$L1$ Probabilités-Statistique pour la santé - $2021 - 2022$

Clément Marteau - Thibault Espinasse

Ce résumé a vocation à être un aide mémoire, à compléter, et éventuellement à corriger. Si vous trouvez des coquilles, merci de me contacter pour les corriger. Il s'agit évidemment d'une version provisoire. Quelques références en suppléments

- Le livre "Statistique mathématique, cours et exercices corrigés" de C. Vial et B. Cadre
- Le polycopié de M. Delarue de Statistique
- Le livre "Statistiques asyptotiques" de M. Van der Vaart [?] *(avancé !)*
- Le site internet Wikistat très riche : `http://wikistat.fr/`
- Le livre "Probabilités, analyse des données et statistique" de Gilbert Saporta

En outre, pour celles et ceux que ça intéresse, voici quelques chaînes youtube de vulgarisation qui peuvent donner un éclairage supplémentaire sur le cours.

1. Science4all
2. ScienceEtonnante
3. Mickaël Launay
4. Hygiène Mentale
5. El Jj
6. Data Gueule
7. 3Blue1Brown
8. La statistique expliquée à mon chat
9. Le chat sceptique
10. François Husson
11. ...

Contents

TODO : regarder les nouveaux programmes de terminale !

1 Introduction

De façon générale, lorsqu'on parle de probabilités, d'événement aléatoire, ou d'hasard, on cherche a prendre en considération *l'imprévisibilité* d'une expérience (ex : lancer de pièce, guérison d'une maladie, durée de vie d'un téléphone portable, ordre des cartes dans un packet mélangé). Que l'aléa soit réel ou modélise notre ignorance à propos du phénomène en question importe peu pour les calculs que l'on souhaite faire.

Pour pouvoir néammoins apprendre d'expériences passées sur ce qui pourrait se passer sur les futures expériences, on parlera souvent de "réaliser de nombreuses fois l'expérience dans des conditions similaires". la probabilité pouvant être comprise comme la *fréquence* (i.e. proportion des expériences) limite d'un résultat, *lorsqu'on réalise l'expérience dans les mêmes conditions un très grand nombres de fois.* On reviendra sur cette notion lorsqu'on évoquera la **Loi des grands nombres** qui décrit mathématiquement cette propriété.[1]

Pour formaliser la notion de probabilité, et de variable aléatoire, en mathématique, on va devoir décrire l'ensemble des résultats possibles d'une expérience, ainsi que la probabilité de chaque résultat. L'objet de la première partie de ce cours, est d'introduire toutes ces notions.

1.1 Univers, événement, et probabilité

Définition 1.1 *On appelle **univers** (et on notera souvent Ω) l'ensemble des résultats possibles d'une expérience aléatoire.*

Remarque : On en reparlera plus tard, mais en réalité, le choix de l'univers est un *choix de modélisation*, et il n'y a pas qu'une seule façon de modéliser une expérience aléatoire. On y reviendra en parlant de variables aléatoires, mais retenez déjà cette remarque.

Exemple 1.1

1. *Pour un lancer d'une pièce, on pourra prendre* $\Omega = \{Pile, Face\}$
2. *Pour le résutat d'un dé, on pourra prendre* $\Omega = \{1, 2, 3, 4, 5, 6\}$
3. *Pour l'âge d'une personne prise au hasard, on pourra prendre* $\Omega = \{0, 1, 2, \cdots, 150\}$*, ou* $\Omega = \mathbb{N}$ *ou encore* $\Omega = \mathbb{R}^+$ *selon si l'on conidère la mesure bornée, entière, ou réelle.*
4. *Pour un lancer de deux pièces, on pourra prendre* $\Omega = \Big\{(Pile, Pile), (Pile, Face), (Face, Pile), (Face, Face)\Big\} = \{Pile, Face\}^2$
5. *Pour un nombre au hasard entre* 0 *et* 1*, on pourra prendre* $\Omega = [0, 1]$

Remarque : Comme expliqué au dessus, on verra qu'en réalité, on aura rarement besoin de connaître Ω explicitement. De plus, quand on introduira la notion de *variable aléatoire*, on verra qu'une même variable aléatoire peut-être obtenue avec des univers différents. Cela est très intuitif : on peut "simuler" un pile ou face équilibré avec un dé équilibré par exemple. L'ensemble Ω est donc un **choix** de modélisation, et une même expérience pourra être étudiée avec des Ω très différents.

Chaque résultat de l'expérience $\omega \in \Omega$ est appelé **événement ponctuel** ou **événement élémentaire**. Un **événement aléatoire** est un ensemble d'événement ponctuel, et correspond donc à une partie $A \subset \Omega$ ($A \in \mathcal{P}(\Omega)$).

[1] Cette vision (réaliser l'expérience un grand nombre de fois dans les mêmes conditions) sera souvent très utile pour aider à se représenter différentes notions. Par exemple dire que l'évenement "le patient est guéri" a une probabilité de $1/4$ pourra se comprendre plus facilement en se représentant 1000 patients similaires pour lesquels on a en moyenne 250 guéris.

Par exemple, pour un lancer de dè à 6 faces, l'événement aléatoire "le dé est tombé sur une face avec un numéro pair" correspond au sous-ensemble $\{2,4,6\}$ de l'ensemble $\Omega = \{1,2,3,4,5,6\}$.

Pour un résultat d'expérience $\omega \in \Omega$, on dira qu'un événement A est **réalisé** si $\omega \in A$. (par exemple, si l'expérience donne $\omega = 4$, alors l'événement

$$A = \text{“le dé est tombé sur une face avec un numéro pair”} = \{2,4,6\},$$

est réalisé car $4 \in \{2,4,6\}$!)

Définition 1.2 *1. L'évenement $\emptyset$ est appelé **événement impossible**.*

*2. L'événement Ω est appelé **événement certain**.*

*3. Pour un événement $A \subset \Omega$, l'évenement A^C (le complémentaire de l'ensemble) est appelé **événement contraire**.*

*4. Pour deux évenements $A, B \subset \Omega$ l'évenement $A \cap B$ sera appelé **A et B***

*5. Pour deux évenements $A, B \subset \Omega$ l'évenement $A \cup B$ sera appelé **A ou B***

*6. Deux événements $A, B \subset \Omega$ sont dits **incompatibles** si $A \cap B = \emptyset$*

Une fois l'univers Ω défini, il ne reste plus qu'à donner la probabilité de chaque évenement. Mathématiquement, certaines propriété d'une "probabilité" sont assez intuitives : la probailité de l'évenement impossible doit être de 0, la probabilité de l'évenement certain doit être de 1, la probabilité de l'union de deux évenements incompatibles doit être la sommes des probabilités de chacun de ces deux événements. Lorsque Ω est fini, donner la probabilité de chacun des évenements ponctuels élémentaires suffit à obtenir les autres par sommation, mais quelques subtilités existent lorsque $\Omega = \mathbb{R}$. Pour les passer sous silence, on considère une tribu $\mathcal{A}$ de l'ensemble des évenements probabilisables (ayez en tête $\mathcal{A} = \mathcal{P}(\Omega)$).

Définition 1.3 *Une **probabilité** sur un univers Ω, muni d'une tribu $\mathcal{A} \subset \mathcal{P}(\Omega)$ (de l'ensemble des événements auquels on peut associer une probabilité) est une application $\mathbb{P}$ de $\mathcal{A}$ dans $[0,1]$ qui vérifie*

1. $\mathbb{P}(\Omega) = 1$

2. Pour toute suite $(A_n)_{n\geq 1} \in \mathcal{A}^{\mathbb{N}}$ d'évenements 2 à 2 incompatibles,

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

On peut déduire de la définition les propriétés suivantes :

Propriétés 1.4 *1. $\mathbb{P}(\emptyset) = 0$*

2. Pour toute suite $(A_n)_{n\geq 1} \in \mathcal{A}^{\mathbb{N}}$ d'évenements,

$$\mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) \leq \sum_{n=1}^{+\infty} \mathbb{P}(A_n).$$

3. $\forall A \in \mathcal{A}, \mathbb{P}(A^C) = 1 - \mathbb{P}(A)$

4. $\forall A, B \in \mathcal{A}$, si $A \subset B$, alors $\mathbb{P}(A) \subset \mathbb{P}(B)$

5. $\forall A, B \in \mathcal{A}, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

6. *Si Ω est fini ou dénombrable, alors, $\forall A \in \mathcal{A}$,*

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

Remarque : à cause de la dernière propriété, lorsque Ω est fini ou dénombrable, on peut se contenter de donner la probabilité de chacun des événements élémentaires seulement. Par exemple :

1. Pour un lancer d'une pièce équilibrée, on pourra prendre $\Omega = \{Pile, Face\}$; $\mathbb{P}(\{\omega\}) = 1/2$
2. Pour le résutat d'un dé équilibré, on pourra prendre $\Omega = \{1, 2, 3, 4, 5, 6\}$; $\mathbb{P}(\{\omega\}) = 1/6$
3. Pour l'âge d'une personne prise au hasard, on pourra prendre $\Omega = \{0, 1, 2, \cdots, 150\}$, la probabilité pourrait être donnée par une pyramide des âges
4. Pour un lancer de deux pièces équilibrées, on pourra prendre $\Omega = \Big\{(Pile, Pile), (Pile, Face), (Face, Pile), (Face, Face)\Big\} = \{Pile, Face\}^2$; $\mathbb{P}(\{\omega\}) = 1/4$

2 Indépendance et conditionnement d'évènements

Maintenant que l'on dispose de la notion d'expérience aléatoire, d'univers et de probabilité, on peut aborder les notions (fondamentales pour la suite) de probabilité conditionnelles et d'indépendance.

2.1 Conditionnement d'évenements

Définition 2.1 *Soient $A, B \in \mathcal{A}$ deux événements, tel que $\mathbb{P}(B) > 0$. On définit la probabilité conditionnelle de A sachant B par*

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

Exemple 2.1

- *On lance un dé équilibré à 6 faces ($\Omega = \{1, 2, 3, 4, 5, 6\}$, $\forall \omega \in \Omega, \mathbb{P}(\omega) = \frac{1}{6}$), on peut alors calculer la probabilité que "le résultat soit pair" ($A = \{2, 4, 6\}$) sachant qu'"il est strictement supérieur à 3" ($B = \{4, 5, 6\}$) :*

$$\mathbb{P}\Big(A \mid B\Big) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}\Big(\{4, 6\}\Big)}{\mathbb{P}\Big(4, 5, 6\Big)} = \frac{2/6}{3/6} = \frac{2}{3}.$$

- *On imagine qu'un virus circule fortement dans le monde. Heureusement il existe un vaccin. Une personne non vaccinée a aujourd'hui 0.05% de chances d'être hospitalisée pour le virus, c'est 10 fois moins pour une personne vacinnée[2]. Si 80% de la population est vaccinée, quelle est la probabilité qu'une personne (tirée uniformément dans la population) soit vaccinée sachant qu'elle est hospitalisée ? En définissant l'évenement H : "la personne tirée est hospitalisée" et V :*

[2] Le vaccin a donc une efficacité de 90% pour les cas graves.

