Table des matières

1 Élèments de probabilités

Définition. Une expérience est *aléatoire* si un observateur peut la répéter dans les mêmes conditions, mais sans pouvoir en prédire le résultat.

Définition. Un *espace échantilion* est un ensemble S des résultats possibles.

Un espace d'échantilion peut être *qualitatif* ou *quantitatif*, ainsi que *discret*, *continu* ou *mixte*. Il peut aussi être *dénombrable* ou *non-dénombrable*.

Définition. Un *évènement* A est un sous-ensemble de S d'intêret à l'observateur.

Définition. Un *évènement élémentaire* A est un résultat particulier, c'est-à-dire, un élèment de S.

La différence entre les deux dernières définitions est que la $\operatorname{card}(A) = 1$ pour un évènement élémentaire tandis que $\operatorname{card}(A) \geq 1$ pour un évènement.

Exemple 1.1. *On observe le résultat du lancer de deux pièces de monnaie. On note P comme un lancer pile et F comme un lancer face. L'ensemble est donc*

$$S = \{PP, FF, PF, FP\}$$

avec chaque résultat ayant $25\,\%$ d'arriver. L'ensemble S est qualitatif, soit pile ou face, et discret.

Exemple 1.2. *On observer la somme obtenue lors du lancer de deux dés à 6 faces. L'ensemble de résultat possible est*

$$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}\,.$$

Il est quantitatif et discret.

Exemple 1.3. *On compte le nombre de lancers d'une pièce pour obtenir une première fois un pile. L'espace échantilion est*

$$S = \{1, 2, 3, 4, 5, \dots\}\,,$$

car il est possible qu'une grande quantité de lancer est effectuée avant d'obtenir un pile. L'ensemble S est quantitatif, discret et infini dénombrable.

Exemple 1.4. *On mesure le temps d'attente à l'arrêt d'autobus. L'espace échantilion est*

$$S = [0, \infty[\,.$$

L'ensemble S est quantitatif, continu et infini non-dénombrable.

Exemple 1.5. *On mesure le temps d'attente à l'arrêt d'autobus ainsi que le nombre de personnes en file à l'arrivée de l'autobus. L'espace échantilion est*

$$S = T \times U,$$

avec

$$T = [0, \infty[$$

et

$$U = \{1, 2, 3, 4, 5, \dots\}\,.$$

L'ensemble S est quantitatif et mixte. Un exemple d'évènement élèmentaire peut être un couple tel que $(4.25\,\mathrm{s}, 4)$.

1.1 Lien entre l'expérience aléatoire et son modèle

Définition. La *fréquence relative* f_A d'un évènement A est le rapport entre le nombre d'observations n_A de l'évèmenent et le nombre n de répétition de l'expérience, c'est-à-dire

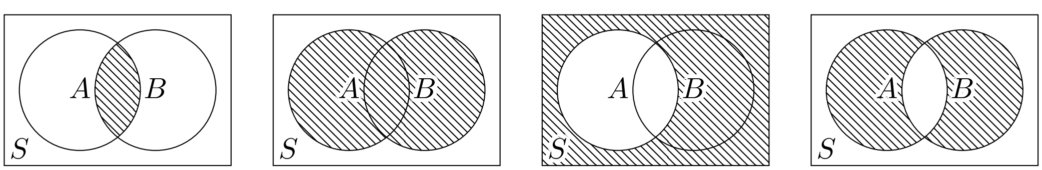
$$f_A = \frac{n_A}{n}.$$

La limite lorsque l'expérience est répétée infiniment est la probabilité de l'évènement A, dénotée

$$\mathbb{P}\,[\,A\,] = \lim_{n \to \infty} f_A.$$

1.2 Opérations sur les ensembles

Soit deux ensembles A et B tel que $A, B \subset S$. La figure 1a montre une intersection tandis que la figure 1b montre une union entre A et B. La figure 1c montre le complémenet de A et la figure 1d montre l'exclusion de deux ensembles.



(a) $A \cap B$ (b) $A \cup B$ (c) A^c (d) $(A \cup B) \backslash (A \cap B)$

FIGURE 1 – Opérations d'ensembles démontrées sur des diagrammes de Venn

1.3 Axiomes fondamentales de la probabilité

Axiome 1. *La probabilité d'un évèment A est plus grand ou égal à 0, c'est-à-dire*

$$\mathbb{P}[\,A\,] \geq 0,$$

pour tout $A \in S$.

Axiome 2. *La probabilité de l'espace d'échantilion S est 1, c'est-à-dire*

$$\mathbb{P}[\,S\,] = 1.$$

Axiome 3. *La probabilité d'un évènement A ou d'un évènement B est équivalent à la somme de leur probabilité, c'est-à-dire*

$$\mathbb{P}[\,A \cup B\,] = \mathbb{P}[\,A\,] + \mathbb{P}[\,B\,],$$

si $A \cup B = \emptyset$.

En général,

$$\mathbb{P}\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} \mathbb{P}[\,A_k\,],$$

si $A_i \cup A_j = \emptyset$, $\forall i, j$.

Théoreme 1.1. $\mathbb{P}[\,A^c\,] = 1 - \mathbb{P}[\,A\,]$.

Démonstration. On sait que $A \cup A^c = S$ et $A \cap A^c = \emptyset$. Hors, $\mathbb{P}[\,A \cup A^c\,] = \mathbb{P}[\,S\,] \Leftrightarrow \mathbb{P}[\,A\,] + \mathbb{P}[\,A^c\,] = 1$, car $A \cup A^c = \emptyset$. En réarrangeant, on obtient que $\mathbb{P}[\,A^c\,] = 1 - \mathbb{P}[\,A\,]$. □

Théoreme 1.2. $\mathbb{P}[\,A\,] \leq 1$.

Démonstration. On sait que $\mathbb{P}[\,A^c\,] \geq 0$ et $\mathbb{P}[\,A^c\,] = 1 - \mathbb{P}[\,A\,]$. En réarrangeant, on obtient que $\mathbb{P}[\,A\,] \leq 1$. □

Théoreme 1.3. $\mathbb{P}[\,\emptyset\,] = 0$.

Démonstration. On sait que $S^c = \emptyset$. Par conséquent, $\mathbb{P}[\,\emptyset\,] = 1 - \mathbb{P}[\,S\,] = 1 - 1 = 0$. □

Théoreme 1.4. $A \subset B \Rightarrow \mathbb{P}[\,A\,] \leq \mathbb{P}[\,B\,]$.

Démonstration. La différence entre A et B est $A^c \cap B$ de sorte qu'on peut écrire $B = A \cup (A^c \cap B)$. Par conséquent, $\mathbb{P}[\,B\,] = \mathbb{P}[\,A\,] + \mathbb{P}[\,A^c \cap B\,] \Leftrightarrow \mathbb{P}[\,B\,] - \mathbb{P}[\,A\,] = \mathbb{P}[\,A^c \cap\ B\,] \geq 0$, car $A \cup (A^c \cap B) = \emptyset$. En réarrangeant, on obtient $\mathbb{P}[\,A\,] \leq \mathbb{P}[\,B\,]$. □

Théoreme 1.5. $\mathbb{P}[\,A \cup B\,] = \mathbb{P}[\,A\,] + \mathbb{P}[\,B\,] - \mathbb{P}[\,A \cap B\,]$.

1.4 Principe d'équiprobabilité

1.4.1 Ensemble fini

On suppose que S est fini, soit $S = \{e_1, e_2, \ldots, e_n\}$. On dit que les résultats d'une expérience aléatoire sont *équiprobables* si

$$\mathbb{P}[\, e_1 \,] = \mathbb{P}[\, e_2 \,] = \cdots = \mathbb{P}[\, e_n \,] = \frac{1}{n}.$$

Dans ce cas, on a que la probabilité d'un évènement $A \subset S$ est

$$\mathbb{P}[\, A \,] = \frac{n_A}{n},$$

où n_A est le nombre d'élément dans A et n celui dans S.

1.4.2 Ensemble infini dénombrable

On suppose que S est infini dénombrable, alors l'équiprobabilité est impossible. Si la probabilité d'un évènement élémentaire est $\mathbb{P}[\, e_i \,] = \epsilon$, on obtient que $\mathbb{P}[\, S \,] = \mathbb{P}[\, e_1 \,] + \mathbb{P}[\, e_2 \,] + \cdots = \infty$, Ce qui est en contradiction avec les axiomes. D'une manière similaire, si $\mathbb{P}[\, e_i \,] = 0$, alors $\mathbb{P}[\, S \,] = 0$, ce qui est aussi en contradiction avec les axiomes.

1.4.3 Ensemble infini non-dénombrable

On suppose que $S = [a, b]$ est infini non-dénombrable. Soit un évènement $A = [c, d] \subset S$. La probabilité de l'évènement est

$$\mathbb{P}[\, A \,] = \frac{l_A}{l} = \frac{d-c}{b-a},$$

où l_A est la longueur de A et l la longueur de S. Par conséquent, la probabilité d'un évènement élémentaire $e \in S$ est

$$\mathbb{P}[\, e \,] = 0,$$

car $l_A = 0$.

Exemple 1.6. *On calcule la somme de deux dés lancés. Quelle est la probabilité d'obtenir chaque somme possible ?*

L'espace échantilion est $S = L \times L$, où $L = \{1, 2, 3, 4, 5, 6\}$ est le résultat possible d'un dé. On peut écrire $S = \{(1,1), (1,2), \ldots, (6,6)\}$. On suppose qu'il y a équiprobabilité de sorte que la probabilité d'obtenir un évènement élémentaire est $1/36$.

Hors, certaines des sommes sont dupliquées de sorte que la probabilité d'obtenir une somme particulière est donnée par la table suivante.

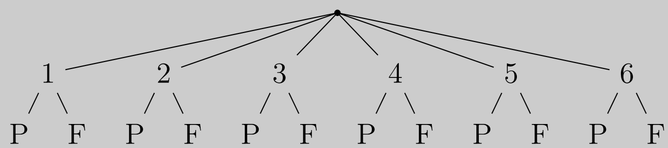
A	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}[\, A \,]$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

2 Analyse combinatoire

2.1 Diagramme en arbre

Exemple 2.1. *On lance un dé, puis une pièce de monnaie. Combien de résultats est-il possible ?*

On peut représenté la situation par l'arbre suivant.



Par le principe de multiplication, il y a $6 \cdot 2 = 12$ *possibilités.*

2.2 Permutations

Définition. Une *permutation* correspond à un choix de k objets parmi n objets distincts. Le choix se fait sans remise et dans un ordre spécifique.

La table 1 résume le principe d'une permutation à l'aide d'une pige d'objet. À chaque pige, la quantité d'objets diminue de 1.

Table 1 – Pige d'objets

Choix d'objet	1	2	...	k
Objets restants	n	$n-1$	...	$n-(n-k)$

À l'aide du principe de multiplication, il y a $n \cdot (n-1) \cdots (n-k+1)$ combinaisons. On dénote le nombre de permutations sans remise de n éléments par

$$\mathcal{P}_n^k = \frac{n!}{(n-k)!}.$$

Lorsqu'il y a remise, le nombre de permutations de n éléments est n^k.

Exemple 2.2. *On dispose de 10 composantes dont 4 défectueuses. On pige 3 composantes au hasard et sans remise. Quelle est la probabilité d'obtenir uniquement des composantes non défectueuses ?*

On suppose qu'il y a équiprobabilité du choix des 3 composantes. La probabilité d'obtenir 3 composantes non défecteuses F *est*

$$\mathbb{P}[F] = \frac{6 \cdot 5 \cdot 4}{10 \cdot 9 \cdot 8} = \frac{\mathcal{P}_6^3}{\mathcal{P}_{10}^3} = \frac{1}{6}.$$

2.3 Combinaisons

Définition. Une *combinaison* correspond au choix de k objets parmis n objets distincts. Le choix se fait sans remise et sans ordre spécifique.

Soit C_k^n le nombre de combinaisons. Le nombre de permutations est égal au nombre de combinaisons multiplié par le nombre d'arrangement possible $k!$, c'est-à-dire

$$\mathcal{P}_n^k = \mathcal{C}_n^k \cdot k!,$$

et en réarrangeant, on obtient

$$\mathcal{C}_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Exemple 2.3. *Combien de codes de deux lettres peut-on former à partir du mot* `OUI` *?*

Avec ordre, il y a $\mathcal{P}_3^2 = 6$ permutations et sans ordre, il y a $\mathcal{C}_3^2 = 3$ combinaisons.

Exemple 2.4. *Quel est la probabilité de gagner le gros lot à la $6/49$?*

Sans ordre, il y a 1 seul cas favorable et $\mathcal{C}_{49}^6$ cas possibles. Par conséquent, la probabilité de gagner G est

$$\mathbb{P}\left[\, G \,\right] = \frac{1}{\mathcal{C}_{49}^6} = \frac{1}{13\,983\,816}.$$

Avec ordre, il y a $6!$ cas favorables et $\mathcal{P}_{49}^6$ cas possibles de sorte que la probabilité est

$$\mathbb{P}\left[\, G \,\right] = \frac{6!}{\mathcal{P}_{49}^6} = \frac{720}{10\,068\,347\,520}.$$

2.3.1 Triangle de Pascal

Une combinaison peut se représenter avec le triangle de Pascal comme le montre la figure 2.

$$\begin{array}{ccccccccc}
 & & & & \mathcal{C}_0^0 & & & & \\
 & & & \mathcal{C}_1^0 & & \mathcal{C}_1^1 & & & \\
 & & \mathcal{C}_2^0 & & \mathcal{C}_2^1 & & \mathcal{C}_2^2 & & \\
 & \mathcal{C}_3^0 & & \mathcal{C}_3^1 & & \mathcal{C}_3^2 & & \mathcal{C}_3^3 & \\
\mathcal{C}_4^0 & & \mathcal{C}_4^1 & & \mathcal{C}_4^2 & & \mathcal{C}_4^3 & & \mathcal{C}_4^4
\end{array}$$

FIGURE 2 – représentation du triangle de Pascal

Théoreme 2.1. $\mathcal{C}_n^k = \mathcal{C}_{n-1}^{k-1} + \mathcal{C}_{n-1}^k$

Démonstration. La démonstration est triviale et est laissée au lecteur. □

2.3.2 Binôme de Newton

La combinaison est souvent appliquée dans le cas de la puissance d'un binôme tel que

$$(a+b)^n = \sum_{k=0}^{n} \mathcal{C}_n^k \cdot a^k \cdot b^{n-k} = \sum_{k=0}^{n} \binom{k}{n} \cdot a^k \cdot b^{n-k},$$

mais elle est plus souvent utilisée avec la deuxième notation.

