléatoires

Soient ξ_1, ξ_2, \dots des v.a.r. i.i.d. Soit N une variable aléatoire à valeurs dans \mathbb{N} de fonction de masse p_N . Nous définissons la somme aléatoire

$$X = \begin{cases} 0 & \text{si } N = 0 \\ \xi_1 + \dots + \xi_N & \text{si } N > 0. \end{cases}$$

On peut abrégé cette formule en $X = \xi_1 + \dots + \xi_N$ (avec la convention que $X = 0$ si $N = 0$).

Exemple 3.2.1. File d'attente. Soit N le nombre de client arrivant à un guichet pendant une certaine période. Soit ξ_i le temps de service requis par le i -ème client. Alors $X = \xi_1 + \dots + \xi_N$ est le temps total de service requis.

3.3 Probabilités conditionnelles dans le cas mélangé

Soient X, N deux variables aléatoires définies conjointement. Supposons que N est à valeurs dans \mathbb{N} .

Définition 3.3.1. La formule (3.1.1) nous permet de définir la fonction de répartition conditionnelle $x \mapsto F_{X|N}(x|n)$ de X sachant $N = n$ par

$$F_{X|N}(x|n) = \frac{\mathbb{P}(X \leq x, N = n)}{\mathbb{P}(N = n)}, \text{ si } \mathbb{P}(N = n) > 0,$$

et cette quantité n'est pas définie si $\mathbb{P}(N = n) = 0$.

Si, de plus, X est une variable continue et que $x \mapsto F_{X|N}(x|n)$ est dérivable en tout x pour tous les n tels que $\mathbb{P}(N = n) > 0$, alors on peut définir la densité conditionnelle $x \mapsto f_{X|N}(x|n)$ de X sachant $N = n$ par

$$f_{X|N}(x|n) = \frac{d}{dx} F_{X|N}(x|n), \text{ si } \mathbb{P}(N = n) > 0.$$

La fonction $x \mapsto f_{X|N}(x|n)$ est une fonction de densité ($\forall n$ tel que $\mathbb{P}(N = n) > 0$).

Proposition 3.3.2. Sous les hypothèse ci-dessus, nous avons

$$\mathbb{P}(a \leq X \leq b, N = n) = \int_a^b f_{X|N}(x|n) dx \times p_N(n), \forall a \leq b.$$

La loi des probabilités totales nous donne la densité f_X de X :

$$f_X(x) = \sum_{n=0}^{+\infty} f_{X|N}(x|n) p_N(n).$$

Définition 3.3.3. Soit g une fonction telle que $\mathbb{E}(|g(X)|) < \infty$. L'espérance conditionnelle de $g(X)$ sachant $N = n$ est définie par

$$\mathbb{E}(g(X)|N = n) = \int_{-\infty}^{+\infty} g(x) f_{X|N}(x|n) dx.$$

Nous définissons l'espérance conditionnelle de $g(X)$ sachant N par

$$\mathbb{E}(g(X)|N) = \text{la v.a.r. qui vaut } \mathbb{E}(g(X)|N = n) \text{ quand } N = n.$$

C'est donc une variable aléatoire définie comme une fonction de N .

Proposition 3.3.4. La loi des probabilités totales nous donne

$$\mathbb{E}(g(X)) = \sum_{n=0}^{+\infty} \mathbb{E}(g(X)|N = n) p_N(n) = \mathbb{E}(\mathbb{E}(g(X)|N)).$$

Démonstration. Nous remarquons que

$$\mathbb{E}(g(X)|N) = \sum_{n=0}^{+\infty} \mathbb{E}(g(X)|N = n) \mathbf{1}_{\{n\}}(N).$$

(Nous notons $\mathbf{1}_{\{n\}}(N)$ la variable aléatoire $\omega \mapsto \mathbf{1}_{\{n\}}(N(\omega))$). L'égalité ci-dessus est une égalité entre variables aléatoires réelles, c'est à dire entre fonctions de Ω dans \mathbb{R} . Pour tout $\omega \in \Omega$,

$$\sum_{n=0}^{+\infty} \mathbb{E}(g(X)|N = n)\mathbf{1}_{\{n\}}(N(\omega)) = \mathbb{E}(g(X)|N = N(\omega))$$

(tous les termes de la somme sont nuls sauf un). Nous en déduisons

$$\begin{aligned} \mathbb{E}(\mathbb{E}(g(X)|N)) &= \mathbb{E}\left(\sum_{n=0}^{+\infty} \mathbb{E}(g(X)|N = n)\mathbf{1}_{\{n\}}(N)\right) \\ (\text{par prop. 1.3.10 de linéarité de } \mathbb{E}) &= \sum_{n=0}^{+\infty} \mathbb{E}(g(X)|N = n)\mathbb{E}(\mathbf{1}_{\{n\}}(N)). \end{aligned}$$

Nous ne justifions pas ici pourquoi nous pouvons appliquer la proposition 1.3.10 à une somme infinie. Remarquons que les quantités $\mathbb{E}(g(X)|N = n)$ sortent de l'espérance parce que ce sont des réels. Nous avons, par le lemme 1.4.4,

$$\begin{aligned} \mathbb{E}(\mathbb{E}(g(X)|N)) &= \sum_{n=0}^{+\infty} \mathbb{E}(g(X)|N = n)p_N(n) \\ &= \sum_{n=0}^{+\infty} \int_{-\infty}^{+\infty} g(x)f_{X|N}(x|n)dx \times p_N(n) \\ (\text{linéarité de } \int) &= \int_{-\infty}^{+\infty} g(x)\left(\sum_{n=0}^{+\infty} f_{X|N}(x|n)p_N(n)\right)dx \\ (\text{prop. 3.3.2}) &= \int_{-\infty}^{+\infty} g(x)f_X(x)dx \\ &= \mathbb{E}(g(X)). \end{aligned}$$

□

3.4 Moments et loi d'une somme aléatoire

On se replace dans le cadre de la section 3.2. Supposons que ξ_1 est une variable discrète. Supposons que ξ_1 et N ont les moments suivants finis

$$\begin{aligned} \mathbb{E}(\xi_1) &= \mu, \quad \text{Var}(\xi_1) = \sigma^2, \\ \mathbb{E}(N) &= \nu, \quad \text{Var}(N) = \tau^2. \end{aligned}$$

Proposition 3.4.1. *Nous avons*

$$\mathbb{E}(X) = \mu\nu, \quad \text{Var}(X) = \nu\sigma^2 + \mu^2\tau^2.$$

Démonstration. Calculons

$$\begin{aligned} \mathbb{E}(X) &= \sum_{n=0}^{+\infty} \mathbb{E}(X|N = n)p_N(n) && (\text{par (3.1.3)}) \\ &= \sum_{n=1}^{+\infty} \mathbb{E}(\xi_1 + \dots + \xi_n|N = n) \\ &= \sum_{n=1}^{+\infty} \mathbb{E}(\xi_1 + \dots + \xi_n|N = n) && (\text{par prop. 3.1.7, 3}) \\ &= \sum_{n=1}^{+\infty} \mathbb{E}(\xi_1 + \dots + \xi_n)p_N(n) && (\text{par prop. 3.1.7, 4}) \\ &= \sum_{n=1}^{+\infty} n\mu p_N(n) \\ &= \mu\nu. \end{aligned}$$

Nous avons

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}((X - \mu\nu)^2) = \mathbb{E}((X - N\mu + N\mu - \nu\mu)^2) \\ &= \mathbb{E}((X - N\mu)^2) + \mathbb{E}((N\mu - \nu\mu)^2) + 2\mathbb{E}(\mu(X - N\mu)(N - \nu)).\end{aligned}$$

Et

$$\begin{aligned}\mathbb{E}((X - N\mu)^2) &= \sum_{n=0}^{+\infty} \mathbb{E}((X - N\mu)^2 | N = n) p_N(n) \\ &= \sum_{n=0}^{+\infty} \mathbb{E}((\xi_1 + \dots + \xi_n - n\mu)^2 | N = n) p_N(n) \\ &= \sum_{n=0}^{+\infty} \mathbb{E}((\xi_1 + \dots + \xi_n - n\mu)^2) p_N(n) \\ &= \sum_{n=0}^{+\infty} \left[\mathbb{E}((\xi_1 - \mu)^2) + \dots + \mathbb{E}((\xi_n - \mu)^2) + \sum_{i,j \in \{1, \dots, n\}, i \neq j} \mathbb{E}((\xi_i - \mu)(\xi_j - \mu)) \right] p_N(n) \\ &= \sum_{n=0}^{+\infty} n\sigma^2 p_N(n) = \sigma^2 \nu,\end{aligned}$$

et

$$\mathbb{E}(\mu^2(N - \nu)^2) = \mu^2 \mathbb{E}((N - \nu)^2) = \mu^2 \tau^2,$$

et

$$\begin{aligned}\mathbb{E}(\mu(X - N\mu)(N - \nu)) &= \mu \sum_{n=0}^{+\infty} \mathbb{E}((X - n\mu)(n - \nu) | N = n) p_N(n) \\ &= \mu \sum_{n=0}^{+\infty} (n - \nu) \mathbb{E}((X - n\mu) | N = n) p_N(n) = 0\end{aligned}$$

car $\forall n \geq 0, \mathbb{E}(X - n\mu | N = n) = \mathbb{E}(\xi_1 + \dots + \xi_n - n\mu) = 0$. Nous avons donc

$$\text{Var}(X) = \nu\sigma^2 + \mu^2\tau^2.$$

□

Définition 3.4.2. Convolution n-ème. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\int_{-\infty}^{+\infty} |f(x)| dx < \infty$. Pour $n \in \mathbb{N}^*$, on définit la convolution n-ème de f par récurrence

$$f^{*1} = f, f^{*n}(x) = \int_{-\infty}^{+\infty} f^{*(n-1)}(x - u) f(u) du, \forall x.$$