"la personne tirée est vaccinée" on peut calculer :

$$\begin{aligned}
\mathbb{P}\big(H \mid \bar{V}\big) &= 0.0005\\
\mathbb{P}\big(H \mid V\big) &= 0.0005(1-0.90) = 0.00005\\
\mathbb{P}(V) &= 0.8\\
\mathbb{P}\big(V \mid H\big) &= \frac{\mathbb{P}\big(H \cap V\big)}{\mathbb{P}(H)}\\
&= \frac{\mathbb{P}\big(H \mid V\big)\mathbb{P}(V)}{\mathbb{P}\big(H \mid V\big)\mathbb{P}(V) + \mathbb{P}\big(H \mid \bar{V}\big)\mathbb{P}(\bar{V})}\\
&= \frac{0.00005 * 0.8}{0.00005 * 0.8 + 0.0005 * 0.2}\\
&= \frac{2}{7} \approx 28\%
\end{aligned}$$

- *J'ai deux enfants, dont au moins un garçon, quelle est la probabilité que l'autre soit un garçon ?*

2.2 Indépendance d'évenements

Définition 2.2 *Soient $A, B \in \mathcal{A}$ deux événements, on dit que A et B sont des événements **indépendants** si*

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

Soient $n \geq 1, A_1, \cdots, A_n \in \mathcal{A}$ des évenements. On dit que les $(A_i)_{i=1,\cdots,n}$ sont

- *Indépendants deux à deux si*

$$\forall 1 \leq i, j \leq n, \mathbb{P}\big(A_i \cap A_j\big) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$

- *Indépendants dans leur ensemble si*

$$\forall 2 \leq k \leq n, \forall 1 \leq i_1, \cdots i_k \leq n, = \mathbb{P}\Big(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}\Big) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_n}).$$

Exemple 2.2

1. *On tire un jour au hasard dans l'année passée. On note A l'évenement "il a fait beau ce jour là" et B l'évenement "il s'agit d'un dimanche". On suppose que $\mathbb{P}(A) = 1/3$ et $\mathbb{P}(B) = 1/7$, et que A et B sont indépendants. On a alors $\mathbb{P}(A \cap B) = \frac{1}{21}$, ce qui est intuitif si on se représente Ω à partir du calendrier.*
2. *On considère un gène formé de 2 allèles A et a. Calculer la probabilité, pour chaque génotype des parents, de chaque génotype des enfants.*
3. *On considère deux gènes possédant chacun deux allèles A, a et B, b, l'un dominant et l'autre récessif, associés à des phénotypes $\phi_{AB}, \phi_{Ab}, \phi_{aB}, \phi_{ab}$. En supposant les gènes indépendants car sur deux chromosomes différents, donner la probabilité pour chaque phénotype d'un enfant issu de deux parents hétérozygotes pour chacun des deux gènes. On peut le faire facilement en considérant*

$$\begin{aligned}
\mathbb{P}\big(AA\big) &= \mathbb{P}\big(BB\big) = 1/4\\
\mathbb{P}\big(Aa\big) &= \mathbb{P}\big(Bb\big) = 1/2\\
\mathbb{P}\big(aa\big) &= \mathbb{P}\big(bb\big) = 1/4\\
\mathbb{P}\big(\phi_A\big) &= \mathbb{P}\big(AA\big) + \mathbb{P}\big(Aa\big) = 3/4 = \mathbb{P}\big(\phi_B\big)\\
\mathbb{P}\big(\phi_a\big) &= \mathbb{P}\big(aa\big) = 1/4 = \mathbb{P}\big(\phi_b\big),
\end{aligned}$$

puis en utilisant l'indépendance des caractères :

$$\begin{aligned}
\mathbb{P}(\phi_{AB}) &= \mathbb{P}(\phi_A)\mathbb{P}(\phi_B) \text{ \textit{par indépendance}} \\
&= 3/4 * 3/4 = \frac{9}{16} \\
\mathbb{P}(\phi_{Ab}) &= \mathbb{P}(\phi_A)\mathbb{P}(\phi_b) \text{ \textit{par indépendance}} \\
&= 3/4 * 1/4 = \frac{3}{16} \\
\mathbb{P}(\phi_{aB}) &= \mathbb{P}(\phi_a)\mathbb{P}(\phi_B) \text{ \textit{par indépendance}} \\
&= 1/4 * 3/4 = \frac{3}{16} \\
\mathbb{P}(\phi_{ab}) &= \mathbb{P}(\phi_a)\mathbb{P}(\phi_b) \text{ \textit{par indépendance}} \\
&= 1/4 * 1/4 = \frac{1}{16}
\end{aligned}$$

3 Variables aléatoires réelles

3.1 Variables aléatoires discrètes : introduction

3.1.1 Motivations

On reprend l'exemple d'un lancer de 2 dés indépendants. Dans ce cas, on peut vouloir s'intéresser à 2 observations différents, par exemple, la somme et le produit des résultats. Pour éviter de devoir changer l'ensemble Ω à chaque fois que l'on veut étudier une variable différente, on peut avoir envie d'introduire une notion pour chaque variable construite à partir d'une *realisation* ω.

C'est le rôle de la notion de **variable aleatoire**. Lorsque la fonction considérée (par exemple S qui prend 2 résultats de dés et renvoie la somme ou P qui renvoie le produit) ne peut prendre qu'un nombre fini de valeurs (ou entier)[3], on parlera de variable aleatoire discrete.

Une variable aléatoire sera donc définie comme une fonction de Ω dans $\mathbb{R}$, par exemple :

$$\begin{aligned}
&\Omega = \{1, \cdots, 6\}^2 \\
X :\ &\Omega \to \mathbb{R} \\
&(\omega_1, \omega_2) \mapsto \omega_1 + \omega_2
\end{aligned}$$

Pour décrire le comportement, **la loi** d'une telle variable, il suffira alors de connaître :

- L'ensemble des valeurs possibles que peut prendre cette variable (ex : $E = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ pour S)
- La probablité pour chaque résultat (ex : $\mathbb{P}(S = 2) = \frac{1}{36}, \mathbb{P}(S = 3) = \frac{2}{36} \cdots, \mathbb{P}(S = 7) = \frac{6}{36}, \mathbb{P}(S = 8) = \frac{5}{36},$
$\cdots, \mathbb{P}(S = 12) = \frac{1}{36}$)

3.1.2 Variables aléatoires discrètes usuelles

On peut enfin introduire les variables aléatoires discrètes usuelles, souvent rencontrées dans des problèmes réels :

[3] la bonne notion sera ici au plus dénombrable

1. Loi uniforme discrète (ex : lancer de dé). X est une variable aléatoire de **loi uniforme discrète** sur $\{1, \cdots, n\}$ (noté $X \sim \mathcal{U}\Big(\{1, \cdots, n\}\Big)$ si X est à valeurs dans $\{1, \cdots, n\}$, et si

$$\forall k \in \{1, \cdots, n\}, \mathbb{P}\Big(X = k\Big) = \frac{1}{n}.$$

2. Loi de Bernoulli (ex : lancer de pièce truqué). Cette loi est utilisée dès que l'on considère une variable ne pouvant prendre que 2 valeurs (ex: Pile ou Face, présence ou absence de maladie, caractère fumeur ou non.... En général, on considère que ces valeurs seront 0 et 1, et on dira que X est une variable aléatoire de **loi de Bernoulli** de paramètre $p \in [0, 1]$ (noté $X \sim \mathcal{B}(p)$ ou $X \sim \mathcal{B}er(p)$)si X peut prendre les valeurs 0 et 1, et que

$$\mathbb{P}\Big(X = 1\Big) = p \ ; \quad \mathbb{P}\Big(X = 0\Big) = 1 - p.$$

   Remarque : Pour tout événement A, la variable aléatoire $X = \mathbb{1}_A$ (i.e. $X(\omega) = \mathbb{1}_{\omega \in A}$) ne prend que les valeurs 0 et 1 par définition de l'indicatrice. C'est donc une variable aléatoire de loi de Bernoulli de paramètre $\mathbb{P}(A)$. Cette propriété est à retenir.

3. Loi Binomiale. Cette loi est utilisée dès que l'on modélise le nombre de "succès" à plusieurs "épreuves de Bernoulli indépendantes", par exemple lorsqu'on compte le nombre d'individus présentant un certain caractère dans une population homogène (ex : nombre de cellules qui se sont divisées sur une certaine unité de temps, dans une population fixée et homogène). X est une variable aléatoire de **loi Binomiale** de paramètres N et p (noté $X \sim \mathcal{B}(N, p)$ ou $X \sim \mathcal{B}in(N, p)$) si X peut prendre les valeurs $\{0, \cdots, N\}$, et

$$\forall k \in \{0, \cdots, N\}, \mathbb{P}(X = k) = \binom{N}{k} p^k (1 - p)^{N-k}.$$

   Remarque : Une variable aléatoire de Bernoulli compte le nombre de succès à une seule épreuve de Bernoulli, donc une Bernoulli de paramètre p correspond à une Binomiale de paramètres 1 et p. Cela sera parfois utile de réécrire les probabilités pour une Bernoulli sous la forme

$$\forall k \in \{0, 1\}, \mathbb{P}(X = k) = p^k (1 - p)^{1-k}.$$

4. Loi de Poisson (ex : nombre de cellules affichant un caractère rare dans une population homogène). Cette loi apparaît naturellement comme limite de la loi Binomiale, quand le nombre d'épreuves tends vers l'infini, tout en conservaznt le nombre moyen de succès.Cela permet de modéliser le nombre d'évenements ponctuels arrivant durant une certaine durée, ou dans une certaine zone spatiale, lorsqu'il y a stationnarité et indépendance des arrivées. X est une variable aléatoire de **loi de Poisson** de paramètre λ (noté $X \sim \mathcal{P}(\lambda)$) si elle est à valeurs dans $\mathbb{N}$ et vérifie

$$\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

5. Loi Géométrique (ex : Nombre de tentative avant de réussir une expérience, réalisée dans les mêmes conditions). Cette loi représente l'instant du premier succès à une suite d'épreuves de Bernoulli indépendantes de paramètre p. X est une variable aléatoire de **loi géométrique** de paramètre p (noté $X \sim \mathcal{G}(p)$) si elle est à valeurs dans $\mathbb{N}^\star$[4], et vérifie

$$\forall k \in \mathbb{N}^\star, \mathbb{P}(X = k) = p(1 - p)^{k-1}.$$

[4]Certains peuvent la définir dans $\mathbb{N}$, en considérant $X - 1$

3.2 Variables aléatoires continues : introduction

3.2.1 Motivations

Le formalisme précédent ne permet pas de modéliser le choix d'un nombre au hasard dans $[0, 1]$. En effet, on ne peut pas donner seulement la probabilité de chaque résultat, puisque n'importe quel nombre a une probablité nulle de sortir.

Il va falloir trouver une autre voie. Pour comprendre cette idée, on peut partir de l'intuition qu'un tel nombre "au hasard" (uniforme) U entre 0 et 1 doit vérifier

$$\forall a < b \in [0,1],\ \mathbb{P}\Big(U \in [a,b]) = b - a.$$

ou de façon équivalente, que

$$\forall t \in [0,1], \mathbb{P}\Big(U \leq t\Big) = t.$$

Connaitre ceci suffit à avoir une idée de comment se comporte la variable U. Pour décrire **la loi** d'une variable X quelconque, il suffit en fait de donner la fonction $F_X : t \mapsto \mathbb{P}(F \leq t)$.

Mais un cas particulier est vraiment intéressant, c'est quand F_X est dérivable et de dérivée continue, auquel cas, en posant $f_X = F'_X$, on peut écrire :

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

Dans ce cas, plus $f_X(x)$ est élevé, plus la variable risque de tomber 'proche' de x. On dira donc qu'une variable aléatoire X est à densité, s'il existe une fonction f_X telle que

$$\forall a \leq b \in \mathbb{R}, \mathbb{P}\Big(X \in [a,b]\Big) = \int_a^b f_X(x)\mathrm{d}x.$$

3.2.2 Variables aléatoires à densité usuelles

On peut aussi introduire les variables à densité usuelles :

6 Loi Uniforme Continue (ex: nombre "au hasard" entre 0 et 1). X est une variable aléatoire de loi Uniforme sur le segment $[a, b], a < b$ (noté $X \sim \mathcal{U}([a, b])$) si elle est à valeurs dans $[a, b]$, et a pour densité la fonction définie par

$$f_X(x) = \frac{1}{b-a}\mathbb{1}_{x\in[a,b]}.$$

7 Loi Exponentielle (ex : temps entre 2 burst d'un gène). Cette loi est l'analogue continue de la loi géométrique. Elle est utilisée pour modéliser des temps inter-arrivée d'évenements ponctuels arrivant, lorsqu'il y a stationnarité et indépendance des arrivées. X est une variable aléatoire de **loi exponentielle** de paramètre $\lambda > 0$ (noté $X \sim \mathcal{E}(\lambda)$) si elle est à valeurs dans $\mathbb{R}^+$, et a pour densité la fonction définie par

$$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{x\in\mathbb{R}^+}.$$

8 Loi Normale (ou Gaussienne) (ex : pression atmosphérique prise en une position et une date aléatoire). Elle peut modéliser de nombreuses variables continues "unimodales" (avec un seul pic), et elle apparaît par exemple comme limite de la loi Binomiale lorsque l'on réalise un grand nombre d'épreuves de Bernoulli. Cette loi apparaît en particulier dans le Théorème Central Limite (que l'on abordera à la fin du cours), ce qui explique sa capacité pour modéliser de nombreux phénomènes réels. En biologie, de nombreuses quantités (titrages...) sont souvent modélisées comme des variables **"log-normales"**, ce qui signifie que le logarithme de ces quantités suit une loi normale. X

est de **loi Normale** de paramètres m et σ^2 (noté $X \sim \mathcal{N}(m, \sigma^2)$) si elle est à valeurs dans $\mathbb{R}$, et a pour densité la fonction définie par

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

3.3 Cas général

3.3.1 Définitions

La notion de **variable aléatoire** est certainement la notion la plus importante de ce chapitre, et souvent la plus difficile à comprendre. La formalisation mathématique d'une variable aléatoire prend en compte, comme indiqué dans l'introduction, que plusieurs *expériences aléatoires* différentes peuvent donner la même *variable aléatoire*. Par exemple, pour jouer à Pile ou Face équilibré, on peut

- Tirer une pièce équilibré
- Lancer un dé à 6 faces équilibré et regarder si le résultat est pair
- Tirer une carte dans un jeu uniformément, et regarder si la couleur est rouge
- Lancer 3 fois une pièce équilibrée et regarder les résultat du second lancer.
- Tirer un nombre uniforme entre 0 et 1, et le comparer à 0.5.

