Variables aléatoires discrètes

4 mars 2017

Table des matières

1 Deux rappels sur la dénombrabilité et la sommabilité **3**

2 Espaces probabilisés **4**
2.1 Espaces probabilisables 4
2.2 Probabilités, cas général 6
2.3 Propriétés, premiers exemples 7

3 Probabilités conditionnelles, indépendance **11**
3.1 Probabilités conditionnelles, trois formules fondamentales 11
3.2 Indépendance, première approche 14

4 Variables aléatoires discrètes **16**
4.1 Exemples de variables aléatoires discrètes 17
4.2 Couples et n-uplets de v. a. discrètes, variables vectorielles 19

5 Moments : espérance, variance et écart-type **22**
5.1 Espérance 22
5.2 Variance, écart-type et covariance 25
5.3 Inégalité de Markov et inégalité de Bienaymé-Tchebychev 28
5.4 Loi faible des grands nombres 29

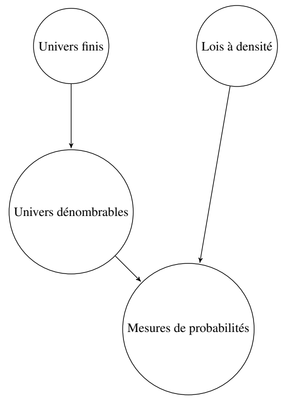
6 Fonctions génératrices **31**
6.1 Définition, premières propriétés 31
6.2 Sommes de variables indépendantes et entières 34

7 Suites de variables aléatoires **36**
7.1 Marches aléatoires 36
7.2 Chaînes de Markov 37

8 Problèmes et simulations **39**
8.1 Sujet Zéro - Mines - 2015 : Files d'attente M/GI/1 39

9 Annexe : La loi de Poisson est un modèle **42**

Vous avez étudié en première année les probabilités sur un ensemble fini. Cela permet de modéliser simplement quelques jeux et quelques applications réelles comme une partie de la théorie des sondages par exemple.
On a pourtant couramment affaire à des situations simples qui font apparaître des ensembles infinis (et pas toujours dénombrables).
Ce chapitre généralise directement ce qui a été vu en MPSI en passant des univers finis à certains univers infinis. L'objet central, comme toujours en probabilité, est la notion de variable aléatoire. Nous nous intéresserons aux variables aléatoires discrètes, c'est à dire dont les valeurs forment un ensemble au plus dénombrable.
Nous ne revenons donc pas sur les notions de probabilités « continues » que vous avez entrevues en terminale[a] comme par exemple la loi uniforme qui permet de modéliser la notion d'incertitude.
En d'autres termes, nous ne traitons que la colonne de gauche dans le graphe situé à droite...



a. Sans toutefois justifier de l'existence d'un espace probabilisable !

Lorsqu'on « travaille » sur un espace probabilisé **fini** (Ω, P), **toute partie de** Ω **est un événement** auquel on sait associer une probabilité :

$$P(A) = \sum_{\omega \in A} P(\{\omega\}). \tag{0.1}$$

Tout le travail de ce chapitre consiste à sortir du cadre fini ; on définit pour cela la notion d'espace probabilisable dans le cadre le plus général, on revient ensuite en arrière en introduisant la notion de variable aléatoire discrète grâce à laquelle, même avec des univers « compliqués, » on peut s'intéresser à **certains événements,** typiquement de la forme, $(X \in A)$ où $X : \Omega \to E$ est une variable discrète (c'est à dire que E est au plus dénombrable), pour lesquels on calculera

$$P(X \in A) = \sum_{k \in A} P(X = k). \tag{0.2}$$

Les exercices (4) et (5) et leurs successeurs directs ((7), (8), (20)), (21) font partie du cheminement pédagogique de ce cours.

1 Deux rappels sur la dénombrabilité et la sommabilité

• *En vue des probabilités discrètes*

Définition 1
• Un E ensemble est **fini** s'il est vide ou s'il existe un entier n tel que E soit en bijection avec $[\![1, n]\!]$.
• On dit qu'un ensemble est **dénombrable** s'il est en bijection avec l'ensemble $\mathbb{N}$ des entiers naturels.
• On dit qu'un ensemble est **au plus dénombrable** s'il est en bijection avec une partie de $\mathbb{N}$. Il est donc fini ou dénombrable par le théorème 1.

Théorème 1
• Une partie de $\mathbb{N}$ est soit finie soit dénombrable.
• Une partie d'un ensemble au plus dénombrable est soit finie soit dénombrable (donc au plus dénombrable).
• Une réunion finie d'ensembles dénombrables est dénombrable.
• Une réunion dénombrable d'ensembles dénombrables est dénombrable.
• Une réunion dénombrable d'ensembles au plus dénombrables est au plus dénombrable.

Exercice 1 $\{0, 1\}^{\mathbb{N}}$ *n'est pas dénombrable*
On considère l'ensemble $E = \{0, 1\}^{\mathbb{N}}$ des suites formées de 0 et de 1. On veut démontrer par l'absurde que cet ensemble n'est pas dénombrable.
On supposera donc qu'il existe une **bijection** $\Phi : \mathbb{N} \to E$.

1. Comprendre les objets, les notations :

(a) Soit $U = (0, 0, 0, 1, 0, 1, ...)$ un élément de E. Au vu de cette description partielle que sont $U_0, U_1, ..., U_5$? Et U_6?

(b) Soit $U = \Phi(p)$ ($p \in \mathbb{N}$ quelconque). Que désignent U_n, $\Phi(p)_n$?

2. On définit une suite V en posant pour $n \in \mathbb{N}$:

$$\begin{cases} V_n & = 0 \text{ si } \Phi(n)_n = 1 \\ V_n & = 1 \text{ si } \Phi(n)_n = 0 \end{cases}$$

(a) Que vaut V_n si $\Phi(n)$ est la suite nulle ? V est elle la suite nulle ?

(b) Que vaut V_n si $\Phi(n)$ est la suite formée uniquement de 1 ? V est elle la suite formée uniquement de 1 ?

(c) Existe-t-il $n \in \mathbb{N}$ tel que $V = \Phi(n)$?

3. Connaissez vous un ensemble non dénombrable ?

• *Un rappel qui nous sera utile pour étudier les propriétés de l'espérance*

Théorème 2 *sommation par paquets, nombres réels*
Si $(I_n)_n$ est une partition de l'ensemble dénombrable I, et si $(a_i)_{\in I}$ est une famille sommable de nombres réels, alors :

1. Chacune des familles $(a_i)_{\in I_n}$ est sommable.
2. La famille $(T_n)_n$ où chaque T_n est défini par

$$T_n = \sum_{i \in I_n} a_i$$

est elle-même sommable.

3. L'égalité qui suit est vérifiée (formule de sommation par paquets) :

$$\sum_{n \in \mathbb{N}} T_n = \sum_{n \in \mathbb{N}} \sum_{i \in I_n} a_i = \sum_{i \in I} a_i \tag{1.1}$$

---

2 Espaces probabilisés

2.1 Espaces probabilisables

Définition 2 *notion de tribu*
Soit Ω un ensemble quelconque. On note, comme à l'accoutumée, $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω . Une tribu $\mathcal{T}$(on dit aussi $\sigma-$algèbre) sur Ω est un sous-ensemble de $\mathcal{P}(\Omega)$ qui vérifie les axiomes suivants :

1. $\mathcal{T}$ est non vide ;

2. $\mathcal{T}$ est stable par passage au complémentaire :

$$A \in \mathcal{T} \Rightarrow \bar{A} = \complement_\Omega A \in \mathcal{T};$$

3. Pour toute famille dénombrable $(A_n)_{n\in\mathbb{N}}$, d'éléments de $\mathcal{T}$, $\underset{n\in\mathbb{N}}{\cup} A_n \in \mathcal{T}$.

Ces propriétés n'ont rien d'étrange, elles constituent un jeu minimal d'axiomes qui permet d'établir que $\mathcal{T}$ vérifie aussi les propriétés suivantes sans lesquelles nous ne pourrions faire de calcul ayant du sens.

Exercice 2 *jouer avec les axiomes, raisonner*

1. Démontrer qu'une tribu $\mathcal{T}$ sur Ω vérifie aussi :
   - Pour toute famille dénombrable $(A_n)_{n\in\mathbb{N}}$, d'éléments de $\mathcal{T}$, $\underset{n\in\mathbb{N}}{\cap} A_n \in \mathcal{T}$;
   - Pour toute famille **finie** $(A_k)_{0\leq k\leq n}$, d'éléments de $\mathcal{T}$, $\overset{n}{\underset{k=0}{\cup}} A_n \in \mathcal{T}$;
   - Pour toute famille **finie** $(A_k)_{0\leq k\leq n}$, d'éléments de $\mathcal{T}$, $\overset{n}{\underset{k=0}{\cap}} A_n \in \mathcal{T}$;
   - $\Omega \in \mathcal{T},\ \emptyset \in \mathcal{T}$;
   - Si $A, B \in \mathcal{T}$, $A \setminus B \in \mathcal{T}$.
2. $\mathcal{P}(\Omega)$ est elle une tribu sur Ω?
3. $\{\emptyset, \Omega\}$ est elle une tribu sur Ω?
4. Une intersection de tribus est elle une tribu ?
5. Soit $\mathcal{B}$ l'intersection de toutes les tribus sur $]0,1[$ contenant les intervalles ouverts contenus dans $]0,1[$. Contient elle les intervalles fermés ?

Définition 3 *vocabulaire de base, extension...*

- Un **espace probabilisable** $(\Omega, \mathcal{T})$ est la donnée d'un ensemble Ω et d'une tribu $\mathcal{T}$ sur Ω;
- Ω est appelé **univers des possibles**, ses éléments des issues, des réalisations ou des possibles.
- Un élément de la tribu $\mathcal{T}$ **est un événement**.
- Si $\mathcal{T} = \mathcal{P}(E)$, les singletons $\{\omega\}$ où $\omega \in \Omega$, sont des évènements et on dit alors que ce sont des **événements élémentaires**. Ce sera le cas pour Ω au plus dénombrable seulement.
- Ω est l'événement **certain** ; $\emptyset$ est l'événement **impossible**.
- Un **système complet d'évènements** est une famille finie ou dénombrable d'éléments de $\mathcal{T}$, deux à deux disjoints et dont la réunion est Ω.
- Deux événements disjoints sont dits **incompatibles,** le complémentaire de A noté $\bar{A}$ quand on fait des probabilités est appelé **évènement contraire** de A.

Exercice 3 *Raisonner et démontrer*

Soit Ω un ensemble quelconque, $(E, \mathcal{T}_E)$ un espace probabilisable et $X : \Omega \to E$ une application.

1. Démontrer que $\{X^{-1}(A)/A \in \mathcal{T}_E\}$ est une tribu sur Ω.
2. On suppose que E est un ensemble fini, $[\![1, n]\!]$ par exemple. Pouvez vous construire une tribu finie sur Ω?

Cas concrets : les exercices (4) et (5).

Exercice 4 *MODéliser*

Il était une fois un pays démocratique dans lequel les électeurs étaient appelés à voter par référendum sur les questions importantes. Quand le résultat du référendum, était de type A, on revotait. S'il était de type B, le résultat devenait définitif.

Nous voulons donc **modéliser** l'expérience qui consiste à lancer une pièce et à recommencer tant que « face » n'apparaît pas. Nous ne nous préoccupons que de définir un univers Ω et une tribu sur Ω dont les éléments ou événements nous intéressent. Nous parlerons de probabilités plus loin (exercice 7).

1. Faire une proposition pour Ω. Pouvez vous proposer un Ω fini et pertinent ?
2. Pouvez vous proposer un Ω dénombrable ?
3. Faire une proposition pour une tribu sur Ω qui permette de représenter (ou contienne) les événements du type « la première « face » apparaît au $n^{ième}$ lancé ». Penser au résultat de l'exercice (3).

Exercice 5 *MODéliser*
Nous voulons **modéliser** l'expérience qui consiste à lancer un dé à six faces et à recommencer tant que 6 n'apparaît pas. Nous ne nous préoccupons que de définir un univers Ω et une tribu sur Ω dont les éléments ou événements nous intéressent. Nous parlerons de probabilités plus loin (exercice 8).

1. Faire une proposition pour Ω.
2. $\star$ Pour le choix proposé, Ω est il dénombrable ?
3. ★ Faire une proposition pour une tribu sur Ω qui permette de représenter (ou contienne) les événements du type « Il y a en tout n lancés. » Penser au résultat de l'exercice (3) [1] .

2.2 Probabilités, cas général

Exercice 6 *échauffements : familles dénombrables d'événements*
On reprend l'espace probabilisable de l'exercice (4).

1. On note A_n l'événement : « Les n premiers lancers amènent tous des nombres autres que 1 ».

   Cette suite est elle monotone pour l'inclusion ? Quelle est l'intersection des A_n?

   *Réponse attendue : langage courant **et** langage ensembliste.*
2. On note A_n l'événement : « Les n premiers lancers amènent au moins une fois un ».

   Cette suite est elle monotone pour l'inclusion ? Quelle est la réunion des A_n?

   *Réponse attendue : langage courant **et** langage ensembliste.*

1. Nous avons toujours plusieurs choix pour définir une tribu. La pertinence de chacun d'eux dépend des événements au sens du langage courant que l'on veut modéliser.

Définition 4 *notion de probabilité*
Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une probabilité sur $(\Omega, \mathcal{T})$ est une application $P : \mathcal{T} \to [0,1]$ qui vérifie les propriétés suivantes :
- $P(\Omega) = 1;$
- Pour toute suite d'événements deux à deux disjoints de $\mathcal{T}$, $(A_n)_n$,

$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} P(A_n) \tag{2.1}$$

Conséquences immédiates :
• $\boxed{P(\emptyset) = 0.}$ Pourquoi ?

• Pour une suite **finie**, $(A_k)_{0 \leq k \leq n}$ d'événements disjoints,

$$\boxed{P(\bigcup_{k=0}^{n} A_k) = \sum_{k=0}^{n} P(A_k)}$$

Il suffit d'appliquer (2.1) à la même suite prolongée par $A_{n+p} = \emptyset$. Cette définition d'une loi de probabilité **généralise** donc celle qui a été donné en première année dans le cadre restreint des univers finis.

