13 Variables aléatoires discrètes

Dans ce chapitre nous étudions les applications définies sur un espace probabilisé dénombrable $(\Omega, \mathcal{T}, P)$, c'est-à-dire des quantités dont la valeur dépend de l'issue d'une expérience aléatoire. Contrairement a ce qui avait été fait en première année, on ne se limite donc plus au cas des variables aléatoires dont l'ensemble des valeurs possibles est fini.

De même que dans le chapitre sur les probabilités, on étudiera des phénomènes aléatoires sans décrire précisément l'espace probabilisé associé: la donnée *a priori* des probabilités des évènements considérés suffit, celles-ci étant données sous forme de loi d'une variable aléatoire.

Dans tout le chapitre, I désigne (sauf mention du contraire) une partie de $\mathbb{N}$, finie ou non. Tout ensemble de la forme $\{x_i : \ i \in I\}$ est alors fini ou dénombrable.
De même, toute famille de la forme $(x_i)_{i\in I}$ est finie ou dénombrable.

On admet qu'il existe une tribu, que l'on notera $\mathcal{B}$, sur $\mathbb{R}$ contenant tous les intervalles, appelée *tribu des boréliens*.

1 Notion de variable aléatoire discrète

1.1 Définition et premières propriétés

Définition 13.1. **Variable aléatoire discrète**

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble, et une application $X : \Omega \to E$. On dit que X est une **variable aléatoire discrète** si, pour tout $x \in E$, on a $X^{-1}(\{x\}) \in \mathcal{T}$ (*i.e.* si $X^{-1}(\{x\})$ est un évènement).
L'image $X(\Omega)$ de X sera appelée l'**univers image** de X ou l'ensemble des valeurs possibles de X.

Remarque 13.1. **Variable aléatoire discrète ou finie**

Soit X une variable aléatoire discrète définie sur un espace probabilisable $(\Omega, \mathcal{T})$.

- ✗ Si $X(\Omega)$ est un ensemble fini $\{x_1, \ldots, x_n\}$, on dit que X est une **variable aléatoire (discrète) finie**.
  Si l'univers Ω est fini, alors $X(\Omega)$ l'est également.
- ✗ Si $X(\Omega)$ est un ensemble infini dénombrable $\{x_i | \ i \in I\}$, on dit que X est une variable aléatoire **discrète infinie dénombrable**.

Exemple 13.1.

Une urne contient des boules numérotées de 1 à n.

- ✗ On pioche une boule au hasard dans cette urne et on note X la variable aléatoire qui prend la valeur du numéro de la boule piochée. On a alors $X(\Omega) = [\![1, n]\!]$ et X est une variable aléatoire finie.
- ✗ On considère la variable aléatoire Y qui prend la valeur 1 si et seulement si la boule piochée a un numéro pair et 0 sinon. Dans ce cas, $Y(\Omega) = \{0, 1\}$ et Y est encore une variable aléatoire finie.
- ✗ On pioche maintenant successivement et avec remise jusqu'à obtenir la boule numérotée 1 et on note Z la variable aléatoire qui prend pour valeur le nombre de pioches effectuées. On peut avoir la boule 1 dès la première pioche, mais on peut piocher arbitrairement longtemps une autre boule donc ici $Z(\Omega) = \mathbb{N}^*$ et cette fois Z est une variable aléatoire infinie dénombrable.

Définition 13.2. **Notations des évènements**

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, $X : \Omega \to E$ une variable aléatoire, et $k \in E$.

✗ L'évènement $X^{-1}(\{k\}) = \{\omega \in \Omega,\ X(\omega) = k\}$ sera noté $[X = k]$.
✗ Si $E \subset \mathbb{R}$, on pose encore:
  – l'évènement $X^{-1}(]-\infty, k]) = \{\omega \in \Omega,\ X(\omega) \leq k\}$ sera noté $[X \leq k]$;
  – l'évènement $X^{-1}(]-\infty, k[) = \{\omega \in \Omega,\ X(\omega) < k\}$ sera noté $[X < k]$.
  On définit de même les évènements $[X \geq k]$ et $[X > k]$.

Proposition 13.1. **Système complet canonique associé à une variable aléatoire**

Soient $(\Omega, \mathcal{T})$ un espace probabilisable, et $X : \Omega \to E$ une variable aléatoire discrète dont l'univers image est $X(\Omega) = \{x_i |\ i \in I\}$. Alors la famille $(X = x_i)_{i \in I}$ est un système complet d'évènements pour Ω.

Cette dernière proposition est importante, car elle permet d'appliquer la formule des probabilités totales en *découpant* l'univers selon les valeurs que prend une variable aléatoire.

Exercice 13.1.

Une urne contient des boules, indiscernables au toucher, numérotées de 1 à n. On pioche successivement et avec remise des boules dans cette urne. On introduit la variable aléatoire X qui prend pour valeur le rang du premier tirage qui permet d'obtenir une boule avec un numéro supérieur ou égal au numéro de la boule piochée au premier au tirage.

1. Que vaut $X(\Omega)$?
2. Déterminer, pour tout $k \in X(\Omega)$, la probabilité $P([X = k])$.

Jusqu'à la fin de ce chapitre, on fixe un espace probabilisé $(\Omega, \mathcal{T}, P)$. Sauf mention du contraire, toutes les variables aléatoires considérées seront discrètes, définies sur $(\Omega, \mathcal{T})$ et à valeurs dans $\mathbb{R}$.

Remarque 13.2.

Par σ-additivité, on déduit alors, de la proposition précédente, $P\left(\bigcup_{i \in I} [X = x_i]\right) = \sum_{i \in I} P(X = x_i) = 1$.

Grâce à la probabilité donnée sur Ω, on peut définir la *loi* d'une variable aléatoire X, qui consiste à associer à chaque valeur que X peut prendre la probabilité que cette valeur soit prise.

Théorème & Définition 13.2. **Loi d'une variable aléatoire**

Soient X une variable aléatoire telle que $X(\Omega) = \{x_i :\ i \in I\}$ et $\mathcal{B} \subset \mathcal{P}(\mathbb{R})$ la tribu des boréliens. Alors, l'application définie par

$$\begin{array}{rccc} \mathbb{P}_X : & \mathcal{B} & \longrightarrow & [0,1] \\ & A & \longmapsto & P(X \in A) \end{array}$$

est une probabilité sur $(\mathbb{R}, \mathcal{B})$, appelée **loi** de X.