Pour chacun de ces cas, l'espace de probablité serait différent :

- Tirer une pièce équilibré : $\Omega = \{Pile, Face\}$.
- Lancer un dé à 6 faces équilibré : $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Tirer une carte dans un jeu uniformément : $\Omega = \{A\diamondsuit, A\clubsuit, A\heartsuit, A\spadesuit, 2\diamondsuit, 2\clubsuit, 2\heartsuit, 2\spadesuit, \cdots, K\diamondsuit, K\clubsuit, K\heartsuit, K\spadesuit\}$.
- Lancer 3 fois une pièce équilibrée : $\Omega = \{P, F\}^3 = \{(P, P, P), (P, P, F), \cdots, (F, F, F)\}$.
- Tirer un nombre uniforme entre 0 et 1 : $\Omega = [0, 1]$

Pour pouvoir introduire la notion de variable aléatoire sur chacun de ces espaces, on se donne la définition suivante :

Définition 3.1 *Une variable aléatoire réelle X sur un univers Ω est une application de Ω dans $\mathbb{R}$:*

$$\begin{aligned} X &: \Omega \to \mathbb{R} \\ \omega \in \Omega &\mapsto X(\omega) \in \mathbb{R} \end{aligned}$$

Exemple 3.1 *On reprend le cadre précédent pour définir la même variable aléatoire (Pile ou Face équilibré) qui vaut 1 si la pièce tombe sur Pile à partir d'espaces de probabilités différents :*

- *Tirer une pièce équilibré*

$$\begin{aligned} \Omega &= \{Pile, Face\} \\ \mathbb{P}(\{\omega\}) &= \frac{1}{2} \\ X &: \omega \mapsto \mathbb{1}_{\omega=Pile} \end{aligned}$$

- *Lancer un dé à 6 faces équilibré et regarder si le résultat est pair*

$$\begin{aligned} \Omega &= \{1,2,3,4,5,6\} \\ \mathbb{P}(\{\omega\}) &= \frac{1}{6} \\ X &: \omega \mapsto \mathbb{1}_{\omega \in \{2,4,6\}} \end{aligned}$$

- *Tirer une carte dans un jeu uniformément, et regarder si la couleur est rouge*

$$\begin{aligned} \Omega &= \{A\diamondsuit, A\clubsuit, A\heartsuit, A\spadesuit, 2\diamondsuit, 2\clubsuit, 2\heartsuit, 2\spadesuit, \cdots, K\diamondsuit, K\clubsuit, K\heartsuit, K\spadesuit\} \\ \mathbb{P}(\{\omega\}) &= \frac{1}{52} \\ X &: \omega \mapsto \mathbb{1}_{\omega \in \{A\diamondsuit, A\heartsuit, 2\diamondsuit, 2\heartsuit, \cdots, K\diamondsuit, K\heartsuit\}} \end{aligned}$$

- *Lancer* 3 *fois une pièce équilibrée et regarder les résultat du second lancer.*

$$\begin{aligned} \Omega &= \{Pile, Face\}^3 \\ \mathbb{P}(\{(\omega_1,\omega_2,\omega_3)\}) &= \frac{1}{8} \\ X &: (\omega_1,\omega_2,\omega_3) \mapsto \mathbb{1}_{\omega_2 = Pile} \end{aligned}$$

- *Tirer un nombre uniforme entre* 0 *et* 1*, et le comparer à* 0.5*.*

$$\begin{aligned} \Omega &= [0,1]; \\ \mathbb{P}([a,b]) &= b-a \\ X &: \omega \mapsto \mathbb{1}_{\omega < 1/2} \end{aligned}$$

Une variable aléatoire X définie une **loi de probabilité** $\mathbb{P}_X$ sur $(\mathbb{R})$ donnée par[5] :

$$\forall A \in \mathcal{P}(\mathbb{R}), \mathbb{P}_X(A) = \mathbb{P}\Bigg(\big\{\omega \in \Omega, X(\omega) \in A\big\}\Bigg).$$

Remarque : à partir de maintenant, on notera tout simplement

$$\mathbb{P}(X \in A) := \mathbb{P}_X(A) \text{ et } \mathbb{P}(X = x) := \mathbb{P}_X\big(\{x\}\big).$$

La loi d'une variable aléatoire réelle est entièrement caractérisée par la fonction de répartition définie au dessous :

Définition 3.2 *La fonction de répartition d'une variable aléatoire réelle X est la fonction*

$$\begin{aligned} F_X &: \mathbb{R} \to [0,1] \\ t &\mapsto \mathbb{P}(X \le t) \end{aligned}$$

Propriétés 3.3 *La fonction de répartition vérifie :*

1. *F_X est une fonction croissante*
2. *F_X tend vers 0 en $-\infty$ et vers 1 en $+\infty$*

[5]En réalité, on ne devrait pas travailler avec l'ensemble des parties de $\mathbb{R}$, mais c'est bien au delà du programme

3. F_X est continue à droite.

4. $\forall a < b \ \in \mathbb{R}, \mathbb{P}\big(a < X \leq b\big) = F_X(b) - F_X(a)$.

Dans ce cours, on s'intéressera à deux types de variables aléatoires :

1. Si X peut prendre un nombre fini ou dénombrable de valeurs $x_k, k \geq 0$, alors F_X est constante par morceaux, on dira que X est **une variable discrete**. Dans ce cas, il suffit de connaître $\mathbb{P}(X = x_k), \forall k \geq 0$ pour connaître la loi de X.

2. Si F_X est dérivable et de dérivée continue, alors, en posant $f_X = F_X$, on a

   - $\forall x \in \mathbb{R}, \mathbb{P}(X = x) = 0$
   - $\forall a < b \in \mathbb{R}, \mathbb{P}\big(a \leq X < b\big) = \mathbb{P}\big(a \leq X \leq b\big) = \mathbb{P}\big(a \leq X \leq b\big) = \mathbb{P}\big(a < X \leq b\big) = \int_a^b f_X(x)\mathrm{d}x$.

   Dans ce cas là, on parle de **variable aléatoire continue** ou **variable aléatoire à densité**

Cela motive la définition suivante :

Définition 3.4 *Soit X une variable aléatoire réelle. S'il existe une fonction f_X vérifiant*

$$\forall a \leq b \in \mathbb{R}, \mathbb{P}\big(a \leq X < b\big) = \int_a^b f_X(x)\mathrm{d}x,$$

alors f_X est appelée ***densité de probabilité*** *de la variable aléatoire X, et F_X est dérivable, de dérivée f_X.*

Remarque : La probabilité que X tombe dans une région donnée correspond alors à l'aire sous la courbe de f_X dans cette région. En particulier, l'aire totale sous la courbe densité vaut 1 :

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

Cette propriété est à retenir !

Définition 3.5 *Soient X, Y deux variables aléatoires réelles. On dit que X et Y ont même loi si elles ont même fonction de répartition : $F_X = F_Y$. Pour des variables discrètes, il suffit que X et Y aient le même ensemble de valeurs $a_k, k \geq 0$ possibles, et que $\mathbb{P}(X = a_k) = \mathbb{P}(Y = a_k), \forall k \geq 0$. Pour des variables à densité, il suffit que $f_X = f_Y$. Si $X_1, \cdots, X_n$ ont même loi, on dira qu'elles sont* ***identiquement distribueées****.*

Comme pour les évenements, des variables aléatoires peuvent être indépendantes ou non (par exemple la taille et le poids d'un individu tiré au hasard ne sont pas indépendantes). Pour définir cette notion, nous avons besoin de parler de la loi jointe de plusieurs variables aléatoires (on parle alors de **vecteur aléatoire**, ou **couple de variable aléatoire** lorsqu'il y en a 2).

Définition 3.6 *Soient $(X_1, \cdots, X_n)$ un vecteur aléatoire réel. La loi jointe de $X_1, \cdots, X_n$ est caractérisée par la fonction*

$$F_{X_1,\cdots,X_n} : (t_1, \cdots, t_n) \mapsto \mathbb{P}\Big(X_1 \leq t_1, \cdots, X_n \leq t_n\Big).$$

Définition 3.7 *Soient $(X_1, \cdots, X_n)$ un vecteur aléatoire réel. On dit que $X_1, \cdots, X_n$ sont indépendantes si*

$$\forall t_1, \cdots, t_n \in \mathbb{R}, \mathbb{P}\Big(X_1 \leq t_1, \cdots, X_n \leq t_n\Big) = \mathbb{P}\Big(X_1 \leq t_1\Big) \cdots \mathbb{P}\Big(X_n \leq t_n\Big).$$

i.e. si

$$\forall t_1, \cdots, t_n \in \mathbb{R}, F_{X_1,\cdots,X_n}(t_1, \cdots, t_n) = F_{X_1}(t_1) \cdots F_{X_n}(t_n).$$

Remarque : Lorsque les X_i sont discrètes, par exemple si X_i peut prendre les valeurs $a_k^{(i)}, k \geq 0$, il suffit de vérifier

$$\forall k_1, \cdots, k_n \geq 0, \mathbb{P}\Big(X_1 = a_{k_1}^{(1)}, \cdots, X_n = a_{k_n}^{(n)}\Big) = \mathbb{P}\Big(X_1 = a_{k_1}^{(1)}\Big) \cdots \mathbb{P}\Big(X_n = a_{k_n}^{(n)}\Big).$$

Dans toute la suite, on aura souvent besoin de parler de suites de variables aléatoires réelles **indépendantes et identiquement distribuées**. On notera v.a.r. i.i.d..

4 Espérance, variance...

La notion d'**d'espérance mathématique** d'une variable aléatoire correspond à la notion intuitive de "moyenne sur le hasard". Attention, il faut dès maintenant être très vigilant·e·s avec une confusion classique entre

- Moyenne d'une échantillon : on fait la moyenne **après** une expérience sur les résultats de l'expérience. Par exemple, je lance 4 fois un dé à 6 faces, et j'obtiens $3, 4, 2, 3$, alors la moyenne de ces résultats est de $\frac{3+4+1+3}{4} = 2.75$
- Espérance d'une variable aléatoire : on peut la calculer **avant** une expérience. Par exemple, pour un dé à 6 faces, l'espérance correpond au résultat moyen que l'on peut espérer avec ce dé, en prenant en compte la probabilité d'obtenir chaque face :

$$E = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = 3.5$$

Notons aussi que si le dé est truqué, alors les probabilités de chaque face changent, et l'espérance change aussi.

Définition 4.1 *Soit X une v.a.r. et h une application de $\mathbb{R}$ dans $\mathbb{R}$. Par définition d'une variable aléatoire, $h(X)$ est aussi une v.a.r.. On définit l'**espérance** de $h(X)$, et on note $\mathbb{E}[h(X)]$, la quantité :*

- *Si X est discrète, et peut prendre les valeurs $a_k, k \geq 0$,*

$$\mathbb{E}[h(X)] := \sum_{k\geq 0} h(a_k)\mathbb{P}(X = a_k),$$

  (à condition que cette série soit convergente)

- *Si X est continue, de densité f_X :*

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

  (à condition que cette intégrale soit convergente)

Propriétés 4.2 *L'espérance mathématique vérifie les propriétés suivantes :*

1. *$\forall a \in \mathbb{R}, \mathbb{E}[a] = a$.*
2. *Linéarité : $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.*
3. *Croissance : Si $X \geq Y$ p.s., $\mathbb{E}[X] \geq \mathbb{E}[Y]$.*
4. *$\forall A \in \mathcal{A}, \mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$*

L'espérance donne une indication sur la "localisation" de la variable : autour de quelle valeur elle tombe (donc autour de quelle valeur est/sont située(s) la ou les "bosses" sur la densité pour une v.a. continue.). Par contre, cette quantité ne donne aucune indication sur la dispersion de la variable (tombe elle "souvent" proche de l'espérance ¿') ni donc sur "l'étalement" de la densité pour une v.a continue.