• Une conséquence : pour tout $A \in \mathcal{T}$, $\boxed{P(\bar{A}) = 1 - P(A)}$

• Une deuxième conséquence : une loi de probabilité est **croissante pour l'inclusion**

$$\boxed{\text{si } A \subset B, \text{ alors, } P(B) = P(A) + P(B \setminus A) \geq P(A).}$$

• **Cas où Ω est au plus dénombrable et $\mathcal{T} = \mathcal{P}(E)$:**

Dans ce cas, toute partie $A \in \Omega$, est finie ou dénombrable et l'on peut écrire $A = \{\omega_i ; i \in I\}$ où I est soit de la forme $[\![1, n]\!]$ soit égal à $\mathbb{N}$. Ainsi $A = \cup_{i \in I}\{\omega_i\}$ et

$$\boxed{P(A) = \sum_{i \in I} P(\{\omega_i\}) \leq 1}$$

En particulier, la famille $(P(\omega))_{\omega \in \Omega}$ est sommable (revoir la définition de la sommabilité) et

$$\sum_{\omega \in \Omega} P(\{\omega\}) = 1 = P(\Omega).$$

2.3 Propriétés, premiers exemples

Théorème 3 *propriétés élémentaires*
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé.

– Continuité croissante : si $(A_n)_{n\in\mathbb{N}}$, est une suite croissante d'événements, alors

$$\lim_{n\to+\infty} P(A_n) = P\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) \tag{2.2}$$

– Continuité décroissante : si $(A_n)_{n\in\mathbb{N}}$, est une suite décroissante d'événements, alors

$$\lim_{n\to+\infty} P(A_n) = P\left(\underset{n\in\mathbb{N}}{\cap} A_n\right) \tag{2.3}$$

– Pour toute suite d'événements $(A_n)_{n\in\mathbb{N}}$, on a

$$P\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) \leq \sum_{n=0}^{\infty} P\left(A_n\right) \tag{2.4}$$

Démonstrations

– Preuve de (2.2) (famille croissante). Faire des dessins !

Observons que

$$\overset{\infty}{\underset{n=0}{\cup}} A_n = A_0 \cup \left(\overset{\infty}{\underset{n=0}{\cup}}(A_{n+1} \setminus A_n)\right).$$

De façon analogue pour la famille finie, $(A_k)_{0\leq k\leq n}$,

$$\overset{n}{\underset{k=0}{\cup}} A_k = A_0 \cup \left(\overset{n-1}{\underset{k=0}{\cup}}(A_{k+1} \setminus A_k)\right).$$

On a donc d'après (2.1),

$$P(\overset{\infty}{\underset{n=0}{\cup}} A_n) = P(A_0) + \sum_{n=0}^{\infty} P(A_{n+1} \setminus \overset{n}{\underset{k=0}{\cup}} A_k)$$

$$P\left(\overset{n}{\underset{k=0}{\cup}} A_k\right) = P(A_0) + \sum_{k=0}^{n-1} P(A_{k+1} \setminus \overset{k}{\underset{j=0}{\cup}} A_j)).$$

En quoi cela prouve-t-il notre assertion ?

– Preuve de (2.3) : *passage au complémentaire.*
– Preuve de (2.4) : cela découle de

$$P(\overset{\infty}{\underset{n=0}{\cup}} A_n) = P(A_0) + P\left(\sum_{n=0}^{\infty}(A_{n+1} \setminus A_n)\right) \leq P(A_0) + \sum_{n=0}^{\infty} P(A_{n+1})$$

□

Exercice 7 *MODéliser*
On reprend l'expérience abordée dans l'exercice (4) qui consiste à lancer une pièce de monnaie et à recommencer tant que « face » n'apparaît pas.
On choisit ici

$$\Omega = \{0\}^{\mathbb{N}} \cup \bigcup_{n=1}^{\infty} \{(0, 0, ..., 0, 1)\}$$

et on choisit pour tribu $\mathcal{T} = \mathcal{P}(\Omega)$.

1. Préciser les $\{(0, 0, ..., 0, 1)\}$ dans la définition un peu cavalière de Ω.
2. Définir une probabilité sur $(\Omega, \mathcal{T})$.
3. Une issue est elle négligeable ?

Définition 5

Soit $(\Omega, \mathcal{T}, P)$ un espace prob**abilisé.**
Un évènement **négligeable (ou quasi-impossible)** est un événement $A \in \mathcal{T}$ tel que $P(A) = 0$.
Un évènement **presque-sûr (ou quasi-certain)** est un événement $A \in \mathcal{T}$ tel que $P(A) = 1$.
Une propriété Q est **presque-sûre** lorsque $A = \{\omega/Q(\omega)\}$ est un évènement presque sûr.

Un **système quasi-complet d'événements** est une famille finie ou dénombrable d'éléments de $\mathcal{T}$, deux à deux disjoints et dont la réunion est un évènement presque-sûr.

Exercice 8 *MODéliser*
On reprend l'expérience abordée dans l'exercice (5) qui consiste à lancer un dé à six faces et à recommencer tant que 6 n'apparaît pas.
On choisit ici

$$\Omega = [\![1, 5]\!]^{\mathbb{N}} \cup \bigcup_{n=1}^{\infty} \{(a_1, a_2, ..., a_{n-1}, 6)/a_i \in [1, 5] \text{ pour } i = 1, ..., n-1\}$$

et on définit la tribu $\mathcal{T}$ comme étant la plus petite tribu contenant les événements $\{(a_1, a_2, ..., a_{n-1}, 6)\}$ avec $a_i \in [1, 5]$...

1. L'ensemble $\bigcup_{n=1}^{\infty} \{(a_1, a_2, ..., a_{n-1}, 6)/a_i \in [1, 5] \text{ pour } i = 1, ..., n-1\}$ est il dénombrable ?
2. $\star$ Pourquoi Ω n'est il pas dénombrable ?
   Le cours sur la dénombrabilité et ses prolongements (certains exercices) doivent être bien connus avant de répondre à cette question.
3. Proposer une probabilité sur $(\Omega, \mathcal{T})$. Que vaut selon votre définition

   $$P\left([\![1, 5]\!]^{\mathbb{N}}\right)?$$

   L'événement du langage courant : « Il ne sort que des 1 indéfiniment » est il représenté (est-il un évènement) dans votre tribu ? Sinon, pourquoi cela vous laisse-t-il de marbre ? Êtes-vous satisfait de votre modèle ?

Exercice 9 *entropie*

La théorie de l'information est un modèle mathématique créé par Claude Shannon en 1948, qui vise à quantifier mathématiquement la notion d'incertitude. Elle a depuis connu des développements aussi bien en statistique qu'en physique théorique ou en théorie du codage.

On se place dans cette partie dans un espace probabilisé $(\Omega, \mathcal{T}, P)$ et on considère un système complet d'événements $(A_1, A_2, ..., A_n)$ de probabilités respectives $(p_1, p_2, ..., p_n)$ toutes non nulles. On définit l'entropie de ce système par le nombre

$$H((A_i)_i) = -\sum_{k=1}^{n} p_k \ln p_k.$$

On veut montrer que l'entropie doit être maximale lorsqu'aucune hypothèse ne peut être privilégiée.

1. On se place dans le cas $n = 4$ et on considère une course de quatre chevaux, A_i étant l'événement « le cheval n° i remporte la course ».
   - Calculer l'entropie lorsqu'il y a équiprobabilité entre les événements (A_i).
   - Calculer cette entropie lorsque $p_1 = p_2 = \dfrac{1}{8}$ et $p_3 = \dfrac{1}{4}$.
2. (a) Vérifier que $f : x \to x \ln x$ est convexe sur $[0, 1]$.
   (b) Conclure.

Exercice 10 *divergence de Kullback Leibler*
En théorie de l'information en théorie de l'apprentissage, en statistique en général, on est souvent appelé à évaluer l'écart entre deux distributions de probabilités (par exemple une distribution théorique dont on ne connait pas les paramètres que l'on veut estimer et une série d'observations empiriques). On utilise souvent la divergence de Kullback-Leibler définie de la façon suivante :

Soient $\Omega = \{\omega_1, ..., \omega_n\}$ un ensemble fini ou dénombrable et P et Q deux lois de probabilités sur $(\Omega, \mathcal{P}(\Omega))$. En notant $p_i = P(\{\omega_i\})$ et $q_i = P(\{\omega_i\})$ et en supposant que, pour tout i, $q_i \neq 0 \Rightarrow p_i \neq 0$,

$$\mathcal{D}(P||Q) = \sum_{i/p_i \neq 0} p_i \ln\left(\frac{q_i}{p_i}\right). \tag{2.5}$$

1. On suppose que Ω est fini et que pour tout $i, p_i > 0$.
   (a) Démontrer que cette expression est positive.
   (b) On pose $f(q_1, ..., q_n) = \sum_{i=1}^{n} p_i \ln\left(\frac{q_i}{p_i}\right)$. Préciser le domaine de définition de cette fonction et montrer qu'elle atteint son minimum en un seul point. Quel est il ?
   (c) On pose $f(q_1, ..., q_n) = \sum_{i/p_i \neq 0} p_i \ln\left(\frac{q_i}{p_i}\right)$. L'article de Wikipedia énonce : $f(q) = 0$ ssi $(p_i = q_i)$ presque sûrement. Qu'est-ce que cela veut dire ? Justifiez l'énoncé une fois que vous l'aurez précisé.

3 Probabilités conditionnelles, indépendance

3.1 Probabilités conditionnelles, trois formules fondamentales

Définition 6 *probabilités conditionnelles*
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et A un événement de probabilité non nulle. On appelle probabilité conditionnelle relative à l'élément A la loi de probabilité (ce qui doit être préalablement vérifié) sur $(\Omega, \mathcal{T})$ définie par

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

On note aussi $P(B|A)$ pour $P_A(B)$.

Interprétation en termes de fréquences

Le sens de cette formule est parfaitement clair sur l'exemple suivant, dans lequel on relève les nombres des événements selon la couleur de la boule, son état d'usure... La probabilité sur Ω (ensemble des boules) étant supposée uniforme,

$$P(U|R) = \frac{P(U \cap R)}{P(R)} = \frac{...}{...} \text{ alors que} P(U|B) = \frac{P(U \cap B)}{P(B)} = \frac{...}{...}$$

	usées	non usées	total
rouges	45	65	110
blanches	48	85	133
total	93	150	243

Théorème 4 *formule des probabilités composées*
Pour toute famille $(A_i)_{1 \leq i \leq n}$ d'évènements d'un espace probabilisé $(\Omega, \mathcal{T}, P)$ telle que $P(\cap_{i=1}^{n-1} A_i) > 0$, on a, avec la convention $\bigcap_{i \in \emptyset} A_i = \Omega$:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{k=1}^{n} P\left(A_k | \bigcap_{i=1}^{k-1} A_i\right)$$

Démonstration.
C'est la définition lorsque $n = 2$. On procède ensuite par récurrence.

$\square$

On sait lire cette formule sur une structure arborescente définie de la façon suivante :
- l'arbre $\mathcal{A}$ a pour racine l'évènement Ω;
- tout nœud de S de hauteur i possède deux fils $S \cap A_i$ et $S \cap \bar{A}_i$.
- à chaque arc (x, y) de l'arbre on associe le poids $P(y|x)$.

Ceci étant fait, la formule des probabilités composées signifie que pour chaque nœud x de l'arbre $P(x)$, est le produit des valeurs de chaque arc du chemin conduisant de la racine à x.

Théorème 5 *formule des probabilités totales*
Soit un $(\Omega, \mathcal{T}, P)$ espace probabilisé et $(A_i)_{i \in I}$ un système quasi-complet d'évènements (ie : $\forall (i,j) \in I^2, A_i \cap A_j = \emptyset$ et $P(\bigcup_{i \in I} A_i) = 1$). Alors, pour tout $X \in \mathcal{T}$,

$$P(X) = \sum_{i \in I} P(X \cap A_i) = \sum_{i \in I} P(X|A_i)P(A_i).$$

Démonstration.

Exercice 11 *d'après CCP (sujet 0)*
Un mobile se déplace aléatoirement dans l'ensemble des sommets d'un triangle ABC de la façon suivante
- si à l'instant n il est sur A, alors à l'instant $n+1$ il est sur B avec une probabilité de 3/4 ou sur C avec une probabilité de 1/4 ;
- si à l'instant n il est sur B, alors à l'instant $n+1$ il est sur A avec une probabilité de 3/4 ou sur C avec une probabilité de 1/4 ;
- si à l'instant n il est sur C alors à l'instant $n+1$ il est sur B.

On note A_n l'évènement « le mobile se trouve en A à l'instant n », B_n l'évènement « le mobile se trouve en B l'instant n » et C_n l'évènement « le mobile se trouve en C à l'instant n »
Le mobile commence son trajet en partant du point A.
On note, pour tout entier naturel n, $a_n = P(A_n), b_n = P(B_n)$ et $c_n = P(C_n)$

1. (a) Calculer $P(C_2)$.
   (b) Déterminer, en les justifiant, des relations de récurrence entre $a_{n+1}, b_{n+1}, c_{n+1}$ et a_n, b_n, c_n. On les écrira sous forme matricielle.
2. Soit $A = \begin{bmatrix} 0 & 3/4 & 0 \\ 3/4 & 0 & 1 \\ 1/4 & 1/4 & 0 \end{bmatrix}$.
   (a) Vérifier que ${}^tA$ est une matrice stochastique (ie : à termes positifs, la somme des termes d'une ligne quelconque étant égale à 1).

(b) Donner une interprétation probabiliste de cette observation.

(c) Montrer que A est diagonalisable et déterminer une matrice inversible P telle que $P^{-1}\,A\,P$ soit diagonale[2]

(d) Déterminer la limite en l'infini des suites $(a_n)_n$, $(b_n)_n$ et (c_n).

3. ★ C'est à peu de chose près (l'ordre des questions) un sujet Zéro CCP. Il ne mentionne pas d'espace probabilisé dans lequel ces calculs seraient valides. Pouvez vous corriger cela ?

Calculs en 1 page 47.

Exercice 12 *loi de succession de Laplace*[3].
On considère des urnes $(\mathcal{U}_k)_{0\leq k\leq N}$. L'urne $\mathcal{U}_k$ contient k boules blanches et $N-k$ boules noires.
On réalise l'expérience qui consiste à :
- choisir une urne au hasard ;
- prélever avec remise successivement n boules dans cette urne.
On note alors K le numéro de l'urne choisie et X_n le nombre de boules blanches prélevées.