2.4 Permutation d'objets semblables

Exemple 2.5. *Combien y a-t-il d'ordres possibles des lettres «ppfff» ?*

Soit 5 cases distinctes représentant l'ordre d'une pige dans les lettres. Il faut choisir les cases où mettre les «p», soit

$$\mathcal{C}_5^2 = \frac{5!}{2!3!} = 10$$

ordres possibles.

Si on échange un «p» et un «f» dans une séquence A, on obtient une séquence B différente de A. Si on échange un «p» avec une autre «p» dans une séquence A, on retrouve la même séquence A.

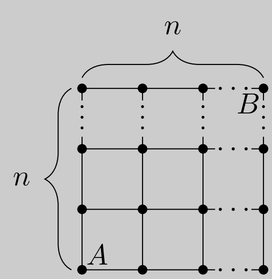
En analysant la formule, le facteur 5! *représente le nombre d'ordres si toutes les lettres étaient différentes. Le facteur* 2! *représente les «p» s'ils étaient différents et le facteur* 3! *représente le nombres des «f» s'ils étaient différents.*

En général, avec n objets objets comprenant n_1 objets de classes 1, n_2 objets de classe 2, $\ldots$, n_k objets à la classe k, on a

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

ordres possibles.

Exemple 2.6. *Soit le graphe $n \times n$ suivant.*



Combien de chemins existe entre A et B suivant, s'il est seulement permis d'aller vers la droite ou vert le haut ?

Exemple 2.6 (suite)**.** *On suppose que tous les chemins possibles sont équiprobables. Peu importe le chemin, il faut avancer $n-1$ vers la droite et $n-1$ vers le haut pour un total de $2n-2$ mouvement.*

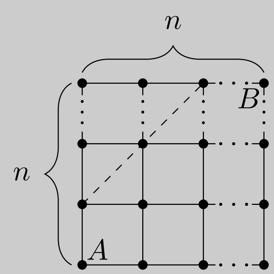
Il suffit de calculer le nombre de permutations de ces mouvements sachant qu'il y des objets semblables. Le nombre chemins est

$$\frac{(2n-2)!}{(n-1)!\,(n-1)!},$$

ce qui est équivalent à $\mathcal{C}_{2n-2}^{n-1}$.

2.5 Équivalence

Exemple 2.7. *Soit le graphe $n \times n$ suivant.*



Quel est la probabilité qu'un chemin choisi au hasard, allant seulement vers la droite ou le haut, ne traverse pas les noeuds au-dessus de la diagonale entre A et B ?

Les «bons» chemins ne passent pas par la ligne critique en pointillée et les «mauvais» chemins passent par la ligne critique. La probabilité d'un bon chemin B peut s'écrire en fonction de son complément, soit

$$\mathbb{P}[\,B\,] = 1 - \mathbb{P}[\,M\,] = 1 - \frac{n_m}{C_{n-1}^{2n-2}},$$

où M est un mauvais chemin et n_m est le nombre de mauvais chemins.

Soit un mauvais chemin M. On définit le point X comme étant le premier point au-dessus de la diagonale que M atteint. On applique une transformation mirroir au chemin suivant X comme la prochaine figure.

Exemple 2.7 (suite)**.** *Il s'avère que tous les mauvais chemins dans le graphe $n \times n$ correspond à un chemin unique dans le graphe $(n-1) \times (n+1)$. C'est-à-dire la transformation est bijective.*

Par conséquent, le problème est équivalent à trouver le nombre de chemins dans un graphe $(n-1)\times(n+1)$. D'une manière similaire à l'exemple 2.6, ce nombre de chemins est donné par

$$\frac{(2n-2)!}{(n-2)!n!},$$

ce qui est équivalent à $\mathcal{C}_{n-2}^{2n-2}$.

Par conséquent, la probabilité d'avoir un bon chemin au hasard est

$$\mathbb{P}\,[\,B\,] = 1 - \frac{\mathcal{C}_{2n-2}^{n-2}}{\mathcal{C}_{2n-2}^{n-1}} = \frac{1}{n}.$$

2.6 Réccurence

Exemple 2.8. *Il est possible de résoudre le problème à l'exemple 2.7 à l'aide de la réccursion.*

En effet, le nombre de «bons» chemins $b_{i,j}$ à partir d'un noeud est la somme des «bons» chemins des noeuds à droite et en haut, c'est-à-dire

$$b_{i,j} = b_{i-1,j} + b_{i,j-1},$$

où i est le nombre de noeuds restants vers la droite et j la nombre de noeuds restants vers le haut.

On sait que $b_{0,0} = 1$, car il ne reste plus de noeud à parcourir fois rendu à B. Aussi, $b_{0,j} = 1$, car il est seulement possible de se rendre à B en allant vers le haut. Sur la diagonale, on a $b_{i,i} = b_{i-1,j}$, car on ne peut pas aller vers le haut. Par conséquent, on obtient le système d'équations à reccurence suivant,

$$b_{i,j} = \begin{cases} 0, & si \quad i = 0, j = 0 \\ 1, & si \quad i = 0 \\ b_{i-1,j}, & si \quad i = j \\ b_{i-1,j} + b_{i,j-1} & sinon \end{cases}$$

Pour une grille 4×4, on peut calculer le nombre de «bons» chemins en développant la récurrence afin d'obtenir $b_{3,3} = 5$ chemins.

3 Probabilité conditionnelles

Définition. Une *probabilité conditionnelle* est la probabilité qu'un évènement A se réalise, si B s'est réalisé.

Mathématiquement, on dénote une probabilité conditionnelle avec

$$\mathbb{P}[\,A|B\,] = \frac{\mathbb{P}[\,A \cap B\,]}{\mathbb{P}[\,B\,]}.$$

où A et B sont des évènements.

Exemple 3.1. *Soit le lancement d'un dé avec les évènements* $A = \{5, 6\}$ *et* $B = \{2, 4, 6\}$*, alors* $\mathbb{P}[\,A\,] = 1/3$ *et*

$$\mathbb{P}[\,A|B\,] = \frac{\mathbb{P}[\,\{6\}\,]}{\mathbb{P}[\,B\,]} = \frac{1}{3}.$$

Si $A = \{6\}$*, alors* $\mathbb{P}[\,A\,] = 1/6$ *et*

$$\mathbb{P}[\,A|B\,] = \frac{\mathbb{P}[\,\{6\}\,]}{\mathbb{P}[\,B\,]} = \frac{1}{3}.$$

3.1 Propriétés

Théoreme 3.1. $\mathbb{P}[\,A \cap B\,] = \mathbb{P}[\,A|B\,] \cdot \mathbb{P}[\,B\,]$.

Théoreme 3.2. $\mathbb{P}[\,A|B\,] = \mathbb{P}[\,B|A\,] = 0$*, si* $A \cap B = \emptyset$.

Théoreme 3.3. $\mathbb{P}[\,A|B\,] \neq \mathbb{P}[\,B|A\,]$*, en général.*

Théoreme 3.4. $\mathbb{P}[\,A|S\,] = \mathbb{P}[\,A\,]$*,* $A \in S$.

Le dernier théorème résulte que toutes probabilités peut s'exprimer sous la forme d'une probabilité conditionnelle.

Exemple 3.2. *On pige sans remise 3 composantes non défecteuses parmis 10 composantes, donc 4 sont défecteuses. Soit* A_i *le* i^e *composante non défecteuses.*

$$\mathbb{P}[\,A_1 \cap A_2 \cap A_3\,] = \mathbb{P}[\,A_3|A_1 \cap A_2\,]\,\mathbb{P}[\,A_2|A_1\,]\,\mathbb{P}[\,A_1\,] = \frac{4}{8} \cdot \frac{5}{9} \cdot \frac{6}{10}$$

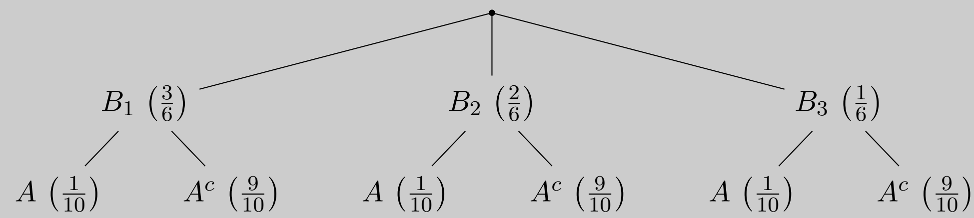
3.2 Probabilités totales

Définition. Les évènements $B_1, B_2, \ldots, B_n$ forment une *partition* si les évènements $B_i \cap B_j = \emptyset$, $\forall i \neq j$, et $B_1 \cup B_2 \cup \cdots \cup B_n = S$.

Avec une partition, on a la règle de probabilités totales, soit

$$\mathbb{P}[\,A\,] = \sum_{i=1}^{n} \mathbb{P}[\,A \cap B_i\,] = \sum_{i=1}^{n} \mathbb{P}[\,A|B_i\,]\,\mathbb{P}[\,B_i\,].$$

Exemple 3.3. *Soit la partition B_1, B_2 et B_3 représenté dans l'arbre suivant.*



À partir de l'arbre, il facile de déterminer les probabilités conditionnelles de A et A^c. Par exemple, $\mathbb{P}[A|B_1] = 3/6 \cdot 1/10 = 1/30$ et $\mathbb{P}[A^c|B_2] = 2/6 \cdot 9/10 = 3/10$.

Théoreme 3.5 (Règle d'inversion)**.** $\mathbb{P}[B|A] = \dfrac{\mathbb{P}[A|B]\,\mathbb{P}[B]}{\mathbb{P}[A]}$

Théoreme 3.6 (Règles de Bayes)**.** $\mathbb{P}[B_j|A] = \dfrac{\mathbb{P}[A|B_j]\,\mathbb{P}[B_j]}{\sum_{i=1}^{n}\mathbb{P}[A|B_i]\,\mathbb{P}[B_i]}$, *où B_i et B_j sont des partitions.*

Exemple 3.4. *On dépiste le cancer du poumon dans une clinique. On sait que*

- 25 % *des individus sont fumeurs*
- 75 % *des individus sont non fumeurs*
- 10 % *des fumeurs développent un cancer*
- 1 % *des non fumeurs développent un cancer*

On détecte un cancer des poumons chez un individu sélectionné au hasard pour dépistage. Quelle est la probabilité que ça soit un fumeur ?

Soit B les fumeurs, B^c les non fumeurs, A la présence de cancer du poumon et A^c son absence. Par conséquent, on cherche $\mathbb{P}[B|A]$. Avec les données, on sait que $\mathbb{P}[A|B] = 0.10$, $\mathbb{P}[A|B^c] = 0.01$, $\mathbb{P}[B] = 0.25$ et $\mathbb{P}[B^c] = 0.75$, de sorte que le théorème de Bayes nous donne

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A|B]\,\mathbb{P}[B]}{\mathbb{P}[A|B]\,\mathbb{P}[B] + \mathbb{P}[A|B^c]\,\mathbb{P}[B^c]} \approx 0.769.$$

3.3 Notion d'indépendance

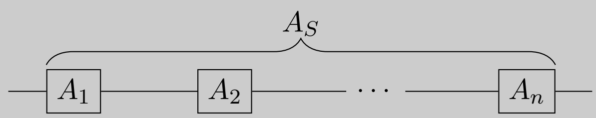
Définition. On dit que les évènements A et B sont *indépendants* si la réalisation d'une n'affecte pas l'autrea.

Mathématiquement parlant, des évènements indépendants A et B sont tels que $\mathbb{P}[A|B] = \mathbb{P}[A]$ et $\mathbb{P}[B|A] = \mathbb{P}[B]$.

Théoreme 3.7. $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] \Leftrightarrow \mathbb{P}[A|B] = \mathbb{P}[A] \wedge \mathbb{P}[B|A] = \mathbb{P}[B]$

Théoreme 3.8. $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] \Leftrightarrow \mathbb{P}[A \cap B^c] = \mathbb{P}[A]\,\mathbb{P}[B^c]$.

Exemple 3.5. *Soit un système en série ayant n composantes comme à la figure suivante.*



Un système en série fonction si et seulement si tous ses composantes fonctionnent. Quelle est la probabilité que le système fonctionne si chaque composante a une probabilité de 75 % de fonctionner ?

Soit l'ensemble échantillion $S = \{F, D\}$ avec F une composante qui fonctionne et D une composante défectueuse. De plus, on définit $A_i = \{F\} \in S$ comme étant la composante i qui fonctionne et $A_i^c = \{D\} \in S$ comme étant la composante i qui est défecteuse.

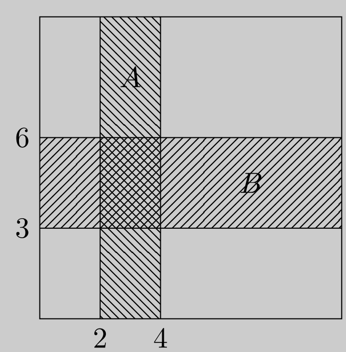
On suppose que les composantes fonctionnent et tombent en panne indépendamment les uns des autres. Le système fonctionne si

$$A_S = \bigcap_{i=1}^{n} A_i,$$

de sorte que

$$\mathbb{P}[\,A_S\,] = \mathbb{P}\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n} \mathbb{P}[\,A_i\,] = 0.75^n.$$

Exemple 3.6. *On génère un point au hasard dans le carré $S = [0, 10] \times [0, 10]$. On définit A comme ayant l'abscisse du point entre 2 et 4, et B comme ayant l'ordonnée du point entre 3 et 6. Est-ce que A et B sont indépendants ?*



Exemple 3.6 (suite). *Puisqu'il y a équiprobabilité continue, alors les probabilités sont données par le rapport entre l'aire de A ou B sur S, alors $\mathbb{P}[A] = 2/10$ et $\mathbb{P}[B] = 3/10$. De plus, $\mathbb{P}[A \cap B] = 6/10$. Puisque $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$, alors les évènements A et B sont indépendants.*

4 Variables Aléatoires

Définition. Une *variable aléatoire* correspond à une expérience aléatoire dont les résultats sont quantitatifs.