Soit $p : \mathbb{Z} \rightarrow [0; 1]$ une fonction de masse. Pour $n \in \mathbb{N}^*$, on définit la convolution n-ème de p par récurrence

$$p^{*1} = p, \forall j \in \mathbb{Z}, p^{*n}(j) = \sum_{i \in \mathbb{Z}} p^{*(n-1)}(j - i) p(i).$$

Proposition 3.4.3. *Si les variables ξ_1, ξ_2, \dots sont continues de densité et $p_N(0) = 0$, alors $Z = \xi_1 + \dots + \xi_N$ est continue, de densité*

$$x \mapsto f_Z(x) = \sum_{n=1}^{+\infty} f^{\star n}(x) p_N(n).$$

En particulier, si N est constante égale à n , la densité de Z est $f^{\star n}$. Si les variables ξ_1, ξ_2, \dots sont discrètes à valeurs dans \mathbb{Z} de fonction de masse p_ξ , alors $Z = \xi_1 + \dots + \xi_N$ est discrète, de fonction de masse \mathbb{Z}

$$j \in \mathbb{Z} \mapsto p_N(0) + \sum_{n=1}^{+\infty} p_\xi^{\star n}(j) p_N(n).$$

En particulier, si N est constante égale à n , la fonction de masse de Z est

$$\begin{cases} j \mapsto \mathbf{1}_{\{0\}}(j) & \text{si } n = 0 \\ p_\xi^{\star n} & \text{si } n > 0. \end{cases}$$

Démonstration. Nous n'écrivons que la démonstration pour le cas discret. Si N est constante égal à 0, alors $Z = 0$, donc la fonction de masse de Z est

$$j \mapsto \begin{cases} 1 & \text{si } j = 0 \\ 0 & \text{sinon} \end{cases} = \mathbf{1}_{\{0\}}(j).$$

Si N est constante égal à $n \in \mathbb{N}^*$, montrons par récurrence que la fonction de masse de Z est $p_\xi^{\star n}$.

- C'est le cas pour $n = 1$.
- Si c'est vrai en $n - 1$. Notons $Y = \xi_1 + \dots + \xi_{n-1}$. La fonction de masse de Y est $p_\xi^{\star(n-1)}$. Par la proposition 1.5.4, $\forall j \in \mathbb{Z}$

$$\begin{aligned} \mathbb{P}(Z = j) &= \sum_{k,r \in \mathbb{Z}: k+r=j} \mathbb{P}(Y = k, \xi_n = j) \\ (\text{car } Y \perp \xi_n) &= \sum_{k,r \in \mathbb{Z}: k+r=j} \mathbb{P}(Y = k) \mathbb{P}(\xi_n = j) = \sum_{i \in \mathbb{Z}} \mathbb{P}(Y = j - i) \mathbb{P}(\xi_n = i) \\ (\text{par hypothèse de récurrence}) &= \sum_{i \in \mathbb{Z}} p_\xi^{\star(n-1)}(j - i) p_\xi(i) = p_\xi^{\star n}(j). \end{aligned}$$

Si N n'est pas une constante, nous avons

$$\mathbb{P}(Z = 0) = p_N(0),$$

et $\forall n \in \mathbb{N}^*$,

$$\begin{aligned} \mathbb{P}(Z = n) &= \sum_{k=0}^{+\infty} \mathbb{P}(Z = n | N = k) p_N(k) \\ (\text{car } \mathbb{P}(Z = n | N = 0) = 0) &= \sum_{k=1}^{+\infty} \mathbb{P}(Z = n | N = k) p_N(k) = \sum_{k=1}^{+\infty} p_\xi^{\star k}(n) p_N(k). \end{aligned}$$

□

Exemple 3.4.4. On se place dans le cadre de la proposition ci-dessus avec des ξ_i continues et

$$\forall z, f(z) = \mathbf{1}_{[0;+\infty[}(z)\lambda e^{-\lambda z},$$

pour un certain $\lambda > 0$,

$$\forall n \in \mathbb{N}^*, p_N(n) = \beta(1 - \beta)^{n-1},$$

pour un certain $\beta \in]0; 1[$. Montrons par récurrence que $f^{*n}(z) = \mathbf{1}_{[0;+\infty[}(z) \frac{\lambda^n}{(n-1)!} z^{n-1} e^{-\lambda z}$, $\forall z \in \mathbb{R}$.

- C'est vrai pour $n = 1$.
- Si c'est vrai en $n - 1$, calculons $\forall z$

$$\begin{aligned} f^{*n}(z) &= \int_{-\infty}^{+\infty} f^{*(n-1)}(z-u)f(u)du \\ &= \int_{-\infty}^{+\infty} \mathbf{1}_{[0;+\infty[}(z-u) \frac{\lambda^{n-1}}{(n-2)!} (z-u)^{n-2} e^{-\lambda(z-u)} \mathbf{1}_{[0;+\infty[}(u) \lambda e^{-\lambda u} du \\ &= \begin{cases} 0 & \text{si } z < 0 \\ \int_0^z \frac{\lambda^{n-1}}{(n-2)!} (z-u)^{n-2} e^{-\lambda(z-u)} \lambda e^{-\lambda u} du & \text{sinon} . \end{cases} \end{aligned}$$

Donc, $\forall z \geq 0$,

$$\begin{aligned} f^{*n}(z) &= \int_0^z \frac{\lambda^n}{(n-2)!} (z-u)^{n-2} e^{-\lambda z} du \\ &= \frac{\lambda^n e^{-\lambda z}}{(n-2)!} \left[\frac{-(z-u)^{n-1}}{n-1} \right]_0^z \\ &= \frac{\lambda^n}{(n-1)!} e^{-\lambda z} z^{n-1}. \end{aligned}$$

Donc la densité de Z est

$$\begin{aligned} z \mapsto f_Z(z) &= \sum_{n=1}^{+\infty} \frac{\lambda^n}{(n-1)!} e^{-\lambda z} z^{n-1} \beta(1-\beta)^{n-1} \\ (\text{somme exponentielle}) &= \lambda e^{-\lambda z} \beta \exp((1-\beta)\lambda z) \\ &= \lambda \beta e^{-\lambda \beta z}. \end{aligned}$$

Donc $Z \sim \mathcal{E}(\lambda\beta)$.

3.5 Conditionnement par une variable continue

Définition 3.5.1. Soient X, Y des v.a.r. continues définies de manière conjointe par une densité de probabilité $(x, y) \mapsto f_{X,Y}(x, y)$. Nous définissons la densité de probabilité conditionnelle $x \mapsto f_{X|Y}(x|y)$ de X sachant $Y = y$ par

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \text{ si } f_Y(y) > 0.$$

La densité conditionnelle n'est pas définie si $f_Y(y) = 0$. Nous définissons la fonction de répartition conditionnelle $x \mapsto F_{X|Y}(x|Y = y)$ de X sachant $Y = y$ par

$$F_{X|Y}(x|Y = y) = \int_{-\infty}^x f_{X|Y}(\xi|y) d\xi, \text{ si } f_Y(y) > 0.$$

Si on prend g telle que $\mathbb{E}(|g(X)|) < \infty$, l'espérance conditionnelle de $g(X)$ sachant $Y = y$ est définie par

$$\mathbb{E}(g(X)|Y = y) = \int_{-\infty}^{+\infty} g(x) f_{X|Y}(x|y) dx, \text{ si } f_Y(y) > 0.$$

Proposition 3.5.2. *Sous les hypothèses, de la définition précédente, nous avons les propriétés suivantes.*

1. Pour tout $a \leq b, c \leq d$

$$\mathbb{P}(a < X < b, c < Y < d) = \int_c^d \left(\int_a^b f_{X|Y}(x|y) dx \right) f_Y(y) dy.$$

2. Les propriétés de la proposition 3.3.2 sont encore vraies, à condition d'adapter le point 6 en

$$\begin{aligned} \mathbb{E}(g(X)h(Y)) &= \mathbb{E}(h(Y)\mathbb{E}(g(X)|Y)) \\ &= \int_{-\infty}^{+\infty} h(y)\mathbb{E}(g(X)|Y = y)f_Y(y) dy. \end{aligned}$$

Dans l cas $h \equiv 1$, nous obtenon

$$\mathbb{E}(g(X)) = \mathbb{E}(\mathbb{E}(g(X)|Y)) = \int_{-\infty}^{+\infty} \mathbb{E}(g(X)|Y = y)f_Y(y) dy.$$

Exemple 3.5.3. Soient X, Y de densité jointe

$$f_{X,Y}(x, y) = \frac{1}{y} e^{-(x/y)-y} \mathbf{1}_{]0;+\infty[}(x) \mathbf{1}_{]0;+\infty[}(y).$$

Calculons la densité de y :

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} \frac{1}{y} e^{-(x/y)-y} \mathbf{1}_{]0;+\infty[}(x) \mathbf{1}_{]0;+\infty[}(y) dx \\ &= \mathbf{1}_{]0;+\infty[}(y) \int_0^{+\infty} \frac{1}{y} e^{-(x/y)-y} dx \\ &= \mathbf{1}_{]0;+\infty[}(y) \frac{e^{-y}}{y} \left[-ye^{-x/y} \right]_0^{+\infty} \\ &= \mathbf{1}_{]0;+\infty[}(y) e^{-y}. \end{aligned}$$

Puis nous calculons la densité conditionnelle (définie seulement pour $y > 0$)

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \mathbf{1}_{]0;+\infty[}(x) \frac{1}{y} e^{-x/y}.$$

Donc, conditionnellement à $Y = y$, X est de loi $\mathcal{E}(1/y)$.