1. Calculer la probabilité de l'événement $(X_n = p)$ (ou la loi de X_n).
2. On se propose de calculer la limite de cette probabilité lorsque N tend vers $+\infty$.
   (a) Monter que
   $$B(p+1, q+1) = \int_0^1 t^p(1-t)^q\,dt = \frac{p!q!}{(p+q+1)!}$$
   (b) Conclure.
3. Sachant que $(X_n = n)$, quelle est la probabilité que $(X_{n+1} = n+1)$?

Théorème 6 *formule de Bayes*
Soit un $(\Omega, \mathcal{T}, P)$ espace probabilisé et $(A_i)_{i\in I}$ un système quasi-complet d'évènements tel que pour tout $i \in I$, $P(A_i) > 0$. Alors, pour tout $X \in \mathcal{T}$ tel que $P(X) > 0$ et tout $j \in I$,
$$P(A_j|X) = \frac{P(X|A_j)P(A_j)}{P(X)} = \frac{P(X|A_j)P(A_j)}{\sum_{i\in I} P(X|A_i)P(A_i)}.$$

Démonstration
Retour à la définition des probabilités conditionnelles et formule des probabilités totales

□

Exercice 13 *banque CCP 2015*
On dispose de 100 dés dont 25 sont pipés de telle sorte que pour ces dés pipés, la probabilité d'obtenir le chiffre 6 lors d'un lancer vaut 1/2.

2. Le sujet CCP suggère d'utiliser une calculette. C'est en fait indispensable !
3. A associer au problème angoissant : le Soleil se lèvera-t-il encore demain ?

1. On choisit un dé au hasard parmi les 100. On le lance et on obtient un 6. Probabilité qu'il soit pipé ?
2. On choisit un dé au hasard parmi les 100. On le lance n fois et on obtient n fois 6. Probabilité p_n qu'il soit pipé ?

Exercice 14
On reprend les dés de l'exercice précédent à ceci près que la probabilité d'obtenir un 6 avec un dé pipé est égale à p. On note X le nombre de lancers amenant un 6 sur un total de n lancers.

1. Calculer $P(X = k)$.
2. A partir de quelle valeur de k est il plus probable que le dé soit pipé ?

3.2 Indépendance, première approche

Définition 7 *événements indépendants*
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé.
On dit que deux événements A et $B \in \mathcal{T}$ sont indépendants ssi $P(A \cap B) = P(A) \times P(B)$.
De la même façon, on dit que $(A_i)_{i\in I} \in \mathcal{T}^I$ est une famille d'événements mutuellement indépendants (ou indépendants dans leur ensemble) ssi pour toute sous-famille finie, $(A_{i_k})_{1\leq k\leq n}$,

$$P\left(\bigcap_{i=k}^{n} A_{i_k}\right) = \prod_{k=1}^{n} A_{i_k}.$$

On observera que la définition de l'indépendance de A et de B revient à dire que, soit l'un des deux événements a une probabilité nulle (et donc également $A \cap B$), soit que $P_A(B) = P(B)$ et $P_B(A) = P(A)$.

Interprétation en termes de fréquences

	usées	non usées	total
rouges	45	65	110
blanches	90		
total			

Exercice 15 *événements indépendants dans leur ensemble et indépendance deux à deux.*
Construire, par exemple dans $\Omega = [\![1, 8]\!]$, une famille d'événements indépendants deux à deux, non mutuellement indépendants.

Exercice 16

1. Justifier que si A et B sont deux événements indépendants, il en va de même pour $\bar{A}$ et B, pour $\bar{A}$ et $\bar{B}$.
2. Généralisation ? Formuler et démontrer.

Exercice 17
On suppose A et C indépendants de même que B et C. Qu'en est il de $A \cup B$ et C?

Exercice 18 *indépendance et probabilités conditionnelles.*
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_i)_{i\in I} \in \mathcal{T}^I$ est une famille de n événements.

1. Monter que si les événements de la famille $(A_i)_{i\in I}$ sont mutuellement indépendants, alors, pour toute sous-famille finie $(A_{i_k})_{1\leq k\leq n}$ telle que $P(\cap_{1\leq k\leq n} A_{i_k}) > 0$ et tout élément A_i en dehors de cette famille finie :
$$P(A_i | A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_n}) = P(A_i).$$
2. On se propose de prouver la réciproque. On suppose donc que pour toute sous-famille finie $(A_{i_k})_{1\leq k\leq n}$ telle que $P(\cap_{1\leq k\leq n} A_{i_k}) > 0$ et tout élément A_i en dehors de cette famille finie :
$$P(A_i | A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_n}) = P(A_i).$$
Considérons alors une sous-famille finie $(B_1, ..., B_m)$ de $(A_i)_{i\in I}$.
   (a) Montrer lorsque $P(B_1 \cap B_2 \cap ... \cap B_m) > 0$ que
   $$P(B_1 \cap B_2 \cap ... \cap B_m) = \prod_{j=1}^{m} P(B_j).$$
   Utiliser la formule des probabilités totales.
   (b) On se place dans le cas où $P(B_1 \cap B_2 \cap ... \cap B_m) = 0$. En raisonnant sur le plus petit indice $p \geq 1$ tel que $P(B_1 \cap B_2 \cap ... \cap B_p) = 0$,

Exercice 19 *indicatrice d'Euler*
On appelle indicatrice d'Euler la fonction $\phi : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ *qui, à* n *associe le nombre des entiers* $1 \leq p \leq n$, *premiers avec* n.

Soit $n \in \mathbb{N}\setminus\{0, 1\}$. On considère l'intervalle $\Omega = [\![1, n]\!]$ muni de la loi de probabilité uniforme. On note :
- $n = \prod_{i=1}^{m} p_i^{\alpha_i}$ la décomposition de n en facteurs premiers ;
- A_p l'événement $A_p = \{x \in \Omega; p|x\}$;
- E l'événement $E = \{x \in \Omega; pgcd(x, n) = 1\}$;

Conseil : on dessinera l'intervalle Ω ce qui permettra d'y représenter l'événement A_{p_i} par exemple sous la forme

$$\boxed{1 \quad ... \quad p_i \quad ... \quad 2p_i \quad ... \qquad (n_i - 1) - p_i \quad ... \quad n_i\, p_i}$$

1. Quel rapport entre $p(E)$ et $\phi(n)$ défini dans le préambule ?
2. Pour $n = 2 \times 11 = 22$, expliciter A_2, A_{11}. Vérifier qu'ils sont indépendants et calculer $P(E)$ ainsi que $\phi(22)$ (formule attendue en prévision de généralisations).
3. Retour au cas général.

(a) Expliciter les événements A_{p_j} et $\cap_{j\in J} A_{p_j}$ pour $(p_j)_{j\in J}$ famille de diviseurs premiers distincts de n.

Montrer que les événements A_{p_j} mutuellement sont indépendants.

(b) Exprimer E en fonction des $\bar{A_{p_i}}$. En déduire $p(E)$ puis l'expression de $\phi(n)$.

Quels théorèmes d'arithmétique interviennent ils ici ?
Corrigé en 2 page 48.

4 Variables aléatoires discrètes

Définition 8 *variable aléatoire discrète*
Soient $(\Omega, \mathcal{T})$ un espace probabili**sable**. Une variable aléatoire discrète sur Ω est une application $X : \Omega \to E$ où E est un ensemble au plus dénombrable, telle que, pour tout élément $x \in E$,

$$X^{-1}(\{x\}) \in \mathcal{T}.$$

Théorème 7 Si $(\Omega, \mathcal{T}, P)$ est un espace probabili**sé** et $X : \omega \to E$ une variable aléatoire discrète. Alors,
– Pour toute partie $B \in \mathcal{P}(E)$, $X^{-1}(B) \in \mathcal{T}$.
– On définit une probabilité sur $(E, \mathcal{P}(E))$ en posant :

$$P_X(B) = P(X \in B).$$

Démonstration Le point crucial est la dénombrabilité de E pour le premier point. Pour le second il y a deux vérifications à faire :
-
-

□

Vocabulaire et notations :
- si $X : (\Omega, \mathcal{T}, P) \to E$ et $Y : (\Omega', \mathcal{T}', P') \to E$ sont deux variables aléatoires discrètes de même loi, c'est à dire que pour tout $x \in E$ (dénombrable), $P(X = x) = P'(Y = x)$, ou encore si $P_X = P_Y$, on note $\boxed{X \sim Y}$.
- On appelle loi de X la loi de probabilité P_X. Lorsqu'on donne un nom ou associe un symbole à cette loi, $\mathcal{L}$ par exemple, on note encore $\boxed{X \sim \mathcal{L}.}$

Avec les exercices de modélisation qui précèdent, ***vous aurez compris que pour construire un modèle, on a une démarche inversée par rapport à cette définition :*** *on se donne Ω, E et $X : \Omega \to E$. On définit alors la tribu $\mathcal{T}$ de telle sorte que X devienne une variable aléatoire. Les probabilités P et P_X sont définies en même temps en posant $P(X \in A) = P_X(A)$. La distinction entre P et P_X est assez formelle dans cette façon de faire. Elle est cruciale lorsqu'on est amené à étudier plusieurs variables aléatoires sur le même espace probabilisé.*

Reprenons nos deux exemples prototypes

Exercice 20 *MODéliser*
On reprend l'expérience abordée dans l'exercice (4) qui consiste à lancer une pièce de monnaie et à recommencer tant que « face » n'apparaît pas.
On choisit ici

$$\Omega = \{0\}^{\mathbb{N}} \cup \bigcup_{n=1}^{\infty} \{(0,0,...,0,1)\}$$

et on choisit pour tribu $\mathcal{T} = \mathcal{P}(\Omega)$.

1. Définir un ensemble dénombrable E, une application $X : \Omega \to E$ et une tribu sur Ω tels que :
   – X soit une variable aléatoire ;
   – l'expérience (« face » sort au $n^{ième}$ lancé) soit représentable sous la forme $(X \in A)$.
2. Définir une probabilité sur $(\Omega, \mathcal{T})$ et préciser la loi de X.

Exercice 21 *MODéliser*
On reprend l'expérience abordée dans l'exercice (5) qui consiste à lancer un dé à six faces et à recommencer tant que 6 n'apparaît pas.
On choisit encore

$$\Omega = [\![1,5]\!]^{\mathbb{N}} \cup \bigcup_{n=1}^{\infty} \{(a_1, a_2, ..., a_{n-1}, 6)/a_i \in [\![1,5]\!] \text{ pour } i = 1, ..., n-1\}.$$

1. Définir un ensemble dénombrable E, une application $X : \Omega \to E$ et une tribu sur Ω tels que :
   – X soit une variable aléatoire ;
   – l'expérience (« 6 » sort au $n^{ième}$ lancé) soit représentable sous la forme $(X \in A)$.
2. Définir une probabilité sur $(\Omega, \mathcal{T})$ et préciser la loi de X.

4.1 Exemples de variables aléatoires discrètes

• **Loi de Bernoulli**. Notation : $\boxed{X \sim b(p).}$

On considère une alternative simple : à chaque répétition soit A soit $\bar{A}$ est réalisé avec pour probabilités respectives p et $1-p$. On définit une variable aléatoire sur $(\Omega = \{A, \bar{A}\}, \mathcal{P}(\Omega))$ en posant $X(A) = 1, X(\bar{A}) = 0$.
La loi de X est alors

$$P(X = 0) = 1 - p, \; P(X = 1) = p.$$

• **Loi binomiale de paramètres** n **et** p. Notation : $\boxed{X \sim \mathcal{B}(n,p).}$

On considère ici n répétitions indépendantes

– On peut représenter les issues de l'expérience par $\Omega = \{0, 1\}^n$, ensemble fini.
– On attribue à chaque événement élémentaire la probabilité $P(\{w\}) = p^k(1-p)^{n-k}$.
– La variable aléatoire X définie sur Ω indiquant le nombre de réalisations de A sur n répétitions a pour distribution (ou loi de probabilité)

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

• **Loi géométrique de paramètre** p. Notation : $\boxed{X \sim \mathcal{G}(p).}$

On considère ici des répétitions indépendantes illimitées de la même expérience de Bernoulli et on considère la variable aléatoire qui renvoie le **numéro d'ordre du premier succès (la première réalisation de** A**)**

$$P(X = k) = (1-p)^{k-1}p.$$

1. Proposer un ensemble probabilisable $(\Omega, \mathcal{T})$. Est il dénombrable ?
2. Proposer une définition de $X : \Omega \to ??$ ainsi que sa distribution de probabilité (ce qui suppose une probabilité sur $(\Omega, \mathcal{T})$).

Théorème 8 *caractérisation de la loi géométrique comme loi sans mémoire*
$X : (\Omega, \mathcal{T}, P) \to \mathbb{N}\backslash\{0\}$ une variable aléatoire telle pour $(n, k) \in (\mathbb{N}\backslash\{0\})^2$,

$$P(X > n + k | X > n) = P(X > k).$$

Alors X suit une loi géométrique de paramètre $p = P(X = 1)$.

Interprétation de l'hypothèse
On peut interpréter X comme le n° d'ordre ou le temps d'attente de la première réalisation d'un événement A (succès) au cours d'une suite d'expériences $E_1, E_2, E_3, ..., E_n, ...$ L'hypothèse suggère alors que si le succès n'est pas au rendez-vous après n tentatives, la probabilité qu'il faille attendre encore k essais est la même qu'à partir de la première réalisation. **On parle alors de loi sans mémoire.**

Démonstration

1. Faites un dessin pour représenter les événements
   $(X > n)$ pour $n \geq 1$.
   Montrer que
   $$P(X > n + m) = P(X > n)P(X > m).$$
2. On note $p = P(X = 1)$ et $\alpha_n = P(X > n)$. Déterminer α_n et en déduire $P(X = n)$. Regardez le dessin !

□

• **Loi de Poisson de paramètre** λ Notation : $\boxed{X \sim \mathcal{P}(\lambda).}$

On dit qu'une variable aléatoire $X : (\Omega, \mathcal{T}, P) \to \mathbb{N}$ suit une loi de Poisson de paramètre $\lambda > 0$ lorsque sa distribution de probabilité est

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

1. Vérifier qu'il s'agit bien d'une loi de probabilité.
2. Calculer $\sum_{k=0}^{\infty} k\, P(X = k)$. Comment interpréteriez-vous cette somme ?

Le théorème qui suit est au programme. Il y est parce que c'est un prototype de situation où la suites des lois d'une suite de variables aléatoires tend vers une autre loi. Il n'est en rien quelque chose de fondamental sur la loi de Poisson qui doit par contre apparaître comme modélisant une situation bien particulière décrite en Annexe 9.