Théorème & Définition 13.3. **Loi et distribution de probabilité**

Soit $\{x_i : i \in I\}$ un ensemble dénombrable de réels.
Une application

$$\begin{array}{ccc} \{x_i :\ i \in I\} & \to & \mathbb{R} \\ x_i & \mapsto & p_i \end{array}$$

est une **distribution de probabilité** si :

$$\forall i \in I,\ p_i \in [0,1] \qquad \text{et} \qquad \sum_{i \in I} p_i = 1.$$

On peut alors lui associer, de manière unique, une loi de probabilité sur $(\mathbb{R}, \mathcal{B})$ en posant

$$\begin{array}{rccc} P : & \mathcal{B} & \longrightarrow & [0,1] \\ & A & \longmapsto & \displaystyle\sum_{i:\ x_i \in A} p_i \end{array}$$

Plus précisément, il existe une variable aléatoire X telle que $\mathbb{P}_X = P$, c'est à dire telle que, pour tout $i \in I$, $P([X = x_i]) = p_i$.

Méthode 13.1. **Déterminer la loi d'une v.a. discrète**

L'application $\mathbb{P}_X$ associe donc à un évènement la probabilité que X y prenne ses valeurs. Si bien entendu la définition de la loi est celle ci-dessus, ce n'est pas exactement cela qu'on va manipuler, mais plutôt la *distribution* des probabilités $x \in X(\Omega) \mapsto P(X = x)$.

En effet, dans le cas d'un univers Ω discret (fini ou infini), la loi d'une variable aléatoire sera entièrement déterminée (comme le précise le **Théorème 13.3** ci-dessus) par la donnée de l'application

$$x \in X(\Omega) \mapsto P(X = x).$$

Ainsi, pour déterminer la loi d'une variable aléatoire, on cherche, pour chacun des éléments de son univers image, la probabilité que la variable aléatoire prenne cette valeur là.

La loi sera la principale information dont on disposera sur une variable aléatoire, l'univers Ω étant le plus souvent implicite et non précisé.

On peut représenter une loi par un histogramme ou un diagramme en bâtons: la hauteur du bâton d'abscisse x_n est alors la probabilité que X prenne la valeur x_n.
Si $x_n \notin X(\Omega)$, alors $P(X = x_n) = 0$.

Remarque 13.3. **Variables aléatoires de même loi**

Il est erroné de penser que deux variables aléatoires qui possèdent la même loi sont égales.
Si on lance deux dés équilibrés, ils ne vont pas forcément tomber sur la même face et pourtant, les variables aléatoires qui prennent pour valeur le résultat de chacune des deux faces suivent bien la même loi (elles prennent les valeurs avec la même probabilité $1/6$).

Lorsque deux variables aléatoires X et Y *suivent* la même loi, c'est à dire lorsque $\mathbb{P}_X = \mathbb{P}_Y$ (ou plus simplement qu'elles ont la même distribution de probabilité) on écrira

$$X \sim Y.$$

Ce théorème permet en particulier de définir une variable aléatoire directement par sa distribution de probabilités, sans définir formellement l'univers ni les évènements associés.

Exercice 13.2.

On considère la suite (u_k) définie par

$$\forall k \geq 2, \qquad u_k = \frac{1}{k-1} - \frac{1}{k}.$$

Montrer qu'il existe une variable aléatoire X telle que, pour tout $k \geq 2$, $P(X = k) = u_k$.

2 Lois usuelles discrètes

2.1 Loi certaine

Une variable aléatoire sera dite certaine si elle prend toujours la même valeur avec une probabilité 1.

Définition 13.3. **Variable aléatoire certaine**

Soit X une variable aléatoire. On dit que X est **certaine** s'il existe a tel que $P(X = a) = 1$.

2.2 Loi uniforme

La loi uniforme est celle des variables aléatoires **finies** qui prennent leurs différentes valeurs de manière équiprobable.

Définition 13.4. **Loi uniforme**

Soit X une variable aléatoire finie telle que $X(\Omega) = \{x_1, \ldots, x_n\}$. On dit que X suit une **loi uniforme** sur $\{x_1, \ldots, x_n\}$, ce qui sera noté $X \sim \mathcal{U}(\{x_1, \ldots, x_n\})$, si

$$\forall i \in [\![1, n]\!],\ P(X = x_i) = \frac{1}{n}.$$

Remarque 13.4.
Dans le cas où $X(\Omega) = \mathbb{N}$, dire que X suit une loi uniforme est absurde (car il n'existe pas de probabilité uniforme sur un ensemble infini).

Exemple 13.2.
Une variable aléatoire donnant le résultat du lancer d'un dé équilibré suit la loi $\mathcal{U}[\![1, 6]\!]$.

En `Python` **.**
La commande `randint(a,b)` (du package `numpy.random`) renvoie une simulation de la loi $\mathcal{U}([\![a, b-1]\!])$.

2.3 Loi de Bernoulli

La loi de Bernoulli intervient à chaque fois qu'il est question d'expérience aléatoire à deux alternatives possibles (souvent appelées succès et échec).

Définition 13.5. **Loi de Bernoulli**
Soit $p \in [0, 1]$, et X une variable aléatoire telle que $X(\Omega) = \{0, 1\}$. On dit que X suit une **loi de Bernoulli** de paramètre p, ce qui sera noté $X \sim \mathcal{B}(1, p)$ ou simplement $X \sim \mathcal{B}(p)$, si

$$P(X = 1) = p \qquad \text{et} \qquad P(X = 0) = 1 - p.$$

Exemple 13.3.
On considère la variable aléatoire qui donne le résultat du lancer d'une pièce de monnaie truquée. Celle-ci suit alors la loi $\mathcal{B}(1, p)$, pour un certain réel p, égal à la probabilité d'obtenir *Pile* (ou *Face*).

Un autre exemple très utilisé de loi de Bernoulli est celui des variables indicatrices (voir **Exercice 13.24**).