Pour contrôler l'écart à l'espérance, la quantité la plus usuelle considère la "moyenne (sur le hasard)" du carré de l'écart à la moyenne. Plus précisemment, $(X - \mathbb{E}[X])^2$ donne une bonne indication de si la variable X est tombée proche ou loin de son espérance, mais cette quantité est aléatoire (car X l'est). On considère donc l'esprance de cette quantité, et cela définit la variance.

Définition 4.3 *La* ***variance*** *d'une v.a.r. X est définie par*

$$Var(X) := \mathbb{E}\left[\Big(X - \mathbb{E}[X]\Big)^2\right].$$

Propriétés 4.4 *La variance vérifie :*

1. $\operatorname{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

2. $\forall a \in \mathbb{R}, \operatorname{Var}(a) = 0$

3. $\forall a, b \in \mathbb{R}, \operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X)$

Définition 4.5 *On appelle* ***écart-type*** *de la v.a.r. X la racine carré de la variance :*

$$\sigma_X = \sqrt{\operatorname{Var}(X)}$$

Remarque : Si X est une v.a.r telle que $\mathbb{E}[X] = m$ et $\operatorname{Var}(X) = \sigma^2$, alors la variable $Y = \frac{X-m}{\sigma}$ est d'espérance nulle, et de variance 1. On dit que Y est centrée (d'espérance nulle) et réduite (de variance 1).

Définition 4.6 *Soit (X, Y) un couple de v.a.r.*

$$\operatorname{Cov}(X, Y) := \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)\big(Y - \mathbb{E}[Y]\big)\Big].$$

À comparer avec la définition de la variance...

Proposition 4.7 *Si X, Y sont indépendantes, alors*

$$\operatorname{Cov}(X, Y) = 0.$$

La réciproque n'est pas vraie !!!!

Propriétés 4.8 *Soient X, Y deux variables aléatoires*[6]

1. $\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

2. La fonction $(X, Y) \mapsto \operatorname{Cov}(X, Y)$ est une forme bilinéaire symmétrique :

- $\forall a_1, a_2 \in \mathbb{R}, \operatorname{Cov}(a_1X_1 + a_2X_2, Y) = a_1 \operatorname{Cov}(X_1, Y) + a_2 \operatorname{Cov}(X_2, Y)$
- $\operatorname{Cov}(X, Y) = \operatorname{Cov}(Y, X)$

3. Conséquence : $\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X, Y)$.

[6] définies sur le même espace de probabilité, on risque de ne plus le préciser, mais c'est sous-entendu

4. Lorsque X, Y sont indépendantes, l'égalité précédente devient : $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Proposition 4.9 (Stabilité des lois normales) *Si X et Y sont deux variables aléatoires **indépendantes** de loi normales, alors toute combinaison affine de X, Y est aussi de loi normale :*

$$\forall a, b, c \in \mathbb{R}, aX + bY + c \text{ est de loi normale.}$$

Attention, cela ne suffit pas à caractériser la loi, puisqu'il reste à calculer espérance et variance. En particulier, en posant $b = 0$, on voit que toute transformation affine d'une loi normale est aussi de loi normale[7]

Quelques calculs d'espérances et de variances (lois classiques):
Savoir faire ces calculs est important, car cela donne une bonne idée des outils de calculs qui peuvent être utilisés.

- Lois discrètes :
  - Bernoulli $\mathcal{B}(p)$

    $$\mathbb{E}[X] = (1-p).0 + p.1 = p.$$

    et

    $$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1-p).$$

  - Binomiale $\mathcal{B}in(n,p)$
    On utilise que $X \sim \sum_{i=1}^n Y_i$ où $Y_i \overset{i.i.d.}{\sim} \mathcal{B}(p)$, et on obtient

    $$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n Y_i\right] = \sum_{i=1}^n \mathbb{E}[Y_i] = np.$$

    et, en utilisant que la variance d'une somme de v.a. indépendante est la somme des variances, on obtient

    $$\mathrm{Var}(X) = np(1-p).$$

  - Géométrique $\mathcal{G}(p)$
    On va utiliser $\forall |x| < 1, \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$, obtenue en dérivant terme à terme l'égalité $\sum x^k = \frac{1}{1-x}$. On obtient

    $$\mathbb{E}[X] = \sum_{k\geq 1} kp(1-p)^{k-1} = \frac{p}{\big(1-(1-p)\big)^2} = \frac{1}{p}.$$

    Pour la variance, on redérive terme à terme, et on obtient $\forall |x| < 1, \sum_{k=1}^{+\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3}$. Ce qui donne

    $$\mathrm{Var}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X^2] = \frac{2p(1-p)}{\big(1-(1-p)\big)^3} + \frac{1}{p} - \frac{1}{p^2} = \frac{2-2p+p-1}{p^2} = \frac{1-p}{p^2}.$$

[7] en réalité cette propriété cache des propriétés beaucoup plus générale, de ce qu'on appelle les "vecteurs gaussiens" définis comme des familles de loi normales (non nécessairement indépendantes) telle que toute combinaison affine reste de loi normale. Mais c'est en dehors du programme de ce cours.

– Uniforme $\mathcal{U}([a,b])$

$$\mathbb{E}[X] = \int_a^b \frac{\mathrm{d}x}{b-a} = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}.$$

et

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \frac{(a+b)^2}{4} = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{4(b^2+ab+a^2)-3(a^2+b^2+2ab)}{12} = \frac{(b-a)^2}{12}.$$

5 Vers l'inférence statistique : théorèmes limites, et intervalles de confiance

Dans cette partie, on esquisse très rapidement le raisonnement nécessaire pour aborder "l'inférence statistique".

5.1 Théorèmes limites

On se basera sur deux théorèmes très utiles :

Théorème 5.1 (Loi des grands nombres) *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables indépendantes et de même loi (i.i.d.) telles que* $\mathbb{E}[|X_n|] < +\infty$. *Alors, avec probabilité* 1,

$$\frac{1}{n}\sum_{i=1}^n X_i \to \mathbb{E}[X_1].$$

Ce théorème correspond à l'intuition que l'on se fait d'une probablité : si on fait l'expérience un grand nombre de fois, alors la moyenne des résultats s'approche de l'espérance de la variable aléatoire. Un cas particulier intéressant est celui d'une variable aléatoire de Bernoulli de paramètre p pour laquelle $\mathbb{E}[X] = p$. Ce résultat dit alors que si l'on regarde la proportion de réussites $\frac{1}{n}\sum_{i=1}^n X_i$ sur un grand nombre d'expérience, alors cette proportion, s'approche de la probabilité de réussite. En lançant une pièce un grand nombre de fois, la proportion de pile s'approchera de la probabilité que la pièce tombe sur pile.

En réalité, cela correspond à l'intuition que l'on se fait de la notion de "probabilité" d'un événement : la proportion du temps où l'évenement se produit, si on réalise l'expérience un grand nombre de fois.

Il reste une question : **à quel point on s'éloigne de cette espérance ?**. C'est l'objet du théorème central limite énoncé ci dessous.

Théorème 5.2 (Théorème Central Limite) *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables indépendantes et de même loi (i.i.d.) telles que* $\mathrm{Var}(X_n) < +\infty$. *Alors, pour tout* $a < b \in \mathbb{R}$,

$$\mathbb{P}\left(\sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}[X_1]}{\sqrt{\mathrm{Var}(X_1)}} \in [a,b]\right) \to \mathbb{P}\Big(Z \in [a,b]\Big),$$

où $Z \sim \mathcal{N}(0,1)$ *désigne une loi normale "centrée réduite" (d'espérance* 0 *et de variance* 1*).*

Ce résultat affirme essentiellement que l'écart à la moyenne se comporte (lorsqu'on a assez d'observations) comme une loi normale.

5.2 Un exemple de raisonnement statistique

Supposons que l'on a administré un médicament à 1000 personnes, en prenant soin de randomiser (voir Exercice 2 "Avancé"). On suppose qu'il existe déjà un médicament pour cette maladie pour lequel l'efficacité est supposée connue et égale à 60% (60% des patients guérissent avec ce médicement). On veut savoir à partir de quel nombre de guérisons sur notre échantillon on peut être confiant de faire mieux que le médicament qui existe déjà. Pour cela, il suffit de modéliser cette expérience comme la réalisation de variables aléatoires $X_1, \cdots, X_{1000}$ i.i.d. Bernoulli de paramètre p (inconnu), qui vallent 1 si le patient correspondant a guéri. Notre objectif est de savoir si p est supérieur) 0.6.

On note $n = 1000$, et $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ la proportions de guéris sur notre échantillon.

La loi des grands nombres affirme que $\bar{X}$ est proche de p, on s'attend donc à ce qu'une "grande" différence entre $\bar{X}$ et 0.6 soit révélateur d'une différence entre p et 0.6. Mais il faut quantifier tout cela pour savoir à partir de quel moment on peut être sereins.

Comme n est assez grands (tout est relatif, mais le théorème central limite donne une très bonne approximation à partir de quelques centaines d'individus), on peut vouloir utiliser ce théoreme. On rappelle, que si $X_1 \sim Ber(p)$, alors $\mathbb{E}[X_1] = p$ et $\mathrm{Var}(X_1) = p(1-p)$. On va alors définir

$$Z = \sqrt{n}\frac{\bar{X} - \mathbb{E}[X_1]}{\sqrt{\mathrm{Var}(X_1)}} = \sqrt{n}\frac{\bar{X} - p}{\sqrt{p(1-p)}}$$

Le théorème central limite stipule que Z est "proche" d'une loi normale $\mathcal{N}(0,1)$, on peut donc approximer $\mathbb{P}(Z \in [a,b])$ par $\mathbb{P}(\mathcal{N}(0,1) \in [a,b])$. On va alors discuter de 2 approches :

Approche par intervalles de confiance

Si on se fixe un seuil de risque α, par exemple $\alpha = 5\%$, on peut calculer numériquement (avec un ordinateur), à partir de l'expression de la densité de la $\mathcal{N}(0,1)$, un nombre q_α tel que $\mathbb{P}(-q_\alpha < \mathcal{N}(0,1) < q_\alpha) = 1 - \alpha$. Par exemple, pour $\alpha = 5\%$, on a $q_\alpha = 1.96$.

On en déduit qu'avec une probabilité d'*environ* $1-\alpha$ (par exemple 95%) on a

$$-q_\alpha < Z < q_\alpha$$

On peut alors modifier cet évenement pour encadrer p inconnu :

$$\begin{aligned}
-q_\alpha < Z < q_\alpha &\Leftrightarrow -q_\alpha < \sqrt{n}\frac{\bar{X} - p}{\sqrt{p(1-p)}} < q_\alpha \\
&\Leftrightarrow -\sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}} < \bar{X} - p < \sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}} \\
&\Leftrightarrow \sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}} > p - \bar{X} > -\sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}} \\
&\Leftrightarrow \bar{X} + \sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}} > p > \bar{X} - \sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}}
\end{aligned}$$

On peut donc affirmer (il faudrait clarifier le $\approx$, mais toute la logique est là) que

$$\mathbb{P}\left(p \in \left[\bar{X} - \sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}}, \bar{X} + \sqrt{p(1-p)}\frac{q_\alpha}{\sqrt{n}}\right]\right) \approx 1 - \alpha.$$

Le problème c'est que l'intervalle donné ici dépend de p, que l'on connaît pas, mais on peut utiliser le fait que $\bar{X}$ est "proche" de p, ce qui donne

$$\mathbb{P}\left(p \in \left[\bar{X} - \sqrt{\bar{X}(1-\bar{X})}\frac{q_\alpha}{\sqrt{n}}, \bar{X} + \sqrt{\bar{X}(1-\bar{X})}\frac{q_\alpha}{\sqrt{n}}\right]\right) \approx 1 - \alpha.$$

Cet intervalle $\left[\bar{X} - \sqrt{\bar{X}(1-\bar{X})}\frac{q_\alpha}{\sqrt{n}}, \bar{X} + \sqrt{\bar{X}(1-\bar{X})}\frac{q_\alpha}{\sqrt{n}}\right]$ est appellé **Intervalle de confiance** au niveau de risque α. (ou niveau de confiance $1-\alpha$).