Théorème 9 *approximation d'une loi binomiale par une loi de Poisson*
On se donne une suite $(p_n)_n \in [0, 1]^{\mathbb{N}\backslash\{0\}}$ et une suite de variables aléatoires $(X_n)_n$ telles que pour tout $n \in \mathbb{N}\backslash\{0\}$, $X_n : (\Omega, \mathcal{T}, P) \to \mathbb{N}$ suit une loi binomiale de paramètres n et $p_n \in [0, 1]$: $X_n \sim \mathcal{B}(n, p_n)$.
On suppose que $n\, p_n = \lambda_n$ a pour limite un réel λ. Alors,

$$\lim_{n \to +\infty} P(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{4.1}$$

Démonstration

1. Interpréter l'hypothèse $n\, p_n = \lambda_n \sim \lambda$.
2. Écrire $P(X_n = k)$ et remplacer p_n par son expression en fonction de n et de λ_n.
3. Organiser et conclure.

4.2 Couples et n-uplets de v. a. discrètes, variables vectorielles

Définition 9
Soient $(\Omega, \mathcal{T}, P)$ un espace probabilisé et X et Y deux variables aléatoires discrètes définies sur cet espace.
• On appelle **loi conjointe du couple** (X, Y) la loi de probabilité définie sur $X(\Omega) \times Y(\Omega)$ par

$$P_{(X,Y)}(A \times B) = P((X \in A) \cap (Y \in B))$$

Comme $X(\Omega) \times Y(\Omega)$ est au plus dénombrable, cette loi est entièrement définie par la donnée des probabilités

$$P(X = x_i \cap Y = y_j), \text{ pour tout } (x_i, y_j) \in X(\Omega) \times Y(\Omega).$$

• Les **lois marginales du couple** (X, Y) sont les lois de X et de Y : elles sont données par

$$P(X = x_i) = \sum_j P(X = x_i \cap Y = y_j) \text{ pour tout } x_i \in Proj_1 X(\Omega) \times Y(\Omega)$$

$$P(Y = y_j) = \sum_j P(X = x_i \cap Y = y_j) \text{ pour tout } y_j \in Proj_2 X(\Omega) \times Y(\Omega)$$

• Pour $x \in X(\omega)$ tel que $P(X = x) \neq 0$ (respectivement $P(X > x) \neq 0$), **la loi conditionnelle de** Y **sachant** $(X = x)$ (respectivement $(X < x)$, etc.) est la loi de probabilité sur $Y(\Omega)$ définie par :

$$P_{(X=x)}(Y = y_j) = P(Y = y_j | X = x) = \frac{P(Y = y_j \cap X = x)}{P(X = x)}$$

respectivement :

$$P_{(X>x)}(Y = y_j) = P(Y = y_j | X > x) = \frac{P(Y = y_j \cap X > x)}{P(X > x)}$$

$X \backslash Y$	y_1	$\dots$	y_q	total
x_1	$P(X = x_1 \cap Y = y_1)$	$\dots$	$P(X = x_1 \cap Y = y_q)$	$P(X = x_1)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_p	$P(X = x_p \cap Y = y_1)$	$\dots$	$P(X = x_p \cap Y = y_q)$	$P(X = x_p)$
total	$P(Y = y_1)$	$\dots$	$P(Y = y_q)$	what else ?

Questions :

1. Pourquoi $X(\Omega) \times Y(\Omega)$ est il au plus dénombrable ?
2. Pourquoi (X, Y) est elle une variable aléatoire discrète ? En quoi $(X \in A) \cap (Y \in B)$ est il un événement ?
3. Montrer que, connaissant les seules lois marginales, on ne peut reconstituer en général, la loi conjointe.

$X \backslash Y$	y_1	y_2	y_3	total
x_1				$\frac{1}{3}$
x_2				$\frac{1}{3}$
x_3				$\frac{1}{3}$
total	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

$X \backslash Y$	y_1	y_2	y_3	total
x_1				$\frac{1}{3}$
x_2				$\frac{1}{3}$
x_3				$\frac{1}{3}$
total	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

Dans ces deux tableaux vous ferez en sorte que les lois conjointes soient distinctes et que les lois marginales soient les mêmes

Définition 10 *n-uplets de variables aléatoires, variables vectorielles...*
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $Z = (X_1, X_2, ..., X_n)$ un n-uplet de v.a définies sur cet espace (on dit aussi une variable aléatoire vectorielle) :.
• On étend les notions précédentes en définissant **la loi conjointe du n-uplet** :

$$P(Z = (x_1, ..., x_n)) = P\left(\bigcap_{i=1}^{n}(X = x_i)\right)$$

• ainsi que les **lois marginales**

$$P(X_n = \omega) = \sum_{(x_1,...,x_{n-1})} = P(X_1 = x_1 \cap ... \cap X_{n-1}x_{n-1} \cap X_n = \omega)$$

Définition 11 *variables aléatoires indépendantes*
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé.
• On dit que deux variables aléatoires discrètes X et Y définies sur cet espace sont indépendantes ssi pour tous $A \subset X(\Omega), B \in Y(\Omega)$,

$$P((X \in A) \cap (Y \in B)) = P(X \in A) \times P(Y \in B).$$

• On dit que des variables aléatoires discrètes $(X_i)_i$ sont mutuellement indépendantes ssi pour toute famille finie d'indices $(i_1, i_2, ..., i_n)$ et toute famille $(A_k)_{i_k}$ avec $A_{i_k} \in X_{i_k}(\Omega)$, on a

$$P\left(\bigcap_{k=1}^{n}(X_{i_k} \in A_{i_k})\right) = \prod_{k=1}^{n} P\left(X_{i_k} \in A_{i_k}\right)$$

Théorème 10 *un résultat fondamental*
Soit $(X_1, X_2, ..., X_n)$ un n-uplet de variables aléatoires discrètes **et mutuellement indépendantes** sur l'espace probabilisé $(\Omega, \mathcal{T}, P)$.
Alors, pour tout couple (f, g) de fonctions $f : X_1(\Omega) \times ... \times X_p(\Omega) \to E = f\left(\prod_{i=1}^{p} X_i(\Omega)\right)$ et $g : X_{p+1}(\Omega) \times ... \times X_n(\Omega) \to F = g\left(\prod_{i=p+1}^{n} X_i(\Omega)\right)$ (avec $1 \leq p < n$) les variables aléatoires discrètes

$$Z = f(X_1, ..., X_p) \text{ et } W = g(X_{p+1}, ..., X_n)$$

sont indépendantes.

Démonstration on ne la fait pas (hors programme).

Remarque : nous ferons un usage systématique de ce résultat dans les circonstances suivantes :
Si $(X_n)_n$ est une suite de variables aléatoires indépendantes, alors :

1. X_{n+1} et la variable vectorielle $Y_n = (X_1, X_2, ..., X_n)$ sont indépendantes ;
2. X_{n+1} et la variable $S_n = \sum_{k=1}^{n} X_k$ sont indépendantes ;
3. X_{n+1}^2 et la variable $T_n = \sum_{k=1}^{n} X_k^2$ sont indépendantes ;

Exercice 22

1. Montrer que des événements sont indépendants ssi leurs **variables indicatrices** sont des variables aléatoires indépendantes (la variable indicatrice de A est définie par $X(\omega) = 1$ si $\omega \in A$, et $X(\omega) = 0$ sinon).
2. Démontrer que si les variables aléatoires discrètes $(X_1, X_2, ..., X_n)$ sont mutuellement indépendantes, alors pour toute famille $(\phi_1, \phi_2, ..., \phi_n)$ de fonctions telles que les variables aléatoires $(\phi_1 \circ X_1, \phi_2 \circ X_2, ..., \phi_n \circ X_n)$ sont définies , celles-ci sont également mutuellement indépendantes.

Exercice 23 *sup et inf de v.a. indépendantes*
On considère des variables aléatoires discrètes $(X_1, X_2, ..., X_n)$ mutuellement indépendantes à valeurs dans $\mathbb{R}$..

1. Donner la fonction de répartition de $\sup(X_i)$;
2. Donner la fonction de répartition de $\inf(X_i)$;

Exercice 24 *somme de v.a. première approche*
On considère deux variables aléatoires discrètes (X_1, X_2) à valeurs dans $\mathbb{N}$.

1. Exprimer la loi de $(X_1 + X_2)$ en fonction des lois conditionnelles de X_2 sachant $(X_1 = k)$ pour $k \in \mathbb{Z}$.
2. Cas de variables indépendantes ?
3. Peut on généraliser à des variables discrètes à valeurs dans $\mathbb{Z}, \mathbb{R}$ ou vectorielles ?

Exercice 25
Un corps radioactif émet un nombre aléatoire N de particules pendant une durée déterminée. On suppose que :
– N suit une loi de Poisson de paramètre λ.
– Chaque particule est détectée avec une probabilité $p \in]0, 1[$.
On note X le nombre de particules détectées.

1. Donner la loi de X.
2. Donner la loi de N sachant que $X = k$.

5 Moments : espérance, variance et écart-type

5.1 Espérance

Définition 12
Soit $X : (\Omega, \mathcal{T}, P) \to \mathbb{R}$ une variable aléatoire discrète et à valeurs réelles. On dit que X admet une **espérance finie** si la famille (au plus dénombrable) $(x \times P(X = x))_{x \in X(\Omega)}$ est sommable. On pose alors

$$E(X) = \sum_{x \in X(\Omega)} x \times P(X = x)$$

Lorsque X prend des valeurs positives, son espérance est la somme

$$E(X) = \sum_{x \in X(\Omega)} x \times P(X = x) \in [0, +\infty].$$

• Exemples de calculs d'espérance :

1. Calculer l'espérance d'une variable aléatoire de loi binomiale de paramètres n et p.

$P(X = k) =$	$E(X) =$

2. Calculer l'espérance d'une variable aléatoire de loi géométrique de paramètre p (à revoir après l'étude des séries entières le cas échéant).

$P(X = k) =$	$E(X) =$

3. Calculer l'espérance d'une variable aléatoire de loi exponentielle de paramètre λ

$P(X = k) =$	$E(X) =$

4. Construire une variable aléatoire à valeurs entières, d'espérance infinie.

$P(X = k) =$	$E(X) =$

Par exemple pour le gain d'un jeu dans lequel le joueur gagne d'autant plus qu'il lance longtemps un dé, le jeu s'arrêtant lorsqu'il obtient un 1 par exemple.

Théorème 11 *comparaison* Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{T}, P)$, à valeurs réelles. Si $|X| \leq |Y|$ et si Y admet une espérance finie, alors X admet aussi une espérance finie.

Démonstration

1. Écrire l'événement $(X = x)$ comme un réunion disjointe (et dénombrable) d'événements de la forme $(X = x) \cap (Y = y)$.
2. Majorer les sommes finies de la forme $\sum x\, P(X = x)$.

3. Conclure avec le théorème de sommation par paquets (1.1).

Théorème 12 *linéarité et positivité*
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{T}, P)$, à valeurs réelles, qui admettent des espérances finies. Alors,
• $X + Y$ est une va. discrète qui admet pour espérance

$$E(X + Y) = E(X) + E(Y).$$

• Pour tout $\lambda \in \mathbb{R}$, λX est une va. discrète qui admet pour espérance

$$E(\lambda X) = \lambda E(X).$$

• Si $X \leq Y$, alors $E(X) \leq E(Y)$

Démonstration on fait la preuve pour l'additivité...

1. Écrire l'événement $(X + Y = z)$ comme réunion disjointe d'événements définis à partir de X ou de Y.
2. Majorer les sommes finies de la forme $\sum |z_k| P(X + Y = z_k)$. En déduire que Z est d'espérance finie.
3. Conclure avec le théorème de sommation par paquets (1.1).

Voir détails en 3 .

□

Théorème 13 *formule de transfert*
Soient X une variable aléatoire discrète sur $(\Omega, \mathcal{T}, P)$, à valeurs réelles, d'espérance finie et f une application de $X(\Omega)$ dans $\mathbb{R}$.
Alors $f \circ X : (\Omega, \mathcal{T}, P) \to \mathbb{R}$ est d'espérance finie ssi la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas

$$E(f(X)) = \sum_{x \in X(\Omega)} f(X) P(X = x) \tag{5.1}$$

Démonstration non faite

□

Théorème 14 *produit de variables aléatoires indépendantes*
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{T}, P)$, à valeurs réelles, qui admettent des espérances. Si X et Y sont **indépendantes**, $X \times Y$ est une va. discrète qui admet pour espérance

$$E(X \times Y) = E(X) \times E(Y).$$

Démonstration

□

5.2 Variance, écart-type et covariance

Définition 13 *moment d'ordre 2*
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{T}, P)$ à valeurs réelles. On dit que X admet un moment d'ordre 2 (ou qu'elle est de carré sommable) ssi $E(|X|^2) < +\infty$ (X^2 admet une espérance finie).

Théorème 15 *variables de carrés sommables*
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{T}, P)$ à valeurs réelles ou complexes.

1. Si $E(|X|^2) < +\infty$, alors $E(|X|) < +\infty$.
2. Si $E(|X|^2) < +\infty$, alors $E\left(|X - E(X)|^2\right) < +\infty$. De plus
$$E\left((X - E(X))^2\right) = E(X^2) - E^2(X) \text{ (formule de Huygens)}$$
3. Si $E(|X|^2) < +\infty$ et $E(|Y|^2) < +\infty$ alors $E(|X\,Y|) < +\infty$. De plus, pour des variables réelles,
$$E\,(X\,Y)^2 \leq E(X^2)\,E(Y^2) \text{ (inégalité de Cauchy-Schwarz)}$$
4. Si $E(|X|^2) < +\infty$ et $E(|Y|^2) < +\infty$ alors $E((|X| + |Y|)^2) < +\infty$.
L'ensemble des v.a. sur $(\Omega, \mathcal{T}, P)$ à valeurs réelles qui admettent un moment d'ordre 2, forme un espace vectoriel.