2.4 Loi binomiale

Théorème & Définition 13.4. **Loi Binomiale**
Soit $n \in \mathbb{N}^*$, $p \in [0, 1]$, et X une variable aléatoire telle que $X(\Omega) = [\![0, n]\!]$. On dit que X suit une **loi binomiale de paramètres** n, p, ce qui sera noté $X \sim \mathcal{B}(n, p)$, si :

$$\forall k \in [\![0, n]\!],\ P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Preuve. La formule du binôme garantit que la formule ci-dessus définit bien une distribution de probabilités. □

Proposition 13.5. **Nombre de succès**
Lorsqu'on répète n fois, de façon indépendante, la même expérience de Bernoulli de paramètre p (disons que p est la probabilité d'un succès), alors la variable aléatoire qui prend la valeur du *nombre de succès* suit la loi $\mathcal{B}(n, p)$.

Exemple 13.4.
On considère une urne contenant des boules, parmi lesquelles il y a une proportion $p \in [0, 1]$ de boules blanches, et une proportion $1 - p$ de boules noires. On effectue n tirages avec remise dans cette urne, et on appelle X la variable aléatoire qui prend pour valeur le nombre de boules blanches obtenues.
Alors $X \sim \mathcal{B}(n, p)$.

Remarque 13.5.
Une loi de Bernoulli de paramètre p est une loi binomiale de paramètres $1, p$, d'où la notation de la loi de Bernoulli.

En `Python` **.**
La commande `binomial(n,p)` (du package `numpy.random`) renvoie une simulation de la loi $\mathcal{B}(n, p)$.

2.5 Loi géométrique

Théorème & Définition 13.6. **Loi géométrique**

Soit $p \in [0,1]$. On dit que X suit **la loi géométrique** de paramètre p, ce qui se note $X \sim \mathcal{G}(p)$, si $X(\Omega) = \mathbb{N}^*$ et si

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

Exercice 13.3.

Démontrer que la formule ci-dessus définit bien une loi de probabilité.

Proposition 13.7. **Temps d'attente du premier succès**

Lorsqu'on répète, de façon indépendante, la même expérience de Bernoulli de paramètre p (disons que p est la probabilité d'un succès) jusqu'à l'obtention d'un succès, alors la variable aléatoire qui prend la valeur du *nombre d'itérations* (ou du rang du premier succès) suit la loi $\mathcal{G}(p)$.

Exemple 13.5.

On lance une infinité de fois un dé équilibré et on note X la variable aléatoire qui prend la valeur du numéro du lancer où apparait, pour la première fois, la face numérotée "6". Alors $X \sim \mathcal{G}(1/6)$.

Exercice 13.4.

Soit $X \sim \mathcal{G}(p)$. Déterminer, pour tout $k \in \mathbb{N}$, $P(X > k)$.

Exercice 13.5. **Coïncidence de lois géométriques indépendantes**

On considère deux lois géométriques de même paramètre p supposées *indépendantes* (on renvoie à la **Définition 13.9**). Déterminer $P(X = Y)$.

Exercice 13.6. **Minimum de deux lois géométriques indépendantes**

On considère deux lois géométriques de même paramètre p supposées *indépendantes*.
Déterminer la loi de $Z = \min(X, Y)$.

La loi géométrique permet aussi de caractériser les lois discrètes *sans mémoire*. On renvoie à l'**Exercice 13.32**).

En `Python` .

La commande `geometric(p)` (du package `numpy.random`) renvoie une simulation de la loi $\mathcal{G}(p)$.

2.6 Loi de Poisson

Théorème & Définition 13.8. **Loi de Poisson**

Soit $\lambda > 0$. On dit que X suit **loi de Poisson** de paramètre λ, ce qui se note $X \sim \mathcal{P}(\lambda)$, si $X(\Omega) = \mathbb{N}$ et si

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

Exercice 13.7.

Démontrer que la formule ci-dessus définit bien une loi de probabilité.

Remarque 13.6.

La loi de Poisson est aussi appelée *loi des évènements rares*, parce que $P(X = k)$ décroît très vite vers 0 lorsque k tend vers $+\infty$. Contrairement aux lois binomiale ou géométrique, on ne peut pas associer une loi de Poisson à la description d'une expérience aléatoire usuelle.
On s'en sert (et c'est dans ce cas une hypothèse faite *a priori*) pour modéliser des évènements associés à des prises de valeurs arbitrairement grandes mais dont la probabilité décroît très vite : nombre de voiture passant à un péage, nombre d'appels dans un standard téléphonique, apparition d'une mutation génétique chez un sujet...

Exercice 13.8. **Stabilité par somme des lois de Poisson**

Soient X et Y deux variables aléatoires *indépendantes* suivant toutes deux des lois de Poisson de paramètres respectifs λ et μ.
Montrer que
$$X + Y \hookrightarrow \mathcal{P}(\lambda+\mu).$$

3 Loi conditionnelle

Définition 13.6. **Loi conditionnelle**

Soit X une variable aléatoire telle que $X(\Omega) = \{x_i :\ i \in I\}$. Soient encore A un évènement non négligeable, et P_A la probabilité conditionnelle sachant A.
On appelle **loi conditionnelle de** X **sachant** A la loi de X relativement à l'espace probabilisé $(\Omega, \mathcal{T}, P_A)$, c'est-à-dire celle associée à la distribution de probabilités
$$\begin{array}{ccc} \{x_i :\ i \in I\} & \to & [0,1] \\ x_i & \mapsto & P_A(X = x_i) \end{array}$$

Exercice 13.9.

On lance une pièce donnant pile avec probabilité $p \in]0;1[$. On note X le rang d'apparition du premier pile et si $[X = n]$, on note Y le nombre de faces obtenus lors de n lancers supplémentaires.

1. Déterminer la loi conditionnelle de Y sachant $[X = n]$.
2. Préciser $Y(\Omega)$ puis montrer que pour tout $k \in \mathbb{N}$,
$$P(Y = k) = \left(\frac{q}{p}\right)^{k-1} \sum_{n=k}^{+\infty} \binom{n}{k} (pq)^n.$$

Exercice 13.10. **Binomiale conditionnée par une Poisson**

Un auto-stoppeur attend au péage d'une autoroute pendant une certaine période. On suppose que le nombre de véhicules franchisant le péage pendant cette période est modélisé par une variable aléatoire suivant la loi de Poisson de paramètre $\lambda > 0$. A chaque fois qu'un véhicule franchit le péage, il lance une pièce truquée qui donne pile avec la probabilité p. On pose $q = 1 - p$.
On note N, X et Y respectivement le nombre de lancers, le nombre de piles obtenus et le nombre de face obtenus.