3.6 Statistiques pour les nuls

On dispose d'un générateur de nombre aléatoires de loi $\mathcal{B}(1/2)$. On veut tester l'hypothèse

(H1) : les tirages sont de loi $\mathcal{B}(1/2)$

contre l'hypothèse

(H2) : les tirages sont de loi $\mathcal{B}(1)$

((H2) veut dire que les variables ne sont pas aléatoires mais constantes égales à 1). On choisit $\alpha \in [0; 1/2]$. La procédure de test est la suivante : on tire X_1, \dots, X_n avec notre générateur (pour un certain n), si $\frac{X_1 + \dots + X_n}{n} \in [0; 1/2 + \alpha/\sqrt{n}]$, on donne le résultat «(H1) est vraie» et sinon on donne le résultat «(H2) est vraie». Soient U_1, U_2, \dots i.i.d. de loi $\mathcal{B}(1/2)$. Nous nous intéressons aux probabilités d'erreur suivantes : $\mathbb{P}(\text{dire (H1)} | (\text{H2}))$ (dire que (H1) est vraie alors que (H2) est vraie) et $\mathbb{P}(\text{dire (H2)} | (\text{H1}))$ (dire que (H2) est vraie alors que (H1) est vraie). Calculons

$$\mathbb{P}(\text{dire (H1)} | (\text{H2})) = \mathbb{P}\left(\frac{n}{n} \leq \frac{1}{2} + \frac{\alpha}{\sqrt{n}}\right) = 0,$$

$$\begin{aligned} \mathbb{P}(\text{dire (H2)} | (\text{H1})) &= \mathbb{P}\left(\frac{U_1 + \dots + U_n}{n} \geq \frac{1}{2} + \frac{\alpha}{\sqrt{n}}\right) \\ &= \mathbb{P}\left(\frac{(U_1 - 1/2) + \dots + (U_n - 1/2)}{n} \geq \frac{\alpha}{\sqrt{n}}\right) \\ &= \mathbb{P}\left(\frac{(U_1 - 1/2) + \dots + (U_n - 1/2)}{\sqrt{n}} \geq \alpha\right). \end{aligned}$$

Par exemple, fixons $\alpha = 1$. Supposons que l'on tire 10 fois avec ce générateur et que l'on obtienne

$$X_1 = \dots = X_{10} = 1.$$

La procédure de test va renvoyer le verdict «(H2) est vraie». Calculons la probabilité d'erreur

$$\begin{aligned} \mathbb{P}(\text{dire (H2)} | (\text{H1})) &= \mathbb{P}\left(\frac{(U_1 - 1/2) + \dots + (U_n - 1/2)}{\sqrt{n}} \geq \alpha\right) \\ (\text{TCL}) &\approx \mathbb{P}(Z \geq \alpha) \end{aligned}$$

avec $Z \sim \mathcal{N}(0, \sigma^2)$ ($\sigma^2 = \text{Var}(U_1)$). Nous avons

$$\begin{aligned} \text{Var}(U_1) &= \mathbb{E}((U_1 - 1/2)^2) \\ &= \frac{1}{4}. \end{aligned}$$

Soit $Y = Z/\sigma$. Par l'exemple 1.6.4, $Y \sim \mathcal{N}(0, 1)$. Nous avons

$$\begin{aligned} \mathbb{P}(Z \geq \alpha) &= \mathbb{P}\left(\frac{Z}{\sigma} \geq 2\right) \\ &= 1 - \mathbb{P}(Y \leq 2) \\ (\text{lecture de la table}) &\approx 1 - 0,9772 \\ &= 0,0228. \end{aligned}$$

3.7 Exercices

3.7.1 Énoncés

1. Écrire un programme permettant de comparer les stratégies dans le problème de Monty Hall (cf. exemple 3.1.5). Justifier la procédure proposée.
2. Soit N le résultat d'un lancer de dé. On tire N fois à pile ou face et on appelle X le nombre de faces obtenues.
 - (a) Calculer $\mathbb{P}(N = 3, X = 2)$.
 - (b) Calculer $\mathbb{P}(X = 5)$.
 - (c) Calculer $\mathbb{E}(X)$.
3. Soit $X \sim \mathcal{P}(\lambda)$ ($\lambda > 0$). Calculer l'espérance de X sachant que X est impair.
4. On jette 4 pièces de 1 euros et 6 pièces de 2 euros. Soit N le nombre total de faces observé. Si $N = 4$, quelle est la probabilité (conditionnelle) que exactement 2 pièces de 1 euro soient tombées sur face ?
5. On jette un dé rouge une seule fois. Puis on jette un dé vert de manière répétitive jusqu'à ce que la somme du dé rouge et du dé vert soit égale à 4 ou à 7, et on arrête alors de jeter le dé vert. Quelle est la probabilité de s'arrêter avec une somme égale à 4 ?
6. On s'intéresse à l'algorithme suivant.

Algorithme 3.1 Simulation de variable aléatoire

```

N=grand(1,1,'uin',1,6); s=0;
for i=1 :N
  U=grand(1,1,'bin',1,1/2); // U est une variable de loi  $\mathcal{B}(1/2)$ 
  s=s+U;
end,
printf('%i \n',s); // affichage de s (au format «entier»)
  
```

Notons Z la variable affichée.

- (a) Calculer $\mathbb{E}(Z)$ et $\text{Var}(Z)$.
 - (b) Calculer la fonction de masse de Z .
7. On s'intéresse à l'algorithme suivant.

Algorithme 3.2 Simulation de variable aléatoire

```

N=0;
for i=1 :6
  N=N+grand(1,1,'bin',1,1/2); // on ajoute à N des variables de loi  $\mathcal{B}(1/2)$ 
end,
Z=0;
for i=1 :N
  Z=Z+grand(1,1,'bin',1,1/2);
end,
printf('%i \n',Z); // affichage de Z
  
```

- (a) Calculer $\mathbb{E}(Z)$, $\text{Var}(Z)$.
 (b) Calculer $\mathbb{P}(Z = 2)$.

8. On s'intéresse à l'algorithme suivant.

Algorithme 3.3 Simulation de variable aléatoire

```
P=grand(1,1,'unf',0,1);
S=0;
for i=1 :3
    U=grand(1,1,'bin',1,P);
    S=S+U;
end,
printf('%i \n',S);
```

Quelle est la probabilité que ce programme affiche 2 ?

9. On s'intéresse à l'algorithme suivant.

Algorithme 3.4 Simulation de variables aléatoires

```
function [z]=expo(t) // on remarque que le résultat de la fonction ne dépend pas de t
    u=grand(1,1,'unf',0,1); // u est de loi  $\mathcal{U}([0; 1])$ 
    z=-log(1-u);
endfunction
b=0; s=0;
while (b==0)
    s=s+1;
    r=expo(1);
    if (r>1) then
        b=1;
    end,
end,
printf("temps total = %i, variable = %f\n",s,r); // affiche s et r
```

- (a) Quelle est la loi de la variable renvoyée par un appel à la fonction expo ?
 (b) Quelle est la loi de la variable s affichée ?
 (c) Quelle est la loi de la variable r affichée ?

3.7.2 Corrigés

1. On se fixe une stratégie (par exemple, changer de porte), on simule un jeu et on note

$$U = \begin{cases} 1 & \text{si on gagne la voiture} \\ 0 & \text{sinon.} \end{cases}$$

Si on recommence l'expérience n fois, générant ainsi des résultats i.i.d. U_1, U_2, \dots la loi des grands nombres nous dit que

$$\frac{U_1 + U_2 + \dots + U_n}{n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbb{E}(U_1) = \mathbb{P}(\text{gagner la voiture}). \quad (3.7.1)$$

Donc la loi des grands nombres nous dit que l'on peut approcher \mathbb{P} (gagner la voiture) par une moyenne empirique. L'algorithme 3.5 simule plusieurs jeux dans lesquels le joueur change de porte et calcule la moyenne empirique de (3.7.1) pour un certain n .

Algorithme 3.5 Simulation du jeu de Monty Hall

```

n=100000 ; // on choisit n «grand»
// les portes sont numérotées 0,1,2
fonction [z]=comp(i,j) // on rentre deux numéros de portes distincts dans cette fonction
// et elle ressort k tel que {i,j,k}={0,1,2}
  b=0 ; k=-1 ;
  while (b=0)
    k=k+1 ;
    if (k==i) then
      b=0 ;
    else
      if (k==j) then
        b=0 ;
      else
        b=1 ;
      end,
    end,
  end,
  z=k ;
endfonction
s=0 ;
for i=1 :n // on recommence la même expérience n fois
  V=grand(1,1,'uin',0,2) ; // V est le numéro de la porte de la voiture, il est tiré au hasard
  //dans {0,1,2}
  S=grand(1,1,'uin',0,2) ; // S est le numéro sélectionné par le joueur (au hasard)
  if (S==V) then // si S=V, le présentateur choisit au hasard parmi les 2 portes restantes
    U=-1+2*grand(1,1,'uin',0,1) ; // U vaut +1 ou -1 avec proba. 1/2
    P=modulo(S+U,3) ; // c'est le reste de la division euclidienne de S+U par 3
  else // si S≠V alors le présentateur n'a pas de choix à faire
    P=comp(S,V) ;
  end, // P est le numéro de la porte ouverte par le présentateur
  F=comp(S,P) ; // le joueur ouvre la seule porte qu'il n'a pas sélectionnée et que le
  // présentateur n'a pas ouverte
  if (F==V) then
    U=1 ;
  else
    U=0 ;
  end, // U vaut 1 si le joueur gagne la voiture et 0 sinon
  s=s+U ;
end,
printf('estimation de la proba. de succes si on change de porte : %f',s/n) ; // affichage
// du résultat