Démonstration

□

Définition 14 *variance, écart-type, covariance*

1. Soit X une variable aléatoire discrète à valeurs réelles définie sur $(\Omega, \mathcal{T}, P)$, la variance et l'écart-type de X sont définis respectivement par
$$var(X) = E((X - E(X))^2) \text{ et } \sigma(X) = var(X)^{1/2}$$
$\boxed{\text{On note aussi } V(X) \text{ pour } var(X).}$
2. Soient X et Y deux variables aléatoires discrètes à valeurs réelles définies sur $(\Omega, \mathcal{T}, P)$, la covariance du couple (X, Y) est définie par :
$$cov(X, Y) = E((X - E(X))(Y - E(Y)).$$

Théorème 16 *propriétés de la variance et de la covariance*

1. *cov* est une forme bilinéaire, symétrique, positive sur le sous-espace des v.a. réelles définies sur $(\Omega, \mathcal{T}, P)$admettant un moment d'ordre 2.
2. $cov(X, Y) = E(X\,Y) - E(X)E(Y)$.
3. $var(X) = cov(X\,X) = E(X^2) - E^2(X)$ (formule de Huygens)

4. $var(\alpha X+\beta)=\alpha^2 var(X)$ pour tout couple $(\alpha,\beta)\in\mathbb{R}^2$.
5. Pour toute famille $(X_1,...,X_n)$ de v.a. réelles définies sur $(\Omega,\mathcal{T},P)$,
$$var\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n var(X_i)+2\sum_{1\le i<j\le n} cov(X_i,X_j).$$
6. Si X et Y sont des v.a. indépendantes, alors
$$cov(X,Y)=0,$$
$$var(X+Y)=var(X)+var(Y).$$
7. Si $(X_1,...,X_n)$ sont mutuellement indépendantes,
$$var\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n var(X_i).$$

Démonstration
Tous les calculs sont faciles et vous devez les faire régulièrement (réflexes à acquérir). Pour le cas des variables indépendantes tout repose sur la relation $E(XY)=E(X)E(Y)$.

□

Exercice 26
Construire deux variables non indépendantes telles que $cov(X,Y)=0$.

Exercice 27
Déterminer l'espérance et la variance de la **moyenne empirique** de variables réelles et **mutuellement indépendantes** $(X_1,...,X_n)$.
Rappel : la moyenne empirique est définie par $S_n=\dfrac{1}{n}\sum_{i=1}^n X_i$.

Définition 15

1. Une v.a. est dite **centrée** lorsque $E(X)=0$ (pour toute v. a. X, $X-E(X)$ est centrée).
2. Une v.a. est dite **réduite** lorsque $V(X)=1$ (pour toute v. a. X, dont la variance n'est pas nulle, $\dfrac{X-E(X)}{\sigma(X)}$ est centrée et réduite).

• Exemples de calculs de variances :

1. Calculer la variance d'une variable aléatoire de loi binomiale de paramètres n et p.

$P(X=k)=$	$E(X)=$	$V(X)=$

2. Calculer la variance d'une variable aléatoire de loi géométrique de paramètre p (à revoir après l'étude des séries entières le cas échéant).

$P(X = k) =$	$E(X) =$	$V(X) =$

3. Calculer la variance d'une variable aléatoire de loi exponentielle de paramètre λ

$P(X = k) =$	$E(X) =$	$V(X) =$

4. Construire une variable aléatoire à valeurs entières, d'espérance finie et de variance infinie.

$P(X = k) =$	$E(X) =$	$V(X) =$

5.3 Inégalité de Markov et inégalité de Bienaymé-Tchebychev

Théorème 17 *inégalité de Markov*
Soit X une variable aléatoire à valeurs réelles et positives, d'espérance finie $E(X)$. Pour tout réel $\lambda > 0$, $P(X \geq \lambda) \leq \dfrac{E(X)}{\lambda}$.

Démonstration
Nous avons $E(X) = \sum_{x \in X(\Omega)} xP(X = x)$. Cela se réécrit,

$$E(X) = \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} xP(X = x) + \sum_{\substack{x \in X(\Omega) \\ x < \lambda}} xP(X = x).$$

Comme X est à valeurs positives,

$$E(X) \geq \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} xP(X = x) \geq \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} \lambda P(X = x) = \lambda P(x \geq \lambda).$$

□

Remarque : On écrit souvent cette inégalité sous la forme $P(X \geq \lambda m) \leq \dfrac{1}{\lambda}$ avec $E(X) = m$.
On l'interprète par exemple en terme de dispersion :
- $\lambda = 2, P(X \geq 2m) \leq \dfrac{1}{2}$: *pour une variable positive, la probabilité que X soit supérieur à deux fois l'espérance est inférieure à* $\dfrac{1}{2}$... ou encore *la médiane ne peut être supérieure à deux fois l'espérance.*
- $\lambda = 4, P(X \geq 4m) \leq \dfrac{1}{4}$: *pour une variable positive, la probabilité que X soit supérieur à quatre fois l'espérance est inférieure à* $\dfrac{1}{4}$...
  ou encore *le quatrième quartile ne peut être supérieur à quatre fois l'espérance.*

Exercice 28
Dans un pays A le salaire moyen est m, le salaire médian est μ.

1. Que dire du salaire médian et du salaire moyen ?
2. Que dire de la proportion de la population qui gagne plus de deux fois, plus de cent fois le salaire moyen ?
3. Que dire des revenus de la population qui se situe dans le dernier décile, dans le dernier centile ?

Données : il y a des best-sellers d'économie scientifique sur la question ($\neq$ de celle des médias).
Ce qui est intéressant, c'est de comparer l'évolution au cours du temps.

Théorème 18 *inégalité de Bienaymé-Tchebychev*
Soit X une variable aléatoire discrète, à valeurs numériques, définie sur un espace probabilisé $(\Omega, \mathcal{T}, P)$.
On suppose que X^2 est d'espérance finie et note $\mu = E(X)$ et $\sigma^2 = E((X-\mu)^2) = E(X^2) - \mu^2$.
Alors, pour tout $\varepsilon > 0$,

$$P\left(|X-\mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Démonstration

Introduisons l'événement $A = (|X-\mu| \geq \varepsilon)$. La variable aléatoire $|X-\mu|^2$ se décompose :

$$|X-\mu|^2 = |X-\mu|^2 \times 1_A + |X-\mu|^2 \times 1_{\bar{A}}.$$

Cela donne :

$$\begin{array}{rcl} |X-\mu|^2 & = & |X-\mu|^2 \times 1_A + |X-\mu|^2 \times 1_{\bar{A}} \\ Var(X) & = & E\left(|X-\mu|^2 \times 1_A\right) + E\left(|X-\mu|^2 \times 1_{\bar{A}}\right) \\ & \geq & \varepsilon^2 E(1_A) + 0 \\ & = & \varepsilon^2 P(A) \end{array}$$

□

5.4 Loi faible des grands nombres

Théorème 19 *loi faible des grands nombres*
Soit $(X_n)_n$ une suite de variables aléatoires indépendantes de même espérance μ et de même variance σ^2. Les moyennes empiriques des $(X_n)_n$ sont les variables aléatoires définie par

$$\bar{X_n} = \frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

1. La moyenne empirique $\bar{X_n}$ a pour espérance μ et pour variance $\dfrac{\sigma^2}{n}$.
2. Pour tout $\varepsilon > 0$,
$$\lim_{n \to +\infty} P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = 0.$$
3. Plus précisément :
$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Démonstration

1. $E(S_n) = \sum_{i=1}^{n} E(X_i) = n\mu$ et $E(S_n) = \mu$.
2. $Var(S_n) = Var(\sum_{i=1}^{n} X_i) = n\sigma^2$ (ici l'hypothèse d'indépendance intervient).
3. L'inégalité de Bienaymé-Tchebychev, pour $\bar{X}_n$ s'exprime :

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

$\square$

Exercice 29 *la plus simple des démonstrations du théorème de Weierstrass*
On rappelle que l'on définit les polynômes de Bernstein de degré n en posant

$$B_k^{(n)}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

et qu'à toute fonction f continue sur $[0,1]$ et à toute subdivision $(0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}, 1)$, on associe le polynôme

$$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_k^{(n)}(x).$$

1. Soit $x \in [0,1]$ et $X_{n,x}$ une variable aléatoire de loi $\mathcal{B}(n,x)$.
   Vérifier que $E\left(f\left(\frac{X_{n,x}}{n}\right)\right) = B_n(f)(x)$.
2. Soit $\varepsilon > 0$.
   (a) Montrer que pour tout $x \in [0,1]$,
   $$|B_n(f)(x) - f(x)| \leq E\left(\left|f\left(\frac{X_{n,x}}{n}\right) - f(x)\right|\right).$$
   (b) Justifier qu'il existe $\alpha_\varepsilon > 0$ tel que
   $$\forall (x,y) \in [0,1]^2, |x-y| \leq \alpha_\varepsilon \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$
   (c) Soit A l'événement défini par $A = \left(\left|\frac{X}{n} - x\right| \leq \alpha_\varepsilon\right)$.
   En écrivant $\left|f\left(\frac{X}{n}\right) - f(x)\right| = \left|f\left(\frac{X}{n}\right) - f(x)\right| 1_A + \left|f\left(\frac{X}{n}\right) - f(x)\right| 1_{\bar{A}}$, montrer qu'il existe un rang à partir duquel
   $$|B_n(f)(x) - f(x)| \leq 2\varepsilon.$$
   (d) Conclusion ?

6 Fonctions génératrices

6.1 Définition, premières propriétés

Définition 16 *fonction génératrice d'une variable entière*
Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, à valeurs dans $\mathbb{N}$. On appelle **fonction génératrice** de X la série entière définie par

$$G_X(z) = \sum_{k=0}^{\infty} P(X = k) z^k = E\left(z^X\right).$$

Remarque : On peut généraliser cette notion définir la fonction génératrice d'une suite numérique $(a_n)_n$ en posant

$$G_{((a_n)_n)}(z) = \sum_{n=0}^{\infty} a_n z^n,$$

dans ce cas G_X apparaît comme la série génératrice associée à la suite $(p_n)_n$ (loi de X). Il s'agit là de l'outil fondamental dans l'étude des dénombrements par exemple.

Théorème 20 *propriétés des fonctions génératrices de variables aléatoires*
Soit X une variable aléatoire entière (à valeurs dans $\mathbb{N}$) de fonction génératrice G_X.

1. La série entière G_X converge normalement sur le disque fermé $\bar{D}(0, 1)$. Elle est en particulier continue sur ce disque.
2. $G_X(1) = 1$ et la série entière G_X a un rayon de convergence supérieur ou égal à 1.
3. G_X caractérise la loi de X : connaissant G_X on déduit $P(X = k) = \dfrac{G_X^{(k)}(0)}{k!}$.
4. X est d'espérance finie ssi G_X est de classe $\mathcal{C}^1$ sur $[0, 1]$. Dans ce cas
$$\boxed{G'_X(1) = E(X).}$$
5. Si G_X est supposée seulement dérivable en 1, alors X est d'espérance finie, G_X est de classe $\mathcal{C}^1$ sur $[-1, 1]$ et $G'(1) = E(X)$.
6. X^n est d'espérance finie ssi G_X est de classe $\mathcal{C}^n$ sur $[0, 1]$. Dans ce cas $G_X^{(n)}(1) = E(X(X-1)...(X_n+1))$.
   En particulier, lorsque X^2 est d'espérance finie,
$$\boxed{V(X) = E\left((X - E(X))^2\right) = E\left(X^2\right) - E\left(X\right)^2 = G"_X(1) + G'_X(1) - G'^2_X(1)}$$

Démonstration On prouvera $(4) \Rightarrow$, puis (5) dont on déduira $(4) \Leftarrow$.

1. Notons $u_n(z) = P(X = n)z^n$, on a $||u_n||_{\underset{[-1,1]}{\infty}} = P(X = n)$ qui est le tg d'une série convergente. De la convergence normale sur l'intervalle fermé on déduit la cv uniforme et la continuité de la somme sur $[-1, 1]$.

2. $G_X(1)$ est défini (convergence simple sur $[-1,1]$) et $G_X(1) = \sum_{n=0}^{\infty} P(X=n) = 1$. On a donc $R \geq 1$.

3. Si $G_X = G_Y$, alors pour tout $n \in \mathbb{N}$, $P(X=n) = \dfrac{G_X^{(n)}(0)}{n!} = \dfrac{G_Y^{(n)}(0)}{n!} = P(Y=n)$.

   Noter que si $R = 1$, le cours sur les séries entières ne nous apporte que des informations sur l'ouvert $]-1,1[$. Ce que allons obtenir dans les propositions suivantes fera donc appel à d'autres techniques (en effet, si nous savons que G_X est de classe $\mathcal{C}^\infty$ sur $]-R,R[$, nous ne savons rien a priori de ce qui se passe sur $[-R,R]$).

4. $\boxed{(4) \Rightarrow}$ On suppose que $\sum nP(X=n)$ converge.
   Rappelons le théorème de dérivation terme à terme d'une série sur un intervalle I
   (a) Si les fonctions u_n sont de classe $\mathcal{C}^1$ sur I,
   (b) s'il existe $a \in I$, tel que $\sum u_n(a)$ converge,
   (c) si $\sum u'_n$ converge uniformément sur tout compact de I,
   Alors, $\sum u_n$ converge uniformément sur tout compact de I, la somme est de classe $\mathcal{C}^1$ et on a
   $$\left(\sum_{n=0}^{\infty} u'n(x)\right)' = \sum_{n=0}^{\infty} u'n(x)$$
   sur I.
   Dans le cas présent avec $u_n(x) = P(X=n)x^n$ et $u'_n(x) = nP(X=n)x^{n-1}$, les deux premiers points sont clairs et le troisième découle de l'hypothèse car
   $$||u'_n||_{\underset{[-1,1]}{\infty}} = n\,P(X=n)$$
   est alors le t.g. d'une série convergente. $\sum u'_n$ converge normalement sur $I = [-1,1]$.
   On a alors sur $[-1,1]$:
   $$G'_X(x)\left(\sum_{n=0}^{\infty} u_n(x)\right)' = \sum_{n=1}^{\infty} k\,P(X=n)x^{n-1}(x)$$
   et, en particulier $G'_X(1) = 1$.