1. Soit $n \in \mathbb{N}$. Déterminer la loi conditionnelle de X sachant $[N = n]$.
2. Montrer que X suit une loi de Poisson de paramètre λp.
3. Sans calcul, que peut-on dire de la loi de Y ?

4 Loi d'une composée $\varphi(X)$

Soit X une variable aléatoire réelle, et une fonction φ à valeurs réelles définie sur $X(\Omega)$. Alors l'application $\varphi \circ X : \Omega \to \mathbb{R}$ sera notée $\varphi(X)$.

Remarque 13.7.

On exprime ici le fait que, considérant une variable aléatoire donnée par sa loi, on peut *oublier* l'espace probabilisé sous-jacent et considérer X comme une variable! Mais cette notation n'est valable que dans le cadre des variables aléatoires: la composée $f \circ g$ de deux applications ne se note pas $f(g)$ en général!

Nous allons montrer que, si X est une variable aléatoire, on peut exprimer la loi de $\varphi(X)$ en fonction de celle de X. On commence par donner un exemple représentatif.

Exercice 13.11.

Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{N}$ et φ la fonction définie par $\varphi(x) = \cos(\pi x)$. Déterminer la loi de $\varphi(X)$.

Plus généralement, on a le résultat ci-dessous, basé sur le principe suivant: $\varphi(X) = y$ si et seulement si X prend une des valeurs x_i que φ envoie sur y.

Proposition 13.9.
Soit X une variable aléatoire, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction définie sur $X(\Omega)$.
Si $\varphi(X)$ est une variable aléatoire, alors son univers image est $\varphi(X(\Omega))$, et sa loi est donnée par :

$$\forall y \in Y(\Omega), \quad P(Y=y) = P(\varphi(X)=y) = \sum_{x \in X(\Omega):\ \varphi(x)=y} P(X=x).$$

Exemple 13.6.
Si $X(\Omega) = \{-1, 0, 1\}$, alors $P(X^2=0) = P(X=0)$ et $P(X^2=1) = P(X=-1) + P(X=1)$.

Théorème 13.10.
Soient X, Y deux variables aléatoires, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction définie sur $X(\Omega)$. Si $X \sim Y$, alors $\varphi(X) \sim \varphi(Y)$.

5 Couples de variables aléatoires. Loi conjointe

5.1 Généralités

On admet qu'il existe une tribu, notée $\mathcal{B}_2$, sur $\mathbb{R}^2$, qui contient toutes les *boules ouvertes*. Elle contient en particulier tous les *rectangles* (produits cartésiens d'intervalles).

Définition 13.7.
On appelle **couple de variables aléatoires** toute fonction $(X, Y) : \Omega \to \mathbb{R}^2$, où X, Y sont deux variables aléatoires réelles.
On appelle alors **loi (conjointe)** du couple (X, Y) l'application

$$\begin{array}{ccc} \mathcal{B}_2 & \to & [0,1] \\ A & \mapsto & P((X,Y) \in A) \end{array}$$

Remarque 13.8.
Un couple de variable aléatoire discrètes est en fait une variable aléatoire à valeurs dans $\mathbb{R}^2$.

La probabilité $P(X = x_i \cap Y = y_j)$ sera notée $P(X = x_i, Y = y_j)$.

Méthode 13.2. **Déterminer la oi d'un couple**

- ✗ Tout comme pour la loi d'une (seule) variable aléatoire, on détermine la loi d'un couple aléatoire en déterminant l'ensemble de ses valeurs possibles $(X, Y)(\Omega)$, puis on calcule $P(X = x, Y = y)$ pour tout $(x, y) \in (X, Y)(\Omega)$.

  La loi conjointe du couple (X, Y) est déterminée par l'application :
  $$\varphi_{X,Y} : \begin{array}{ccc} X(\Omega) \times Y(\Omega) & \to & [0;1] \\ (i,j) & \mapsto & P\left(X=i,\ Y=j\right). \end{array}$$
- ✗ On observera que $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$ mais qu'il n'y a pas nécessairement égalité.
- ✗ On ne cherchera (donc) pas systématiquement à déterminer l'ensemble des valeurs possibles de (X, Y): on pourra se contenter de donner $P(X = x, Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

Exercice 13.12.
Une urne contient 3 boules numérotées de 1 à 3. On tire successivement et sans remise deux boules dans cette urne et on note X le numéro de la première boule et Y le numéro de la deuxième boule.

Déterminer la loi conjointe de (X, Y).

Exercice 13.13.

On lance une pièce qui, à chaque lancer, donne *Pile* avec probabilité p et *Face* avec probabilité $q = 1 - p$ (avec $p \in]0;1[$). On note X le rang d'apparition du premier *Pile*, Y le rang d'apparition du premier *Face* et Z le rang d'apparition du second *Pile*.

1. Déterminer la loi conjointe de (X, Z).
2. Déterminer la loi conjointe de (X, Y).

Proposition 13.11.

On a :

$$\sum_{i \in X(\Omega)} \sum_{j \in Y(\Omega)} P\left([X = i] \cap [Y = j]\right) = \sum_{j \in Y(\Omega)} \sum_{i \in X(\Omega)} P\left([X = i] \cap [Y = j]\right) = 1.$$

Dans la suite de cette section, J désigne une partie de $\mathbb{N}$.

Définition 13.8. **Lois marginales**

Soit (X, Y) un couple de variables aléatoires tel que $X(\Omega) = \{x_i : \ i \in I\}$ et $Y(\Omega) = \{y_j : \ j \in J\}$. On appelle **lois marginales** du couple (X, Y) les lois des variables aléatoires X et Y.