On peut décrire une variable aléatoire par sa

- fonction de répartition
- fonction de masse (variable aléatoire discrète)
- fonction de densité (variable aléatoire continue)

Exemple 4.1. *Soit X le nombre de défaults de soudure d'un transistor à 3 pattes. Quel est l'espace échantillion de X ?*

L'espace échantillion est $S_X = \{0, 1, 2, 3\}$, alors X est une variable aléatoire discrète.

Exemple 4.2. *Soit X un nombre réel choisie au hasard dans l'intervalle $[0, 2]$. Quel est l'espace échantillion de X ?*

L'espace échantillion est $S_X = [0, 2]$, alors X est une variable aléatoire continue.

Exemple 4.3. *Soit X le temps d'attente d'un client au guichet automatique. Quel est l'espace d'échantillion de X ?*

L'espace échantillion est $S_X = 0 \cup [0, \infty[$. Il y a une partie discrète et continue, alors X est une variable aléatoire mixte.

4.1 Fonction de répartition

Définition. Un *fonction de répartition* $F_X(x)$ est égal à la probabilité qu'une variable aléatoire X soit plus petite qu'une valeur x, c'est-à-dire $F_X(x) = \mathbb{P}[X \leq x]$, $\forall x \in \mathbb{R}$.

Théoreme 4.1. $0 \leq F_X(x) \leq 1$.

Théoreme 4.2. $\lim_{X \to -\infty} F_X(x) = 0$.

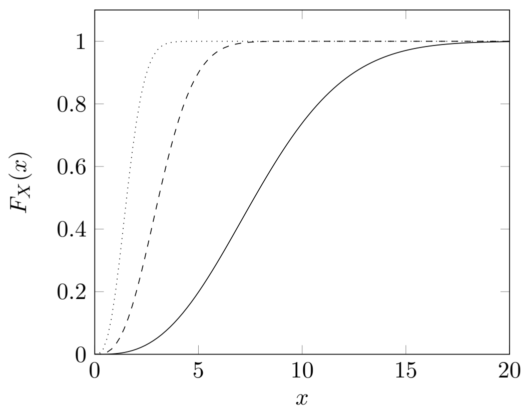
Théoreme 4.3. $\lim_{X \to \infty} F_X(x) = 1$.

Théoreme 4.4 (non décroissance). $x_0 < x_1 \Leftrightarrow F_X(x_0) < F_X(x_1)$.

Théoreme 4.5 (continuité à droite)**.** $F_X(x^+) = F_X(x)$.

Il en résulte que toutes fonctions de répartition $F_X(x)$ sont croissantes, mais peuvent contenir des discontinuités. Elles peuvent représenté des variables aléatoires discrètes, continues ou mixtes. La figure 3 est une exemple de la fonction de répartition d'une distribution de Maxwell-Boltzmann pour certains paramètres différents.

FIGURE 3 – $F_X(x)$ de certaines distributions de Maxwell-Boltzmann



4.2 Fonction de masse

Définition. Une *fonction de masse* $P_X(x_k)$ est égal à la probabilité qu'une variable aléatoire X soit égal à une valeur discrète $x_k \in S_X$, c'est-à-dire $P_X(x_k) = \mathbb{P}[\, X = x_k \,]$ avec $S_X = \{x_1, x_2, \ldots, x_k | x_1 < x_2 < \cdots < x_k\}$

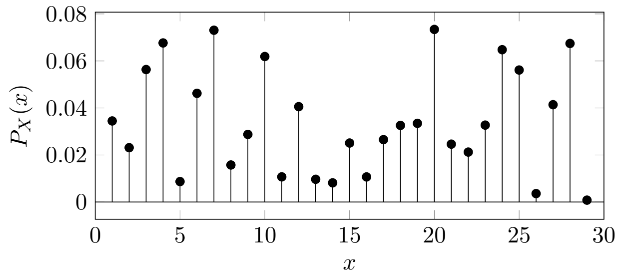
Théoreme 4.6. $P_X(x_k) \geq 0$.

Théoreme 4.7. $\displaystyle\sum_{a < x_k \leq b} P_X(x_k) = \mathbb{P}[\, a < X \leq b \,]$.

Théoreme 4.8. $\displaystyle\sum_{k=1}^{\infty} P_X(x_k) = 1$

Une fonction de masse est nulle en tout point sauf aux valeurs discrètes possibles. De plus, la somme de toutes les valeurs discrètes donne 1. La figure 4 montre un exemple d'une fonction de masse.

FIGURE 4 – exemple d'une fonction de masse



4.3 Fonction de densité

Définition. Une *fonction de densité* $f_X(x)$ est égal à la probabilité de qu'une variable aléatoire X soit autour d'une valeur x, c'est-à-dire

$$f_X(x) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathbb{P}\left[x - \frac{\epsilon}{2} \leq X \leq x + \frac{\epsilon}{2} \right],$$

où ϵ est positif.

Théoreme 4.9. $f_X(x) \geq 0$

Théoreme 4.10. $\displaystyle\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$

4.4 Règles de calcul fondamentale

Théoreme 4.11. $\mathbb{P}[\, a < X \leq b \,] = F_X(b) - F_X(a)$

Démonstration. Soit les ensembles $A = \{X \leq a\}$, $B = \{X \leq b\}$ et $C = \{a < X \leq b\}$, avec $a < b$, comme à la figure 5.

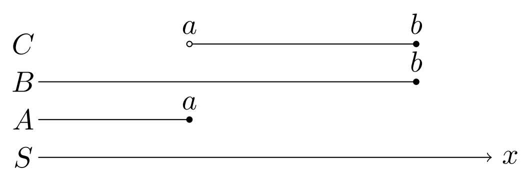


FIGURE 5 – représentation sur une droite numérique

On sait que $A \cap C = \emptyset$ et $A \cup C = B$. Par conséquent, $\mathbb{P}[\, B \,] = \mathbb{P}[\, A \,] + \mathbb{P}[\, C \,] \Leftrightarrow \mathbb{P}[\, C \,] = \mathbb{P}[\, B \,] - \mathbb{P}[\, A \,] = F_X(b) - F_X(a)$. □

Théoreme 4.12. $\mathbb{P}[\, X = x \,] = F_X(x) - F_X(x^-)$.

Démonstration. Soit $a = x - \epsilon$ et $b = x$, où ϵ est positif. Selon le théorème 4.11, on a

$$\mathbb{P}[\, x - \epsilon < X \leq x \,] = F_X(x) - F_X(x - \epsilon).$$

En prenant la limite lorsque $\epsilon \to 0$, on obtient

$$\lim_{\epsilon \to 0} \mathbb{P}\left[\, x - \epsilon < X \leq x \,\right] = F_X(x) - \lim_{\epsilon \to 0} F_X(x - \epsilon).$$

Avec $F_X(x^-) = \lim_{\epsilon \to 0} F_X(x-\epsilon)$ et $(x - \epsilon < X \leq x) \equiv (X = x)$ lorsque $\epsilon \to 0$, on obtient

$$\mathbb{P}\left[\, X = x \,\right] = F_X(x) - F_X(x^-). \qquad \square$$

4.5 Liens entre les différentes fonctions

Toutes variables aléatoires peuvent être décrites par une fonction de répartition. Lorsque la variable aléatoire est continue, alors elle peut aussi être décrite pas une fonction de densité de probabilité. Si la variable aléatoire est discrète, alors elle peut être décrite pas une fonction de masse.

Théoreme 4.13. $f_X(x) = \dfrac{\mathrm{d}}{\mathrm{d}x} F_X(x)$ *si* X *est continue.*

Théoreme 4.14. $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$ *si* X *est continue.*

Théoreme 4.15. $F_X(x) = \displaystyle\sum_{x_k \leq x} P_X(x_k)$ *si* X *est discrète.*

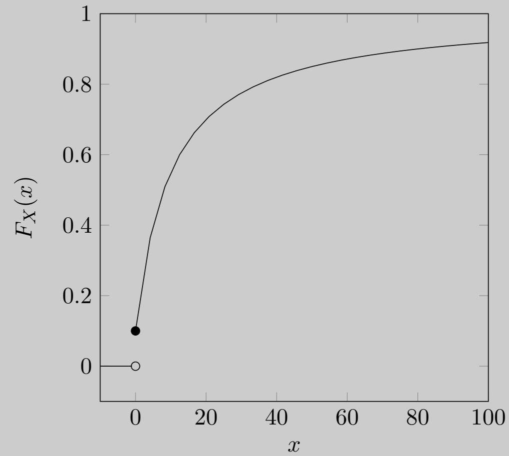
Théoreme 4.16. $P_X(x_k) = \begin{cases} F_X(x_1), & k = 1 \\ F_X(x_k) - F_X(x_{k-1}), & k \neq 1 \end{cases}.$

Exemple 4.4. *Soit* X *le temps d'attente d'un client au guichet automatique. On suppose que* $10\,\%$ *des visites sont sans attente. Quel est la fonction de répartition ?*

La variable aléatoire est mixte. On sait que $F_X(0) = 1/10$ *et que* $F_X(x) = 0$ *si* $x < 0$. *La fonction de répartition est*

$$F_X(k) = \begin{cases} 0, & x < 0 \\ \dfrac{x+1}{x+10}, & x \geq 0 \end{cases}$$

Le graphique suivant montre la fonction de répartition $F_X(x)$.


1
0.8
0.6
0.4
0.2
0
$F_X(x)$
0
20
40
60
80
100
x

Exemple 4.5. *Soit X le nombre de défaults de soudure d'un transistor à 3 pattes. La fonction de masse de X est donnée par*

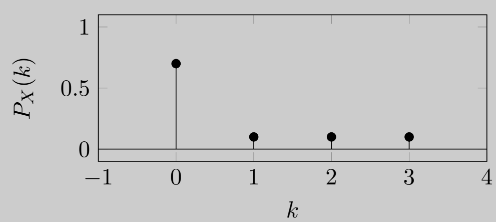
$$P_X(k) = \begin{cases} 0.7, & k = 0 \\ 0.1, & k = 1 \\ 0.1, & k = 2 \\ 0.1, & k = 3 \end{cases}$$

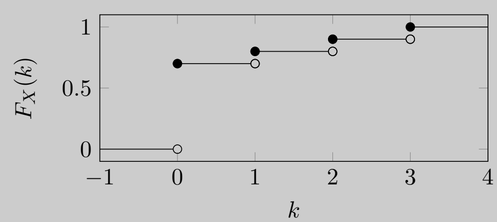
Quelle est la fonction de répartition de $P_X(k)$?

On applique le théorème 4.15 pour obtenir la fonction de répartition. Lorsque $x < 0$, alors $F_X(x) = 0$. Lorsque $0 \leq x < 1$, alors $F_X(x) = 0.7$. Lorsque $1 \leq x < 2$, alors $F_X(x) = 0.8$. En continuant, on obtient

$$F_X(k) = \begin{cases} 0.0, & k < 0 \\ 0.7, & 0 \leq k < 1 \\ 0.8, & 1 \leq k < 2 \\ 0.9, & 2 \leq k < 3 \\ 1.0, & 3 \leq k \end{cases}$$

Les fonctions de masse et de répartition sont montrées dans les figures suivantes.





Exemple 4.6. *Soit X un nombrbe réel choisie au hasard dans l'intervalle $[0, 2]$. Quelle est la fonction de répartition de X ?*

On sait que $F_X(x) = \mathbb{P}[\, X \leq x \,] = x/2$ lorsque $0 \leq x \leq 2$, $F_X(x) = 0$ si $x < 0$ et $F_X(x) = 1$ si $x > 2$. On obtient donc la fonction de répartition

$$F_X(k) = \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 \leq x \leq 2 \\ 1, & 2 \leq x \end{cases}$$

4.6 Fonction conditionnelle

Définition. Une *fonction conditionnelle* est la probabilité qu'une variable aléatoire X prenne une valeur plus petit ou égal à x sachant un évènement A.

Théoreme 4.17. $F_X(x|A) = \dfrac{\mathbb{P}[\, \{X \leq x\} \cap A \,]}{\mathbb{P}[\, A \,]}$.

Théoreme 4.18. $f_X(x|A) = \dfrac{\mathrm{d}}{\mathrm{d}x} F_X(x|A)$ *si* X *est continue.*

Théoreme 4.19. $P_X(x_k|A) = \begin{cases} \dfrac{P_X(x_k)}{\mathbb{P}[\,A\,]}, & x_k \in A \\ 0, & x_k \notin A \end{cases}$ *si* X *est discrète.*

4.7 Médiane et quantile

Définition. La médiane d'une variable aléatoire X continue est le nombre réel $x_{1/2}$ tel que $F_X(x_{1/2}) = {}^1\!/_2$.

Définition. Le *quantile* d'ordre p d'une variable aléatoire X continue est le nombre réel x_p tel que $F_X(x_p) = p$.

Exemple 4.7. *Dans une certaine population, la taille* X *d'un adulte choisi au hasard possède la fonction de répartition*

$$F_X(x) = \begin{cases} 0, & x < 1.2 \\ 1.5x - 1.8, & 1.2 \leq x < 1.7 \\ 0.5x - 0.1, & 1.7 \leq x < 2.2 \\ 1, & 2.2 \leq x \end{cases}$$

Calculer la médiane et le quantile d'ordre 95.

Puisque $F_X(1.7) = 0.75$, *alors le quantile d'ordre 95 est dans la tranche* $1.7 \leq x < 2.2$. *Il suffit de résoudre* $0.95 = 0.5x_{0.95} - 0.1$ *et on obtient que* $x_{0.95} = 2.1$.

4.8 Lois de probabilités discrètes

4.8.1 Loi de Bernoulli

Une expérience aléatoire où la variable aléatoire X peut être un *succès* ou un *échec* suit une loi de Bernoulli de paramètre p dénotée

$$X \sim \mathcal{B}\left(1, p\right),$$

où p est la probabilité de succès. On dénote un succès comme $X = 1$ et un échec comme $X = 0$. Par conséquent, la probabilité d'un succès est $\mathbb{P}[\,X = 1\,] = p$ tandis que celle d'un échec est $\mathbb{P}[\,X = 0\,] = q$ où $q = 1 - p$.