5. On suppose **seulement** que $G'_X(1)$ existe. On écrit le taux de variation
   $$\begin{aligned}\frac{G_X(1) - G_X(x)}{1-x} &= \frac{1}{1-x}\sum_{n=0}^{\infty} P(X=n)(1-x^n) \\ &= \frac{1}{1-x}\sum_{n=0}^{\infty} P(X=n)(1-x)(1+x+x^2+...+x^{n-1}) \\ &= \sum_{n=0}^{\infty} P(X=n)(1+x+x^2+...+x^{n-1})\end{aligned}$$

On voit bien que le terme $nP(X)n$ va apparaître en faisant tendre x vers 1 dans $(1 + x + ... + x^{n-1})$ mais le problème est que nous avons affaire à une somme de série (une limite sur n.) Observons que la fonction

$$x \to \sum_{n=0}^{\infty} P(X = n)(1 + x + x^2 + ... + x^{n-1})$$

est croissante sur $[0, 1[$ (chaque terme est une fonction croissante) donc inférieure à sa limite qui est $\ell = G'_X(1)$.
Comme c'est une série à termes positifs sur $[0, 1]$, on a aussi pour tout N :

$$\begin{aligned}\sum_{n=0}^{N} P(X = n)(1 + x + x^2 + ... + x^{n-1}) &\leq \sum_{n=0}^{\infty} P(X = n)(1 + x + x^2 + ... + x^{n-1}) \\ &\leq \ell = G'_X(1)\end{aligned}$$

La première somme est finie, on fait tendre x vers 1, cela donne, pour tout N,

$$\sum_{n=0}^{N} nP(X = n) \leq \ell...$$

Le reste suit.
Cela fait on retourne à la réciproque $\boxed{(4) \Leftarrow:}$ si G_X est de classe $\mathcal{C}^1$ sur $[-1, 1]$, elle est dérivable en 1 et bla bla bla...

6. On se contente de faire les calculs. On mémorise la formule avec l'exercice qui suit.

□

Fonctions Génératrices de lois usuelles.

Loi	paramètre	$P(X = n)$	G_X	$E(X)$	$V(X)$
de Bernoulli	$b(p)$	$P(X = k) = \begin{cases} 1 - p & \text{si } k = 0 \\ p & \text{si } k = 1 \end{cases}$	$1 - p + pz$	p	$p(1 - p)$
binomiale	$\mathcal{B}(n, p)$	$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$	$(1 - p + pz)^n$	np	$np(1 - p)$
géométrique	$\mathcal{G}(p)$	$P(X = k) = (1 - p)^{k-1} p$	$\dfrac{pz}{1 - (1 - p)z}$	$\dfrac{1}{p}$	$\dfrac{1 - p}{p^2}$

de Poisson	$\mathcal{P}(\lambda)$	$P(X=n) = e^{-\lambda}\dfrac{\lambda^n}{n!}$	$e^{\lambda(z-1)}$	λ	λ

Exercice 30

Retrouver espérance et variance des lois usuelles à l'aide de leurs série génératrices

6.2 Sommes de variables indépendantes et entières

Théorème 21
Si $(X_1, X_2, ..., X_n)$ sont des variables entières **mutuellement indépendantes,** alors,

1. La somme $S_2 = X_1 + X_2$ a pour fonction génératrice le produit des fonctions génératrices de X_1 et de X_2 : $G_{X_1}(z) \times G_{X_2}(z)$.
2. La loi de $X_1 + X_2$ est le produit de convolution (ou produit de Cauchy) des lois de X_1 et de X_2.
3. La somme $S_n = X_1 + X_2 + ... + X_n$ a pour fonction génératrice le produit des fonctions génératrices $\prod_{k=1}^{n} G_{X_k}(z)$.
4. La loi de $S_n = X_1 + X_2 + ... + X_n$ est le produit de convolution (ou produit de Cauchy) des lois de $X_1, X_2, ...$ et X_n.

Démonstration

□

Exercice 31

1. Somme de lois de Poisson indépendantes de paramètres λ et μ?
2. Somme de lois binomiales indépendantes de paramètres (n, p) et (m, p)?
3. Somme de lois binomiales indépendantes de paramètres (n, p) et (n, q)?

Exercice 32 ⋆ *convergence d'une suite de fonctions génératrices*
Pour faire cet exercice, il faut disposer du théorème de convergence dominée...

1. Soit $f(z) = \displaystyle\sum_{k=0}^{\infty} a_k z^k$ une fonction définie par une série entière de rayon $R \geq 1$. On pose, pour $r \in [0, 1[$ et $n_0 \in \mathbb{N}$, $\phi_r(\theta) = e^{-in_0\theta} f\left(re^{i\theta}\right)$.
   (a) Justifier que ϕ_r est somme d'une série de fonctions normalement convergente sur $\mathbb{R}$ dont on précisera les termes généraux.
   (b) En déduire $\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \phi_r(t)\, dt$.

(c) On suppose que la série entière $\sum_n a_n z^n$ converge normalement sur $\bar{D}(0,1)$. Que peut on dire de $\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \phi_1(t)\,dt$?

2. On considère une suite de fonctions $(f_n)_n$ définies par $f_n(z) = \displaystyle\sum_{k=0}^{\infty} a_{n,k} z^k$. On suppose que

   - chacune de ces séries converge normalement sur $\bar{D}(0,1)$.
   - $(f_n)_n$ converge **simplement** vers une fonction f sur $\bar{D}(0,1)$.

   Que peut on dire de la suite des intégrales $\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} e^{-in_0\theta} f_n\left(e^{-it}\right)\,dt$?

3. Soit $(X_n)_n$ une suite de v.a. entières de fonctions génératrices G_{X_n}.

   (a) On suppose que cette suite de fonctions génératrices converge simplement sur $\bar{D}(0,1)$ vers la fonction génératrice g d'une variable entière Y.
   Montrer que pour tout $k \in \mathbb{N}$, $\lim_{n\to+\infty} P(X_n = k) = P(Y = k)$.

   (b) **Illustration** : Soit $(X_n)_n$ une suite de variables aléatoires de lois $\mathcal{B}\left(n, \dfrac{\lambda_n}{n}\right)$ avec $\lim \lambda_n = \lambda$. Montrer que $\lim_{n\to+\infty} P(X_n = k) = P(Y = k)$ ou Y suit une loi de Poisson de paramètre λ.

7 Suites de variables aléatoires

7.1 Marches aléatoires

Définition 17 *marches aléatoires*
Soit $(\Omega, \mathcal{T}, P)$un espace probabilisé et $(S_n)_{n\in\mathbb{N}}$ une suite de va.a.discrètes

$$S_n : (\Omega, \mathcal{T}) \to (E, \mathcal{P}(E))$$

où E est **dénombrable et pourvu d'une structure de groupe** $(E, +)$. On dit que la suite $(S_n)_n$ est une **marche aléatoire** sur E ssi la famille $(X_n)_n$ dans laquelle

$$\begin{cases} X_0 & = S_0 \\ X_{n+1} & = S_{n+} - S_n \end{cases}$$

est une suite de v.a. (mutuellement) indépendantes et de même loi.

Exercice 33 *marches aléatoires.* On se propose d'étudier une marche aléatoire sur $\mathbb{Z}^2$ c'est à dire le mouvement d'une particule qui se déplace par sauts sur des points de coordonnées entières. On considère donc une suite de variables aléatoires $(Z_n)_{n\in\mathbb{N}}$ $\left(Z_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix} \in \mathbb{Z}^2\right)$, définies sur une espace probabilisé $(\Omega, \mathcal{T}, P)$, telle que
– pour tout $n \in \mathbb{N}$, les deux variables aléatoires x_n et y_n sont indépendantes ;
– les variables aléatoires (les sauts) $(Z_{n+1} - Z_n)_n$ sont mutuellement indépendant(e)s ;
– pour tout $n \in \mathbb{N}$, $Z_{n+1} - Z_n$ suit une loi uniforme sur l'ensemble $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$
On suppose que $Z_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

1. Représenter les points de $\mathbb{Z}^2$ de coordonnées $|x| \le 3$ et $|y| \le 3$ que la particule peut atteindre lors d'une telle marche.
2. Montrer que l'on peut exprimer $x_{n+1} - x_n$ à l'aide d'une de variable suivant une loi de Bernoulli de paramètre que l'on précisera.
   En déduire la loi de x_n.
   Indication : on observera que $x_n = \sum_{k=0}^{n-1}(x_{k+1} - x_k)$.
3. Calculer pour $(x, y) \in \mathbb{Z}^2$ la probabilité $P\left(Z_n = \begin{bmatrix} x \\ y \end{bmatrix}\right)$.
4. On note O_n la variable aléatoire égale à 1 si la particule est à l'origine au temps $t = n$ et égale à 0 sinon, et U_n la variable égale au nombre de passages à l'origine entre les instants 1 et $2n$.
   (a) Exprimer U_n en fonction des O_k.
   (b) En déduire l'espérance de U_n. On l'exprimera comme somme partielle d'une série $\sum_{j=1}^{n} v_j$.

(c) Étudier la convergence de cette série. En déduire un équivalent de $E(U_n)$.

Exercice 34 *marche aléatoire dans* $\mathbb{Z}$
On considère une suite de variables de Bernoulli, $(b_n)_n$ indépendantes, de même paramètre p et définies sur $(\Omega, \mathcal{T}, P)$. On pose pour $n \in \mathbb{N}$, $X_n = 2b_n - 1$ puis $S_0 = 0, S_n = \sum_{k=1}^{n} X_n$.

1. Interpréter tout cela en termes de déplacements d'une particule sur un axe. Faire un dessin et préciser S_{14} lorsque

$$(X_n)_n = (+1, +1, +1, -1, -1, +1, -1, +1, -1, -1, -1, -1, +1, -1, ...).$$

2. Donner la loi de S_n.
3. Calculer la probabilité de l'événement $(S_n = 0)$. Donner un équivalent de $P(S_{2n} = 0)$.
4. **Loi du premier instant d'arrivée au point 1.**
   Soit $r \in \mathbb{N}^*$.
   On note T_r le premier indice n pour lequel $S_n = r$ lorsqu'il existe, sinon $T_r = +\infty$.
   (a) Exprimer $P(T_2 = n + m | T_1 = n)$ en fonction de la loi de T_1 (lorsque $P(T_1 = n) > 0$).
   (b) On note ϕ_i la fonction génératrice de T_i. Montrer que $\phi_2(z) = \phi_1(z)^2$.
   (c) Exprimer $P(T_1 = n)$ en fonction de la loi de T_2. En déduire une expression de ϕ_1 en fonction de ϕ_2 puis une équation vérifiée par ϕ_1.
   (d) Calculer la loi de T_1 préciser en particulier $T_1(\infty)$.
   (e) Quelle est la loi de T_{-1}?
5. **Loi du premier retour à l'origine.**
   On note T_0 le premier indice $n \geq 1$ pour lequel $S_n = 0$ s'il existe, sinon on pose $T_0 = +\infty$ comme de bien entendu ; T_0 est donc le temps d'attente du premier **retour** à 0.
   (a) Déduire des questions précédentes la loi de T_0.
   (b) Montrer que l'événement (*Il existe un temps n fini pour lequel $T_0 = n$*) ou encore (*la particule revient à son point de départ en un temps fini*) est **presque-sûr.**

7.2 Chaînes de Markov

Définition 18 *chaînes de Markov*
Une suite $(X_n)_n$ de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ à valeurs dans E, un ensemble discret, est une **chaîne de Markov** ssi : pour tous $n \in \mathbb{N}, y, x_0, x_1, ..., x_n \in E$, on a

$$P(X_{n+1} = y | X_n = x_n \cap X_{n-1} = x_{n-1} \cap ... \cap X_0 = x_0) = P(X_{n+1} = y | X_n = x_n) \quad (7.1)$$

dès lors que $P(X_n = x_n n..., X_0 = x_0) > 0$.
E est l'**espace des états** de la chaîne.

8 Problèmes et simulations

8.1 Sujet Zéro - Mines - 2015 : Files d'attente M/GI/1

Epreuve 0 de probabilités

On considère la file d'attente à une caisse de supermarché. Il y a un serveur et un nombre de places infini. Les clients sont servis selon la discipline « premier arrivé, premier servi ». On appelle « système », l'ensemble des clients en attente et du client en service. On considère $(A_n,\, n \geq 1)$ la suite de variables aléatoires à valeurs dans $\mathbf{N}$ où A_n représente le nombre de clients arrivés pendant le service du client n.

On définit la suite $(X_n,\, n \geq 1)$ comme suit

$$X_0 = 0 \text{ et } X_{n+1} = \begin{cases} A_{n+1} & \text{si } X_n = 0, \\ X_n - 1 + A_{n+1} & \text{si } X_n > 0. \end{cases} \tag{1}$$

On suppose que les variables aléatoires $(A_n,\, n \geq 1)$ sont indépendantes et de même loi, de loi commune celle d'une variable aléatoire A.

Hypothèse : On suppose que

– A est à valeurs entières,

– $\mathbf{P}(A \geq n) > 0$ pour tout entier n,

– A a une espérance finie, on note $\rho = \mathbf{E}[A]$.

---

1 Fonction caractéristique

Dans cette section X représente une variable aléatoire quelconque à valeurs dans $\mathbf{N}$. On définit sa fonction caractéristique ϕ_X par

$$\begin{aligned} \phi_X \,:\, \mathbf{R} &\longrightarrow \mathbf{C} \\ t &\longmapsto \mathbf{E}\left[e^{itX}\right]. \end{aligned}$$

1

1. Montrer que ϕ_X est continue sur $\mathbf{R}$ et périodique.

2. Soit X et Y deux variables aléatoires à valeurs dans $\mathbf{N}$ telles que $\phi_X = \phi_Y$. Montrer que X et Y ont même loi.

   Indication : on pourra considérer les intégrales

   $$I_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_X(t) e^{-ikt} \, dt.$$

   pour tout entier k.

3. Si $\mathbf{E}[X] < +\infty$, montrer que ϕ_X est dérivable sur $\mathbf{R}$ et calculer $\phi'_X(0)$.

4. Calculer la fonction caractéristique d'une variable aléatoire $Z = Y + 1$ où Y est de loi géométrique de paramètre p.

2 Remarques préliminaires

5. Etablir que X_n représente le nombre de clients dans le système au moment du départ du client n.

6. Existe-il $M > 0$ tel que $\mathbf{P}(X_n \leq M) = 1$ pour tout $n \geq 0$?

7. Montrer que pour tout $n \geq 1$, $X_{n+1} - X_n \geq -1$.

8. Pour tout $n \geq 0$, montrer que les variables aléatoires X_n et A_{n+1} sont indépendantes.

3 Convergence

9. Etablir l'identité suivant pour X une variable aléatoire à valeurs entières :

   $$\mathbf{E}\left[e^{itX} \mathbf{1}_{\{X>0\}}\right] = \phi_X(t) - \mathbf{P}(X = 0)$$

10. Pour tout entier n, établir la relation suivante :

    $$\phi_{X_{n+1}}(t) = \phi_A(t) \left[e^{-it} \phi_{X_n}(t) + (1 - e^{-it}) \mathbf{P}(X_n = 0)\right].$$

On suppose dorénavant que $0 < \rho < 1$.