Méthode 13.3. **Déterminer les lois marginales**

Déterminer les lois marginales du coupe (X, Y), c'est déterminer les applications

$$\begin{array}{ccc} X(\Omega) & \to & \mathbb{R} \\ x_i & \mapsto & P(X = x_i) \end{array} \quad \text{et} \quad \begin{array}{ccc} Y(\Omega) & \to & \mathbb{R} \\ y_j & \mapsto & P(Y = y_j) \end{array}$$

Pour ce faire, on utilise la **formule des probabilités totales**. Plus précisément,

✗ On obtient la loi de X grâce à la formule des probabilités totales avec le s.c.e $\{[Y = y_j] : j \in J\}$.

$$\begin{aligned} P\left([X = x_i]\right) &= \sum_{j \in J} P\left(X = x_i, \ Y = y_j\right) \\ &= \sum_{j \in J} P([Y = y_j]) P_{[Y = y_j]}([X = x_i]) \end{aligned}$$

✗ On obtient la loi de Y grâce à la formule des probabilités totales avec le s.c.e $\{[X = x_i] : i \in I\}$.

$$\begin{aligned} P\left([Y = y_j]\right) &= \sum_{i \in I} P\left(X = x_i, \ Y = y_j\right) \\ &= \sum_{i \in I} P([X = x_i]) P_{[X = x_i]}([Y = y_j]) \end{aligned}$$

Si la loi du couple est résumée dans un tableau à double entrée, les lois marginales s'obtiennent en sommant les éléments de chaque ligne et de chaque colonne.

Remarque 13.9.

Si on peut déduire les lois marginales à partir de la loi conjointe, on ne peut pas déduire la loi conjointe d'un couple en fonction des lois marginales: il y a en effet dans un couple une troisième information, qui consiste en l'interaction entre les deux variables le composant.
En d'autres termes : X_1 peut avoir la même loi que X_2 et Y_1 avoir la même loi que Y_2 sans que les couples (X_1, Y_1) et (X_2, Y_2) aient la même loi.

5.2 Notion d'indépendance

Définition 13.9. **Variables aléatoires indépendantes**

Soient X, Y deux variables aléatoires. On dit que X et Y sont **indépendantes**, et on note $X \perp\!\!\!\perp Y$ si, quel que soit $(A, B) \in \mathcal{P}\left(X(\Omega)\right) \times \mathcal{P}\left(Y(\Omega)\right)$, les évènements $[X \in A]$ et $[Y \in B]$ sont indépendants.

Remarque 13.10.

- ✗ Dans le cas de deux variables aléatoires indépendantes X et Y, la donnée de la loi du couple (X, Y) est donc équivalente à la donnée des lois de X et de Y.
- ✗ L'indépendance de deux variables aléatoires ne présage en rien de leur loi. En particulier, deux variables aléatoires peuvent être indépendantes et avoir même loi.

En pratique, c'est toujours la caractérisation suivante que l'on utilise.

Proposition 13.12. **Caractérisation de l'indépendance de deux v.a.d**

Soient X, Y deux variables aléatoires. Alors X et Y sont indépendantes si et seulement si, quel que soit $(x, y) \in X(\Omega) \times Y(\Omega)$, les évènements $[X = x]$ et $[Y = y]$ sont indépendants. Soit:

$$X \perp\!\!\!\perp Y \iff \forall (x, y) \in X(\Omega) \times Y(\Omega),\ P(X=x, Y=y) = P(X=x)P(Y=y).$$

La notion d'indépendance s'étend à un nombre quelconque de variables aléatoires, en considérant des évènements mutuellement indépendants.

Définition 13.10.

Soit $(X_i)_{1 \le i \le m}$ une famille finie de variables aléatoires. Celles-ci sont dites **mutuellement indépendantes** si, quel que soit $k \in [\![1, m]\!]$, les entiers distincts $i_1, \dots, i_k$ appartenant à $[\![1, m]\!]$, et $(A_1, \dots, A_k) \in \prod_{j=1}^{k} \mathcal{P}\left(X_{i_j}(\Omega)\right)$ on a:

$$P(X_{i_1} \in A_1, \dots, X_{i_k} \in A_k) = P(X_{i_1} \in A_1) \times \cdots \times P(X_{i_k} \in A_k).$$

Remarque 13.11.

- ✗ $X_1, \dots, X_m$ sont mutuellement indépendantes si et seulement si quel que soit $(A_1, \dots, A_m) \in \prod_{j=1}^{m} \mathcal{P}\left(X_j(\Omega)\right)$, les évènements $X_1 \in A_1, \dots, X_m \in A_m$ sont mutuellement indépendants.
- ✗ Cette définition s'étend aux familles infinies dénombrables, qui sont dites mutuellement indépendantes si toute sous-famille finie est mutuellement indépendante.
- ✗ Une suite (X_i) de variables aléatoires (mutuellement) indépendantes de même loi est appelée une suite de *variables aléatoires indépendantes identiquement distribuées*, abrégé en *v.a.i.i.d.*

Exemple 13.7.

On modélise le plus souvent une suite de n expériences aléatoires indépendantes (comme par exemple n lancers de dés) par une suite (X_k) de n v.a.i.i.d (donnant par exemple le résultat de chacun des n lancers).

Proposition 13.13.

Soit $(X_i)_{1 \le i \le m}$ une famille finie de variables aléatoires. Celles-ci sont mutuellement indépendantes si et seulement si, quel que soit $k \in [\![1, m]\!]$, les entiers distincts $i_1, \dots, i_k$ appartenant à $[\![1, m]\!]$, et $(x_1, \dots, x_k) \in X_{i_1}(\Omega) \times \cdots \times X_{i_k}(\Omega)$ on a:

$$P(X_{i_1} = x_1, \dots, X_{i_k} = x_k) = P(X_{i_1} = x_1) \times \cdots \times P(X_{i_k} = x_k).$$

Théorème 13.14. **Lemme des coalitions**

Soient $1 \le m < n$ des entiers, ainsi que des variables aléatoires $X_1, \dots, X_n$. On considère encore des fonctions $f_1, \dots, f_n$ définies sur $\mathbb{R}$, ainsi que deux fonctions de plusieurs variables $F : \mathbb{R}^m \to \mathbb{R}$ et $G : \mathbb{R}^{n-m} \to \mathbb{R}$. Si $X_1, \dots, X_n$ sont mutuellement indépendantes, alors les trois suites de variables aléatoires suivantes le sont également:

i. $X_1, \dots, X_m$;
ii. $f_1(X_1), \dots, f_n(X_n)$;
iii. $F(X_1, \dots, X_m), G(X_{m+1}, \dots, X_n)$.