Théoreme 4.20. $P_X(k) = \begin{cases} q & si \quad k = 0 \\ p & si \quad k = 1 \end{cases}.$

4.8.2 Loi binomiale

Une expérience aléatoire où la variable aléatoire X est le nombre de *succès* obtenu en n essaie suit une loi binomiale dénotée

$$X \sim \mathcal{B}\left(n, p\right),$$

où p est la probabilité d'un succès individuel. L'espace échantillion est $S_X = \{0, 1, 2, \ldots, n\}$.

Théoreme 4.21. $P_X(k) = \mathcal{C}_n^k p^k q^{n-k}$ *où* $k = 0, 1, 2, 3, \ldots, n$.

Démonstration. Soit A_i le i-ème succès et A_i^c le i-ème échec dans une expérience aléatoire à n essai. La probabilité d'obtenir k succès est donnée par

$$\mathbb{P}\left[\underbrace{A_1 \cap A_2 \cap \cdots \cap A_k}_{k \text{ fois}} \cap \underbrace{A_{k+1}^c \cap A_{k+1}^c \cap \cdots \cap A_n^c}_{n-k \text{ fois}}\right],$$

où $k = 0, 1, 2, 3, \ldots, n$.

Hors, la probabilité d'un succès est $\mathbb{P}[\,A_i\,] = p$ et celle d'un échec est $\mathbb{P}[\,A_i^c\,] = q$, où $q = 1 - p$. Puisque chaque essai est indépendant des autres, la probabilité peut s'écrire

$$\underbrace{\mathbb{P}[\,A_1\,]\,\mathbb{P}[\,A_2\,]\cdots P(A_k)}_{k \text{ fois}} \cdot \underbrace{\mathbb{P}\left[\,A_{k+1}^c\,\right]\mathbb{P}\left[\,A_{k+2}^c\,\right]\cdots P(A_n^c)}_{n-k \text{ fois}} = p^k q^{n-k}.$$

De plus, les succès peuvent être à n'importe quel essai. Par conséquent, il faut choisir k essaies parmis les n essaies de sorte que la probabilité d'obtenir k succès en tout est $P_X(k) = \mathbb{P}[\,X = k\,] = \mathcal{C}_n^k p^k q^{n-k}$. □

Proposition. $\mathbb{P}[\,S_X\,] = 1$.

Démonstration. Par définition, on a

$$\mathbb{P}[\,S_X\,] = \sum_{k=0}^{n} \mathcal{C}_n^k p^k q^{n-k}$$

Selon le théorème binomial, soit

$$(a+b)^n = \sum_{k=0}^{n} \mathcal{C}_n^k p^k q^{n-k},$$

et que $q = 1 - p$, on peut simplifier de sorte à obtenir

$$\mathbb{P}[\,S_X\,] = (p+q)^k = 1^k = 1.$$ □

Exemple 4.8. *Vous achetez un billet de 6/49 à chaque semaine depuis vos 18 ans. Quelle est la probabilité de gagner le gros lot au plus tard à votre* 98^e *anniversaire ?*

Ce problème se décrit à l'aide d'une loi binomiale, soit

$$X \sim \mathcal{B}\left(n, \frac{1}{13\,983\,816}\right),$$

où n est le nombre d'essaies et X le nombre de gros lots gagnés.

Exemple 4.8 (suite)**.** *Puisqu'on joue chaque semaine de 18 ans à 98 ans, un total de* $n = 52 \cdot (98 - 18) = 4160$ *billets sont achetés. La probabilité d'obtenir que des échecs est données par*

$$\left(1 - \frac{1}{13\,983\,816}\right)^{4160} = \left(\frac{13\,983\,815}{13\,983\,816}\right)^{4160},$$

de sorte que la probabilité d'obtenir au moins 1 billet gagnant est donnée par le complément, soit

$$\mathbb{P}\,[\,X \geq 1\,] = 1 - \left(\frac{13\,983\,815}{13\,983\,816}\right)^{4160} \approx 0.000\,297.$$

4.8.3 Loi géométrique

Une expérience aléatoire où la variable aléatoire X est le nombre d'essais requis pour obtenir un premier succès suit une loi géométrique dénotée

$$X \sim \mathcal{G}\,(p)\,,$$

où p est la probabilité d'avoir un succès à un essai. L'espace échantillion est $S_X = \{1, 2, \dots\}$.

Théoreme 4.22. $p_X(k) = q^{k-1}p$.

Démonstration. Soit A_i le i-ième succès et A_i^c le i-ème échec dans un expérience aléatoire où il faut k essai pour obtenir un succès. La probabilité d'obtenir le succès au k-ième essai est donnée par

$$\mathbb{P}\left[\underbrace{A_1^c \cap A_2^c \cap \cdots \cap A_{k-1}^c}_{k-1 \text{ fois}} \cap A_k\right],$$

où $k = 1, 2, \dots$

Hors, la probabilité d'un succès est $\mathbb{P}\,[\,A_i\,] = p$ et celle d'un échec est $\mathbb{P}\,[\,A_i^c\,] = q$, où $q = 1 - p$. Puisque chaque essai est indépendant des autres, la probabilité peut s'écrire

$$p_X(k) = \underbrace{\mathbb{P}\,[\,A_1^c\,]\,\mathbb{P}\,[\,A_2^c\,] \cdots \mathbb{P}\,[\,A_{k-1}\,]}_{k-1 \text{ fois}} \cdot \mathbb{P}\,[\,A_k\,] = q^{k-1}p.$$ □

Théoreme 4.23. $F_X(n) = 1 - q^n$ *où* $n = 1, 2, \dots$

Démonstration. Selon une variante du théorème 4.15, on a

$$F_X(n) = \sum_{k=1}^{n} p_X(k) = \sum_{k=1}^{n} q^{k-1}p = p\sum_{k=1}^{n} q^{k-1}.$$

On remarque que la somme est une série géométrique, alors

$$\sum_{k=1}^{n} q^{k-1} = \frac{1 - q^n}{1 - q}.$$

Puisque $q = 1 - p \Leftrightarrow p = 1 - q$, on obtient

$$F_X(n) = (1-q)\frac{1-q^n}{1-q} = 1 - q^n. \qquad \square$$

Théoreme 4.24. $\mathbb{P}\left[\, S_X \,\right] = 1$.

Démonstration. En utilisant à nouveau la série géométrique, on a

$$\mathbb{P}\left[\, S_X \,\right] = \sum_{k=1}^{\infty} p_X(k) = p \sum_{k=1}^{\infty} q^{k-1} = p\frac{1}{1-q} = (1-q)\frac{1}{1-q} = 1,$$

puisque $q = 1 - p \Leftrightarrow p = 1 - q$. $\square$

Théoreme 4.25 (absence de mémoire)**.** $\mathbb{P}\left[\, X > k + j | X > j \,\right] = \mathbb{P}\left[\, X > k \,\right]$ *où* $k, j \in \mathbb{Z}$*.*

Démonstration. On sait que

$$\mathbb{P}\left[\, X > k + j | X > j \,\right] = \frac{\mathbb{P}\left[\, \{X > k + j\} \cap \{X > j\} \,\right]}{\mathbb{P}\left[\, X > j \,\right]}.$$

Puisque $k \in \mathbb{Z}$, alors $\{X > k + j\} \cap \{X > j\} = \{X > k + 1\}$ de sorte que

$$\mathbb{P}\left[\, X > k + j | X > j \,\right] = \frac{\mathbb{P}\left[\, X > k + j \,\right]}{\mathbb{P}\left[\, X > j \,\right]}.$$

Avec le complément du théorème 4.23, on a

$$\mathbb{P}\left[\, X > k + j | X > j \,\right] = \frac{q^{k+j}}{q^j} = q^k = \mathbb{P}\left[\, X > k \,\right]. \qquad \square$$

4.8.4 Loi de Poisson

Cette loi est un cas limite de la loi binomiale lorsque $n \to \infty$. On pose $p = \alpha/n$ où α est un nombre positif, de sorte que la probabilité lorsque n est grand devient nulle. On dénote une loi de Poisson par

$$X \sim \mathrm{Poi}\,(\alpha) = \lim_{n \to \infty} \mathcal{B}\left(n, \frac{\alpha}{n}\right).$$

Par conséquent, la probabilité que X prend une valeur proche de α est élevée. Puisqu'elle est basée sur une loi binomiale, alors l'espace échantillion est $S_X = \{0, 1, 2, \dots\}$.

Théoreme 4.26. $p_X(k) = \dfrac{\alpha^k}{k!}\mathrm{e}^{-\alpha}$.

Théoreme 4.27. $\mathbb{P}\left[\, S_X \,\right] = 1$.

Démonstration. En utilisant le développement en série de la fonction exponentielle, on a

$$\mathbb{P}\left[\, S_X \,\right] = \sum_{k=0}^{\infty} p_X(k) = \mathrm{e}^{-\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} = \mathrm{e}^{-\alpha}\mathrm{e}^{\alpha} = 1. \qquad \square$$

4.8.5 Approximation par une loi de Poisson

Soit une loi binomiale $X \sim \mathcal{B}(n,p)$. Si p est près de 0, alors $X \approx \mathrm{Poi}(np)$. Par conséquent,

$$\mathbb{P}[\,X = k\,] \approx \frac{(np)^k}{k!}\mathrm{e}^{-np}.$$

En général, l'approximation est bonne si $n \geq 30$ et $p \leq 0.05$. Si p est proche de 1, alors on considère les échecs au lieu des succès. Dans ce cas,

$$\mathbb{P}[\,X = k\,] \approx \frac{(nq)^{n-k}}{(n-k)!}\mathrm{e}^{-nq}.$$

4.9 Loi des probabilités continues

4.9.1 Loi uniforme (continue)

Une expérience aléatoire où la variable aléatoire X est le choix d'un nombre réel dans un intervalle $[a,b]$ suit une loi uniforme dénotée

$$X \sim \mathcal{U}(a,b)\,,$$

où l'ensemble échantillion est $S_X = [a,b]$.

Théoreme 4.28. $f_X(x) = \begin{cases} \dfrac{1}{b-a} & si\ a \leq x \leq b \\ 0 & sinon \end{cases}.$

Théoreme 4.29. $F_X(x) = \begin{cases} 0 & si \quad x < a \\ \dfrac{x-a}{b-a} & si \quad a \leq x \leq b \\ 1 & si \quad b < x \end{cases}.$

4.9.2 Loi exponentielle

Une expérience aléatoire où la variable aléatoire X suit la loi exponentielle dénotée

$$X \sim \mathrm{Exp}(\lambda)\,,$$

où $\lambda > 0$, a l'espace échantillion $S_X = [0,\infty[$.

Théoreme 4.30. $f_X(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & si \quad x \geq 0 \\ 0 & si \quad x \leq 0 \end{cases}.$

Théoreme 4.31. $F_X(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x} & si \quad x \geq 0 \\ 0 & si \quad x \leq 0 \end{cases}.$

Théoreme 4.32 (absence de mémoire)**.** $\mathbb{P}[\,X > s+t | X > t\,] = \mathbb{P}[\,X > t\,]$ *où* $s,t \in [0,\infty[$.

4.9.3 Loi gamma

Une expérience aléatoire où la variable aléatoire X suit la loi gamma dénotée

$$X \sim \mathrm{Gam}\,(\alpha, \lambda)\,,$$

où $\alpha > 0$ et $\lambda > 0$, a l'espace échantillion $S_X = [0, \infty[$.

Théoreme 4.33. $f_X(x) = \dfrac{(\lambda x)^{\alpha-1}\lambda \mathrm{e}^{-\lambda x}}{\Gamma(\alpha)}.$

Théoreme 4.34. $F_X(x) = 1 - \displaystyle\sum_{k=0}^{n-1} \frac{(\lambda x)^k \mathrm{e}^{-\lambda x}}{k!}$ *si* $\alpha = n = 1, 2, 3, \ldots$

Théoreme 4.35. $\mathbb{P}\,[\,S_X\,] = 1.$

Démonstration. Il suffit d'intégrer la fonction de densité de probabilité, soit

$$\int_0^\infty \frac{(\lambda x)^{\alpha-1}\lambda \mathrm{e}^{-\lambda x}}{\Gamma(\alpha)}\,\mathrm{d}x = \frac{1}{\Gamma(\alpha)} \int_0^\infty (\lambda x)^{\alpha-1}\lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x,$$

et en posant $y = \lambda x$, on obtient

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1}\mathrm{e}^{-y}\,\mathrm{d}y = \frac{1}{\Gamma(\alpha)}\Gamma(\alpha) = 1. \qquad \square$$

4.9.4 Loi normale

On modèle souvent les erreurs d'observation et l'addition des résultats de plusieurs expérience aléatoires d'une variable aléatoire X par une loi normale dénotée

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right),$$

où $\mu \in \mathbb{R}$ est la moyenne et $\sigma^2 > 0$ la variance. On appel aussi *sigma* comme l'écart-type. L'espace échantillion est $S_X = \mathbb{R}$.

Théoreme 4.36. $f_X(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\dfrac{1}{2\sigma^2}(x-\mu)^2\right).$

Théoreme 4.37. $F_X(x) = \displaystyle\int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\mathrm{d}z$ *où* $z = \dfrac{t-\mu}{\sigma}.$

Théoreme 4.38. $\displaystyle\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\,\mathrm{d}x = 1$

Démonstration. Soit

$$I = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\,\mathrm{d}x$$

et

$$I^2 = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\,\mathrm{d}x \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\,\mathrm{d}y\,.$$

Ensuite,

$$I^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)/2}\,\mathrm{d}x\mathrm{d}y\,.$$

En utilisant les coordonnées polaires, on a

$$I^2 = \frac{1}{2\pi}\int_{0}^{2\pi}\int_{0}^{\infty} \mathrm{e}^{-r^2/2} r\,\mathrm{d}r\mathrm{d}\theta,$$

de sorte qu'en posant posant $u = r^2$, on peut obtenir

$$I^2 = -\frac{1}{2\pi}\int_{0}^{2\pi} \mathrm{e}^{-r^2/2}\Big|_0^{\infty}\,\mathrm{d}\theta = -\frac{1}{2\pi}\int_{0}^{2\pi}\mathrm{d}\theta = 1$$

de sorte que $I = 1$. □

4.10 Fonction d'une variable aléatoire

Exemple 4.9. *On suppose que X est la valeur d'ampliture d'un signal an temps t. Le signal numérisé peut s'écrire*

$$Y = \mathrm{signe}(X)\cdot\Delta\cdot\mathrm{part}\left(\frac{|X|}{\Delta}+\frac{1}{2}\right),$$

où Δ est le pas de quantification.