On admet qu'alors la suite $(\mathbf{P}(X_n = 0),\ n \geq 1)$ converge vers une limite, notée α.

On suppose que A n'est pas *arithmétique*, c'est-à-dire que $|\phi_A(t)| < 1$ pour $t \notin \{2k\pi,\ k \in \mathbb{Z}\}$.

On pose

$$\begin{aligned} \theta\ :\ [-\pi,\ \pi] &\longrightarrow \mathbf{C} \\ 0 &\longmapsto 1 \\ t &\longmapsto \alpha\ \frac{\phi_A(t)(1-e^{-it})}{1-\phi_A(t)e^{-it}} \text{ pour } t \neq 0. \end{aligned}$$

11. Etablir le développement limité à l'ordre 1, de ϕ_A au voisinage de 0.

12. Que doit valoir α pour que θ soit continue en 0 ?

13. On fixe $\epsilon > 0$. Pour tout $t \in [-\pi, \pi]\backslash\{0\}$, identifier $\beta_t \in]0,\, 1[$ tel que pour tout entier n suffisamment grand, on ait l'identité suivante :

$$\left|\phi_{X_{n+1}}(t) - \theta(t)\right| \leq \beta_t\, |\phi_{X_n}(t) - \theta(t)| + \epsilon.$$

14. Montrer que la suite de fonctions $(\phi_{X_n},\ n \geq 1)$ converge simplement vers θ.

4 Application

On suppose que

$$\phi_A(t) = \frac{1}{1+\rho-\rho e^{it}}.$$

15. Identifier la loi de A.

16. Montrer que ϕ_A satisfait les hypothèses requises.

17. Calculer θ et identifier la loi de Y.

9 Annexe : La loi de Poisson est un modèle

On propose ici d'introduire la loi de Poisson qui apparaît naturellement dans la modélisation de certains processus de comptage : émission de particules radioactives, appels téléphoniques dans un standard, arrivées dans une file d'attente... Les débordements au delà du programme sont mineurs (lois continues) mais néanmoins réels.

• On considère donc une expérience aléatoire qui consiste à dénombrer des événements ponctuels (et jamais simultanés) pendant un intervalle de temps de longueur t. On note X_t le nombre d'occurrences de ces événements au cours de l'intervalle $[0, t[$ (et on convient de poser $X_0 = 0$). Ainsi, pour tout $t > 0$, X_t est une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathcal{T}, P)$ à valeurs dans $\mathbb{N}$.

• On peut associer à ce processus de comptage la famille $(T_n)_n$ où chaque variable aléatoire T_n est le temps d'attente de la $n^{ième}$ observation.

$$T_n(\omega) = \inf\{t > 0; X_t(\omega) = n\} \tag{9.1}$$

T_n est donc une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathcal{T}, P)$ à valeurs dans $(\mathbb{R}_+^*, \mathcal{B})$.

• On fera sur le processus les trois hypothèses suivantes :

H1 stationnarité
La probabilité d'observer n occurrences du phénomène entre t_1 est t_2 ne dépend que de $|t_2 - t_1|$.

$$\forall (t_1, t_2) \in \mathbb{R}^{*+}, (X_{t_1+s} - X_{t_1}) \text{ et } (X_{t_2+s} - X_{t_2}) \text{ suivent la même loi} \tag{9.2}$$

H2 absence de mémoire
Les nombres d'observations sur des intervalles disjoints sont indépendants

$$\text{si } t_1 < t_2 \leq t_3 < t_4, \text{ alors } (X_{t_2} - X_{t_1}) \text{ et } (X_{t_4} - X_{t_3}) \text{ sont indépendantes} \tag{9.3}$$

H3 espacement
La probabilité d'avoir plus d'une observation dans un intervalle de longueur h est négligeable devant h lorsque h tend vers 0

$$P(X_{t+h} - X_t > 1) \underset{h \to 0}{=} o(h) \tag{9.4}$$

Théorème 22 *lois de Poisson et loi exponentielle*
Soient $(X_t)_{t \in \mathbb{R}^{*+}}$ un processus de comptage et $(T_n)_n$ la famille des temps d'attente associée par la relation (9.1). On suppose que les hypothèses **H1, H2** et **H3** sont vérifiées. Alors,
– pour tout $t > 0$, la loi de X_t est la loi de Poisson de paramètre λt, ie :

$$P(X_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t};$$

– la loi de T_1 (temps d'attente de la première observation) est une loi exponentielle de paramètre $\lambda > 0$, ie :

$$P(T_1 \leq t) = \lambda \int_0^t e^{-\lambda s}\, ds$$

– pour tout $n \in \mathbb{N}$, la loi de T_n (temps d'attente de la $n^{i\grave{e}me}$ observation) est :

$$P(T_n \leq t) = \frac{\lambda^n}{(n-1)!} \int_0^t s^{n-1}\, e^{-\lambda s}\, ds$$

Démonstration
• Notons $p(n,s) = P(X_s = n)$ et commençons par évaluer $p(0,t)$. Avec l'hypothèse H1 (9.2),

$$p(n,s) = P(X_s = 0) = P(X_{t+s} - X_t = 0)$$

pour tout $t \geq 0$. D'autre part, en observant que $(X_{t+s} = 0) \subset (X_t = 0)$ et avec l'hypothèse d'indépendance (9.3), il vient :

$$\begin{aligned} P(X_{t+s} = 0) &= P(X_{t+s} - X_t = 0 \cap X_t - X_0 = 0) \\ &= P(X_s = 0) P(X_t = 0). \end{aligned}$$

Cela nous conduit à l'équation fonctionnelle :

$$p(0, s+t) = p(0,t)\, p(0,s).$$

Si nous observons que $s \to p(0,s)$ est continue à droite et que les seules fonctions continues à droite, vérifiant l'équation $f(x+y) = f(x)f(y)$ sont les fonctions $x \to e^{ax}$, $a \in \mathbb{R}$ (démonstration proposée avec l'exercice 38), nous avons $p(0,s) = e^{as}$. Comme $p(0,s) \leq 1$, il existe $\lambda > 0$ tel que

$$p(0,s) = P(X_s = 0) = e^{-\lambda s}.$$

• Revenons à $p(n,t) = P(X_t = n)$:

$$\begin{aligned} p(n, t+s) &= P(X_{t+s} = n \cap X_t = n) \\ &+ P(X_{t+s} = n \cap X_t = n-1) \\ &+ P(X_{t+s} = n \cap X_t < n-1) \end{aligned}$$

$$\begin{aligned} p(n, t+s) &= P(X_t = n)\, P(X_s = 0) \\ &+ P(X_s = 1)\, P(X_t = n-1) \\ &+ P(X_{t+s} = n \cap X_t < n-1) \end{aligned}$$

$$
\begin{aligned}
\frac{p(n,t+s)-p(n,t)}{s} &= P(X_t=n)\,\frac{P(X_s=0)-1}{s} && (9.5)\\
&+ \frac{P(X_s=1)}{s}\,P(X_t=n-1) && (9.6)\\
&+ \frac{P(X_{t+s}=n\cap X_t<n-1)}{s} && (9.7)
\end{aligned}
$$

Comme $P(X_s=0)=e^{-\lambda s}$ et $P(X_s>1)=o(s)$, on a :

$$
\begin{aligned}
\lim_{s\to 0}\frac{P(X_s=0)-1}{s} &= \frac{e^{-\lambda\, s}-1}{s} = -\lambda && (9.8)\\
\lim_{s\to 0}\frac{P(X_t=1)}{s} &= \lim_{s\to 0}\frac{1-P(X_t=0)-P(X_t>1)}{s} = \lambda && (9.9)\\
\frac{P(X_{t+s}=n\cap X_t<n-1)}{s} &\leq \frac{P(X_s>1)}{s} = o(1) && (9.10)
\end{aligned}
$$

Ce qui donne : $\dfrac{d\,p(n,t)}{dt}=-\lambda\,p(n,t)+\lambda\,p(n-1,t)$.

Posons, pour $n\in\mathbb{N}^*$, $P(t)=\begin{bmatrix}p(0,t)\\ p(1,t)\\ \vdots\\ p(n,t)\end{bmatrix}$. $P(t)$ est la solution du problème de Cauchy :

$$
\frac{dP(t)}{dt}=\begin{bmatrix}-\lambda & 0 & 0 & \dots & 0\\ \lambda & -\lambda & 0 & & 0\\ 0 & \lambda & -\lambda & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \dots & 0 & \lambda & -\lambda\end{bmatrix} P(t),\ \ P(0)=\begin{bmatrix}1\\0\\ \vdots\\ 0\end{bmatrix}
$$

On vérifie que le vecteur de coordonnées $\boxed{P_k(t)=\dfrac{(\lambda\,t)^n}{n!}\,e^{-\lambda\,t}}$ est **la** solution du problème.

• **Loi du temps d'attente (lois continues)** :
On observe que $T_1\leq t$ ssi $X_t\neq 0$. Il vient donc

$$
P(T_1\leq t)=1-P(X_t=0)=1-e^{-\lambda\,t}=\lambda\int_0^t e^{-\lambda\,u}\,du
$$

La loi du temps d'attente de la première observation est donc donnée par la fonction densité $f(t)=H(t)\,\lambda\,e^{-\lambda\,t}$ où $H(t)$ est la fonction de Heaviside. C'est une loi exponentielle de paramètre λ.

• **Loi du temps d'attente de la n^{ième} observation** :

On observe encore que $T_n \leq t$ ssi $X_t \geq n$. On a donc

$$P(T_n \leq t) = 1 - e^{-\lambda\, t} \sum_{k=0}^{n-1} \frac{(\lambda\, t)^k}{k!}.$$

Dérivons pour déterminer la densité de probabilité :

$$\begin{cases} \dfrac{d}{dt} P(T_n \leq t) & = 0 \text{ si } t < 0; \\ \dfrac{d}{dt} P(T_n \leq t) & = e^{-\lambda\, t} \left(\lambda \sum_{k=0}^{n-1} \dfrac{(\lambda\, t)^k}{k!} - \lambda \sum_{k=1}^{n-1} \dfrac{(\lambda\, t)^{k-1}}{(k-1)!} \right) \text{ si } t > 0; \end{cases}$$

La densité de T_n est donc $H(t) \dfrac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda\, t}$ (loi de Erlang de paramètres λ et n.

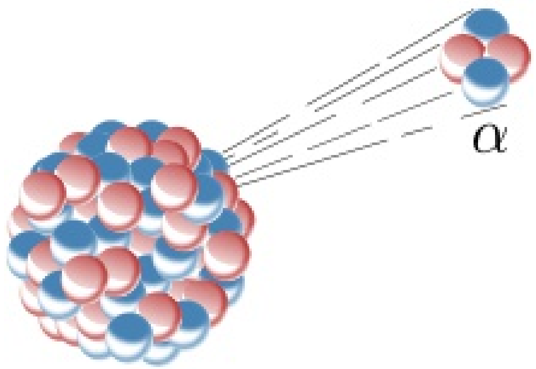
□

Exercice 35 *équation fonctionnelle*

1. Justifier que sous les hypothèses H1,H2,H3 (ou (9.2), (9.3), (9.4)), la fonction $f(t) = P(X_t = 0)$ est continue à droite.
2. Démontrer que les seules fonctions continues à droite, non nulles, vérifiant l'équation $f(x+y) = f(x)f(y)$ sont les fonctions $x \to e^{ax}$, $a \in \mathbb{R}$.
   Indication : on commencera par montrer que pour $r = \dfrac{p}{q} \in \mathbb{Q}$, $f(r) = f(1)^{p/q}$.

corrigé en (4)

Exercice 36 *générateur aléatoire physique*



1. **Loi de désintégration radioactive**
   On note $N(t)$ le nombre de radionucléides d'une espèce donnée présents dans un échantillon radioactif à un instant t. On considère que la probabilité de désintégration d'un de ces radionucléides est toujours la même pour l'espèce considérée, quel que soit l'environnement. Le nombre total de désintégrations pendant un intervalle de temps donné est alors proportionnel à $N(t)$ ce que l'on traduit par :

$$\frac{dN(t)}{dt} = -\lambda N(t) \text{ ou } N(t) = N(t_0) e^{-\lambda(t-t_0)}$$

Sous quelles conditions pensez vous que X_t étant le nombre de désintégrations entre les instants T et $T+t$ les hypothèses (9.2),(9.3) et (9.4) puissent être considérées comme acceptables. Vont elles d'elles-mêmes ?

2. On suppose que l'on dispose d'un dispositif de comptage du nombre X_1 de désintégrations par seconde et que le paramètre $\lambda\, t$ de la loi de X_t est connu.
   (a) Soit $N \in \mathbb{N}^*$ et $\Phi : \mathbb{N} \to \{0, 1, 2, ...N\}$, une fonction croissante. Exprimer $P(\phi(X_t) = n)$ en fonction de valeurs $P(X_t = k)$ pertinentes.
   (b) On observe que la suite des fréquences cumulées $s_k = \sum_{i=0}^{k} P(X_t = i)$, est strictement croissante de limite 1. Pour chaque entier $k \in \mathbb{N}$, on pose
   $$\Phi(k) = n \text{ si et seulement si } \frac{n}{N} \leq s_k < \frac{n+1}{N}$$
      i. Montrer que Φ est croissante. Quelles sont les valeurs possibles de $\phi(k)$?
      ii. Est on sûr que Φ soit surjective sur $[0, N-1]$? Donner une CNS portant sur t pour que cela soit le cas.
      iii. On suppose dorénavant cette condition remplie. Soit n un entier tel que $0 \leq n \leq N-1$. On note k_n le plus petit entier tel que $\frac{n}{N} \leq s_{k_n}$ et k_n^* le plus grand entier tel que $s_{k_n^*} < \frac{n+1}{N}$. Exprimer $P(\Phi(X_t) = n)$.
      iv. Soit $\varepsilon > 0$. Donner une condition sur (on suppose λ fixé) pour que $P(\phi(X_t) = n) = \frac{1}{N} \pm \varepsilon$.

Corrigé en 5.