Une fois l'expérience effectuée, on peut faire l'application numérique. Par exemple, si 650 patient·e·s ont guéri dans notre échantillon, et pour un risque de $\alpha = 5\%$, on aura alors un intervalle de confiance donné par :

$$ic = [\frac{650}{100} - \sqrt{\frac{650}{100}(1-\frac{650}{100})}\frac{1.96}{\sqrt{1000}}, \frac{650}{100} - \sqrt{\frac{650}{100}(1-\frac{650}{100})}\frac{1.96}{\sqrt{1000}}] \approx [0.62, 0.68].$$

Remarque : Attention, une fois cet intervalle calculé, on NE peut PAS affirmer "p a 95% de chances d'être entre 0.62 et 0.68, vu que p N'est PAS aléatoire. C'est l'intervalle lui même qui est aléatoire, et va changer pour chaque expérience (alors que p ne changera pas), et 95% **des expériences** donneront un intervalle de confiance qui contient la vraie valeur.

Enfin, si l'on veut répondre à la question, on pourrait dire que comme 0.6 ne tombe pas dans cet intervalle, on peut conclure de façon statistiquement significative, avec un risque de 5%, que le médicament est meilleur que celui qui est déjà sur le marché. *Dans cet exemple, ce raisonnement a une petite faille, car une partie 5% de risque est utilisée pour prévoir les cas ou notre médicament est pire. Pour faire les choses correctement, il faudrait un intervalle de la forme $[a, 1]$, pour lequel on ne s'occupe que de la valeur minimale.*

Approche par tests d'hypothèses

Le raisonnement est un peu différent dans ce cadre. On va faire une hypothèse par défaut appellée **hypothèse nulle** notée H_0, qui correspond à celle que l'on veut rejeter. Ici, on va supposer que notre médicament n'est pas meilleur que l'ancien : $H_0 : p \leq 0.6$. On appelle **hypothèse alternative** (notée H_1) celle que l'on veut montrer, notre médicament est meilleur que l'ancien : $H_1 : p > 0.6$.

Pour faire les calculs, on va supposer H_0 vraie, et considérer le cas où l'on a le plus de risque de conclure à l'efficacité du médicament (H_1) à tord. On suppose donc à partir de maintenant que $p = 0.6$. On va alors chercher un évenement improbable si H_0 est vraie, mais que l'on attend d'observer si H_1 est vraie. Cet évenement prendra la forme $Z > s$, pour s bien choisie (qui correspond aussi à $\bar{X} > s'$ où s' s'exprime en fonction de s. Il ne reste plus qu'à choisir s, et pour cela, il suffit d'utiliser le TCL qui affirme $\mathbb{P}(Z > s) \approx \mathbb{P}\Big(\mathcal{N}(0,1) > s\Big)$. Comme pour un intervalle de confiance, on peut se fixer un risque α, et trouver un seuil s_α qui vérifie $\mathbb{P}\Big(\mathcal{N}(0,1) > s_\alpha\Big) = \alpha$ (en utilisant un ordinateur pour calculer les intégrales de la densité par exemple). Par exemple, pour $\alpha = 5\%$, on a $s_\alpha \approx 1.65$.

On en déduit que, si H_0 est vraie, et n assez grand, alors

$$\mathbb{P}(Z > s_\alpha) < \alpha.$$

Il suffit ensuite de comparer la valeur de Z calculée sur l'expérience, et en supposant H_0 vraie avec s_α, et de conclure en fonction :

- Si la valeur z donnée sur l'expérience, et en supposant $p = 0.6$ vérifie $z < s_\alpha$, alors cet évenement n'est pas improbable si H_0 est vraie, et on ne peut rien conclure. On ne peut pas **rejeter** H_0, et on ne peut pas affirmer que notre médicament est meilleur.
- Si la valeur z donnée sur l'expérience, et en supposant $p = 0.6$ vérifie $z > s_\alpha$, alors cet évenement est improbable si H_0 est vraie, et on peut conclure de façon statistiquement significative, avec un risque de α que le médicament est meilleur. On peut **rejeter** H_0.

Sur l'exemple précédent, si on prend un risque de 5% (alors $s_{0.05} = 1.65$), et que l'on suppose avoir observé 650 guérisons dans l'échantillon, alors si H_0 est vraie, on a

$$z = \sqrt{1000}\frac{\frac{650}{1000} - 0.6}{\sqrt{0.6(1-0.6)}} \approx 3.22 > 1.65.$$

On peut alors rejeter le test, et conclure de façon statistiquement significative, au risque de 5%, que notre médicament est meilleur que celui qui est sur le marché.

Enfin, on peut calculer la p-valeur, qui correspond à la probabilité, si H_0 est vraie, d'avoir un résultat au moins aussi surprenant que celui de notre expérience. Ici, cela correspond (il faut un ordinateur pour l'évaluer) à :

$$p - value \approx \mathbb{P}(\mathcal{N}(0,1) > 3.22) \approx 0.069\%$$

6 Aspects numériques

6.1 Bases de R et simulations de variables aléatoires

b) Aspects numériques : Bases d'utilisation de R, manipulation de vecteurs et de matrices, simulation de variables aléatoires, observation de la Loi des Grands Nombres et du Théorème Central Limite, principe de Monte-Carlo

6.2 Représentation de données, corrélation et causalité

b) Aspects numériques : Représentation des données, indices numériques de position et de dispersion, corrélation et exemple de "corrélation n'est pas causalité"

7 Exercices

7.1 Espace de probabilité

Exercice 1 :

On lance 1 dé à 6 faces.

1. Proposer un ensemble Ω, et définir $\mathbb{P}$ sur $\mathcal{P}(\Omega)$ pour décrire l'expérience.
2. Expliciter l'évenement A : "le résultat du dé est pair" et B : "le résultat est strictement supérieur à quatre", et donner les probabilités de ces évenements.
3. Expliciter $A \cap B$, $A \cup B$, A^C et B^C, et donner les probabilités de ces évenements.

Exercice 2 :

On lance (indépendamment) 2 dé à 6 faces.

1. Proposer un ensemble Ω, et définir $\mathbb{P}$ sur $\mathcal{P}(\Omega)$ pour décrire l'expérience.
2. Expliciter l'évenement A : "la somme des résultats des dés vaut six" et B : "le résultat de chaque dé est impair", et donner les probabilités de ces évenements.
3. Expliciter $A \cap B$, A^C et B^C, et $\left(A^C \cup B^C\right)^C$, et donner les probabilités de ces évenements.

Exercice 3 :

On lance (indépendamment) 2 dé à 6 faces. On pose $\Omega = \{1, 2, \cdots, 6\}^2$ et $\forall \omega \in \Omega, \mathbb{P}\big(\{\omega\}\big) = \frac{1}{36}$.

1. Pour tout k entre 2 et 12, expliciter l'évenement S_k : "la somme des résultats des dès vaut k", et donner la probabilité de ces évenements.
2. Expliciter l'évenement D_1 : "le résultat du premier dé est pair", D_2 : "les deux dés ont même parité, et $D_1 \cap D_2$ et donner les probabilités de ces évenements.
3. Expliciter l'évenement A : "la somme des résultats des dés est supérieure ou égale à dix", B : "Au moins un des résultats de et $A \cap B$ et donner les probabilités de ces évenements.

7.2 Conditionnement

Exercice 1 : *(CC 2020)*

Lors d'un QCM, on interroge des étudiants sur la signification du sigle ADN. A chaque personne, on propose trois réponses différentes: 1, 2 ou 3, la réponse correcte étant la réponse 1. Tout étudiant connaissant la réponse correcte la donne, sinon il choisit au hasard une des trois réponses proposées. On suppose que la probabilité qu'un étudiant connaisse la bonne réponse est égale à $3/4$. Soit R l'événement "*l'étudiant connait la bonne réponse*" et B l'événement "*l'étudiant donne la bonne réponse*".

1. Déduire de l'enoncé que $\mathbb{P}(B/R) = 1$ et que $\mathbb{P}(B/R^c) = 1/3$.
2. Calculer la valeur de $\mathbb{P}(R/B)$.

Exercice 2 :

De nombreux logiciels de triche existent pour les jeux vidéos. A tel point qu'une proportion conséquente des joueurs peuvent se mettre à en utiliser. On suppose que pour un certain jeu, 20% des joueurs utilisent un logiciel de triche. De plus, la probablilté de gagner une partie en utilisant ce logiciel, est de 50%, alors qu'elle est seulement de 5% sans ce logiciel.

Quelle est la probablité qu'un gagnant soit en réalité un tricheur ?

Exercice 3 :

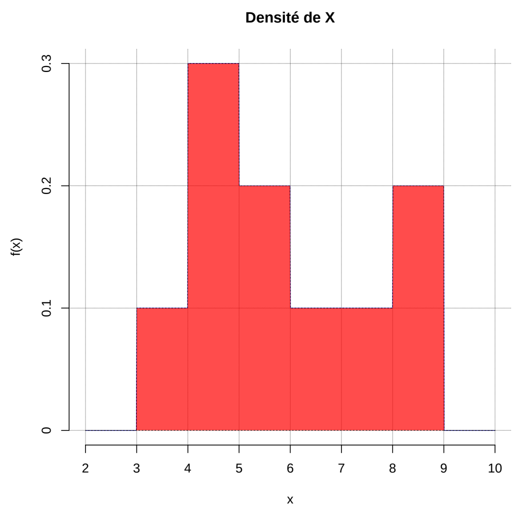
Une maladie rare touche 1% de la population. Il existe un test pour la détecter, qui détecte avec 99% de chances la maladie chez un patient malade, mais se trompe 5% du temps pour un patient non porteur de la maladie (le test est un faux positif).

1. Donner la probabilité d'être malade sachant qu'on a reçu un test positif.
2. Que devient la probablité si 2 tests successifs sortent positifs ? (en supposant l'indépendance entre les deux résultats sachant le statut de maladie de la personne).

7.3 Variables aléatoires à densité

Exercice 1

Soit X une variable aléatoire de densité f représentée ci-dessous. Le quadrillage est représenté seulement pour aider à la lecture de la courbe.



1. Quelle est l'aire totale sous la courbe ?

2. Que vaut $\mathbb{P}(X = 5.5)$?
3. Donner, en lisant le graphique, $\mathbb{P}(X \leq 4)$.
4. Donner, en lisant le graphique, $\mathbb{P}(X \in [4, 7])$.
5. Donner, en lisant le graphique, $\mathbb{P}(X \leq 5 \mid X \leq 8)$.

Exercice 2

Soit X une variable aléatoire de loi uniforme sur $[2, 4]$. On rappelle que X a alors pour densité la fonction $f_X : x \mapsto f(x) = \frac{1}{2}\mathbb{1}_{x \in [2,4]}$.

1. Calculer $\mathbb{P}(X \leq 3)$
2. Calculer $\mathbb{E}[X]$.
3. Calculer $\mathbb{E}[X^2]$.
4. En déduire $\mathrm{Var}(X)$.

Exercice 3

Soit X une variable aléatoire de densité donnée par $f_X : x \mapsto f(x) = 2x\mathbb{1}_{x \in [0,1]}$.

1. Pour $t \in [0, 1]$, calculer $\mathbb{P}(X \leq t)$ en fonction de t.
2. Calculer $\mathbb{E}[X]$.
3. Calculer $\mathbb{E}[X^2]$.
4. En déduire $\mathrm{Var}(X)$.

7.4 Avancé

Exercice 1 :

On considère un gène formé de 2 allèles A et a. On tire deux personnes au hasard dans la population (on supposera donc l'indépendance de leur génotypes), et on définit les évenements $AA_1, AA_2, aa_1, aa_2, Aa_1, Aa_2$ par "AA_1 : la personne 1 possède 2 allèles A"... (de même pour les autres).

Si ces deux personnes ont un enfant, alors on considère que cet enfant hérite aléatoirement (et uniformément) d'un allèle de chacun de ces parents. On note AA_3, Aa_3, aa_3 les évenements : "AA_3 : l'enfant possède 2 allèles A"... (de même pour les autres).