4.10.1 X et Y sont des variables aléatoires discrètes

Exemple 4.10. *Soit $X \sim \mathcal{B}(2, 1/4)$ et $Y = (X-1)^2$. Quelle est la fonction de masse de Y ?*

On sait que $\mathbb{P}[\,Y=0\,] = \mathbb{P}[\,X=1\,]$ et $\mathbb{P}[\,Y=1\,] = \mathbb{P}[\,X=0\,]+\mathbb{P}[\,X=2\,]$. On sait que $\mathbb{P}[\,X=0\,] = 9/16$, $\mathbb{P}[\,X=1\,] = 6/16$ et $\mathbb{P}[\,X=0\,] = 1/16$ de sorte que $\mathbb{P}[\,Y=0\,] = 6/16$ et $\mathbb{P}[\,Y=1\,] = 10/16$.

4.10.2 X et Y sont des variables aléatoires discrète et continue

Exemple 4.11. *Soit $X \sim \mathcal{N}(0,1)$ et*

$$Y = \begin{cases} -1 & si \quad X < -1/2 \\ 0 & si \quad -1/2 \le X < 1/2 \\ 1 & si \quad 1/2 \le X \end{cases}.$$

Quelle est la probabilité que $Y = 1$?

On cherche $\mathbb{P}[\,X \le -1/2\,] = \Phi(-1/2) = \Phi(1/2) \approx 0.3085$. De plus, $\mathbb{P}[\,-1/2 \le X \le 1/2\,] = \Phi(1/2) - \Phi(-1/2) \approx 0.3830$. Finalement, $\mathbb{P}[\,X < -1/2\,] \approx 0.3085$.

4.10.3 X et Y sont de svariables aléatoires continues

Exemple 4.12. *Soit* $X \sim \mathcal{U}(-1, 2)$ *et* $Y = X^2$. *Quelle est la fonction de répartition de* Y *?*

4.11 Espérance mathématique

Définition. Une *espérance* d'une variable aléatoire X, dénotée $\mathrm{E}[X]$, est la somme des valeur possibles de X pondérées par leur probabilité.

Théoreme 4.39. $\mathrm{E}[X] = \sum_{k=1}^{\infty} x_k p_X(x_k)$ *si* X *est discrète.*

Théoreme 4.40. $\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x$ *si* X *est continue.*

Exemple 4.13. *Quelle est l'espérance d'un lancer d'un dé ?*

On calcule l'espérance d'une variable discrète, soit

$$\mathrm{E}[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5.$$

Exemple 4.14. *Soit* $X \sim \mathrm{Poi}(\alpha)$. *Quelle est l'espérance de* X *?*

Par définition, on calcule l'espérance avec

$$\mathrm{E}[X] = \sum_{k=0}^{\infty} k \cdot p_X(k) = \sum_{k=0}^{\infty} k \cdot \frac{\mathrm{e}^{-\alpha}\alpha^k}{k!} = \sum_{k=1}^{\infty} k \cdot \frac{\mathrm{e}^{-\alpha}\alpha^k}{k!},$$

car le premier terme à $k = 0$ *est nul. Par conséquent,*

$$\mathrm{E}[X] = \sum_{k=1}^{\infty} \frac{\mathrm{e}^{-\alpha}\alpha^k}{(k-1)!} = \mathrm{e}^{-\alpha} \sum_{k=1}^{\infty} \frac{\alpha^k}{(k-1)!} = \mathrm{e}^{-\alpha} \sum_{i=0}^{\infty} \frac{\alpha^{i+1}}{i!} = \mathrm{e}^{-\alpha}\alpha \sum_{i=0}^{\infty} \frac{\alpha^i}{i!}.$$

Hors, la somme est le développement en séries de la fonction exponentielle, alors

$$\mathrm{E}[X] = \mathrm{e}^{-\alpha}\alpha\mathrm{e}^{\alpha} = \alpha.$$

Exemple 4.15. *Soit* $X \sim \mathrm{Exp}(\lambda)$. *Quelle est l'espérance de* X *?*

Par définition, on calcule l'espérance avec

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \int_{0}^{\infty} x\lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x\,.$$

En posant $u = x$ *et* $\mathrm{d}v = \lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x$, *on obtient*

$$\mathrm{E}[X] = -x\mathrm{e}^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} \mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}\int_0^{\infty} \lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda},$$

car l'aire sous la fonction de densité de probabilités est égal à 1.

Soit $Y = g(X)$, où X et Y sont des variables aléatoires et g une tranformation. Si X et Y discrètes, alors

$$\mathrm{E}[Y] = \mathrm{E}[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_X(x_k),$$

et si X et Y sont continues, alors

$$\mathrm{E}[Y] = \mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x\,.$$

Théoreme 4.41. $\mathrm{E}[c] = c$, *où* c *est une constante.*

Théoreme 4.42. $\mathrm{E}[aX+b] = a\mathrm{E}[X] + b$, *où* a *et* b *des constantes.*

Théoreme 4.43. $\mathrm{E}[X|A] = \sum_{k=1}^{\infty} x_k p_X(x_k|A)$, *où* X *discrète.*

Théoreme 4.44. $\mathrm{E}[X|A] = \int_{-\infty}^{\infty} x f_X(x|A)$, *où* X *continue.*

Théoreme 4.45. $\mathrm{E}[X] = \sum_{i=1}^{n} E(X|B_i)\mathbb{P}[B_i]$, *où* $B_1, \ldots, B_n$ *des partitions de* S_X.

Exemple 4.16. *Soit* X *une variable aléatoire mixte. Quelle est la forme de l'espérance de* X *?*

On définie C *comme l'événement où* X *prend une valeur continue et* D *lorsque* X *prend une valeur discrète. Par conséquent, on a*

$$\mathrm{E}[X] = \underbrace{\mathrm{E}[X|C]}_{\int}\mathbb{P}[C] + \underbrace{\mathrm{E}[X|D]}_{\Sigma}\mathbb{P}[D]\,.$$

4.12 Variance

Définition. La *variance* d'une variable aléatoire X, dénotée $\mathrm{Var}[X]$, est définie comme

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

En développant le carré de la variance, on obtient

$$\mathrm{Var}[X] = \mathrm{E}\left[X^2 - 2X\mathrm{E}[X] + \mathrm{E}[X]^2\right] = \mathrm{E}\left[X^2\right] - 2\mathrm{E}[X\mathrm{E}[X]] + \mathrm{E}\left[\mathrm{E}[X]^2\right],$$

et en simplifiant les constantes, on a

$$\mathrm{Var}[X] = \mathrm{E}\left[X^2\right] - 2\mathrm{E}[X]\,\mathrm{E}[X] + \mathrm{E}[X]^2 = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$$

Exemple 4.17. *Soit $X \sim \mathcal{B}(1,p)$. Quelle est la variance de X ?*

On sait que l'espérance de X est

$$\mathrm{E}[X] = 0 \cdot (1-p) + 1 \cdot p = p,$$

et celle de X^2 est

$$\mathrm{E}\left[X^2\right] = 0^2 \cdot (1-p) + 1^2 \cdot p = p.$$

Par conséquent, la variance est

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

Théoreme 4.46. $\mathrm{Var}[c] = c$, *où c est une constante.*

Théoreme 4.47. $\mathrm{Var}[aX+b] = a^2\mathrm{Var}[X]$, *où a et b des constantes.*

Théoreme 4.48. $\mathrm{Std}[X] = \sqrt{\mathrm{Var}[X]}$.

Théoreme 4.49. $\mathrm{Var}[X|A] = \mathrm{E}\left[X^2|A\right] - \mathrm{E}[X|A]^2$.

4.13 Inégalité de Markov

Soit X une variable aléatoire prenant des valeurs non négatives. On peut montrer que

$$\mathbb{P}[X \geq a] \leq \frac{\mathrm{E}[X]}{a}, \forall a > 0.$$

4.14 Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire dont la moyenne $\mathrm{E}[X] = \mu_X$ et la variance $\mathrm{Var}[X] = \sigma_X^2$ existent. On peut montrer que

$$\mathbb{P}[|X - \mu_X| \geq a] \leq \frac{\sigma_X^2}{a^2}, \forall a > 0.$$

Exemple 4.18. *Soit le lancer d'une pièce de monnaie avec $X \sim \mathcal{B}(n, 1/2)$. On calcule la moyenne avec*

$$\mathrm{E}\left[\frac{X}{n}\right] = \frac{1}{n}\mathrm{E}[X] = \frac{1}{n}\cdot n\cdot\frac{1}{2} = \frac{1}{2}$$

et la variance avec

$$\mathrm{Var}\left[\frac{X}{n}\right] = \frac{1}{n^2}\mathrm{Var}[X] = \frac{1}{n^2}\cdot n\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4n}.$$

Selon l'inégalité de Bienaymé-Tchebychev, on a

$$\mathbb{P}\left[\left|\frac{X}{n} - \frac{1}{2}\right| \geq 0.01\right] \geq \frac{1/4n}{(0.01)^2} = \frac{10000}{4n} = \frac{2500}{n}.$$

4.15 Fonction caractéristique

Définition. Une fonction *caractéristique*, dénoté $\phi_X(\omega)$, est l'espérance d'une variable aléatoire X tel que

$$\phi_X(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega X}\right],$$

où j est le nombre imaginaire tel que $j^2 = -1$.

Théoreme 4.50. $\phi_X(\omega) = \sum_{k=1}^{\infty} \mathrm{e}^{j\omega x_k}\cdot p_X(x_k)$, *si X est discrète.*

Théoreme 4.51. $\phi_X(\omega) = \int_{-\infty}^{\infty} \mathrm{e}^{j\omega X}\cdot f_X(x)\,\mathrm{d}x$, *si X est continue.*

Théoreme 4.52. $\phi_X(0) = 1$.

Théoreme 4.53. $\mathrm{E}[X^n] = (-j)^n \left[\frac{\mathrm{d}^n}{\mathrm{d}\omega^n}\phi_X(\omega)\right]_{\omega=0}$.

Exemple 4.19. *Soit $X \sim \mathcal{B}(n,p)$. Quelle est la fonction caractéristique de X ?*

La fonction caractéristique de la loi binomiale est donnée par

$$\phi_X(\omega) = \sum_{k=0}^{n} \mathrm{e}^{j\omega k}\cdot \mathcal{C}_k^n p^k q^{n-k} = \sum_{k=0}^{n} \mathcal{C}_k^n \left(p\mathrm{e}^{j\omega}\right)^k q^{n-k} = (p\mathrm{e}^{j\omega} + q)^n,$$

selon le binôme de Newton.

Exemple 4.20. *Soit* $X \sim \mathcal{N}(\mu, \sigma^2)$. *On peut montrer que*

$$\phi_X(\omega) = \exp\left(j\omega\mu - \frac{1}{2}\omega^2\sigma^2\right).$$

Soit $Y = aX + b$. *Quelle est la fonction caractéristique de* Y *? Par définition, on a*

$$\phi_Y(\omega) = \mathrm{E}\left[\, \mathrm{e}^{j\omega Y} \,\right] = \mathrm{E}\left[\mathrm{e}^{j\omega(aX+b)}\right] = \mathrm{e}^{j\omega b}\mathrm{E}\left[\, \mathrm{e}^{j\omega aX} \,\right] = \mathrm{e}^{j\omega b} \cdot \exp\left(j\omega a\mu + \frac{1}{2}\omega^2 a^2\sigma^2\right)$$

de sorte à obtenir

$$\phi_Y(\omega) = \exp\left[j\omega \underbrace{(a\mu + b)}_{\mu_Y} - \frac{1}{2}\omega^2 \underbrace{(a^2\sigma)^2}_{\sigma_Y^2}\right].$$

La fonction $\phi_Y(\omega)$ *est de la même forme que* $\phi_X(\omega)$, *alors on peut en déduire que* Y *suit aussi une loi normale.*

4.16 Fiabilité

On s'intéresse à la durée de vie des systèmes et de leurs composantes. Soit T une variable aléatoire, continue et non-négative, représentant une durée de vie. En général, l'espace d'échantillion est $S_T = [0, \infty[$.

On définie la fonction de fiabilité $R(T)$ tel que

$$R(t) = \mathbb{P}\left[\, T > t \,\right] = 1 - F_T(t)$$

et le taux de défaillance $r(t)$ comme

$$r(t) = \lim_{s \to t} f_T(s|T > t) = \frac{f_T(t)}{1 - F_T(t)} = -\frac{R\prime(t)}{R_(t)}$$

Sachant $r(t)$, il est possible de déterminer la fonction de fiabilité en trouvant la solution de l'équation différentielle suivante, soit

$$r = -\frac{R\prime}{R},$$

avec $R(0) = 1$, de sorte à obtenir

$$R(t) = \exp\left(-\int_0^t r(s)\,\mathrm{d}s\right).$$

Exemple 4.21. *Soit $T \sim \operatorname{Exp}(\lambda)$. Quel est le taux de défaillance ?*

On calcule la fonction de fiabilité avec

$$R(t) = 1 - F_T(t) = 1 - (1 - \mathrm{e}^{-\lambda t} = \mathrm{e}^{-\lambda t}$$

de sorte à obtenir le taux de défaillance

$$r(t) = \frac{\lambda \mathrm{e}^{-\lambda t}}{\mathrm{e}^{-\lambda t}} = \lambda.$$

On remarque que si $T \sim \operatorname{Exp}(\lambda)$, alors T a la propriété de non-vieillissement.