10 Corrigés

Corrigé n° 1 *corrigé de l'exercice* **??**. *Une feuille de calcul MAPLE, vous devez savoir faire la même chose avec votre calculette. Ce sujet Zéro n'était pas gérable sans elle !*

restart;
with(*LinearAlgebra*) :

$A := Matrix\left(\left[\left[0, \frac{3}{4}, 0\right], \left[\frac{3}{4}, 0, 1\right], \left[\frac{1}{4}, \frac{1}{4}, 0\right]\right]\right);$
Eigenvectors(*A*);
P := %[2];

$$\begin{bmatrix} 0 & \frac{3}{4} & 0 \\ \frac{3}{4} & 0 & 1 \\ \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix}$$

$$\begin{bmatrix} -\frac{1}{4} \\ 1 \\ -\frac{3}{4} \end{bmatrix}, \begin{bmatrix} -\frac{3}{2} & \frac{12}{7} & -1 \\ \frac{1}{2} & \frac{16}{7} & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -\frac{3}{2} & \frac{12}{7} & -1 \\ \frac{1}{2} & \frac{16}{7} & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \mathbf{(1)}$$

$P.DiagonalMatrix\left(\left[-\frac{1}{4^n}, 1, -\frac{3^n}{4^n}\right]\right).P^{(-1)};$

$$\begin{bmatrix} -\frac{3}{10\,4^n} + \frac{12}{35} - \frac{5}{14}\frac{3^n}{4^n} & -\frac{3}{10\,4^n} + \frac{12}{35} + \frac{9}{14}\frac{3^n}{4^n} & \frac{6}{5\,4^n} + \frac{12}{35} - \frac{6}{7}\frac{3^n}{4^n} \\ \frac{1}{10\,4^n} + \frac{16}{35} + \frac{5}{14}\frac{3^n}{4^n} & \frac{1}{10\,4^n} + \frac{16}{35} - \frac{9}{14}\frac{3^n}{4^n} & -\frac{2}{5\,4^n} + \frac{16}{35} + \frac{6}{7}\frac{3^n}{4^n} \\ \frac{1}{5\,4^n} + \frac{1}{5} & \frac{1}{5\,4^n} + \frac{1}{5} & -\frac{4}{5\,4^n} + \frac{1}{5} \end{bmatrix} \qquad \mathbf{(2)}$$

Corrigé n° 2 *corrigé de l'exercice 19.* On note :
- $n = \prod_{i=1}^{m} p_i^{\alpha_i}$ la décomposition de n en facteurs premiers ;
- A_p l'événement $A_p = \{x \in \Omega; p|x\}$;
- E l'événement $E = \{x \in \Omega; pgcd(x, n) = 1\}$;

1. ...
2. • On commence avec **deux** événements pour y voir clair.

   Par le théorème de Gauss $x \in A_{p_i} \cap A_{p_j}$ ssi $p_i\, p_j|x$. En effet, si $p_i|x$ alors $x = k\, p_i$ et si $p_j|x = k\, p_i$, $p_j|k$ qui s'écrit $p_j\, k'$... Donc $\boxed{A_{p_i} \cap A_{p_j} = \{x \in \Omega; p_i\, p_j|x\}.}$

   Dessinons A_{p_i} $\boxed{1 \quad ... \quad p_i \quad ... \quad 2p_i \quad ... \quad (n - p_i) \quad n}$

   puis $A_{p_i} \cap A_{p_j}$: $\boxed{1 \quad ... \quad p_ip_j \quad ... \quad 2p_ip_j \quad ... \quad (n - p_ip_j) \quad n}$

   On a $Card(A_{p_j}) = \dfrac{n}{p_j}$, $Card(A_{p_i}) = \dfrac{n}{p_i}$, $Card(A_{p_i} \cap A_{p_j}) = \dfrac{n}{p_j\, p_i}$
   d'où l'on tire (équiprobabilité) : $p(A_{p_i} \cap A_{p_j}) = p(A_{p_i}) \times p(A_{p_j})$
   • **Généralisons à une famille finie quelconque**
   Par le théorème de Gauss
   $x \in \cap_j A_{p_j}$ ssi $\prod_j p_j|x$.
   Donc $\boxed{\cap_j A_{p_j} = \{x \in \Omega; \prod_j p_j|x\}.}$
   Comme précédemment, en notant $p = \prod_j p_j$ et $n_J = \dfrac{n}{p}$, il vient

   $\boxed{\cap_j A_{p_j} = \{1, p, 2p, ..., (n_J - 1)p, n_J\, p = n\}}$

   On a donc $P\left(\cap_{j\in J} A_{p_j}\right) = \dfrac{n_J}{n} = \prod_{j\in J} \dfrac{1}{p_j} = \prod_{j\in J} P(A_{p_j})$.
3. $x \in \Omega$ est premier avec n ssi aucun des p_i ne divise x.
   Donc $E = \cap_{i=1}^{m} \bar{A_{p_i}}$ et les $\bar{A_{p_i}}$ étant indépendants :

   $$p(E) = \frac{\phi(n)}{n} = \prod_{i=1}^{m} (1 - p(A_{p_i})) = \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right)$$

   De cela on déduit la formule pour $\phi(n)$.

Quels théorèmes d'arithmétique interviennent ils ici ? Le théorème fondamental (décomposition unique en facteurs premiers) et le lemme de Gauss. Point.

Corrigé n° 3 ***corrigé de la démonstration du théorème 11***

1. On écrira, sous forme de réunion disjointe et dénombrable

$$
\begin{aligned}
(X+Y=z) &= \bigcup_{x\in X(\Omega)} P(X+Y=z\cap X=x)\\
&= \bigcup_{x\in X(\Omega)} P(Y=z-x\cap X=x)\\
&= \bigcup_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ x+y=z}} P(X=x\cap Y=y)
\end{aligned}
$$

2. Cela nous permet de réécrire les sommes **au départ finies**

$$
\begin{aligned}
\sum_{k=1}^{n}|z_k|P(X+Y=z_k) &= \sum_{k=1}^{n}|z_k|\sum_{x\in X(\Omega)} P(X=x\cap Y=z_k-x)\\
&= \sum_{k=1}^{n}|z_k|\sum_{x+y=z_k} P(X=x\cap Y=y)\\
&= \sum_{k=1}^{n}\sum_{x+y=z_k}|z_k|P(X=x\cap Y=y)\\
&= \sum_{x+y=z_k}|x+y|P(X=x\cap Y=y)\\
&\le \sum_{k=1}^{n}\sum_{x+y=z_k}(|x|+|y|)P(X=x\cap Y=y)
\end{aligned}
$$

Ce dernier terme se réécrit à son tour :

$$
\sum_{k=1}^{n}\sum_{x+y=z_k}|x|P(X=x\cap Y=y)\quad +\sum_{k=1}^{n}\sum_{x+y=z_k}|y|P(X=x\cap Y=y \qquad (10.1)
$$

Comme $(|x|P(X=x)$ est sommable on a :

$$
\begin{aligned}
\sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ \exists k, x+y=z_k}}|x|P(X=x\cap Y=y) &= \sum_{x\in X(\Omega)}|x|\sum_{\substack{y\in X(\Omega)\\ \exists k, x+y=z_k}}P(X=x\cap Y=y)\\
&\le \sum_{x\in X(\Omega)}|x|P(X=x)=E(|X|).
\end{aligned}
$$

Nous avons donc montré que $\sum_{k=1}^{n}|z_k|P(X+Y=z_k)$ est sommable et majorée par $E(|X|)+E(|Y|)$.

3. Puisque la famille est sommable, nous pouvons réécrire les sommes par paquets :

$$\begin{aligned}
\sum_z zP(X+Y=z) &= \sum_z \sum_{x+y=z} (x+y)P(X=x \cap Y=y) \\
&= \sum_z \sum_{x+y=z} xP(X=x \cap Y=y) + \sum_z \sum_{x+y=z} yP(X=x \cap Y=y) \\
&= \sum_x x \sum_y P(X=x \cap Y=y) + \sum_y y \sum_x P(X=x \cap Y=y) \\
&= E(X) + E(Y)
\end{aligned}$$

Corrigé n° 4 de l'exercice 38

1. Montrons que la fonction $f(t) = P(X_t = 0)$ est continue à droite. En observant que $(X_{t+s} = 0) \subset (X_t = 0)$ et avec l'hypothèse d'indépendance H2-(9.3), il vient :

$$\begin{aligned} P(X_{t+s} = 0) &= P(X_{t+s} = 0 \cap X_t = 0) \\ &= P(X_{t+s} - X_t = 0 | X_t = 0) P(X_t = 0) \\ &= P(X_s = 0) P(X_t = 0). \end{aligned}$$

   Avec l'hypothèse H3-(9.4),

$$\begin{aligned} f(t+s) - f(t) &= P(X_{t+s} = 0) - P(X_t = 0) \\ &= P(X_t = 0)(P(X_s = 0) - 1) \\ &= -P(X_t = 0) P(X_s > 0) \\ &\underset{s \to 0^+}{=} o(1) \end{aligned}$$

2. Soit f non nulle et continue à droite vérifiant $f(x+y) = f(x)f(y)$ pour tout $(x, y) \in \mathbb{R}^2$.

   (a) Commençons par observer que $f(0) = 1$: il existe $x \in \mathbb{R}$, tel que $f(x) \neq 0$ et $f(x) = f(x+0) = f(0)f(x)$.

   (b) Pour tous $n \in \mathbb{N}, t \in \mathbb{R}$ $f(n.t) = f(t)^n$. En effet,
   – cela est vérifié pour $n = 0$ car $f(0) = 1$.
   – si $f(n.t) = f(t)^n$, alors, $f((n+1)t) = f(n\,t).f(t) = f(t)^{n+1}$...

   (c) Pour tout $m \in Z$, $f(m\,t) = f(t)^m$; en effet, $f(x-x) = f(x)f(-x) = f(0) = 1$ donc $f(x) \neq 0$ et $f(-x) = f(x)^{-1}$. Donc si $m \in \mathbb{Z}^-$, $f(m\,t) = f(|m|\,t)^{-1} = f(t)^m$.

   (d) Montrons enfin que pour tout $r = \dfrac{p}{q} \in \mathbb{Q}$, $f(r) = f(1)^{p/q}$:
   – On observe tout d'abord que

$$f(p) = f\left(q\frac{p}{q}\right) = f\left(\frac{p}{q}\right)^q \text{ soit } f\left(\frac{p}{q}\right) = f(p)^{1/q} = f(1)^{p/q}$$

   Cela a un sens car la fonction f est positive $\left(f(x) = f\left(\frac{x}{2} + \frac{x}{2}\right) = f\left(\frac{x}{2}\right)^2\right)$ et ainsi, $f(x)^{1/q}$ est toujours défini si $q \neq 0$.

   (e) On montre enfin que pour $x \in \mathbb{R}$, comme f est continue à droite, pour toute suite de réels $(r_n)_n$ qui converge vers x avec $r_n > x$,

$$\begin{aligned} \lim_{n \to +\infty} f(r_n) &= \lim_{r \to x^+} f(r) = f(x) \\ \lim_{n \to +\infty} f(r_n) &= \lim_{n \to +\infty} f(1)^{r_n} = f(1)^x \end{aligned}$$

   En posant $f(1) = e^a$, le tour est joué.

Corrigé n° 5 ***corrigé de l'exercice 39***

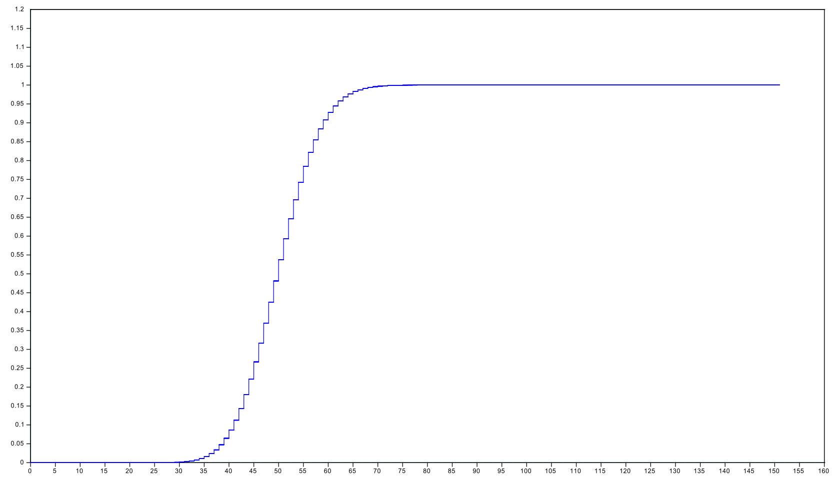
1. Rappelons nos trois hypothèses :

$$\forall (t_1, t_2) \in \mathbb{R}^{*+}, (X_{t_1+s} - X_{t_1}) \text{ et } (X_{t_2+s} - X_{t_2}) \text{ suivent la même loi} \tag{10.2}$$

Cela suppose que le nombre moyen de désintégrations ne varie pas au cours du temps. Or la probabilité de désintégration est proportionnelle au nombre de particules. On doit donc considérer une durée d'observation au cours de laquelle $N(t) = e^{-\lambda t}$ varie peu.

$$\text{si } t_1 < t_2 \leq t_3 < t_4, \text{ alors } (X_{t_2} - X_{t_1}) \text{ et } (X_{t_4} - X_{t_3}) \text{ sont indépendantes} \tag{10.3}$$

$$P(X_{t+h} - X_t > 1) \underset{h \to 0}{=} o(h) \tag{10.4}$$



Pour un tel dispositif dans le commerce voir http ://www.aw-el.com/
consultée novembre 2012

2.

Index

ecart-type, 25

centrée
    variable, 26
certain
    évènement, 5
chaînes de Markov, 38
comptage
    processus, 42
conjointe
    loi, 19
covariance, 25

Erlang
    loi de, 45
évènement, 5
    élémentaire, 5
    certain, 5
    impossible, 5
    négligeable, 9
    presque-sûr, 9
    quasi-certain , 9
    quasi-impossible, 9
    quasi-négligeable, 9
Exponentielle
    loi, 44

impossible
    évènement, 5
information, 10

loi
    conjointe, 19
    de Erlang, 45
    de Poisson, 42
    Exponentielle, 44

marches aléatoires, 36
Markov
    chaînes de, 38

n-uplet
    de variables aléatoires, 21

négligeable
    évènement, 9

Poisson
    loi, 42
presque-sûr
    évènement, 9
processus
    de comptage, 42

quasi-certain
    évènement, 9
quasi-impossible
    évènement, 9

réduite
    variable, 26

Shannon, 10
système
    complet, 5, 9

temps
    d'attente, 42

variable
    centrée, 26
    réduite, 26
variable aléatoire
    vectorielle, 21
variance, 25