On suppose que

$$\begin{aligned}\mathbb{P}(AA_1) = \mathbb{P}(AA_2) &= p_0\\ \mathbb{P}(Aa_1) = \mathbb{P}(Aa_2) &= r_0\\ \mathbb{P}(aa_1) = \mathbb{P}(aa_2) &= q_0 = 1 - p_0 - r_0\end{aligned}$$

1. Donner $p_1 = \mathbb{P}(AA_3)$ en fonction de p_0, r_0 et q_0.
2. Donner $r_1 = \mathbb{P}(aa_3)$ en fonction de p_0, r_0 et q_0.
3. Donner $q_1 = \mathbb{P}(Aa_3)$ en fonction de p_0, r_0 et q_0.
4. On suppose qu'une population a une répartition à la génération 0 de p_0, r_0, q_0 entre les 3 phénotypes AA, Aa, aa. En supposant que les couples se forment indépendamment des phénotypes, quelle sera la répartition à la génération 1 ? à la génération 2 ? Commenter.

Exercice 2 : On donne les probabilités conditionnelles de guérison de deux médicaments agissant sur les calculs rénaux, en fonction de la taille des calculs. On note P, L les évenements "petits calculs" et "larges calculs", A, B

les évenements “on a administré le médicament A ou B, et G, G^C les évenements ”le/la patient·e a “guéri” ou “pas guéri”.

$$\mathbb{P}(G \mid A, P) = 0.8$$
$$\mathbb{P}(G \mid B, P) = 0.7$$
$$\mathbb{P}(G \mid A, L) = 0.6$$
$$\mathbb{P}(G \mid B, L) = 0.5$$

1. Pour un premier plan de traitement, les médecins choisissent l'administration suivante des médicamments A et B :

   $$\mathbb{P}(A \mid P) = 0.2 = 1 - \mathbb{P}(B \mid P)$$
   $$\mathbb{P}(A \mid L) = 0.8 = 1 - \mathbb{P}(B \mid L)$$

   De plus, on choisit la définition de “petits calculs” et “gros calculs” en prenant les 50% les plus petits ou plus larges, de sorte qu'on a

   $$\mathbb{P}(P) = 0.5 = \mathbb{P}(L)$$

   Calculer $\mathbb{P}(G \mid A)$ et $\mathbb{P}(G \mid B)$. Quel médicament semblerait le meilleur au vue de ce seul résultat ? Commenter.
2. Pour une étude **randomisée**, on choisira aléatoirement le médicament qu'on administrera à chaque patient, de sorte d'avoir

   $$\mathbb{P}(A \mid P) = \mathbb{P}(A \mid L) = 1 - \mathbb{P}(B \mid P) = 1 - \mathbb{P}(B \mid L)$$

   De plus, on fixe arbitrairement cette probabilité à 1/2. Calculer de nouveau $\mathbb{P}(G \mid A)$ et $\mathbb{P}(G \mid B)$, et commenter.

Exercice 3 :

On modélise le temps d'efficacité d'une anesthésie locale (en heures) par une variable aléatoire T de densité $f_T(x) = x\mathbb{1}_{x\in[0,1]} + (2-x)\mathbb{1}_{x\in[1,2]}$.

1. Quelle est la probabilité que l'anesthésie soit efficace moins de 30 min ?
2. Quelle est la probabilité que l'anesthésie soit efficace plus de $1h30$?
3. Que vaut le temps moyen d'efficacité de l'anesthésie ?

8 Corrections

8.1 Espace de probabilité

Exercice 1 :

1. On peut par exemple prendre $\Omega = \{1, 2, 3, 4, 5, 6\}$, et définir $\forall \omega \in \Omega, \mathbb{P}\Big(\{\omega\}\Big) = \frac{1}{6}$, d'où on peut déduire $\forall A \in \mathcal{P}(\Omega), \mathbb{P}\Big(A\Big) = \frac{Card(A)}{6}$.
2. On a $A = \{2, 4, 6\}$ et $B = \{5, 6\}$, d'où $\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$ et $\mathbb{P}(B) = \frac{2}{6} = \frac{1}{3}$.
3. On en déduit $A \cap B = \{6\}$, $A \cup B = \{2, 4, 5, 6\}$, $A^C = \{1, 3, 5\}$ et $B^C = \{1, 2, 3, 4\}$, et $\mathbb{P}(A \cap B) = \frac{1}{6}$, $\mathbb{P}(A \cup B) = \frac{4}{6} = \frac{2}{3}$, $\mathbb{P}(A^C) = \frac{3}{6} = \frac{1}{2}$ et $\mathbb{P}(B^C) = \frac{4}{6} = \frac{2}{3}$.

Exercice 2 :

1. On peut par exemple prendre $\Omega = \{1, 2, 3, 4, 5, 6\}^2$, et définir $\forall \omega \in \Omega, \mathbb{P}\Big(\{\omega\}\Big) = \frac{1}{36}$, d'où on peut déduire $\forall A \in \mathcal{P}(\Omega), \mathbb{P}\Big(A\Big) = \frac{Card(A)}{36}$.

2. Expliciter l'évenement $A = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$ et $B : \{(1,1),(1,3),(1,5),(3,1),(3,3),\cdots,(5,3),(5,5)\}$, et on en déduit

$$\mathbb{P}(A) = \frac{5}{36}$$
$$\mathbb{P}(B) = \frac{9}{36} = \frac{1}{4}$$

3. On a

$$\begin{aligned}
A \cap B &= \{(1,5),(3,3),(5,1)\} \\
A^C &= \text{“la somme est différente de six”} \\
B^C &= \text{“au moins un des dé est pair”} \\
A^C \cup B^C &= \text{“la somme est différente de six ou au moins un des dés est pair”} \\
\left(A^C \cup B^C\right)^C &= \text{“la somme est égale à six et aucun des dés n'est pair”} = A \cap B
\end{aligned}$$

On en déduit

$$\begin{aligned}
\mathbb{P}\Big(A \cap B\Big) &= \frac{3}{36} = \frac{1}{12} \\
\mathbb{P}\Big(A^C\Big) &= 1 - \mathbb{P}(A) = 1 - \frac{5}{36} = \frac{31}{36} \\
\mathbb{P}\Big(B^C\Big) &= 1 - \mathbb{P}(B) = 1 - \frac{1}{4} = \frac{3}{4} \\
\mathbb{P}\Big(\big(A^C \cup B^C\big)^C\Big) &= \mathbb{P}(A \cap B) = \frac{1}{12}
\end{aligned}$$

Exercice 3 :

On lance (indépendamment) 2 dé à 6 faces. On pose $\Omega = \{1,2,\cdots,6\}^2$ et $\forall \omega \in \Omega, \mathbb{P}\big(\{\omega\}\big) = \frac{1}{36}$.

1. On a

$$\begin{aligned}
S_2 &= \{(1,1)\} \\
S_3 &= \{(1,2),(2,1)\} \\
S_4 &= \{(1,3),(2,2),(3,1)\} \\
&\vdots \\
S_6 &= \{(1,5),(2,4),(3,3),(4,2),(5,1)\} \\
S_7 &= \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \\
S_8 &= \{(2,6),(3,5),(4,4),(5,3),(6,2)\} \\
&\vdots \\
S_1 2 &= \{(6,6)\}
\end{aligned}$$

On en déduit

$$\begin{aligned}
\mathbb{P}(S_2) &= \frac{1}{36} \\
\mathbb{P}(S_3) &= \frac{2}{36} \\
\mathbb{P}(S_4) &= \frac{3}{36} \\
&\vdots \\
\mathbb{P}(S_6) &= \frac{5}{36} \\
\mathbb{P}(S_7) &= \frac{6}{36} \\
\mathbb{P}(S_8) &= \frac{5}{36} \\
&\vdots \\
\mathbb{P}(S_12) &= \frac{1}{36}
\end{aligned}$$

2.

$$\begin{aligned}
D_1 &= \begin{Bmatrix} (2,1),(2,2),\cdots,(2,5),(2,6) \\ (4,1),(4,2),\cdots,(4,5),(4,6) \\ (6,1),(6,2),\cdots,(6,5),(6,6) \end{Bmatrix} \\
D_2 &= \begin{Bmatrix} (1,1),(1,3),(1,5) \\ (2,2),(2,4),(2,6) \\ \vdots \\ (5,1),(5,3),(5,5) \\ (6,2),(6,4),(6,6) \end{Bmatrix} \\
D_1 \cap D_2 &= \begin{Bmatrix} (2,2),(2,4),(2,6) \\ (4,2),(4,4),(4,6) \\ (6,2),(6,4),(6,6) \end{Bmatrix}
\end{aligned}$$

On en déduit

$$\mathbb{P}(D_1) = \frac{18}{36} = \frac{1}{2}\mathbb{P}(D_2) \qquad = \frac{18}{36} = \frac{1}{2}\mathbb{P}(D_1 \cap D_2) = \frac{9}{36} = \frac{1}{4}$$

3.

$$\begin{aligned}
A &= \{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\} \\
B &= \{(1,6),(2,6),(3,6),(4,6),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\} \\
A \cap B &= \{(4,6),(5,6),(6,4),(6,5),(6,6)\}
\end{aligned}$$

On en déduit

$$\mathbb{P}(A) = \frac{6}{36} = \frac{1}{6}\mathbb{P}(B) \qquad = \frac{11}{36}\mathbb{P}(A \cap B) = \frac{5}{36}$$

On remarque que $\mathbb{P}(D_1 \cap D_2) = \mathbb{P}(D_1)\mathbb{P}(D_2)$, alors que $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B)$. Donc D_1, D_2 sont des évenements indépendants, alors que A et B ne sont pas indépendants. On peut l'intuiter de la façon suivante :

- Connaîre la parité d'un dé n'apporte pas d'information sur le fait que les deux dés aient même parité
- Savoir que la somme des dés est plus grande que 10 augmente la probabilité d'avoir au moins un 6.

8.2 Conditionnement

Exercice 1 :

1. Sachant qu'un étudiant connait la bonne réponse (R est réalisé) on sait qu'il va la donner. On a donc bien $\mathbb{P}(B/R) = 1$. Par ailleurs, un étudiant ne connaissant pas la bonne réponse, en choisit une au hasard. Vu qu'il y a trois bonnes réponses disponibles, il a donc une chance sur trois de trouver la bonne. Autrement dit $\mathbb{P}(B/R^c) = 1/3$.

2. En appliquant successivement la formule de probabilité conditionnelle puis la formule des probabilités totales, on a

$$\mathbb{P}(R/B) = \frac{\mathbb{P}(R \cap B)}{\mathbb{P}(R)} = \frac{\mathbb{P}(B/R)\mathbb{P}(R)}{\mathbb{P}(B/R)\mathbb{P}(R) + \mathbb{P}(B/R^c)\mathbb{P}(R^c)} = \frac{1 \times 3/4}{1 \times 3/4 + 1/3 \times 1/4} = \frac{9}{10}.$$

Exercice 2 :

On va considérer que l'on tire un joueur au hasard. On peut noter T l'évenement "le joueur utilise le logiciel de triche", et G l'évenement "le joueur gagne la partie". L'énoncé donne alors les probabilités suivantes :

$$\begin{aligned}\mathbb{P}(T) &= 0.2\\ \mathbb{P}(G \mid T) &= 0.5\\ \mathbb{P}(G \mid T^C) &= 0.05\end{aligned}$$

On peut alors en déduire

$$\begin{aligned}\mathbb{P}(T \mid G) &= \frac{\mathbb{P}(T)\mathbb{P}(G \mid T)}{\mathbb{P}(T)\mathbb{P}(G \mid T) + \mathbb{P}(T^C)\mathbb{P}(G \mid T^C)}\\ &= \frac{0.2 * 0.5}{0.2 * 0.5 + 0.8 * 0.05} = \frac{0.1}{0.1 + 0.04}\\ &= \frac{10}{14} = \frac{5}{7}\end{aligned}$$

La probabilité qu'un gagnant soit en réalité un tricheur est donc de 5 chances sur 7.