```

- (a) Pour chaque tirage de pile ou face, on note $\xi_i = 1$ si le résultat est face et $\xi_i = 0$ sinon. Nous avons donc $\xi_1 + \dots + \xi_n$ est le nombre total de faces sur n tirages ($\forall n$). Nous avons $X = \xi_1 + \dots + \xi_N$ et donc

$$\begin{aligned}
 \mathbb{P}(N = 3, X = 2) &= \mathbb{P}(X = 2 | N = 3) \mathbb{P}(N = 3) \\
 &= \mathbb{P}(\xi_1 + \dots + \xi_N = 2 | N = 3) \mathbb{P}(N = 3) \\
 &= \mathbb{P}(\xi_1 + \xi_2 + \xi_3 = 2 | N = 3) \mathbb{P}(N = 3) \\
 &= \mathbb{P}(\xi_1 + \xi_2 + \xi_3 = 2) \frac{1}{6} \\
 &= C_3^2 \left(\frac{1}{2}\right)^3 \frac{1}{6} \\
 &= 3 \frac{1}{8} \frac{1}{6} = \frac{1}{16}.
 \end{aligned}$$

- (b) Calculons

$$\begin{aligned}
 \mathbb{P}(X = 5) &= \sum_{j=1}^6 \mathbb{P}(X = 5 | N = j) \mathbb{P}(N = j) \\
 &= \sum_{j=1}^6 \mathbf{1}_{\{5,6\}}(j) C_j^5 \left(\frac{1}{2}\right)^5 \frac{1}{6} \\
 &= \left(\frac{1}{2}\right)^5 \frac{1}{6} [C_6^5 + C_6^6] \\
 &= \left(\frac{1}{2}\right)^5 \frac{1}{6} (6 + 1) = \left(\frac{1}{2}\right)^5 \frac{7}{6}.
 \end{aligned}$$

- (c) Nous pouvons utiliser la proposition sur les moments d'une somme aléatoire

$$\begin{aligned}
 \mathbb{E}(X) &= \mathbb{E}(\xi_1) \mathbb{E}(N) \\
 &= \frac{1}{2} \frac{1}{6} (1 + 2 + \dots + 6) \\
 &= \frac{1}{12} \frac{6 \times 7}{2} = \frac{21}{12} = \frac{7}{4}.
 \end{aligned}$$

3. Notons Y la variable qui vaut 1 si X est impair et 0 sinon. Nous avons

$$\begin{aligned}
 \mathbb{P}(Y = 1) &= \mathbb{P}(\cup_{i=0}^{+\infty} \{X = 2i + 1\}) \\
 &= \sum_{i=0}^{+\infty} \mathbb{P}(X = 2i + 1) \\
 &= \sum_{i=0}^{+\infty} \frac{\lambda^{2i+1}}{(2i+1)!} e^{-\lambda} \\
 &= \sinh(\lambda) e^{-\lambda}.
 \end{aligned}$$

Nous avons donc

$$\begin{aligned} \forall x, p_{X|Y}(x, 1) &= \frac{p_{X,Y}(X = x, Y = 1)}{p_Y(Y = 1)} \\ &= \begin{cases} 0 & \text{si } x \text{ pair} \\ \frac{\mathbb{P}(X=x)}{p_Y(Y=1)} & \text{si } x \text{ impair.} \end{cases} \end{aligned}$$

Donc, pour x impair,

$$p_{X|Y}(x, 1) = \frac{1}{\sinh(\lambda)e^{-\lambda}} \frac{\lambda^x}{x!} e^{-\lambda} = \frac{\lambda^x}{\sinh(\lambda)}.$$

Donc

$$\begin{aligned} \mathbb{E}(X|Y = 1) &= \sum_{j=0}^{+\infty} p_{X|Y}(2j+1, 1)(2j+1) \\ &= \sum_{j=0}^{+\infty} \frac{\lambda^{2j+1}}{\sinh(\lambda)} (2j+1) \\ &= \frac{\lambda}{\sinh(\lambda)} \sum_{j=0}^{+\infty} (2j+1)\lambda^{2j} \\ \text{("technique de la série dérivée")} &= \frac{\lambda}{\sinh(\lambda)} \frac{d}{d\lambda} \left(\sum_{j=0}^{+\infty} \lambda^{2j+1} \right) \\ \text{(série géométrique)} &= \frac{\lambda}{\sinh(\lambda)} \frac{d}{d\lambda} \left(\frac{\lambda}{1-\lambda^2} \right) \\ &= \frac{\lambda}{\sinh(\lambda)} \left(\frac{1-\lambda^2 + 2\lambda \times \lambda}{(1-\lambda^2)^2} \right) \\ &= \frac{\lambda}{\sinh(\lambda)} \times \frac{1+\lambda^2}{(1-\lambda^2)^2}. \end{aligned}$$

Nous ne justifions pas cette «technique de la série dérivée». Elle est utile dans de nombreux exercices.

4. Notons X_1, X_2, X_3, X_4 les variables aléatoires (à valeurs dans $\{P, F\}$) résultats du lancer des 4 pièces de 1 euro. Notons X_7, \dots, X_{10} les variables aléatoires (à valeurs dans $\{P, F\}$) résultats du lancer des 6 pièces de 2 euros. Les variables X_1, \dots, X_{10} sont i.i.d. (avec $\mathbb{P}(X_1 = F) = 1/2$). Calculons

$$\begin{aligned} \mathbb{P}(N = 4) &= \mathbb{P}(\sqcup_{\{i_1, \dots, i_4\} \subset \{1, \dots, 10\}} \{X_i = F, \forall i \in \{i_1, \dots, i_4\} \text{ et } X_i = P, \forall i \notin \{i_1, \dots, i_4\}\}) \\ &= \sum_{\{i_1, \dots, i_4\} \subset \{1, \dots, 10\}} \mathbb{P}(\{X_i = F, \forall i \in \{i_1, \dots, i_4\} \text{ et } X_i = P, \forall i \notin \{i_1, \dots, i_4\}\}) \\ &\quad \text{(par indépendance des } X_i) \\ &= \sum_{\{i_1, \dots, i_4\} \subset \{1, \dots, 10\}} \left(\frac{1}{2} \right)^{10} \\ &= C_{10}^4 \left(\frac{1}{2} \right)^{10}, \end{aligned}$$

et

$$\begin{aligned}
 & \mathbb{P}(\#\{i \in \{1, \dots, 4\} : X_i = F\} = 2, N = 4) \\
 &= \mathbb{P}(\#\{i \in \{1, \dots, 4\} : X_i = F\} = 2, \#\{i \in \{5, \dots, 10\} : X_i = F\} = 2) \\
 &= \mathbb{P}(\sqcup_{\{i_1, i_2\} \subset \{1, \dots, 4\}, \{j_1, j_2\} \subset \{5, \dots, 10\}} \{X_i = F \text{ ssi } i \in \{i_1, i_2, j_1, j_2\}\}) \\
 &= \sum_{\{i_1, i_2\} \subset \{1, \dots, 4\}, \{j_1, j_2\} \subset \{5, \dots, 10\}} \mathbb{P}(\{X_i = F \text{ ssi } i \in \{i_1, i_2, j_1, j_2\}\}) \\
 &= C_4^2 C_6^2 \left(\frac{1}{2}\right)^{10}.
 \end{aligned}$$

Donc la probabilité recherchée est

$$\frac{4!6!4!6!}{2!2!4!2!10!} = \frac{1}{8 \times 10 \times 3 \times 7} = \frac{1}{1680}.$$

5. Notons X_0 la résultat du dé rouge et X_1, X_2, \dots les résultats successifs du dé vert. Nous avons

$$\begin{aligned}
 \{\text{arrêt sur un 4}\} &= \sqcup_{i \geq 1} (\{1 \leq X_0 \leq 3\} \cup \{\text{arrêt en } i\}) \\
 &= \sqcup_{i \geq 1} (\{1 \leq X_0 \leq 3\} \\
 &\quad \cup \{X_j \notin \{7 - X_1, 4 - X_1\}, \forall j \in \{1, \dots, i-1\}\} \cup \{X_i = 4 - X_1\}). \quad (3.7.2)
 \end{aligned}$$

Donc¹ (en utilisant l'indépendance des X_i)

$$\begin{aligned}
 \mathbb{P}(\{\text{arrêt sur un 4}\}) &= \sum_{i=1}^{+\infty} \left(\frac{3}{6}\right) \left(\frac{4}{6}\right)^{i-1} \left(\frac{1}{6}\right) = \frac{3}{36} \sum_{i=1}^{+\infty} \left(\frac{4}{6}\right)^{i-1} \\
 &(\text{somme géométrique}) = \frac{3}{36} \times \frac{1}{\left(1 - \frac{4}{6}\right)} = \frac{3}{36} \times \frac{6}{2} = \frac{1}{4}. \quad (3.7.3)
 \end{aligned}$$

6.

- (a) Notons U_1, U_2, \dots les variables tirées successivement dans la boucle de l'algorithme. Les variables N, U_1, U_2, \dots sont indépendantes. Nous avons $Z = U_1 + \dots + U_N$. Calculons

$$\begin{aligned}
 \mathbb{E}(U_1) &= \frac{1}{2}, \quad \text{Var}(U_1) = \frac{1}{2}, \\
 \mathbb{E}(N) &= \frac{1}{6}(1 + \dots + 6) = \frac{7}{2}, \\
 \mathbb{E}(N^2) &= \frac{1}{6}(1 + 2^2 + \dots + 6^2) = \frac{6 \times 7 \times 13}{6 \times 6} = \frac{91}{6}, \\
 \text{Var}(N) &= \frac{91}{6} - \frac{49}{4} = \frac{182 - 156}{12} = \frac{26}{12} = \frac{13}{6}.
 \end{aligned}$$

Nous avons donc

$$\begin{aligned}
 \mathbb{E}(Z) &= \frac{1}{2} \times \frac{7}{2} = \frac{7}{4}, \\
 \text{Var}(Z) &= \frac{7}{2} \times \frac{1}{2} + \frac{1}{4} \times \frac{13}{6} = \frac{7}{4} + \frac{13}{24} = \frac{21 + 13}{24} = \frac{34}{24} = \frac{17}{12}.
 \end{aligned}$$

¹ Attention, astuce : dans (3.7.2), l'ensemble $\{1, \dots, i-1\}$ est vide si $i = 1$. Il faut donc calculer à part la probabilité $\mathbb{P}(1 \leq X_0 \leq 3, X_1 = 4 - X_1) = \frac{3}{6} \times \frac{1}{6}$. C'est bien le premier terme de la somme dans (3.7.3) car $\left(\frac{4}{6}\right)^0 = 1$.