4.16.1 Durée de vie moyenne

La *durée de vie moyenne* est l'espérance de T et peut se calculer selon

$$\mathrm{E}[\,T\,] = \int_0^\infty R(t)\,\mathrm{d}t\,.$$

Démonstration. Par définition, on a

$$\mathrm{E}[\,T\,] = \int_0^\infty t \cdot f_T(t)\,\mathrm{d}t\,.$$

Hors, on peut exprimé t comme

$$t = \int_0^t \mathrm{d}s$$

de sorte à obtenir

$$\mathrm{E}[\,T\,] = \int_0^\infty \int_0^t f_T(t)\,\mathrm{d}s\,dt = \int_0^\infty \underbrace{\int_s^\infty f_T(t)\,\mathrm{d}t}_{\mathbb{P}[\,T>s\,]}\,\mathrm{d}s = \int_0^\infty R(t)\,\mathrm{d}t,$$

en inversant l'ordre d'intégration. □

Exemple 4.22. *Soit un système en série à n composantes. On suppose que les composantes fonctionnent indépendaments les un des autres. Quelle est la fonction de fiabilité du système ?*

La probabilité que le système fonctionne après un temps t est équivalent à la probabilité que tous les composantes fonctionnents après un temps t, soit

$$\mathbb{P}[\,T > t\,] = \mathbb{P}[\,\{T_1 > t\} \cap \cdots \cap \{T_N > t\}\,] = \mathbb{P}[\,T_1 > t\,] \cdots \mathbb{P}[\,T_n > t\,]\,,$$

car les événements sont indépendants.

Exemple 4.22 (suite). *Par conséquent, on obtient*

$$R(t) = \prod_{k=1}^{n} R_k(t).$$

Exemple 4.23. *Soit un système en parrallèle à n composantes. On suppose que les composantes fonctionnents indépendaments les un de sautres. Quelle est la fonction de fiablité dus système ?*

La probabilité que le système soit en panne est équivalent à la probabilité que tous les composantes soit en panne après un temps t, soit

$$\mathbb{P}\,[\,T \leq t\,] = \mathbb{P}\,[\,\{T_1 \leq t\} \cap \cdots \cap \{T_n \leq t\}\,] = \mathbb{P}\,[\,T_1 \leq t\,] \cdots \mathbb{P}\,[\,T_n \leq t\,]\,,$$

car les événements sont indépendants. Par conséquent, on obtient

$$R(t) = 1 - \prod_{k=1}^{n} \Big[1 - R_k(t)\Big].$$

5 Vecteurs aléatoires

Définition. Un *vecteur aléatoire* est formé de plusieurs variables aléatoires observées lors d'une même expérience.

Exemple 5.1. *Soit le lancer de 2 dés.*

On pose X comme étant le résultat du premier dé et Y comme étant le résultat du deuxième dé. On défini le vecteur aléatoire (X, Y). L'espace échantillion est $S_{X,Y} = \{(1,1), (1,2), \ldots, (6,5), (6,6)\}$.

La probabilité d'avoir un évènement est

$$\mathbb{P}\,[\,\{X = j\}] \cap \{Y = i\}\,] = \frac{1}{36},$$

pour $i = 1, \ldots, 6$ et $j = 1, \ldots, 6$.

Exemple 5.2. *On génère un point au hasard dans le triangle T de coordonnées $(0,0)$, $(1,1)$ et $(0,1)$.*

Chaque point dans le triangle peut être considéré comme une vecteur aléatoire (X,Y) où X est l'abscisse et Y l'ordonnée. Soit $f_{X,Y}(x,y)$ la fonction de masse du vecteur tel que

$$f_{X,Y}(x,y) = \begin{cases} c & \text{si } (x,y) \text{ est dans le triangle} \\ 0 & \text{sinon} \end{cases},$$

où la constante c est donnée par

$$\iint_T f_{X,Y}(x,y)\,\mathrm{d}A = 1 \Leftrightarrow c = 2.$$

5.1 Vecteur aléatoire discret

5.1.1 Fonction de masse conjointe

Théoreme 5.1. $p_{X,Y}(x_j, y_k) = \mathbb{P}\left[\,\{X = x_j\} \cap \{Y = y_k\}\,\right]$.

Théoreme 5.2. $p_{X,Y}(x_j, y_k) \geq 0$.

Théoreme 5.3. $\displaystyle\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} p_{X,Y}(x_j, y_k) = 1$.

5.1.2 Fonction de masse marginale

Théoreme 5.4. $\displaystyle p_X(x_j) = \sum_{k=0}^{\infty} p_{X,Y}(x_j, y_k)$.

Théoreme 5.5. $\displaystyle p_Y(y_k) = \sum_{j=0}^{\infty} p_{X,Y}(x_j, y_k)$.

5.1.3 Fonction de masse conditionnelle

Théoreme 5.6. $\displaystyle p_{Y|X}(y_k|x_j) = \frac{p_{X,Y}(x_j, y_k)}{p_X(x_j)}$.

Théoreme 5.7. $\displaystyle p_{X|Y}(x_j|y_k) = \frac{p_{X,Y}(x_j, y_k)}{p_Y(y_k)}$.

Théoreme 5.8. $\displaystyle \mathbb{P}\left[\,A\,\right] = \sum\sum_{A} p_{X,Y}(x_j, y_k)$.

5.2 Vecteur aléatoire continu

5.2.1 Fonction de densité conjointe

Théoreme 5.9. $f_{X,Y}(x,y) = \dfrac{1}{\delta\epsilon}\mathbb{P}\left[\left\{x - \dfrac{\delta}{2} \leq X \leq x + \dfrac{\delta}{2}\right\} \cap \left\{y - \dfrac{\epsilon}{2} \leq Y \leq y + \dfrac{\epsilon}{2}\right\}\right]$

Théoreme 5.10. $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y = 1.$

5.2.2 Fonction de densité marginale

Théoreme 5.11. $f_X(x) = \displaystyle\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y.$

Théoreme 5.12. $f_Y(y) = \displaystyle\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x.$

5.2.3 Fonction de densité conditionnelle

Théoreme 5.13. $f_{Y|X}(y|x) = \dfrac{f_{X,Y}(x,y)}{f_X(x)}.$

Théoreme 5.14. $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}.$

Théoreme 5.15. $\mathbb{P}[\,A\,] = \displaystyle\iint_A f_{X,Y}(x,y)\,\mathrm{d}A.$

5.3 Fonction de répartition conjointe

Théoreme 5.16. $F_{X,Y}(x,y) = \mathbb{P}[\,\{X \leq x\} \cap \{Y \leq y\}\,].$

Théoreme 5.17. $F_{X,Y}(x,y) = \displaystyle\sum_{x_j \leq x}\sum_{y_k \leq y} p_{X,Y}(x_j, y_k)$ *dans le cas discret.*

Théoreme 5.18. $F_{X,Y}(x,y) = \displaystyle\int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\,\mathrm{d}t\mathrm{d}s$ *dans le cas continue.*

Théoreme 5.19. $f_{X,Y}(x,y) = \dfrac{\partial^2}{\partial x\,\partial y}F_{X,Y}(x,y).$

5.4 Fonction de répartition marginale

Théoreme 5.20. $F_X(x) = \displaystyle\lim_{y\to\infty} F_{X,Y}(x,y).$

Théoreme 5.21. $F_Y(y) = \displaystyle\lim_{x\to\infty} F_{X,Y}(x,y).$

Exemple 5.3. *Quelle est la probabilité que (X,Y) soit dans un rectangle R ? On suppose qu'on connait $F_{X,Y}(x,y) = \mathbb{P}[\{X \leq x\} \cap \{Y \leq y\}]$.*

$$\mathbb{P}[R] = F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + 2F_{X,Y}(a,c).$$

5.5 Indépendance dans un vecteur

Définition. Si X et Y sont des variables aléatoires *indépendantes*, il faut que

1. $S_{X,Y} = S_X \times S_Y$
2. $F_{X,Y}(x,y) = F_X(x)F_Y(y)$
   $\Rightarrow\ p_{X,Y}(x_j,y_k) = p_X(x_j)p_Y(y_k)$
   $\Rightarrow\ f_{X,Y}(x,y) = f_X(x)f_Y(y)$

Si on veut déterminer s'il y a indépendance ou non entre deux variables aléatoires X et Y, il faut vérifier les deux conditions. Si elles sont respectées, alors X et Y sont des variables aléatoires indépendantes.

Exemple 5.4. *La veille d'un examen, un professeur estime qu'il recevra X questions par courriel, où $X \sim \mathrm{Poi}(\alpha)$. Le professeur répond à chaque question avec une probabilité p, et ce indépendamment d'une question à l'autre. Soit Y le nombre de réponses du professeur. Déterminer $p_Y(y)$.*

L'espace échantillion du vecteur aléatoire discrète (X,Y) est

$$S_{X,Y} = \{(0,0),(1,0),(1,1),\ldots,(j,0),\ldots,(j,j)\},$$

soit un triangle, où j est le nombre de questions reçus.

$$p_{X|Y}(j|k) = \underbrace{p_{Y|X}(k,j)}_{\mathcal{B}(j,p)}\underbrace{p_X(j)}_{\mathrm{Poi}(\alpha)} = \mathcal{C}_j^k p^k q^{j-k}\frac{\mathrm{e}^{-\alpha}}{j!}\alpha^j$$

$$p_Y(k) = \sum_{j=k}^{\infty} p_{X,Y}(j,k) = \frac{\mathrm{e}^{-\alpha}p^k}{k!}\sum_{k=0}^{\infty}\frac{\alpha^j q^{j-k}}{(j-k)!} = \frac{\mathrm{e}^{-\alpha}(\alpha p)^k\mathrm{e}^{\alpha q}}{k!} = \frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}$$

5.6 Espérance conditionelle

Théoreme 5.22. $\mathrm{E}[Y|X=x_j] = \displaystyle\sum_{k=1}^{\infty} y_k p_{Y|X}(y_k|x_j)$ *dans un cas discret.*

Théoreme 5.23. $\mathrm{E}[Y|X=x_j] = \displaystyle\int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,\mathrm{d}x$ *dans un cas continu.*

En général, $\mathrm{E}[Y|X=x] = g_1(x)$ et $\mathrm{E}[X|Y=y] = g_2(y)$. On a donc, $\mathrm{E}[Y|X] = g_1(X)$ et $\mathrm{E}[X|Y] = g_2(Y)$, c'est-à-dire que l'espérance conditionelle est une variable aléatoire.

Théoreme 5.24. $\mathrm{E}[Y] = \mathrm{E}[\mathrm{E}[Y|X]]$.

Démonstration. On sait que

$$\mathrm{E}[\mathrm{E}[Y|X])] = \int_{-\infty}^{\infty} \mathrm{E}[Y|X=x] f_X(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,\mathrm{d}y \cdot f_X(x)\,\mathrm{d}x$$

$$\mathrm{E}[\mathrm{E}[Y|X]] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \cdot \underbrace{f_{Y|X}(y|x) f_X(x)}_{f_{X,Y}(x,y)}\,\mathrm{d}y\mathrm{d}x = \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y$$

$$\mathrm{E}[\mathrm{E}[Y|X]] = \int_{-\infty}^{\infty} y f_Y(y)\,\mathrm{d}y = \mathrm{E}[Y]. \qquad \square$$

Théoreme 5.25. $\mathrm{E}[h(Y)] = \mathrm{E}[\mathrm{E}[h(y)|X]]$.

5.7 Variance conditionnelle

Théoreme 5.26. $\mathrm{Var}[Y|X] = \mathrm{E}\left[Y^2\middle|X\right] - (\mathrm{E}[Y|X])^2$.

Théoreme 5.27. $\mathrm{Var}[Y] = \mathrm{E}[\mathrm{Var}[Y|X]] + \mathrm{Var}[\mathrm{E}[Y|X]]$.

5.8 Covariance et corrélation

Théoreme 5.28. $\mathrm{E}[g(X,Y)] = \displaystyle\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} g(x_j, y_k) p_{X,Y}(x_j, y_k)$ *dans un cas discret.*

Théoreme 5.29. $\mathrm{E}[g(X,Y)] = \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y$ *dans un cas discret.*

Exemple 5.5. *Soit la fonction de masse conjointe de X et Y suivante.*

	0	1	2
0	$1/6$	$1/6$	$1/6$
1	0	$1/6$	$1/6$
2	0	0	$1/6$

$$\mathrm{E}\left[X^2Y\right] = \sum_{j=0}^{2}\sum_{k=0}^{2} j^2 k p_{X,Y}(j,k) = \frac{13}{6}$$

Exemple 5.6. *Soit $X \sim \mathcal{N}(1,1)$ et $Y \sim \mathcal{N}(2,1)$ des variables aléatoires indépendates de sorte que*

$$f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

Calculer $\mathrm{E}[X+Y]$ *et* $\mathrm{E}[XY]$.

Par définition, on a

$$\mathrm{E}[X+Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y,$$

et avec la linéarité de l'espérance, on obtient

$$\mathrm{E}[X+Y] = \underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y}_{\mathrm{E}[X]} + \underbrace{\int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y}_{\mathrm{E}[Y]}$$

de sorte à obtenir que $\mathrm{E}[X+Y] = 3$*. L'indépendance des variables n'est pas importante.*

Par définition, on a

$$\mathrm{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy)f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y,$$

et avec l'indépendance des variables, on obtient

$$\mathrm{E}[XY] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x \int_{-\infty}^{\infty} y f_Y(y)\,\mathrm{d}y = \mathrm{E}[X]\,\mathrm{E}[Y]$$

de sorte à obtenir $\mathrm{E}[XY] = 2$.

Théoreme 5.30. $\mathrm{E}[aX+bY] = a\mathrm{E}[X] + b\mathrm{E}[Y]$.

Théoreme 5.31. $\mathrm{E}[g(X,Y)] = \mathrm{E}[g_1(X)]\,\mathrm{E}[g_2(Y)]$ *si* X *et* Y *indépendants et* $g(X,Y) = g_1(X)g_2(Y)$.

Définition. La *corrélation* de deux variables aléatoires X et Y est l'espérance de leur produit, soit $\mathrm{E}[XY]$.

Lorsque $\mathrm{E}[XY] = 0$, on dit que X et Y sont orthogonales.

Définition. La *covariance* de deux variables aléatoires X et Y est donnée par

$$\mathrm{Cov}[X,Y] = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y].$$

Théoreme 5.32. $\mathrm{Cov}[X,X] = \mathrm{Var}[X]$.

Définition. Le *coefficient de corrélation* $\rho_{X,Y}$ de deux variables aléatoires X et Y est donné par

$$\rho_{X,Y} = \frac{\mathrm{Cov}[X,Y]}{\mathrm{Std}[X]\,\mathrm{Std}[Y]}.$$

Théoreme 5.33. $-1 \leq \rho_{X,Y} \leq 1$.