Exercice 3 :

1. On suppose que l'on tire une personne au hasard dans la population. On note M l'évenement "la personne est malade", et T l'évenement "le test est positif". L'énoncé donne les probabilités suivantes :

$$\begin{aligned}\mathbb{P}(M) &= 0.01\\ \mathbb{P}(T \mid M) &= 0.99\\ \mathbb{P}(T \mid M^C) &= 0.05\end{aligned}$$

On a alors

$$\begin{aligned}\mathbb{P}(M \mid T) &= \frac{\mathbb{P}(M)\mathbb{P}(T \mid M)}{\mathbb{P}(M)\mathbb{P}(T \mid M) + \mathbb{P}(M^C)\mathbb{P}(T \mid M^C)}\\ &= \frac{0.01 * 0.99}{0.01 * 0.99 + 0.99 * 0.05}\\ &= \frac{0.01}{0.01 + 0.05}\\ &= \frac{1}{6}\end{aligned}$$

La probabilité d'être malade sachant qu'on a reçu un test positif est donc de une chance sur 6.

2. On veut connaître la probabilités d'être malade sachant que deux tests successifs sont sortis positifs. Il suffit de faire le même calcul avec l'évenement T^2 : "Deux tests successifs sont sortis positifs". On peut utiliser l'indépendance (sachant le statut de maladie de la personne) pour calculer

$$\begin{aligned}\mathbb{P}(T^2 \mid M) &= \mathbb{P}(T \mid M)^2 \\ &= 0.99^2 \\ \mathbb{P}(T^2 \mid M^C) &= \mathbb{P}(T \mid M^C)^2 \\ &= 0.05^2\end{aligned}$$

On peut alors faire le calcul

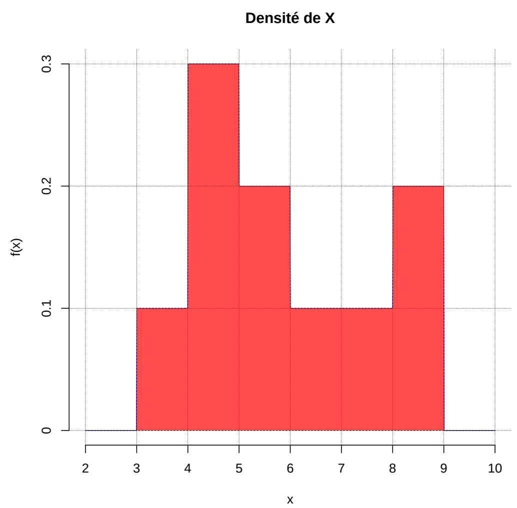
$$\begin{aligned}\mathbb{P}(M \mid T^2) &= \frac{\mathbb{P}(M)\mathbb{P}(T^2 \mid M)}{\mathbb{P}(M)\mathbb{P}(T^2 \mid M) + \mathbb{P}(M^C)\mathbb{P}(T^2 \mid M^C)} \\ &= \frac{0.01 * 0.99^2}{0.01 * 0.99^2 + 0.99 * 0.05^2} \\ &\approx 0.80\end{aligned}$$

Après deux tests successifs positifs, la probablité d'être malade monte à environ 80% (au lieu d'environ 17%).

8.3 Variables aléatoires à densité

Exercice 1

Soit X une variable aléatoire de densité f représentée ci-dessous. Le quadrillage est représenté seulement pour aider à la lecture de la courbe.



1. Par définition d'une densité, l'aire totale sous la courbe vaut 1. On peut le vérifier en comptant les petits rectangles représentés sur le quadrillage, qui ont chacun une aire de $0.1 * 1 = 0.1$, et il y en a bien 10.
2. C'est une des propriétés d'une variable aléaoire à densité, les probabilités $\mathbb{P}(X = x)$ sont nulles pour tout x. Donc
$$\mathbb{P}(X = 5.5) = 0.$$
3. Cette probabilité correpond à l'aire sous la courbe avant 4, il n'y a qu'un rectangle d'aire 0.1, donc
$$\mathbb{P}(X \leq 4) = 0.1$$

4. Cela correspond à l'aire sous la courbe entre 4 et 7. Il y a 6 rectangles d'aire 0.1, ce qui donne

$$\mathbb{P}(X \in [4,7]) = 0.6.$$

5. On peut appliquer la formule pour le conditionnement, puis lire les aires correspondantes sur le graphique. On obtient

$$\begin{aligned}
\mathbb{P}(X \leq 5 \mid X \leq 8) &= \frac{\mathbb{P}(X \leq 5 \cap X \leq 8)}{\mathbb{P}(X \leq 8)} \\
&= \frac{\mathbb{P}(X \leq 5)}{\mathbb{P}(X \leq 8)} \\
&= \frac{0.4}{0.8} \\
&= 0.5
\end{aligned}$$

Exercice 2

Soit X une variable aléatoire de loi uniforme sur $[2,4]$. On rappelle que X a alors pour densité la fonction $f_X : x \mapsto f(x) = \frac{1}{2}\mathbb{1}_{x\in[2,4]}$.

1. On peut poser le calcul, on trouve

$$\begin{aligned}
\mathbb{P}(X \leq 3) &= \int_{-\infty}^{3} f(x)\mathrm{d}x \\
&= \int_{-\infty}^{3} \frac{1}{2}\mathbb{1}_{x\in[2,4]}\mathrm{d}x \\
&= \int_{2}^{3} \frac{1}{2}\mathrm{d}x \\
&= [\frac{x}{2}]_2^3 \\
&= \frac{3}{2} - \frac{2}{2} \\
&= \frac{1}{2}
\end{aligned}$$

2. On peut poser le calcul, on trouve

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x\frac{1}{2}\mathbb{1}_{x\in[2,4]}\mathrm{d}x \\
&= \int_{2}^{4} \frac{x}{2}\mathrm{d}x \\
&= [\frac{x^2}{4}]_2^4 \\
&= \frac{4^2}{4} - \frac{2^2}{4} \\
&= \frac{12}{4} \\
&= 3
\end{aligned}$$

3. On peut poser le calcul, on trouve

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2 \frac{1}{2} \mathbb{1}_{x\in[2,4]}\mathrm{d}x \\
&= \int_2^4 \frac{x^2}{2}\mathrm{d}x \\
&= [\frac{x^3}{6}]_2^4 \\
&= \frac{4^3}{6} - \frac{2^3}{6} \\
&= \frac{56}{6} \\
&= \frac{28}{3}
\end{aligned}$$

4. On peut en déduire

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{28}{3} - 3^2 \\
&= \frac{1}{3}
\end{aligned}$$

Exercice 3

Soit X une variable aléatoire de densité donnée par $f_X : x \mapsto f(x) = 2x\mathbb{1}_{x\in[0,1]}$.

1. Pour $t \in [0,1]$, on a

$$\begin{aligned}
\mathbb{P}(X \leq t) &= \int -\infty^t f_X(x)\mathrm{d}x \\
&= \int -\infty^t 2x\mathbb{1}_{x\in[0,1]}\mathrm{d}x \\
&= \int_0^t 2x\mathrm{d}x \\
&= \left[2\frac{x^2}{2}\right]_0^t \\
&= t^2 - 0 = t^2
\end{aligned}$$

2. On peut calculer

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x.2x\mathbb{1}_{x\in[0,1]}\mathrm{d}x \\
&= \int_0^1 2x^2\mathrm{d}x \\
&= \left[\frac{2x^3}{3}\right]_0^1 \\
&= \frac{2.1^3}{3} - \frac{2.0^3}{3} \\
&= \frac{2}{3}
\end{aligned}$$

3. On peut calculer

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2.2x\mathbb{1}_{x\in[0,1]}\mathrm{d}x \\
&= \int_0^1 2x^3\mathrm{d}x \\
&= \left[\frac{2x^4}{4}\right]_0^1 \\
&= \frac{2.1^4}{4} - \frac{2.0^4}{4} \\
&= \frac{2}{4} = \frac{1}{2}
\end{aligned}$$

4. On peut en déduire

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{1}{2} - \left(\frac{2}{3}\right)^2 \\
&= \frac{9}{18} - \frac{8}{18} \\
&= \frac{1}{18}
\end{aligned}$$

8.4 Avancé

Exercice 1 :

On considère un gène formé de 2 allèles A et a. On tire deux personnes au hasard dans la population, et on définit les évenements $AA_1, AA_2, aa_1, aa_2, Aa_1, Aa_2$ par "AA_1 : la personne 1 possède 2 allèles A"... (de même pour les autres).

Si ces deux personnes ont un enfant, alors on considère que cet enfant hérite aléatoirement (et uniformément) d'un allèle de chacun de ces parents. On note AA_3, Aa_3, aa_3 les évenements : "AA_3 : l'enfant possède 2 allèles A"... (de même pour les autres).

1. En supposant le choix de chaque allèle uniforme, on a

$$\begin{aligned}
&\mathbb{P}\Big(AA_3 \mid AA_1, AA_2\Big) = 1 \\
&\mathbb{P}\Big(AA_3 \mid AA_1, Aa_2\Big) = \mathbb{P}\Big(AA_3 \mid Aa_1, AA_2\Big) = \frac{1}{2} \\
&\mathbb{P}\Big(AA_3 \mid Aa_1, Aa_2\Big) = \frac{1}{4} \\
&\mathbb{P}\Big(AA_3 \mid AA_1, aa_2\Big) = \mathbb{P}\Big(AA_3 \mid Aa_1, aa_2\Big) = \mathbb{P}\Big(AA_3 \mid aa_1, aa_2\Big) = 0
\end{aligned}$$

On en déduit, en supposant l'indépendance des génotype des parents :

$$\begin{aligned}
p_1 = \mathbb{P}(AA_3) &= \mathbb{P}\Big(AA_3 \mid AA_1, AA_2\Big)\mathbb{P}(AA_1)\mathbb{P}(AA_2) + \mathbb{P}\Big(AA_3 \mid AA_1, Aa_2\Big)\mathbb{P}(AA_1)\mathbb{P}(Aa_2) \\
&\quad + \mathbb{P}\Big(AA_3 \mid Aa_1, AA_2\Big)\mathbb{P}(Aa_1)\mathbb{P}(AA_2) + \mathbb{P}\Big(AA_3 \mid Aa_1, Aa_2\Big)\mathbb{P}(Aa_1)\mathbb{P}(Aa_2) \\
&= 1 * p_0 * p_0 + \frac{1}{2} * p_0 * r_0 + \frac{1}{2} * r_0 * p_0 + \frac{1}{4} r_0 * r_0 \\
&= p_0^2 + p_0 r_0 + \frac{r_0^2}{4} \\
&= \left(p_0 + \frac{r_0}{2}\right)^2
\end{aligned}$$

2. Par symétrie, on a

$$q_1 = \mathbb{P}(aa_3) = \left(q_0 + \frac{r_0}{2}\right)^2$$

3. On peut en déduire

$$\begin{aligned} r_1 = \mathbb{P}(Aa_3) &= 1 - \mathbb{P}\Big(AA_3\Big) - \mathbb{P}\Big(Aa_3\Big) \\ &= 1 - \left(p_0 + \frac{r_0}{2}\right)^2 - \left(q_0 + \frac{r_0}{2}\right)^2 \end{aligned}$$

4. Pour la génération 1, il s'agit exactement du calcul qu'on a fait au dessus, on aura donc une répartition de (p_1, r_1, q_1) entre les 3 phénotypes.

   Pour la génération 2, le même type de raisonnement nous donnera une répartition de (p_2, r_2, q_2) donnés par

$$\begin{aligned} p_2 &= \left(p_1 + \frac{r_1}{2}\right)^2 \\ q_2 &= \left(q_1 + \frac{r_1}{2}\right)^2 \\ r_2 &= 1 - \left(p_1 + \frac{r_1}{2}\right)^2 - \left(q_1 + \frac{r_1}{2}\right)^2 \end{aligned}$$

   On peut l'exprimer en fonction de (p_0, r_0, q_0) :

$$\begin{aligned} p_2 &= \left(\left(p_0 + \frac{r_0}{2}\right)^2 + \frac{1}{2}\left(1 - \left(p_0 + \frac{r_0}{2}\right)^2 - \left(q_0 + \frac{r_0}{2}\right)^2\right)\right)^2 \\ &= \left(\frac{1}{2}\left(p_0 + \frac{r_0}{2}\right)^2 + \frac{1}{2} - \frac{1}{2}\left(1 - p_0 - r_0 + \frac{r_0}{2}\right)^2\right)\right)^2 \\ &= \left(\frac{1}{2}\left(p_0 + \frac{r_0}{2}\right)^2 + \frac{1}{2} - \frac{1}{2}\left(p_0 + \frac{r_0}{2}\right)^2 - \frac{1}{2} + \left(p_0 + \frac{r_0}{2}\right)\right)^2 \\ &= \left(p_0 + \frac{r_0}{2}\right)^2 = p_1 \end{aligned}$$

   De même, par symétrie, on aura $q_2 = q_1$ et par conséquent $r_2 = r_1$. Sous ces hypothèses, on a donc un équilibre dès la première génération, et la proportion de chaque allèle dans la population ne bouge plus.