(b) Calculons, pour $n \in \mathbb{N}^*$ et $k \in \{0, \dots, n\}$,

$$\mathbb{P}(U_1 + \dots + U_n = k) = C_n^k \left(\frac{1}{2}\right)^n.$$

La variable Z est à valeurs dans $\{0, \dots, 6\}$. Calculons, pour $k \in \{0, \dots, 6\}$,

$$\begin{aligned} \mathbb{P}(Z = k) &= \sum_{n=1}^{+\infty} \mathbb{P}(Z = k | N = n) \mathbb{P}(N = n) \\ &= \sum_{n=1}^6 \mathbb{P}(Z = k | N = n) \mathbb{P}(N = n), \end{aligned}$$

or $\mathbb{P}(Z = k | N = n) = 0$ si $n < k$, donc

$$\mathbb{P}(Z = k) = \sum_{n=\max(k,1)}^6 C_n^k \left(\frac{1}{2}\right)^n \frac{1}{6}.$$

7. Notons U_1, \dots, U_6 les variables $\sim \mathcal{B}(1/2)$ utilisées dans la première boucle et notons $N = U_1 + \dots + U_6$. Notons V_1, V_2, \dots les variables $\sim \mathcal{B}(1/2)$ utilisées dans la deuxième boucle et notons $Z = V_1 + \dots + V_N$. La variable Z est celle affichée par le programme.

(a) Nous avons

$$\mathbb{E}(V_1) = \frac{1}{2}, \quad \text{Var}(V_1) = \frac{1}{2},$$

$$\mathbb{E}(N) = \frac{6}{2} = 3,$$

$$\begin{aligned} \mathbb{E}(N^2) &= \sum_{1 \leq i, j \leq 6} \mathbb{E}(U_i U_j) \\ \text{(les } U_i \text{ sont i.i.d.)} &= 6\mathbb{E}(U_1^2) + 30\mathbb{E}(U_1)\mathbb{E}(U_2) \\ &= 3 + \frac{30}{4} = \frac{42}{4} = \frac{21}{2}, \end{aligned}$$

$$\text{Var}(N) = \frac{21}{2} - 9 = \frac{3}{2}.$$

Donc

$$\mathbb{E}(Z) = \frac{1}{2} \times 3 = \frac{3}{2},$$

$$\text{Var}(Z) = \frac{3}{2} + \frac{1}{4} \times \frac{3}{2} = \frac{12+3}{8} = \frac{15}{8}.$$

(b) Calculons

$$\begin{aligned}
\mathbb{P}(Z = 2) &= \sum_{n=0}^6 \mathbb{P}(Z = 2|N = n)\mathbb{P}(N = n) \\
(\mathbb{P}(Z = 2|N = n) = 0 \text{ si } n \leq 1) &= \sum_{n=2}^6 \mathbb{P}(Z = 2|N = n)C_6^n \left(\frac{1}{2}\right)^n \\
&= \sum_{n=2}^6 C_n^2 C_6^n \left(\frac{1}{2}\right)^n \\
&= \sum_{n=2}^6 \frac{n!6!}{2!(n-2)!n!(6-n)!} \left(\frac{1}{2}\right)^n \\
&= \frac{6 \times 5}{8} \sum_{n=2}^6 C_4^{n-2} \left(\frac{1}{2}\right)^{n-2} \\
(\text{changement d'indice } k = n - 2) &= \frac{30}{8} \sum_{k=0}^4 C_4^k \left(\frac{1}{2}\right)^k \\
(\text{formule du binôme}) &= \frac{30}{8} \left(1 + \frac{1}{2}\right)^4 \\
&= \frac{30 \times 3^4}{8 \times 2^4}.
\end{aligned}$$

8. Notons U_1, U_2, U_3 les variables successives générées par la boucle. La variable P suit une loi $\mathcal{U}([0; 1])$. Conditionnellement à P , les U_i suivent la loi $\mathcal{B}(P)$. Notons $S = U_1 + U_2 + U_3$. Conditionnellement à P , S suit une loi $\mathcal{B}(3, P)$. Calculons

$$\begin{aligned}
\mathbb{P}(S = 2) &= \int_{-\infty}^{+\infty} \mathbb{P}(S = 2|P = y)\mathbf{1}_{[0;1]}(y)dy \\
&= \int_{-\infty}^{+\infty}
\end{aligned}$$

9.

- (a) La variable u tirée lors d'un appel à la fonction expo est de loi $\mathcal{U}([0; 1])$ (de densité $u \in \mathbb{R} \mapsto \mathbf{1}_{[0;1]}(u)$). Notons R la variable $-\log(1 - u)$. La fonction

$$g : t \in]0; 1] \mapsto -\log(1 - t) \in \mathbb{R}^+$$

est bijective décroissante, d'inverse

$$g^{-1} : y \in [0; +\infty[\mapsto 1 - e^{-y}.$$

Nous avons $R(\omega) \in [0; +\infty]$ (pour tout ω). Notons f la densité de R . Nous avons donc $f(y) = 0$ si $y < 0$. Par la formule de changement de variable, nous avons pour $y \geq 0$,

$$\begin{aligned}
f(y) &= -\frac{\mathbf{1}_{[0;1]}(g^{-1}(y))}{g'(g^{-1}(y))} \\
&= -\frac{1}{\left(\frac{1}{1-(1-\exp(-y))}\right)} = e^{-y}.
\end{aligned}$$

Donc la densité de R est la fonction

$$y \in \mathbb{R} \mapsto \mathbf{1}_{[0;+\infty[}(y)e^{-y},$$

(Même si ce n'est pas demandé, on peut remarquer que R est de loi $\mathcal{E}(1)$).

- (b) Notons R_1, R_2, \dots les variables successivement simulées dans la boucle «while». Notons T la variable s affichée à la fin du programme. La boucle «while» s'arrête dès que l'on tire une variable $R_i > 1$. Donc², $\forall k \in \mathbb{N}^*$,

$$\begin{aligned} \mathbb{P}(T = k) &= \mathbb{P}(\{R_i \leq 1, \forall i \in \{1, \dots, k-1\}\} \cap \{R_k > 1\}) \\ (\text{les } R_j \text{ sont indépendants}) &= \prod_{i=1}^{k-1} \mathbb{P}(R_i \leq 1) \times \mathbb{P}(R_k > 1). \end{aligned}$$

Calculons

$$\begin{aligned} \mathbb{P}(R_1 \leq 1) &= \int_{-\infty}^1 \mathbf{1}_{[0;+\infty[}(y)e^{-y} dy \\ &= \int_0^1 e^{-y} dy \\ &= [-e^{-y}]_0^1 = 1 - e^{-1}, \end{aligned}$$

$$\begin{aligned} \mathbb{P}(R_1 > 1) &= \int_1^{+\infty} \mathbf{1}_{[0;+\infty[}(y)e^{-y} dy \\ &= \int_1^{+\infty} e^{-y} dy \\ &= [-e^{-y}]_1^{+\infty} = e^{-1}. \end{aligned}$$

Les R_1, R_2, \dots sont de même loi donc

$$\mathbb{P}(T = k) = (1 - e^{-1})^{k-1} \times e^{-1}.$$

(On peut remarquer, même si ce n'est pas demandé, que $T \sim \mathcal{G}(e^{-1})$.)

- (c) La variable r (que nous noterons R) affichée à la fin du programme est toujours > 1 . Donc, $\forall t \leq 1$, $\mathbb{P}(R \leq t) = 0$. Calculons, pour un $t > 1$,

$$\begin{aligned} \mathbb{P}(R \leq t) &= \mathbb{P}(\bigsqcup_{k=1}^{+\infty} (\{T = k\} \cap \{R \leq t\})) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}(T = k, R \leq t) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}(T = k, R_k \leq t) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}(R_1 \leq 1, \dots, R_{k-1} \leq 1, R_k \in]1; t]) \\ (\text{les } R_j \text{ sont i.i.d.}) &= \sum_{k=1}^{+\infty} \mathbb{P}(R_1 \leq 1)^{k-1} \mathbb{P}(1 < R_1 \leq t). \end{aligned}$$

²Attention, nous utilisons dans le calcul suivant diverses conventions dans le cas $k = 1$. Les étudiants peu à l'aise avec ces conventions peuvent toujours faire un calcul à part pour $k = 1$, le résultat sera le même.

Rappelons la fonction de répartition de la loi $\mathcal{E}(1)$:

$$F : x \in \mathbb{R} \mapsto \mathbf{1}_{[0;+\infty[}(x)(1 - e^{-x}).$$

Nous avons

$$\mathbb{P}(R_1 \leq 1) = F(1), \quad \mathbb{P}(1 < R_1 \leq t) = F(t) - F(1).$$

Donc

$$\begin{aligned} \mathbb{P}(R \leq t) &= \sum_{k=1}^{+\infty} (1 - e^{-1})^{k-1} (e^{-1} - e^{-t}) \\ \text{(somme géométrique)} &= (e^{-1} - e^{-t}) \times \frac{1}{1 - (1 - e^{-1})} \\ &= \frac{e^{-1} - e^{-t}}{e^{-1}} = e^{-(t-1)}. \end{aligned}$$

(Ce qui définit complètement la loi de R .)