Théoreme 5.34. $\rho_{X,Y} = 1 \Leftrightarrow Y = aX + b$ *où* $a > 0$.

Théoreme 5.35. $\rho_{X,Y} = -1 \Leftrightarrow Y = aX + b$ *où* $a < 0$.

Théoreme 5.36. $\rho_{X,Y} = 0$ *si* X *et* Y *sont indépendants.*

5.9 Loi binormale

Soit un vecteur aléatoire (X, Y), où la fonction de probabilité de densité suit l'équation

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{x-\mu_Y}{\sigma_Y}\right)^2 + 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\},$$

suit une loi binormale dénotée

$$(X, Y) \sim \mathcal{N}\left(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho\right).$$

où $\rho = \rho_{X,Y}$.

Il en résulte que $X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$ $Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$. Si $\rho = 0$, alors on voit que $f_{X,Y} = f_X(x)f_Y(y)$ de sorte que X et Y sont indépendants.

La conditionelle de cette loi est une loi normale, soit

$$X|\{Y = y\} \sim \mathcal{N}\left(\mu, \sigma^2\right),$$

où

$$\mu = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y)$$

et

$$\sigma^2 = \sigma_X^2(1-\rho^2).$$

5.10 Estimation

On cherche à prévoir la valeur inconnue de Y à partir de X. Soit $g(X)$ la fonction estimateur de Y. On cherche à minimiser l'erreur quadratique moyenne définie par

$$\mathrm{E}\left[(Y - g(X))^2\right].$$

5.10.1 Estimateur constant

Dans un cas où $g(X) = c$, le problème se résume à

$$\min \mathrm{E}\left[(Y - c)^2\right].$$

En dérivant, on obtient la constante optimale

$$\hat{c} = \mathrm{E}[Y],$$

de sorte à obtenir l'erreur d'estimation

$$\mathrm{E}\left[(Y - \mathrm{E}[Y])^2\right] = \mathrm{Var}[Y].$$

5.10.2 Estimateur linéaire

Dans un cas où $g(X) = aX + b$, le problème se résume à

$$\min \mathrm{E}\left[(Y - aX - b)^2 \right]$$

de sorte à obtenir

$$\hat{a} = \frac{\mathrm{Cov}\,[X, Y]}{\mathrm{Var}\,[\,X\,]} = \frac{\mathrm{Std}\,[\,Y\,]}{\mathrm{Std}\,[\,X\,]}\rho_{X,Y}$$

et

$$\hat{b} = \mathrm{E}\,[\,Y\,] - \hat{a}\mathrm{E}\,[\,X\,]\,.$$

Par conséquent, le meilleur estimateur linéaire de Y en fonction de X est

$$g(X) = \hat{a}X + \hat{b}$$

avec une erreur d'estimation donnée par

$$\mathrm{Var}\,[\,Y\,]\left(1 - \rho_{X,Y}^2\right).$$

5.10.3 Estimateur non linéaire

$$\min \mathrm{E}\left[(Y - g(X))^2 \right]$$

$$\min \mathrm{E}\left[\mathrm{E}\left[(Y - g(X))^2 \middle| X \right] \right]$$

$$\min \sum_{j=1}^{\infty} \underbrace{\mathrm{E}\,[\,Y - g(X) | X = x_j\,]}_{\text{fonction positive à min } \forall x_j} p_X(x_j)$$

On minimise pour chaque x_j avec une constante différente. On retrouve un cas $g(X) = c$, mais conditionnel à $X = x_j$. On retrouve que le meilleur estimateur de Y en fonction de X est

$$g(X) = \mathrm{E}\,[\,Y|X\,]\,.$$

5.10.4 Estimateur d'une binormale

Lorsque X et Y suivents des distributions normales, alors on retrouve que

$$\mathrm{E}\,[\,Y|X\,] = \hat{a}X + \hat{b}.$$

Exemple 5.7. *Soit* $(X, Y) \sim \mathcal{N}\left(\mu_X = 3, \mu_Y = 1, \sigma_X^2 = 4, \sigma_Y^2 = 9, \rho = 1/4\right)$, *quel est le meilleur estimateur de* Y *en fonction de* X *?*

$$\hat{a} = \frac{\mathrm{Std}\,[\,Y\,]}{\mathrm{Std}\,[\,X\,]}\rho = \frac{3}{8}$$

$$\hat{b} = \mathrm{E}[Y] - \hat{a}\mathrm{E}[X] = -\frac{1}{8}$$

$$g(X) = \frac{3}{8}X - \frac{1}{8}$$

5.11 Combinaisons linéaires

Définition. Soit un vecteur aléatoire $(X_1, X_2, \ldots, X_n)$, alors la *combinaison linéaire* Z des variables aléatoires est

$$Z = a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n = a_0 + \sum_{i=1}^{n} a_iX_i.$$

Théoreme 5.37. $\mathrm{E}[Z] = a_0 + \sum_{i=1}^{n} a_i\mathrm{E}[X_i]$.

Théoreme 5.38. $\mathrm{Var}[Z] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j\mathrm{Cov}[X_i, X_j]$.

Démonstration. Soit $Y_i = a_iX_i$. Par définition, on a

$$\mathrm{Var}[Z] = \mathrm{Var}\left[\sum_{i=1}^{n} Y_i\right] = \mathrm{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^2\right] - \left(\sum_{i=1}^{n} \mathrm{E}[Y_i]\right)^2.$$

En écrivant explicitement le carré, on a

$$\mathrm{Var}[Z] = \mathrm{E}\left[\left(\sum_{i=1}^{n} Y_i\right)\left(\sum_{j=1}^{n} Y_j\right)\right] - \left(\sum_{i=1}^{n} \mathrm{E}[Y_i]\right)\left(\sum_{j=1}^{n} \mathrm{E}[Y_j]\right)$$

$$\mathrm{Var}[Z] = \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} Y_iY_j\right] - \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[Y_i]\,\mathrm{E}[Y_j]$$

$$\mathrm{Var}[Z] = \sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\mathrm{E}[Y_iY_j] - \mathrm{E}[Y_i]\,\mathrm{E}[Y_j]\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j\mathrm{Cov}[X_i, X_j] \qquad \square$$

Théoreme 5.39. $\mathrm{Var}[Z] = \sum_{i=1}^{n} a_i^2\mathrm{Var}[X_i] + 2\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}}_{i<j} a_ia_j\mathrm{Cov}[X_i, X_j]$.

5.12 Variables aléatoires indépendantes

Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes.

Théoreme 5.40. $\mathrm{Cov}\,[X_i, X_j] = 0$.

Théoreme 5.41. $\mathrm{E}\,[\,Z\,] = a_0 + \sum_{i=1}^{n} \mathrm{E}\,[\,X_i\,]$.

Théoreme 5.42. $\mathrm{Var}\,[\,Z\,] = \sum_{i=1}^{n} a_i^2 \mathrm{Var}\,[\,X_i\,]$.

5.13 Variables aléatoires indépendantes et identiquement distributées

Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes et identiquement distribuées de sorte que $Z = X_1 + X_2 + \cdots + X_n$.

Théoreme 5.43. $\mathrm{E}\,[\,S_n\,] = n\mathrm{E}\,[\,X_1\,]$.

Théoreme 5.44. $\mathrm{Var}\,[\,S_n\,] = n\mathrm{Var}\,[\,X_1\,]$.

Exemple 5.8. *Si* $X_1 \sim \mathcal{B}\,(n_1, p)$ *et* $X_2 \sim \mathcal{B}\,(n_2, p)$ *sont des variables aléatoires indépendantes, alors on a*

$$X_1 + X_2 \sim \mathcal{B}\,(n_1 + n_2, p)\,.$$

Exemple 5.9. *Si* $X_i \sim \mathrm{Poi}\,(\alpha_i)$ *sont des variables aléatoires indépendantes, alors on a*

$$X_1 + X_2 + \cdots + X_n \sim \mathrm{Poi}\,(\alpha)\,,$$

où $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

Exemple 5.10. *Si* $X_i \sim \mathrm{Exp}\,(\lambda)$ *sont des variables aléatoires indépendantes, alors on a*

$$X_1 + X_2 + \cdots + X_n \sim \mathrm{Gam}\,(n, \lambda)\,.$$

Exemple 5.11. *Si* $X_i \sim \mathrm{Gam}\,(\alpha_i, \lambda)$ *sont des variables aléatoires indépendantes, alors on a*

$$X_1 + X_2 + \cdots + X_n \sim \mathrm{Gam}\,(\alpha, \lambda)\,,$$

où $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

Exemple 5.12. *Si $X_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$ sont des variables aléatoires indépendantes, alors on a*

$$a_0 + \sum_{i=1}^{n} a_i X_i \sim \mathcal{N}\left(\mu, \sigma^2\right),$$

où

$$\mu = a_0 + \sum_{i=1}^{n} a_i \mu_i \quad \textit{et} \quad \sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \mathrm{Cov}\left[X_i, X_j\right].$$

5.14 Loi faible des grands nombres

Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes et indentiquement distribuées de moyenne μ. Soit $S_n = X_1 + X_2 + \cdots + X_n$, alors on a

$$\lim_{n \to \infty} \mathbb{P}\left[\left|\frac{S_n}{n} - \mu\right| < c\right] = 1,$$

pour tout $c > 0$.

5.15 Loi forte des grands nombres

Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes et indentiquement distribuées de moyenne μ et de variance σ^2. Soit $S_n = X_1 + X_2 + \cdots + X_n$, alors on a

$$\mathbb{P}\left[\lim_{n \to \infty} \frac{S_n}{n} = \mu\right] = 1.$$

5.16 Théorème central limite

Soit $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes et identiquement distribuées de moyenne μ et de variance σ^2. Si n est *grand*, on a que $S_n \approx \mathcal{N}\left(n\mu, n\sigma^2\right)$ ou encore

$$\frac{S_n}{n} \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right),$$

et

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \approx \mathcal{N}(0,1).$$

En général, le théorème central limite est applicable si $n \geq 30$.

Exemple 5.13. *Soit $X \sim \mathcal{B}(2000, 1/2)$. Calculez $\mathbb{P}[\,990 \leq X \leq 1100\,]$.*

On approxime la loi binomiale par la loi normale. Par conséquent,

$$X \approx \mathcal{N}(np, npq)$$

et

$$\mathbb{P}[\,X = k\,] \approx \frac{1}{\sqrt{2\pi \cdot npq}} \exp\left\{-\frac{(k-np)^2}{2npq}\right\}$$

Puisqu'on approxime une loi discrète par une loi continue, on applique une correction de continuité, c'est-à-dire

$$\mathbb{P}[\,a \le X \le b\,] \approx \Phi\left(\frac{b + 1/2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - 1/2 - np}{\sqrt{npq}}\right)$$

$$\mathbb{P}[\,a < X \le b\,] \approx \Phi\left(\frac{b + 1/2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a + 1/2 - np}{\sqrt{npq}}\right)$$

Exemple 5.14. *Soit la fonction de probabilité de densité suivante*

$$f_{X,Y}(x,y) = \frac{3}{4},$$

si $0 < y < 1$ *et* $x^2 < y$. *Calculez* $f_X(x)$ *et* $f_Y(y)$, *ainsi que* $\mathbb{P}[\,Y \ge X\,]$.

Pour un x, *on a*

$$f_X(x) = \int_{x^2}^{1} f_{X,Y}(x,y)\,\mathrm{d}y = \frac{3}{4}(1 - x^2),$$

si $-1 < x < 1$. *Pour un* y, *on a*

$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} f_{X,Y}(x,y)\,\mathrm{d}x = \frac{3}{2}\sqrt{y},$$

si $0 < y < 1$.

Par définition, on a

$$\mathbb{P}[\,Y \ge X\,] = \int_0^1 \int_{-\sqrt{y}}^{\sqrt{y}} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y = \frac{7}{8}.$$

6 Processus Stochastique

Définition. Un *processus stochastique* est une expérience aléatoire qui ce déroule dans le temps.

On note $X(t)$ la variable aléatoire correspondant à l'observation du processus au un temps $t \ge 0$. Les valeurs possibles pour ces observations forment *l'espace des états* du processus. Les événements élémentaires du processus sont ce qu'on appel des *trajectoires* dans l'espace des états.

Un processus stochastique est défini par l'ensemble $\{X(t) | t \in T\}$, avec $T = \{0, 1, 2, \dots\}$ dans un cas discret ou $T = [0, \infty[$ dans un cas continue.

Il existe 4 types de processus stochastique :

1. processus stochastique à temps discret et état discret (PSTDED)

   Exemple. *Le nombre de personnes en classe à chaque jour.*

2. processus stochastique à temps discret et état continu (PSTDEC)

   Exemple. *La valeur d'une action à la fermmeture de la bourse.*

3. processus stochastique à temps continu et étant discret (PSTCED)

   Exemple. *La longueur de la file d'attente à la cafétéria.*

4. processus stochastique à temps discret et état discret (PSTCEC)

   Exemple. *La valeur d'une action en temps réel.*

6.1 Caractéristique des processus stochastiques

6.1.1 PSTCEC

Dans le cas d'un processus stochastique à temps continu et état continu,

Dans le cas d'un PSTCEC, il y a une fonction de répartition du 1er ordre, soit $F(x;t) = \mathbb{P}\left[\, X(t) \leq x \,\right]$, une fonction de répartition du 2e ordre, soit $F(x_1, X_2; t_1, t_2) = \mathbb{P}\left[\, X(t_1) \leq x_1, X(t_2) \leq x_2 \,\right]$.