Exercice 2 : On donne les probabilités conditionnelles de guérison de deux médicaments agissant sur les calculs rénaux, en fonction de la taille des calculs. On note P, L les évenements "petits calculs" et "larges calculs", A, B les évenements "on a administré le médicament A ou B, et G, G^C les évenements "le/la patient·e a "guéri" ou "pas guéri".

$$\begin{aligned} \mathbb{P}(G \mid A, P) = 0.8 \\ \mathbb{P}(G \mid B, P) = 0.7 \\ \mathbb{P}(G \mid A, L) = 0.2 \\ \mathbb{P}(G \mid B, L) = 0.1 \end{aligned}$$

1. Il faut ici appliquer la formule de Bayes pour se ramener aux probabilités données dans l'énoncé. On a

$$\mathbb{P}(G \mid A) = \frac{\mathbb{P}\Big(G \cap A\Big)}{\mathbb{P}(A)}$$

avec

$$\begin{aligned}\mathbb{P}\Big(G\cap A\Big) &= \mathbb{P}\Bigg((G\cap A\cap P)\cup(G\cap A\cap L)\Bigg)\\ &= \mathbb{P}\Bigg(G\cap A\cap P\Bigg)+\mathbb{P}\Bigg(G\cap A\cap L\Bigg) \text{ car les évenements sont incompatibles.}\end{aligned}$$

or,

$$\begin{aligned}\mathbb{P}\Bigg(G\cap A\cap P\Bigg) &= \mathbb{P}\Big(A\cap P\Big)\mathbb{P}\Big(G\mid A\cap P\Big)\\ &= \mathbb{P}\Big(P\Big)\mathbb{P}\Big(A\mid P\Big)\mathbb{P}\Big(G\mid A\cap P\Big)\\ \text{de même } \mathbb{P}\Bigg(G\cap A\cap L\Bigg) &= \mathbb{P}\Big(L\Big)\mathbb{P}\Big(A\mid L\Big)\mathbb{P}\Big(G\mid A\cap L\Big).\end{aligned}$$

Cela donne donc

$$\mathbb{P}\Big(G\cap A\Big)=\mathbb{P}\Big(P\Big)\mathbb{P}\Big(A\mid P\Big)\mathbb{P}\Big(G\mid A\cap P\Big)+\mathbb{P}\Big(L\Big)\mathbb{P}\Big(A\mid L\Big)\mathbb{P}\Big(G\mid A\cap L\Big).$$

Et enfin

$$\mathbb{P}(A)=\mathbb{P}(A\cap P)+\mathbb{P}(A\cap L)=\mathbb{P}(P)\mathbb{P}(A\mid P)+\mathbb{P}(L)\mathbb{P}(A\mid L).$$

On obtient alors la formule suivante :

$$\mathbb{P}(G\mid A)=\frac{\mathbb{P}\Big(P\Big)\mathbb{P}\Big(A\mid P\Big)\mathbb{P}\Big(G\mid A\cap P\Big)+\mathbb{P}\Big(L\Big)\mathbb{P}\Big(A\mid L\Big)\mathbb{P}\Big(G\mid A\cap L\Big)}{\mathbb{P}(P)\mathbb{P}(A\mid P)+\mathbb{P}(L)\mathbb{P}(A\mid L)}$$

et de même

$$\mathbb{P}(G\mid B)=\frac{\mathbb{P}\Big(P\Big)\mathbb{P}\Big(B\mid P\Big)\mathbb{P}\Big(G\mid B\cap P\Big)+\mathbb{P}\Big(L\Big)\mathbb{P}\Big(B\mid L\Big)\mathbb{P}\Big(G\mid B\cap L\Big)}{\mathbb{P}(P)\mathbb{P}(B\mid P)+\mathbb{P}(L)\mathbb{P}(B\mid L)}$$

L'application numérique donne alors

$$\begin{aligned}\mathbb{P}(G\mid A)&=\frac{0.5*0.2*0.8+0.5*0.8*0.2}{0.5*0.2+0.5*0.8}=0.32\\ \mathbb{P}(G\mid B)&=\frac{0.5*0.8*0.7+0.5*0.2*0.1}{0.5*0.8+0.5*0.2}=0.58\end{aligned}$$

Au vu de ces résultats, le médicament B guérira 58% des patient·e·s, contre 32% pour le médicament A. Le médicement B semblerait meilleur, si l'on ne regarde que ce résultat. Ce n'est évidemment pas le cas (vu que le médicament traite mieux à la fois les petits et les larges calculs), et on voit l'importance de prendre en compte les "facteurs de confusion", lorsque l'étude n'est pas randomisée. Cet exemple constitue un paradoxe (on peut faire mieux dans chaque sous cas, et moins bien globalement), qui est appelé "paradoxe de Simpson".

2. Il suffit de reprendre l'application numérique en changeant les valeurs de ces probabilités. On obtient alors

$$\mathbb{P}(G\mid A)=\frac{0.5*0.5*0.8+0.5*0.5*0.2}{0.5*0.5+0.5*0.5}=0.5\mathbb{P}(G\mid B) \qquad =\frac{0.5*0.5*0.7+0.5*0.5*0.1}{0.5*0.5+0.5*0.5}=0.4$$

Comme attendu, on observe que le médicament A est meilleur. Cet exercice illustre l'importance des études randomisée, ou de prendre en compte les facteurs de confusion dans l'étude.

Exercice 3 :

Pour tous les calculs, il faudra souvent séparer les intégrales en 2 pour traiter chaque terme à part.

1. On veut calculer

$$\begin{aligned}
\mathbb{P}(T \leq 0.5) &= \int_{-\infty}^{0.5} f_T(x)\mathrm{d}x \\
&= \int_{-\infty}^{0.5} x\mathbb{1}_{x\in[0,1]} + (2-x)\mathbb{1}_{x\in[1,2]}\mathrm{d}x \\
&= \int_{-\infty}^{0.5} x\mathbb{1}_{x\in[0,1]}\mathrm{d}x + \int_{-\infty}^{0.5}(2-x)\mathbb{1}_{x\in[1,2]}\mathrm{d}x \\
&= \int_{0}^{0.5} x\mathrm{d}x + 0 \\
&= [\frac{x^2}{2}]_0^{0.5} \\
&= \frac{0.5^2}{2} - \frac{0^2}{2} \\
&= \frac{1}{8}
\end{aligned}$$

   Pour 1/8 des patients, l'anesthésie sera efficace moins de 30 min.

2. On peut s'y prendre de même :

$$\begin{aligned}
\mathbb{P}(T \geq 1.5) &= \int_{1.5}^{+\infty} f_T(x)\mathrm{d}x \\
&= \int_{1.5}^{+\infty} x\mathbb{1}_{x\in[0,1]} + (2-x)\mathbb{1}_{x\in[1,2]}\mathrm{d}x \\
&= \int_{1.5}^{+\infty} x\mathbb{1}_{x\in[0,1]}\mathrm{d}x + \int_{1.5}^{+\infty}(2-x)\mathbb{1}_{x\in[1,2]}\mathrm{d}x \\
&= 0 + \int_{1.5}^{2}(2-x)\mathrm{d}x \\
&= \left[\frac{-(2-x)^2}{2}\right]_{1.5}^{2} \\
&= \frac{-(2-2)^2}{2} - \frac{-(2-1.5)^2}{2} \\
&= \frac{1}{8}
\end{aligned}$$

   Pour 1/8 des patients, l'anesthésie sera efficace plus de $1h30$ min.

3. On veut calculer

$$
\begin{aligned}
\mathbb{E}[T] &= \int_{-\infty}^{+\infty} x f_T(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2 \mathbb{1}_{x\in[0,1]} + x(2-x)\mathbb{1}_{x\in[1,2]}\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2 \mathbb{1}_{x\in[0,1]} + \int_{-\infty}^{+\infty} x(2-x)\mathbb{1}_{x\in[1,2]}\mathrm{d}x \\
&= \int_0^1 x^2 + \int_1^2 x(2-x)\mathrm{d}x \\
&= \left[\frac{x^3}{3}\right]_0^1 + \int_1^2 2x\mathrm{d}x - \int_1^2 x^2\mathrm{d}x \\
&= \frac{1^3}{3} - \frac{0^3}{3} + \left[2\frac{x^2}{2}\right]_1^2 - \left[\frac{x^3}{3}\right]_1^2 \\
&= \frac{1}{3} - 0 + \frac{2.2^2}{2} - \frac{2.1^2}{2} - \frac{2^3}{3} + \frac{1^3}{3} \\
&= \frac{1}{3} + 4 - 1 - \frac{8}{3} + \frac{1}{3} \\
&= 1
\end{aligned}
$$

En moyenne, l'anesthésie sera efficace $1h$.

9 Rappel sur quelques bases de mathématiques

9.1 Notations

Lors d'une définition, il est possible de rencontrer le symbole := qui signifie seulement "égal par définition à".

On rappelle les quantificateurs utilisés dans ce poly : $\forall$ signifie "Quel que soit" ou "pour tout", et $\exists$ signifie "il existe". On pourra réfléchir à l'importance de l'ordre de ces quantificateurs, en méditant sur la différence entre les phrases

- "Il existe une solution à tous les problèmes" *(le chocolat ?)*
- "Pour tout problème, il existe une solution" *(adaptée au problème)*

Dans tout ce polycopié, nous utiliserons abondamment la notation $\mathbb{1}$ pour désigner la fonction indicatrice :

$$\mathbb{1}_{x\in A} := \begin{cases} 0 \text{ si } x \notin A \\ 1 \text{ si} x \in A \end{cases}$$

9.2 Rappel sur les ensembles

Un ensemble E est dit au plus dénombrable si l'on peut numéroter ses éléments, c'est à dire s'il existe une **injection** de E dans $\mathbb{N}$. Les ensembles finis, $\mathbb{N}, \mathbb{Z}$ et $\mathbb{Q}$ par exemple sont "au plus dénombrables", mais $\mathbb{R}$ n'est pas dénombrable.

On note $\cup$ l'union de deux ensembles, et $\cap$ leur intersection. On rappelle

$$
\begin{aligned}
x \in A \cup B &\Leftrightarrow x \in A \text{ ou } x \in B \\
x \in A \cap B &\Leftrightarrow x \in A \text{ et } x \in B
\end{aligned}
$$

9.3 Calculs d'intégrales

On ne rappelle pas ici les fondements mathématique de l'intégrale, mais seulement quelques propriétés :

- Intégrale d'une constante sur un segment : $\forall a, b, C \in \mathbb{R}, \int_a^b C\mathrm{d}x = C(b-a)$.

- Intégrale d'un monôme sur un segment : $\forall a, b \in \mathbb{R}, \forall n \geq 2, \int_a^b x^n \mathrm{d}x = \frac{b^{n-1}-a^{n-1}}{n-1}$.
- Intégrale de la fonction exponentielle sur un segment : $\forall a, b, \alpha \in \mathbb{R}, \int_a^b e^{\alpha x} \mathrm{d}x = \frac{e^{\alpha b}-e^{\alpha a}}{\alpha}$
- Pour intrégrer une fonction positive sur un intervalle infini, on peut prendre la limite en "poussant" la borne à l'infini :
$$\int_a^{+\infty} f(x)\mathrm{d}x = \lim_{M\to\infty} \int_a^M f(x)\mathrm{d}x.$$
- Pour intégrer une fonction f quelconque sur un intervalle infini, on peut séparer la partie positive f^+ et négative$-f^-$ (si les intégrales existent) :
$$\int_a^{+\infty} f(x)\mathrm{d}x = \int_a^{+\infty} f^+(x)\mathrm{d}x - \int_a^{+\infty} f^-(x)\mathrm{d}x$$

9.4 Séries

On rencontrera dans ce cours des sommes infinies.

- Lorsque tous les termes sont positifs, $\sum_{i=0}^{+\infty} a_i = \lim_{M\to\infty} \sum_{i=0}^{M} a_i$.