Remarque 13.12.
Le deuxième point ci-dessus implique par exemple que, si $X \perp\!\!\!\perp Y$ alors $f(X) \perp\!\!\!\perp g(Y)$, quelles que soient les fonctions f et g. Intuitivement cela signifie que, quels que soient les procédés déterministes employés (ici des fonctions de la variable réelle), on ne peut créer une dépendance là où il n'y en avait pas.

Exemple 13.8.

- ✗ Si X, Y, Z sont mutuellement indépendantes, alors $X+Y$ et Z sont indépendantes.
- ✗ Si X_1, X_2, X_3, X_4, X_5 sont 5 variables aléatoires discrètes discrètes mutuellement indépendantes alors les variables $X_1 - X_2 + 2X_4^2$ et $X_3 - X_5^2$ sont indépendantes.
- ✗ Si X^2 et Y^2 ne sont pas indépendantes, alors X et Y ne le sont pas non plus.

5.3 Loi d'une somme de deux variables aléatoires

Proposition 13.15. **Loi de la somme**
Soient X, Y deux variables aléatoires telle que $X(\Omega) = \{x_i| : i \in I\}$ et $Y(\Omega) = \{y_j : j \in J\}$. Alors la loi de $Z = X+Y$ est donnée par

$$Z(\Omega) = \{x_i + y_j \ \ i \in I, j \in J\}$$

et

$$\forall z \in Z(\Omega),\ P(X+Y=z) = \sum_{(i,j):\ x_i+y_j=z} P((X=x_i) \cap (Y=y_j))$$

Remarque 13.13.
Si X et Y prennent leurs valeurs dans $\mathbb{N}$, on obtient:

$$P(X+Y=n) = \sum_{i+j=n} P(X=i, Y=j) = \sum_{i=0}^{n} P(X=i, Y=n-i).$$

C'est le plus souvent cette formule que l'on rencontre en exercice.

Exercice 13.14.
Soient X et Y deux variables indépendantes suivant toutes deux la loi géométrique $\mathcal{G}(p)$.
1. Déterminer, en utilisant un s.c.e associé à X, $P(X=Y)$.
2. Montrer, en utilisant un s.c.e associé à X que pour tout $n \geq 2$: $P(X+Y=n) = (n-1)p^2q^{n-2}$.

Théorème 13.16. **Stabilité par somme de certaines lois usuelles**
Soient $X_1, X_2, \ldots, X_m$ des variables aléatoires mutuellement indépendantes. On a les résultats suivants.

i. Si $X_1, \ldots, X_m \sim \mathcal{B}(1,p)$, alors $X_1 + \cdots + X_m \sim \mathcal{B}(m,p)$.
ii. Si $X_1 \sim \mathcal{B}(n_1,p)$ et $X_2 \sim \mathcal{B}(n_2,p)$, alors $X_1 + X_2 \sim \mathcal{B}(n_1+n_2,p)$.
iii. Si $X_1 \sim \mathcal{P}(\lambda_1), \ldots, X_m \sim \mathcal{P}(\lambda_m)$, alors $X_1 + \cdots + X_m \sim \mathcal{P}(\lambda_1 + \ldots + \lambda_m)$.

6 Espérance et variance

6.1 Variables aléatoires admettant une espérance

Définition 13.11. **Espérance d'une variable aléatoire discrète**
Soit X une variable aléatoire telle que $X(\Omega) = \{x_i : \ \ i \in I\}$. On dit que X **admet une espérance** si la série $\sum x_i P(X=x_i)$ converge **absolument**.
Lorsque c'est le cas, sa somme est appelée espérance de Xs et notée $E(X)$.

Remarque 13.14.
L'hypothèse de convergence absolue permet de démontrer que la somme $\sum_{i\in I} x_i P(X = x_i)$ ne dépend pas de l'ordre dans lequel on somme les termes.

Exemple 13.9.
Toute variable aléatoire définie sur un espace probabilisé fini admet une espérance, et celle-ci est une somme finie.

Exercice 13.15.
Montrer qu'il existe une variable aléatoire X telle que $X(\Omega) = \mathbb{N}^*$ et $P(X = n) = \dfrac{1}{n(n+1)}$ pour tout $n \in \mathbb{N}^*$.
Montrer ensuite que X n'admet pas d'espérance.

Théorème 13.17. **Espérance des lois usuelles**
Soit X une variable aléatoire.

i. Si X suit une loi certaine, elle prend une unique valeur a, et alors $E(X) = a$.
ii. Si $X \sim \mathcal{U}(\llbracket 1, n\rrbracket)$, alors $E(X) = \dfrac{n+1}{2}$.
iii. Si $X \sim \mathcal{B}(1, p)$, alors $E(X) = p$.
iv. Si $X \sim \mathcal{B}(n, p)$, alors $E(X) = np$.
v. Si $X \sim \mathcal{G}(p)$, alors X admet une espérance et $E(X) = \dfrac{1}{p}$.
vi. Si $X \sim \mathcal{P}(\lambda)$, alors X admet une espérance et $E(X) = \lambda$.

Proposition 13.18. **Propriétés de l'espérance**
Soient $\lambda, \mu \in \mathbb{R}$ et X, Y deux variables aléatoires **admettant** une espérance. On a les résultats suivants.

i. **Linéarité.** $\lambda X + \mu Y$ admet une espérance, et on a $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$.
ii. **Positivité.** Si X est à valeurs positives, on a $E(X) \geq 0$. Et sous cette hypothèses, on a $E(X) = 0$ si et seulement si $P(X = 0) = 1$.
iii. **Croissance.** Si $X(\omega) \geq Y(\omega)$ pour tout $\omega \in \Omega$, alors on a $E(X) \geq E(Y)$.

Proposition 13.19.
Soient X, Y deux variables aléatoires. Si Y admet une espérance, et si $|X| \leq |Y|$, alors X admet une espérance et $E(|X|) \leq E(|Y|)$.