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = \mathbb{P}\left[\, \{X(t_1) \leq x_1\} \cap \dots \cap \{X(t_n) \leq x_n\} \cap \,\right]$$

$$f(x_1, \dots, x_n; t_1, \dots, t_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F(x_1, \dots, x_n; t_1, \dots, t_n)$$

Exemple 6.1. *Soit un patient qui recoit une dose de médicament Y. Au temps $t = 0$, on a $Y \sim \mathcal{U}\left(0, 1\right)$. La quantité active X après t unités de temps est définie par $X(t) = Y\mathrm{e}^{-t}$. Quel est la fonction de répartition du premier ordre, soit $F(x;t)$?*

Par définition, on a $F(x;t) = \mathbb{P}\left[\, X(t) \leq x \,\right] = \mathbb{P}\left[\, Y \,\right] \leq x\mathrm{e}^t$. Pour une loi uniforme, on a

$$F(x;t) = \begin{cases} 0 & si \quad x\mathrm{e}^t < 0 \\ x\mathrm{e}^t & si \quad 0 \leq x \leq e^{-t} \\ 1 & si \quad \mathrm{e}^{-t} < x \end{cases}.$$

6.1.2 Moyenne d'un processus stochastique

Définition. La *moyenne* d'un processus stochastique, dénotée $m_X(t)$, est donnée par

$$m_X(t) = \mathrm{E}\left[\, X(t) \,\right] = \int_{-\infty}^{\infty} x f(x;t) \, \mathrm{d}x \,.$$

Exemple 6.2. *Quel est la moyenne de $X(t) = Y\mathrm{e}^{-t}$, où $Y \sim \mathcal{U}(0,1)$?*

Avec l'exemple précèdent, on obtient que la fonction de densité de probabilité est donnée par

$$f(x;t) = \frac{\partial}{\partial x} F(x;t) = \mathrm{e}^{t},$$

pour $0 \leq x \leq e^{-t}$. La moyenne est donc

$$m_X(t) = \int_0^{e^{-t}} x\mathrm{e}^{t}\,\mathrm{d}x = \frac{e^{-t}}{2}.$$

6.1.3 Fonction d'autocorrélation

Définition. La *fonction d'autocorrélation*, dénotée $R_X(t_1,t_2)$, est la corrélation d'un processus stochastique $X(t)$ avec lui-même à deux temps différents. Par conséquent, on a

$$R_X(t_1,t_2) = \mathrm{E}\left[\, X(t_1)X(t_2) \,\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1,x_2;t_1,t_2)\,\mathrm{d}x_1\mathrm{d}x_2\,.$$

6.1.4 Fonction d'autocovariance

Définition. La *fonction d'autocovariance*, dénotée $C_X(t_1,t_2)$, est la covariance d'un processus stochastique $X(t)$ avec lui-même à deux temps différents. Par conséquent, on a

$$C_X(t_2,t_2) = R_X(t_1,t_2) - m_X(t_1)m_X(t_2).$$

6.1.5 Processus stochastique stationnaire au sens large

Définition. Si $m_X(t) = c$, où c est une constante, et si $R_X(t_1,t_2) = h(t_2 - t_1)$, où h est une fonction, alors on dit que le processus stochastique est *stationnaire au sens large* (SSL).

Exemple 6.3. *Soit un signal aléatoire $X(t) = Y\sin(t+Z)$, où Y et Z sont des variables aléatoires indépendantes et $Z \sim \mathcal{U}(0,2\pi)$. Montrer que ce signal est SSL.*

Par définition, la moyenne est donnée par

$$m_X(t) = \mathrm{E}\left[\, Y\sin(t+Z) \,\right] = \mathrm{E}\left[\, Y \,\right] \underbrace{\mathrm{E}\left[\, \sin(t+Z) \,\right]}_{\int_0^{2\pi}\sin(t+z)\frac{1}{2\pi}=0} \mathrm{d}z = 0,$$

ce qui est une constante. De plus, la fonction d'autocorrélation est donnée par

$$R_X(t_1,t_2) = \mathrm{E}\left[\, Y\sin(t_1+Z)Y\sin(t_2+Z) \,\right] = \mathrm{E}\left[\, Y^2 \,\right]\mathrm{E}\left[\, \sin(t_1+Z)\sin(t_2+Z) \,\right],$$

et avec une identité trigonométrique, on obtient

$$R_X(t_1, t_2) = \frac{1}{2}\mathrm{E}\left[\,Y^2\,\right]\left(\mathrm{E}\left[\cos(t_1 - t_2)\right] - \underbrace{\mathrm{E}\left[\cos(t_1 - t_2 + 2Z)\right]}_{0}\right)$$

$$= \frac{1}{2}\mathrm{E}\left[\,Y^2\,\right]\cos(t_1 - t_2) = h(t_2 - t_1).$$

Par conséquent, ce signal est stationnaire au sens large.

6.2 Chaîne de Markov

Définition. Une *chaîne de Markov* est un processus stochastique à temps discret et état continu où :

1. l'espace des états est un ensemble de nombres entiers
2. on dénote X_n l'état au temps $n = 0, 1, \ldots$
3. la probabilité de passer de l'état i à l'état j suivant ne dépend ni de n ni de la trajectoire passée

On dénote $p_{i,j}$ la probabilité de passer de l'état i vers l'état j comme étant la probabilité conditionnelle suivante, soit

$$p_{i,j} = \mathbb{P}\left[\,X_{n+1} = j | X_n = i\,\right].$$

6.2.1 Représentation graphique

Une chaîne de Markov peut se représenter comme un graphe, où les noeuds sont les états et les segments sont les probabilité $p_{i,j}$. La figure 6 représente le graphe d'une chaîne de Markov ayant 4 états.

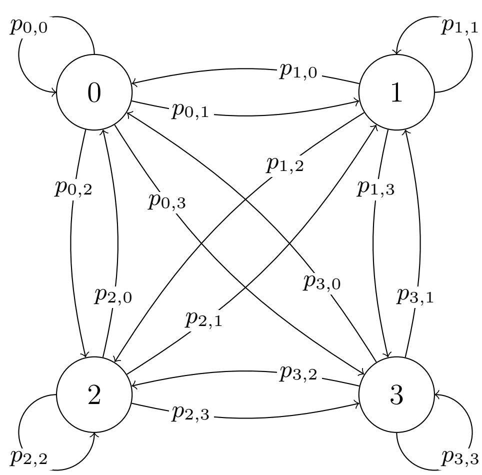


FIGURE 6 – Chaîne de Markov à 4 états

6.2.2 Représentation matricielle

Soit P la matrice des probabilités de transition en une étape où chaque cellule $p_{i,j}$ sont les probabilités de passer de l'état i vers l'état j tel que

$$P = \begin{bmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,n} \\ p_{1,0} & p_{1,1} & \cdots & p_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,0} & p_{n,1} & \cdots & p_{n,n} \end{bmatrix}.$$

Exemple 6.4. *On établie des prévisions météorologiques. On défini 3 états suivants : une journée ensoleillée (état 0), une journée nuageuse (état 1) et une journée pluvieuse (état 2). Les probabilités de transition sont définies par la matrice suivante :*

$$P = \begin{bmatrix} 0.40 & 0.35 & 0.25 \\ 0.30 & 0.20 & 0.50 \\ 0.75 & 0.10 & 0.15 \end{bmatrix}$$

Quelle est la probabilité qu'il pleuve le lendemain d'une journée ensoleillée ? Quelle est la probabilité que trois journées ensoleillées suivent une journée nuageuse ? Quelle est la probabilité qu'il ne pleuve pas pendant les deux jours suivant une journée pluvieuse ?

On cherche $p_{0,2}$ dans la matrice, alors la probabilité est de 25%.

On cherche $p_{1,0}p_{0,0}p_{0,0}$, alors la probabilité est de 4.8%.

Il faut calculer toutes les trajectoires possibles. Il y a 4 cas possibles On cherche $p_{2,0}p_{0,0} + p_{2,0}p_{0,1} + p_{2,1}p_{1,0} + p_{2,1}p_{1,1}$, alors la probabilité est 61.25%.

6.2.3 Probabilité de transition en n étapes

La probabilité d'atteindre un état j à partir d'un état i en n étapes, dénotée $p^n_{i,j} = \mathbb{P}\left[X_{m+n} = j | X_m = i \right]$, est donnée en multipliant n fois la matrice des probabilités P, c'est à dire

$$P^{(n)} = \underbrace{P \cdot P \cdots P}_{n \text{ fois}}.$$

6.3 Processus de Poisson

Définition. Un processus de poisson est un processus stochastique à temps continue et état discret. On dénote $N(t)$ le nombre d'événements qui se sont produits dans l'intervalle de temps $[0, t]$.

Trois conditions fondamentales définissent le processus de Poisson :

1. $N(0) = 0$
2. Soit les intervalles $]t_1, t_2]$ et $]t_3, t_4]$, avec $t_1 < t_2 \leq t_3 < t_4$, et les variables aléatoires indépendantes $X = N(t_2) - N(t_1)$ et $Y = N(t_4) - N(t_3)$. On dit que les *accroissements* du processus de Poisson sont indépendants.

3. Soit $X = N(\tau + t) - N(\tau)$ l'accroissement dans l'intervalle $]\tau, \tau + t]$. Alors, on a que $X \sim \text{Poi}(\lambda t)$, où λ est le taux du processus de poisson. On dit que les accroissements sont *stationnaire*.

Théoreme 6.1. $C_N(t_1, t_2) = \lambda \cdot min(t_1, t_2)$.

Exemple 6.5. *Soit un processus de Poisson de taux* $\lambda = 2$ *avec* $t_1 = 3$, $t_2 = 7$, $t_3 = 8$ *et* $t_4 = 10$. *Quelle est la probabilité d'avoir 1 événement dans l'intervalle* $]t_1, t_2]$ *et plus de 1 événement dans l'intervalle* $]t_3, t_4]$ *?*

On définie X *et* Y *comme étant le nombre d'événements dans* $]t_1, t_2]$ *et* $]t_3, t_4]$ *respectivement, alors* $X \sim \text{Poi}(\lambda t)$ *et* $Y \sim \text{Poi}(\lambda t)$. *Par conséquent, on cherche*

$$\mathbb{P}[\{X = 1\} \cap \{Y \geq 1\}] = \mathbb{P}[X = 1]\,\mathbb{P}[Y \geq 1] = 8\mathrm{e}^{-8}\left(1 - \mathrm{e}^{-4}\right).$$

6.3.1 Temps d'arrivé

Soit T_1 le temps d'arrivée du premier événement. On cherche

$$\mathbb{P}[T_1 > t] = \mathbb{P}[N(t) = 0] = \mathrm{e}^{-\lambda t}$$

Par conséquent,

$$T_1 \sim \text{Exp}(\lambda)$$

Soit T_2 le temps entre le premier et le deuxième événement. On a que $T_1 = t_1$ lorsque le premier événement ce produit. On définit

$$M(t) = N(t + t_1) - \underbrace{N(t_1)}_{1},$$

le nombre d'événements dans $[t_1, t]$. On remarque que $M(t)$ est aussi un processus de Poisson de sorte que $T_2 \sim \text{Exp}(\lambda)$.

Soit T_k le temps entre $(k-1)$-ième et le k-ième événement. Par conséquent, $T_k \sim \text{Exp}(\lambda)$ avec $T_1, T_2, \ldots$ sont des variables aléatoires indépendantes et identiquement distributées.

Soit

$$S_n = \sum_{k=1}^{n} T_k,$$

le temps d'arrivée du n-ième événement. Par conséquent,

$$S_n \sim \text{Gam}(n, \lambda).$$

Exemple 6.6. *En observant un processus de Poisson de taux* λ, *on remarque qu'un seul événement s'est produit avant le temps* t. *Quelle est la distribution du temps d'arrivée* T *de cet événement ?*

Le problème s'écrit en terme d'une conditionnelle, soit

$$\begin{aligned}\mathbb{P}[\,T \le \tau | N(t) = 1\,] = \mathbb{P}[\,N(\tau) = 1 | N(t) = 1\,] &= \frac{\mathbb{P}[\,\{N(\tau) = 1\} \cap \{N(t) = 1\}\,]}{\mathbb{P}[\,N(t) = 1\,]} \\ &= \frac{\mathbb{P}[\,\{X = 1\} \cap \{Y = 0\}\,]}{\mathbb{P}[\,N(t) = 1\,]} = \frac{\tau}{t}\end{aligned}$$

$N(\tau) = X$, $N(t) = X + Y$, $X \sim \text{Poi}\,(\lambda\tau)$, $Y \sim \text{Poi}\,(\lambda(t-\tau))$

6.4 Processus de Wiener

Définition. Un *processus de Wiener* est un processus stochastique à temps continu et état continu où $W(t)$ représente la position aléatoire d'une particule à un temps t le long d'un axe.

On construit ce processus à partir de la position $X(t)$ de la particule soumise à une marche aléatoire avec des pas de $\pm\epsilon$ de distance à chaque δ unité de temps. On pose $\epsilon = \sigma\sqrt{\delta}$, où σ est une constante, on obtient

$$W(t) = \lim_{\delta \to 0^+} X(t)$$

On définit 3 conditions pour qu'un processus stochastique soit un processus de Wiener :

1. $W(0) = 0$
2. les accroissements du processus sont indépendants et stationnaires
3. $W(t) \sim \mathcal{N}\left(0, \sigma^2 t\right)$

On appel σ^2 le *coefficient de diffusion*. Si $\sigma^2 = 1$, on appel ce processus le *mouvement brownien standard* et on dénote parfois $B(t) = W(t)$. Les trajectoires de ce processus sont des fonctions continues, mais dérivables nulle part, comme une fractale. La *dérivée généralisée* de ce processus s'appele *bruit blanc gaussien.*

Exemple 6.7. *Calculer la fonction de densité du 2-ième ordre $f(w_1, w_2, t_1, t_2)$ du processus de Wiener. On suppose que $t_1 < t_2$.*

On peut écrire la densité de l'événement suivant en terme d'accroissements, soit

$$\{W(t_1) = w_1, W(t_2) = w_2\} \equiv \{W(t_1) = w_1, W(t_2) - W(t_1) = w_2 - w_1\}\,.$$

On définie $X = W(t_1) \sim \mathcal{N}\left(0, \sigma^2 t_1\right)$ et $Y = W(t_2) - W(t_1) \sim \mathcal{N}\left(0, \sigma^2 t_2\right)$. Par conséquent, on a

$$\begin{aligned}f(w_1, w_2; t_1, t_2) &= f_X(w_1) f_Y(w_2 - w_1) \\ &= \frac{1}{\sqrt{2\pi\sigma^2 t_1}} \exp -\frac{(w_1^2)}{2\sigma^2 t_1} \cdot \frac{1}{\sqrt{2\pi\sigma^2 (t_2 - t_1)}} \exp -\frac{(w_2 - w_1)^2}{2\sigma^2 (t_2 - t_1)}.\end{aligned}$$

6.4.1 Fonction d'autocovariance

$$C_w(t_1, t_2) = \sigma^2 \cdot \min t_1 t_2$$