Chapitre 4

Estimation et test d'hypothèse

4.1 Estimation

4.1.1 Estimation d'une moyenne par intervalle de confiance

Théorème 4.1.1. Soient X_1, X_2, \dots des variables i.i.d. telles que $\mathbb{E}(X_1^2) = \sigma^2 < \infty$, $\mathbb{E}(X_1) = \mu$. Soit $\alpha \in]0; 1[$. Soit n «grand». Notons $\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ (la moyenne empirique). Alors :

$$\mathbb{P}\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) \approx 1 - \alpha, \quad (4.1.1)$$

où $z_{\alpha/2}$ est la valeur telle que, si $Z \sim \mathcal{N}(0, 1)$, $\mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$. Si $n \geq 30$, soit $\widehat{\sigma}_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2}$, alors :

$$\mathbb{P}\left(\bar{X}_n - z_{\alpha/2} \frac{\widehat{\sigma}_n}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{\alpha/2} \frac{\widehat{\sigma}_n}{\sqrt{n}}\right) \approx 1 - \alpha. \quad (4.1.2)$$

Les symboles « \approx » veulent dire ici que l'on considère que les termes de part et d'autre des ces égalités sont raisonnablement proches.

Démonstration. Il s'agit d'une application du TCL. Calculons :

$$\begin{aligned} \mathbb{P}\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) &= \mathbb{P}\left(\mu - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \bar{X}_n \leq \mu + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) \\ &= \mathbb{P}\left(-z_{\alpha/2} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq z_{\alpha/2}\right) \\ (\text{par le TCL}) &\approx 1 - \alpha. \end{aligned}$$

Par la loi des grands nombres, $\bar{X}_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mu$ et $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \sigma^2$. On admet alors (sans démonstration) que $\widehat{\sigma}_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \sigma$. On considère que $\widehat{\sigma}_n \approx \sigma$ pour $n \geq 30$. Nous avons alors :

$$\mathbb{P}\left(\bar{X}_n - z_{\alpha/2} \frac{\widehat{\sigma}_n}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{\alpha/2} \frac{\widehat{\sigma}_n}{\sqrt{n}}\right) \approx 1 - \alpha.$$

□

Définition 4.1.2. La quantité \bar{X}_n s'appelle la moyenne empirique. La quantité $\widehat{\sigma}_n$ s'appelle l'écart-type empirique ($\widehat{\sigma}_n^2$ s'appelle la variance empirique). La quantité σ s'appelle l'écart-type.

Exemple 4.1.3. On interroge des entreprises d'informatique sur le temps moyen de recouvrement d'achats de produits informatiques. Les résultats sont les suivants.

- Nombre d'entreprises : 225.
- Temps moyen (empirique) de recouvrement : 2,6 ans.
- Écart-type empirique du temps de recouvrement (c'est à dire s_n) : 0,6 ans.

On suppose que les réponses données sont des variables i.i.d (notées X_1, X_2, \dots). On voudrait estimer leur moyenne μ avec un niveau de confiance de 95%. On pose donc $\alpha = 0,05$. Soit $Z \sim \mathcal{N}(0, 1)$. Calculons $z_{\alpha/2}$, ce nombre vérifie $\mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = \alpha$, c'est à dire $\mathbb{P}(Z \leq z_{\alpha/2}) = \mathbb{P}(Z < -z_{\alpha/2}) + \mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = \frac{1 - \mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2})}{2} + \mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 0,5 - \frac{0,05}{2} = 0,475$ (nous utilisons ici la symétrie de la gaussienne). D'après la table de la loi gaussienne, $z_{\alpha/2} = 1,96$. Le théorème précédent nous dit donc que

$$\mathbb{P}\left(\mu \in \left[2,6 - 1,96 \times \frac{0,6}{\sqrt{225}}; 2,6 + 1,96 \times \frac{0,6}{\sqrt{225}}\right] = [2,52; 2,68]\right) \approx 0,95.$$

On attribue à l'intervalle $[2,52; 2,68]$ un niveau de confiance de 95% de contenir la vraie valeur moyenne de la période de recouvrement de achats informatiques.

Remarque 4.1.4. • Plus le niveau de confiance est élevé, plus l'amplitude de l'intervalle est grande. Pour la même taille d'échantillon, on perd donc de la précision en gagnant de la confiance.

- L'intervalle de confiance peut être différent pour des échantillons différents puisqu'il est centré sur la moyenne empirique qui dépende de l'échantillon.
- Dans les cas où la variance σ est inconnue (comme dans l'exemple ci-dessus), il faut appliquer la formule (4.1.2) du théorème ci-dessus (si toutefois $n \geq 30$) et donc l'amplitude de l'intervalle de confiance peut varier suivant les échantillons puisqu'elle dépend de l'écart-type empirique.

4.1.2 Marge d'erreur associée à l'estimation de la moyenne et taille d'échantillon requise pour ne pas excéder la marge d'erreur

On se place dans le cadre du théorème 4.1.1. La marge d'erreur est la quantité $|\bar{X}_n - \mu|$. D'après le théorème précédent, pour un niveau de confiance $1 - \alpha$, la marge d'erreur est inférieure à $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$. On pourrait aussi vouloir fixer la marge d'erreur et chercher la taille minimale de l'échantillon requise (α étant toujours fixé).

Proposition 4.1.5. Dans le cadre du th. 4.1.1, la taille minimale de l'échantillon requise pour atteindre la marge d'erreur E est

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2.$$

Cette taille d'échantillon nous assure que la marge d'erreur dans l'estimation de μ par \bar{X}_n est au plus E . Une fois la taille d'échantillon calculée avec la formule ci-dessus, on peut directement affirmer

$$\bar{X}_n - E \leq \mu \leq \bar{X}_n + E,$$

avec une probabilité $\geq 1 - \alpha$.

4.1.3 Estimation d'une moyenne par intervalle de confiance dans le cas d'un petit échantillon ($n < 30$)

On suppose toujours que l'on a un échantillon X_1, X_2, \dots, X_n de variables réelles i.i.d. telles que $\mathbb{E}(X_1) = \mu$. On note $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$, $\widehat{\sigma}_n = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{n}}$. Dans le cas $n < 30$, l'estimation de σ^2 par la variance empirique $\widehat{\sigma}_n^2$ n'est plus fiable. On applique alors le théorème ci-dessous.

Définition 4.1.6. La loi de Student à k degrés de liberté ($k \geq 1$) a pour densité sur \mathbb{R}

$$f_{T,k}(t) = \frac{1}{\sqrt{k\pi}} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{1}{\left(1 + \frac{t^2}{k}\right)^{\frac{k+1}{2}}},$$

où Γ est la fonction Γ d'Euler ($\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$). L'expression exacte de $f_{T,k}$ a peu d'importance, il nous suffit de connaître ses propriétés. Si X est de loi de densité $f_{T,k}$ ($k \geq 1$), alors :

- X est à valeurs dans $]-\infty; +\infty[$
- La densité de X a le dessin de la figure 4.1.6.

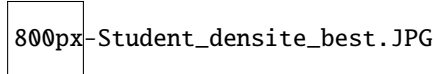


FIGURE 4.1 – Densité de la loi de Student pour différents degrés de liberté.

- $f_{T,k}$ est une fonction paire
- $\mathbb{E}(X) = 0$, $\text{Var}(X) = \frac{k}{k-2}$
- Pour tout k , les queues de distribution de $f_{T,k}$ sont plus lourdes que celles de la gaussienne.
- Quand $k \rightarrow +\infty$, $f_{T,k}$ se rapproche de la gaussienne $\mathcal{N}(0, 1)$.
- On pourra se servir de la table donnée dans l'annexe B. La table se lit de la manière suivante. Pour v =degrés de libertés et P une probabilité, on lit dans la table t tel que $\mathbb{P}(|X| > t) = P$. Par exemple, avec $k = 11$, $\mathbb{P}(|X| \geq 2, 201) = 0, 05$.

Théorème 4.1.7. Sous les hypothèse ci-dessus, la variable $\frac{\bar{X}_n - \mu}{(\widehat{\sigma}_n / \sqrt{n})}$ suit (approximativement) une loi de Student à $n - 1$ degrés de liberté. Pour $\alpha \in [0; 1]$,

$$\mathbb{P}(\mu \in [\bar{X}_n - t_{\alpha/2, n-1} \frac{\widehat{\sigma}_n}{\sqrt{n}}; \bar{X}_n + t_{\alpha/2, n-1} \frac{\widehat{\sigma}_n}{\sqrt{n}}]) \approx 1 - \alpha,$$

où $t_{\alpha/2, n-1}$ est tel que, si X de loi de densité $f_{T, n-1}$, $\mathbb{P}(|X| \leq t_{\alpha/2, n-1}) = 1 - \alpha$. Et donc la marge d'erreur statistique au niveau de confiance $100(1 - \alpha)\%$ est

$$E = t_{\alpha/2, n-1} \frac{\widehat{\sigma}_n}{\sqrt{n}}.$$

Exemple 4.1.8. On s'intéresse à l'âge moyen des dirigeants de menuiserie. Une étude nous donne le résultats suivants : taille de l'échantillon (n)=16, moyenne empirique (\bar{X}_n)=48,4 ans, écart-type empirique ($\widehat{\sigma}_n$)=3,5 ans. Pour $\alpha \in [0; 1]$, l'intervalle de confiance de niveau $(1 - \alpha)$ est, d'après le théorème ci-dessus

$$[\bar{X}_n - t_{\alpha/2,15} \frac{\widehat{\sigma}_n}{\sqrt{n}}; \bar{X}_n + t_{\alpha/2,15} \frac{\widehat{\sigma}_n}{\sqrt{n}}].$$

On obtient les résultats suivants.