Proposition 13.20.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ et admettant une espérance. On a alors

$$E(X) = \sum_{n=1}^{+\infty} P(X \geq n).$$

Définition 13.12. **Variable aléatoire centrée**
Une variable aléatoire X est dite centrée si l'on a $E(X) = 0$.

Si la loi d'une composée ne s'exprime pas aisément, il n'en est pas de même de son espérance, comme le montre le résultat suivant, extrêmement utile.

Théorème 13.21. **Théorème de transfert**
Soit X une variable aléatoire telle que $X(\Omega) = \{x_i : \ i \in I\}$, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction définie sur $X(\Omega)$. Si $\varphi(X)$ est une variable aléatoire, alors on a les résultats suivants:

i. $\varphi(X)$ admet une espérance si et seulement si $\sum \varphi(x_i) P(X = x_i)$ converge absolument;
ii. lorsque l'une des deux conditions ci-dessus est vérifiée on a $E(\varphi(X)) = \sum_{i\in I} \varphi(x_i) P(X = x_i)$.

Exercice 13.16.
Soient X une loi de Poisson de paramètre $\lambda > 0$ et $Y = 1/(X+1)$. Montrer, à l'aide du théorème de transfert que Y admet une espérance et la déterminer.

Exercice 13.17.
Soit $X \sim \mathcal{B}(n,p)$. Le but est de calculer la probabilité $a = P(A)$ où A est l'évènement réalisé si et seulement si X prend une valeur paire.

1. On introduit $Y = (-1)^X$. Quelle est la loi de Y? Expliciter son espérance en fonction de a.
2. Calculer, à l'aide du théorème de transfert, l'espérance de Y en fonction de n et de p. Conclure.

Remarque 13.15.

- ✗ Ce résultat montre que pour connaitre l'espérance de $\varphi(X)$, il n'est pas nécessaire de connaitre sa loi: celle de X suffit.
- ✗ On utilise ce résultat pour calculer $E(X^2)$ sans déterminer au préalable la loi de X^2.
- ✗ Le théorème de transfert se généralise aux couples, et aux n-uplets de variables aléatoires.

Théorème 13.22.
Soient X, Y deux variables aléatoires réelles indépendantes admettant une espérance. Alors XY admet également une espérance, et on a $E(XY) = E(X)E(Y)$.

6.2 Espérance d'un produit

Remarque 13.16.
Soit (X,Y) un couple de deux variables aléatoires discrètes. Alors, le théorème de transfert permet d'écrire, sous réserve de convergence absolue,
$$E(XY) = \sum_{i \in X(\Omega)} \sum_{j \in Y(\Omega)} ijP([X=i] \cap [Y=j]).$$

Proposition 13.23.
Si X et Y sont **indépendantes** et admettent une espérance, alors XY admet une espérance, et on a
$$E(XY) = E(X)E(Y).$$

On peut utiliser le théorème précédent pour montrer que deux variables aléatoires discrètes **ne sont pas** indépendantes en vérifiant que $E(XY) \neq E(X)E(Y)$.

Exercice 13.18.
Soient X_1, X_2, et X_3, indépendantes, suivant des lois de Bernoulli de paramètre p. Montrer que les variables aléatoires $Y_1 = X_1X_2$ et $Y_2 = X_2X_3$ ne sont pas indépendantes.

7 Variables aléatoires admettant une variance

7.1 Notion de moment d'ordre quelconque

Définition 13.13. **Moment d'ordre r**
Soit X une variable aléatoire et $r \in \mathbb{N}^*$. On dit que X **admet un moment d'ordre** r si la variable aléatoire X^r admet une espérance. Lorsque c'est le cas on définit son moment d'ordre r, noté $m_r(X)$, par $m_r(X) = E(X^r)$.

Remarque 13.17.
Si X admet un moment d'ordre $r \in \mathbb{N}^*$, le théorème de transfert donne $m_r(X) = \displaystyle\sum_{x \in X(\Omega)} x^r P(X = x)$.

Proposition 13.24.
Soit X une variable aléatoire admettant un moment d'ordre 2. Alors X et $(X - E(X))^2$ admettent une espérance.

Preuve. Le premier point se déduit de $2|x| \leq x^2 + 1$. □

Remarque 13.18.
Plus généralement, si X possède un moment d'ordre $r \in \mathbb{N}^*$, alors X possède aussi un moment d'ordre $s \in \mathbb{N}^*$ pour tout $s \leq r$.

7.2 Variance

Définition 13.14.
Soit X une variable aléatoire. Si X admet un moment d'ordre 2, alors $E((X - E(X))^2)$ est appelée **variance** de X, et est notée $V(X)$.
On définit encore **l'écart-type** de X, noté $\sigma(X)$, en posant $\sigma(X) = \sqrt{V(X)}$.

Remarque 13.19.

- ✗ L'espérance d'une variable aléatoire donne sa valeur moyenne. Pour obtenir une description plus précise de la loi d'une variable aléatoire, on cherche à donner une quantité synthétisant la dispersion de celle-ci autour de sa moyenne: on utilise les notions de variance et d'écart-type.
- ✗ Plaçons-nous dans le cadre d'une expérimentation où la grandeur X mesure plusieurs fois de suite un même phénomène naturel. Même si X mesure un phénomène déterministe, on doit le considérer comme étant une variable aléatoire, ne serait-ce que parce que chaque mesure est entachée d'erreurs elles-même aléatoires. Il alors est naturel d'approximer la valeur théorique de X par la moyenne des valeurs mesurées, c'est-à-dire par son espérance $E(X)$. L'écart-type est alors une approximation de *l'erreur moyenne de mesure par rapport à la valeur théorique.*
- ✗ Si X est exprimé dans une certaine unité u, alors $V(X)$ est exprimé en u^2. C'est pour cette raison qu'en sciences physiques on considère plutôt l'écart type σ, qui est lui exprimé en u. En mathématiques on préfère la variance, qui possède des propriétés algébriques plus simples et plus naturelles.

☞ En pratique, on calcule souvent la variance à l'aide du théorème de transfert et de la formule suivante.