Niveau de confiance	$t_{\alpha/2,15}$	$t_{\alpha/2,15} \frac{\widehat{\sigma}_n}{\sqrt{n}}$	$\bar{X}_n - t_{\alpha/2,15} \frac{\widehat{\sigma}_n}{\sqrt{n}}$	$\bar{X}_n + t_{\alpha/2,15} \frac{\widehat{\sigma}_n}{\sqrt{n}}$
90%	1,7531	1,534	46,87	49,93
95%	2,1315	1,865	46,54	50,26
99%	2,9467	2,578	45,82	50,98

4.1.4 Estimation d'une proportion par intervalle de confiance

Si on dispose d'un échantillon X_1, \dots, X_n de variable i.i.d. et que l'on note, $\forall k, U_k = \mathbf{1}_A(X_k)$ pour un certain ensemble A . Alors les variables U_k sont i.i.d. de loi $\mathcal{B}(p)$ avec $p = \mathbb{P}(X_1 \in A) = \mathbb{E}(U_1)$. Si on veut estimer p , on peut appliquer les théorèmes précédents à l'échantillon U_1, \dots, U_n . Notons $\widehat{p}_n = \frac{U_1 + \dots + U_n}{n}$.

Proposition 4.1.9. Sous les hypothèses ci-dessus,

$$\mathbb{P}(p \in [\widehat{p}_n - z_{\alpha/2} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}; \widehat{p}_n + z_{\alpha/2} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}]) \approx 1 - \alpha.$$

Démonstration. Nous avons $\mathbb{E}(U_1^2) = 1^2 \times p + 0^2 \times (1-p) = p$. Donc $\text{Var}(U_1^2) = p - p^2 = p(1-p)$. Donc on considère que $\widehat{p}_n(1 - \widehat{p}_n)$ est proche de $\text{Var}(U_1^2)$ et on applique le TCL. \square

Exemple 4.1.10. Un sondage révèle que 218 personnes sur 509 personnes interrogées déclarent qu'elles payent pour avoir accès à internet. On cherche à déterminer la proportion p de personnes payent pour avoir accès internet. On fait l'hypothèse que les réponses des sondés sont indépendantes et sont aléatoires de telle manière que, pour chaque personne, la probabilité de payer pour avoir accès à internet est p . Ici $n = 509$, $\widehat{p}_n = \frac{218}{509}$. Voici un tableau récapitulatif les bornes de l'intervalle de confiance pour différents niveaux de confiance :

Niveau de confiance	$z_{\frac{\alpha}{2}}$	$z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}$	$\widehat{p}_n - z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}$	$\widehat{p}_n + z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}$
90%	1,645	0,036	0,392	0,464
95%	1,96	0,043	0,385	0,471
99%	2,576	0,057	0,372	0,485

TABLE 4.1 – Tableau récapitulatif

4.1.5 Marge d'erreur associée à l'estimation de p et taille d'échantillon requise

En appliquant les résultats précédents, on obtient la proposition suivante.

Proposition 4.1.11. Lorsque \widehat{p}_n est utilisé comme estimation de p alors, pour un niveau de confiance $100(1 - \alpha)\%$, la marge d'erreur $E = |\widehat{p}_n - p|$ vérifie

$$E \leq z_{\alpha/2} \sqrt{\frac{\widehat{p}_n(1 - \widehat{p}_n)}{n}}.$$

On peut aussi chercher à déterminer la taille minimale de l'échantillon requise pour une marge d'erreur fixée à l'avance. Il s'agit de résoudre l'équation en n : $E = z_{\alpha/2} \sqrt{\frac{\widehat{p}_n(1 - \widehat{p}_n)}{n}}$ (attention, il y a du n à de nombreux endroits dans cette égalité).

Proposition 4.1.12. 1. Si on connaît une valeur \widehat{p}_{n_0} approchant p obtenue par un sondage préalable sur un petit échantillon (par exemple avec $n_0 = 30$). Pour une marge d'erreur E désirée et un niveau de confiance $100(1 - \alpha)\%$, la taille d'échantillon minimale requises est

$$n = \frac{z_{\alpha/2}^2 \widehat{p}_{n_0}(1 - \widehat{p}_{n_0})}{E^2}. \quad (4.1.3)$$

2. Si on ne dispose pas d'une valeur approximative de p , on la «fixe» à 0,5. Cette valeur représente le cas le plus défavorable, c'est à dire la valeur qui donne le plus grand écart-type possible. La taille d'échantillon requise pour un niveau de confiance $100(1 - \alpha)\%$ et une marge d'erreur E est alors

$$n = \frac{z_{\alpha/2}^2}{4E^2}.$$

Si on veut le niveau de confiance 95,44%, il se trouve que $z_{0,0228} = 2$ et donc on trouve

$$n = \frac{1}{E^2}.$$

Démonstration. 1. Ce résultat est une application de la proposition 4.1.11.

2. D'après la proposition 4.1.11, la marge d'erreur vérifie $|p - \widehat{p}_n| \leq z_{\alpha/2} \sqrt{\frac{\widehat{p}_n(1 - \widehat{p}_n)}{n}}$ (pour tout estimateur raisonnable \widehat{p}_n de p). On sait que $\forall x \in [0; 1], 0 \leq x(1 - x) \leq 1/4$, donc $E \leq z_{\alpha/2} \frac{1}{2\sqrt{n}}$. Donc, à α et E fixé, si on veut que $|p - \widehat{p}_n| \leq E$, il suffit de prendre

$$n \geq z_{\alpha/2}^2 \frac{1}{4E^2}.$$

□

Exemple 4.1.13. Un sondage révèle que sur 100 dirigeants de P.M.E., 82 se disent en faveur de la mondialisation. Pour un niveau de confiance de 95%, la marge d'erreur est

$$z_{0,025} \sqrt{\frac{\widehat{p}_n(1 - \widehat{p}_n)}{n}} = 1,96 \times \sqrt{\frac{0,82 \times 0,18}{100}} = 0,075.$$

4.2 Exercices

4.2.1 Énoncés des exercices

1. Supposons que la dépense moyenne par ménage pour les biens culturels est de 664 €. On suppose que ces dépenses suivent une loi normale de variance 4096. On prélève un échantillon aléatoire de taille 25 de ces dépenses (on note X_1, X_2, \dots les dépenses).

Nombre d'observations :	40
Moyenne :	69,3
Écart-type :	9,06

TABLE 4.2 – Tableau récapitulatif

- (a) Calculer $\mathbb{P}(\bar{X}_n \leq 650)$ pour $n = 25$.
- (b) Calculer $\mathbb{P}(638,40 \leq \bar{X}_n \leq 689,60)$ pour $n = 25$.
- (c) Sur 100000 ménages, combien ont une dépense supérieure à 760 € ? (Justifier la réponse.)
2. On observe les temps d'assistance requis dans un service d'assistance par téléphone (que nous noterons X_1, X_2, \dots). On fait $n = 40$ observations. Voici un tableau récapitulatif de ces observations :
- (a) Déterminez un intervalle de confiance pour le vrai temps moyen avec un niveau de confiance de
i) 90% , ii) 95% , iii) 99% .
- (b) Quel est la taille de l'échantillon qu'il faudrait avoir pour avoir une marge d'erreur de 1 avec un niveau de confiance de 90% ?
3. On s'intéresse au coût par P.M.E. du passage à la norme ISO9000. On obtient les résultats suivant sur un échantillon de 20 P.M.E.

Moyenne empirique	19682,00 €
Écart-type empirique	1713,80 €

TABLE 4.3 – Tableau des résultats

- (a) Déterminez, pour un niveau de confiance de 95%, la marge d'erreur statistique dans l'estimation du coût moyen en supposant que l'on peut appliquer le TCL.
- (b) Déterminez, pour un niveau de confiance de 95%, la marge d'erreur statistique dans l'estimation du coût moyen en utilisant le th. 4.1.7.
4. Un organisme de sondage veut effectuer un sondage auprès des T.P.E. (les très petites entreprise) pour déterminer leur degré d'informatisation. On veut déterminer la taille de l'échantillon requise pour estimer le pourcentage de T.P.E. qui sont informatisées avec une marge d'erreur n'excédant pas 4% et un niveau de confiance de 95%. D'après un sondage effectué il y a deux ans, 46% des T.P.E. étaient informatisées.
- (a) Quelle est la marge d'erreur que l'on ne veut pas excéder ?
- (b) Quelle expression est requise pour déterminer la taille de l'échantillon ?
- (c) Combien de T.P.E. devrait-on sonder, selon les exigences requise ici ?
5. Le vérificateur interne d'une entreprise a effectué une vérification par échantillonnage aléatoire sur 28 transactions et a obtenu une erreur moyenne d'écriture de 225 € avec un écart-type empirique de 20 €.

- (a) En supposant que le montant associé aux erreurs d'écriture est distribué normalement, estimez par intervalle de confiance l'erreur moyenne réelle avec un niveau de confiance de 95%.
- (b) Quelle est la marge d'erreur statistique dans l'estimation de l'erreur moyenne d'écriture, pour le niveau de confiance précisé en 5a ?

4.2.2 Corrigés des exercices

1.

- (a) $\mathbb{P}(X_n \leq 650) = \mathbb{P}\left(\frac{5(\bar{X}_n - 664)}{64} \leq \frac{5(650 - 664)}{64}\right) \approx \mathbb{P}(Z \leq -1.09)$ (par le TCL) avec $Z \sim \mathcal{N}(0, 1)$ et $\mathbb{P}(Z \leq -1, 09) = \mathbb{P}(Z \leq 1, 09) = 1 - \mathbb{P}(Z \leq 1, 09) = 1 - 0, 8621 = 0, 1379$

(b)

$$\begin{aligned} \mathbb{P}(638, 40 \leq \bar{X}_n \leq 689, 60) &= \mathbb{P}\left(\frac{5 \times (638, 40 - 664)}{64} \leq \frac{5 \times (\bar{X}_n - 664)}{64} \leq \frac{5 \times (689, 60 - 664)}{64}\right) \\ &\stackrel{\text{(TCL)}}{=} \mathbb{P}(-2 \leq Z \leq 2) = \mathbb{P}(Z \leq 2) - \mathbb{P}(Z \leq -2) \\ &= \mathbb{P}(Z \leq 2) - \mathbb{P}(Z \geq 2) = 2\mathbb{P}(Z \leq 2) - 1 \\ &= 2 \times 0, 9772 - 1 = 0, 9544 \end{aligned}$$

- (c) Pour tout n on note $U_n = \mathbf{1}_{[760; +\infty[}(X_n)$. Les variables (X_n) sont indépendantes donc les variables (U_n) sont indépendantes. Par la loi des grands nombres : $\bar{U}_n := \frac{U_1 + \dots + U_n}{n} \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(U_n) = \mathbb{P}(X_n \geq 760)$. Donc $\bar{U}_{100000} = \mathbb{P}(X_n \geq 760) = \mathbb{P}\left(\frac{X_n - 664}{64} \geq \frac{760 - 664}{64}\right) \approx$ (par le TCL) $\mathbb{P}(Z \geq 1, 5) = 1 - \mathbb{P}(Z \leq 1, 5) = 1 - 0, 9332 = 0, 0668$. Donc sur 100000 personnes, environ 67000 dépensent plus de 760 € par an en biens culturels.

2.

- (a) On cherche $z_{\alpha/2} : \mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$ (pour $Z \in \mathcal{N}(0, 1)$, $\alpha \in [0; 1]$). Nous avons

$$\begin{aligned} \mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) &= \mathbb{P}(Z \leq z_{\alpha/2}) - \mathbb{P}(Z \leq -z_{\alpha/2}) = \mathbb{P}(Z \leq z_{\alpha/2}) - \mathbb{P}(Z \geq z_{\alpha/2}) \\ &= \mathbb{P}(Z \leq z_{\alpha/2}) - (1 - \mathbb{P}(Z \leq z_{\alpha/2})) = -1 + 2\mathbb{P}(Z \leq z_{\alpha/2}). \end{aligned}$$

- i. On veut $-1 + 2\mathbb{P}(Z \leq z_{\alpha/2}) = 0, 9$ donc $\mathbb{P}(Z \leq z_{\alpha/2}) = 0, 85$ donc $z_{\alpha/2} = 1, 04$ donc l'intervalle de confiance est $[67, 8; 70, 8]$.
- ii. De même, on veut ici $\mathbb{P}(Z \leq z_{\alpha/2}) = 0, 975$ donc $z_{\alpha/2} = 1, 96$ donc l'intervalle de confiance est $[66, 5; 72, 1]$.
- iii. De même, on veut ici $\mathbb{P}(Z \leq z_{\alpha/2}) = 0, 995$ donc $z_{\alpha/2} = 2, 57$ donc l'intervalle de confiance est $[65, 6; 73]$.
- (b) On applique le résultat du cours

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 = \left(\frac{1, 04 \times 9, 06}{1}\right)^2 \approx 89.$$

3.

- (a) Soit $\alpha = 0,05$. Nous avons $\mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2})$ ($Z \sim \mathcal{N}(0, 1)$) donc $\mathbb{P}(Z \leq z_{\alpha/2}) = (1+\alpha)/2$ (calcul déjà fait dans l'exercice précédent). On lit sur la table : $z_{\alpha/2} = 1,96$. L'intervalle de confiance est

$$[\bar{X}_n - z_{\alpha/2} \frac{\widehat{\sigma}_n}{\sqrt{n}}; \bar{X}_n + z_{\alpha/2} \frac{\widehat{\sigma}_n}{\sqrt{n}}] = [19869, 78; 19494, 22].$$

- (b) Soit X de loi de Student à 19 degrés de liberté. On lit sur la table : $t_{\frac{\alpha}{2}; 19} = 2,093$. D'après le th. 4.1.7, nous avons donc l'intervalle de confiance suivant (toujours pour le niveau 95%)

$$[\bar{X}_n - t_{\frac{\alpha}{2}; 19} \frac{\widehat{\sigma}_n}{\sqrt{n}}; \bar{X}_n + t_{\frac{\alpha}{2}; 19} \frac{\widehat{\sigma}_n}{\sqrt{n}}] = [19502, 65; 19681, 35].$$

4.

- (a) Appelons \widehat{p} l'estimée que l'on va faire, on veut que $\frac{|\widehat{p}-p|}{p} \leq 0,04$ (où p est la vraie proportion de T.P.E. informatisées). On peut se contenter de réutiliser l'ancien estimateur en disant que l'on veut $|\widehat{p}-p| \leq 0,04 \times 0,46$.
- (b) La formule (4.1.3) nous donne la taille de l'échantillon minimale pour répondre aux contraintes imposées :

$$n = \left\lceil \frac{z_{\alpha/2}^2 \widehat{p}_{n_0} (1 - \widehat{p}_{n_0})}{E^2} \right\rceil,$$

avec donc $E = 0,04 \times 0,46 = 0,0184$, $\alpha = 0,05$, et \widehat{p}_{n_0} l'estimation provenant de la précédente enquête ($\widehat{p}_{n_0} = 0,46$). Attention, on veut n entier, c'est pourquoi on prend la partie entière supérieure ($\lceil \dots \rceil$).

- (c) On calcule : $z_{\alpha/2} = 1,96$, $n = 2819$.

5.

- (a) Soit $\alpha = 0,05$. Soit $n = 28$. Soit \bar{X}_n la moyenne empirique (= 225). Soit $\widehat{\sigma}_n$ l'écart-type empirique (= 20). Soit μ la moyenne réelle. On suppose que $\sqrt{n} \frac{(\bar{X}_n - \mu)}{\widehat{\sigma}_n}$ est de loi $\mathcal{N}(0, 1)$. On cherche E tel que $\mathbb{P}(|\bar{X}_n - \mu| \leq E) \geq 1 - \alpha$. Nous avons

$$\begin{aligned} \mathbb{P}(|\bar{X}_n - \mu| \leq E) &= \mathbb{P}\left(\sqrt{n} \frac{(\bar{X}_n - \mu)}{\widehat{\sigma}_n} \geq \frac{\sqrt{n}E}{\widehat{\sigma}_n}\right) + \mathbb{P}\left(\sqrt{n} \frac{(\bar{X}_n - \mu)}{\widehat{\sigma}_n} \leq -\frac{\sqrt{n}E}{\widehat{\sigma}_n}\right) \\ &= 2\mathbb{P}\left(\sqrt{n} \frac{(\bar{X}_n - \mu)}{\widehat{\sigma}_n} \geq \frac{\sqrt{n}E}{\widehat{\sigma}_n}\right) = 2 - 2\mathbb{P}\left(\sqrt{n} \frac{(\bar{X}_n - \mu)}{\widehat{\sigma}_n} \leq \frac{\sqrt{n}E}{\widehat{\sigma}_n}\right). \end{aligned}$$

On veut donc $\mathbb{P}\left(\sqrt{n} \frac{(\bar{X}_n - \mu)}{\widehat{\sigma}_n} \leq \frac{\sqrt{n}E}{\widehat{\sigma}_n}\right) \geq 1 - \frac{\alpha}{2} = 0,925$. Donc on prend $\frac{\sqrt{n}E}{\widehat{\sigma}_n} = 1,96$, donc $E = 1,96 \times \frac{\widehat{\sigma}_n}{\sqrt{n}} = 363,84$. On peut donc dire que $\mu \in [-137,85; 587,85]$ avec probabilité 0,95. (Ce n'est pas un résultat très intéressant mais nous n'aurions pas dû appliquer le TCL puisque $n < 30$.)

- (b) Comme $n < 30$, on veut se servir de la loi de Student. La marge d'erreur est $E = t_{\alpha/2; n-1} \frac{\widehat{\sigma}_n}{\sqrt{n}} = 2,048 \times \frac{20}{\sqrt{28}} = 214,66$. Donc on peut dire que $\mu \in [10,34; 439,66]$ avec probabilité 0,95.

Appendix A

Table de la loi normale

Figure A.1: Table de la loi normale

Appendix B

Table de la loi de Student

Figure B.1: Table de la loi de Student

Appendix C

Fonctions, intégrales et sommes usuelles

Nous rappelons ($\forall a, b$)

$$\int_a^b e^{-x} dx = [-e^{-x}]_a^b = -e^{-b} + e^{-a}.$$

Par intégration par parties, nous avons

$$\int_a^b x e^{-x} dx = [-x e^{-x}]_a^b + \int_a^b e^{-x} dx.$$

Et $\forall \alpha$,

$$\int_a^b x^{\alpha-1} e^{-x} dx = [-x^{\alpha-1} e^{-x}]_a^b + (\alpha-1) \int_a^b x^{\alpha-2} e^{-x} dx.$$

Formule du binôme (de Newton) :

$$\forall x, y \in \mathbb{R}, \forall 0 \leq k \leq n (k, n \in \mathbb{N}), (x+y)^n = \sum_{k=0}^n C_n^k x^k y^{n-k},$$

où $C_n^k = \frac{n!}{k!(n-k)!}$ (c'est nombre de parties à k éléments dans un ensemble à n éléments) ($n! = 1 \times 2 \times 3 \times \dots \times n$).

Somme géométrique :

$$\forall p \in \mathbb{R}, \forall n \in \mathbb{N}, \sum_{k=0}^n p^k = \frac{1-p^{n+1}}{1-p},$$

$$\text{si, de plus, } |p| < 1, \sum_{k=0}^{+\infty} p^k = \frac{1}{1-p}.$$

Somme arithmétique

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Somme de carrés

$$1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Exponentielle :

$$\forall x \in \mathbb{R}, e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

Cosinus hyperbolique :

$$\forall x \in \mathbb{R}, \cosh(x) = \frac{e^x + e^{-x}}{2} = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}.$$

Sinus hyperbolique :

$$\forall x \in \mathbb{R}, \sinh(x) = \frac{e^x - e^{-x}}{2} = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}.$$