Proposition 13.25. **Formule de Kœnig-Huygens**
Soit X une variable aléatoire admettant un moment d'ordre 2. On a alors $V(X) = E(X^2) - E(X)^2$.

Proposition 13.26. **Variance d'une transformation affine**
Soit X une variable aléatoire admettant un moment d'ordre 2. Alors quels que soient les réels a et b, $aX + b$ admet une variance, et $V(aX + b) = a^2V(X)$.

Définition 13.15. **Variable aléatoire réduite**
Une variable aléatoire X admettant un moment d'ordre 2 est dite **réduite** si $V(X) = 1$.

Proposition 13.27.
Soit X une variable aléatoire admettant un moment d'ordre 2 dont la variance est non nulle.
Alors la variable aléatoire $X^* = \dfrac{1}{\sigma(X)}(X - E(X))$ est centrée et réduite et s'appelle la variable aléatoire centrée-réduite associée à X.

Proposition 13.28.
Soit X une variable aléatoire admettant un moment d'ordre 2. On a les résultats suivants.

i. $V(X) \geq 0$;
ii. $V(X) = 0 \iff P(X = E(X)) = 1$.

En d'autres termes, la variance est toujours positive, et s'annule si et seulement si la variable aléatoire est certaine.

Preuve. Il suffit d'appliquer la positivité de l'espérance à la variable aléatoire $(X - E(X))^2$. □

Théorème 13.29. **Variance des lois usuelles**

Soit X une variable aléatoire. On a alors les résultats suivants.

i. Si X suit une loi certaine, alors $V(X) = 0$.
ii. Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$, alors $V(X) = \dfrac{n^2-1}{12}$.
iii. Si $X \sim \mathcal{B}(1, p)$, alors $V(X) = p(1-p)$.
iv. Si $X \sim \mathcal{B}(n, p)$, alors $V(X) = np(1-p)$.
v. Si $X \sim \mathcal{G}(p)$, alors X admet une variance et $V(X) = \dfrac{1-p}{p^2}$.
vi. Si $X \sim \mathcal{P}(\lambda)$, alors X admet une variance et $V(X) = \lambda$.

Lois usuelles : synthèse

Loi de X	$X(\Omega)$	$P(X=k)$	$E(X)$	$V(X)$
$\mathcal{U}(\llbracket 1, n \rrbracket)$	$\llbracket 1, n \rrbracket$	$\dfrac{1}{n}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$
$\mathcal{B}(1,p)$	$\{0,1\}$	$\begin{cases} p & \text{si } k=1 \\ 1-p & \text{si } k=0 \end{cases}$	p	$p(1-p)$
$\mathcal{B}(n,p)$	$\llbracket 0, n \rrbracket$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
$\mathcal{G}(p)$	$\mathbb{N}^*$	$p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
$\mathcal{P}(\lambda)$	$\mathbb{N}$	$e^{-\lambda}\dfrac{\lambda^k}{k!}$	λ	λ

8 Covariance de deux variables aléatoires

La covariance se présente au départ comme un outil introduit pour calculer la variance d'une somme de variables aléatoires. On verra ensuite qu'elle possède une utilité propre.

Proposition 13.30.

Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Alors $[X - E(X)] \times [Y - E(Y)]$ admet une espérance, et l'on a

$$E\Big([X - E(X)] \times [Y - E(Y)]\Big) = E(XY) - E(X)E(Y).$$

Définition 13.16. **Covariance**

Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. On appelle **covariance** de X et de Y la quantité définie par

$$\operatorname{cov}(X, Y) = E\Big([X - E(X)] \times [Y - E(Y)]\Big) = E(XY) - E(X)E(Y).$$

Intuitivement, la covariance de X et de Y est une mesure relative (ou couplée) des dispersions de X et de Y autour de leurs moyennes. Elle se calcule grâce au théorème de transfert.

Remarque 13.20.

On déduit directement de la définition la formule $V(X) = \operatorname{cov}(X, X)$.

Théorème 13.31.

Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Si X, Y sont indépendantes, alors on a $\mathrm{cov}(X, Y) = 0$.

Remarque 13.21.

Attention! La réciproque du résultat ci-dessus est fausse.

Définition 13.17.

Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Si $\mathrm{cov}(X, Y) = 0$, on dit que X et Y dont **décorelées**.

Proposition 13.32. **Propriétés élémentaires de la covariance**

Soient X, X', Y, Y' des variables aléatoires admettant un moment d'ordre 2, et $\lambda, \mu \in \mathbb{R}$. On a les résultats suivants.

i. **Variance d'une somme**: $X + Y$ admet une variance et l'on a $V(X + Y) = V(X) + V(Y) + 2\mathrm{cov}(X, Y)$.
ii. **Symétrie de la covariance** : $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$.
iii. **Bilinéarité de la covariance**: $\mathrm{cov}(\lambda X + \mu X', Y) = \lambda \cdot \mathrm{cov}(X, Y) + \mu \cdot \mathrm{cov}(X', Y)$ et $\mathrm{cov}(X, \lambda Y + \mu Y') = \lambda \cdot \mathrm{cov}(X, Y) + \mu \cdot \mathrm{cov}(X, Y')$.

Remarque 13.22.

- ✗ On se place sur l'espace vectoriel des variables aléatoires à valeurs réelles définies sur un espace probabilisé donné. La covariance **ressemble** à un produit scalaire (elle est bilinéaire, symétrique et positive), et l'écart-type à la norme associée. Mais la covariance n'est pas définie positive, car $V(X) = 0$ n'implique pas $X = 0$.
  Par analogie, on peut toutefois considérer que la covariance permet de mesurer le *degré d'indépendance linéaire* existant entre deux variables aléatoires.

- ✗ Dans le cas d'une étude statistique, la variance d'une variable aléatoire X est l'erreur moyenne commise lors des mesures de la grandeur associée. La formule $\mathrm{cov}(X, Y) = \frac{1}{2}(V(X + Y) - V(X) - V(Y))$ indique que $\mathrm{cov}(X, Y)$ donne la variation des erreurs commises lorsque l'on passe de (X, Y) à $X + Y$. Celles-ci pouvant s'ajouter ou se compenser, la covariance est donc une donnée précieuse.

Exercice 13.19.

Soient X, Y, Z trois variables aléatoires mutuellement indépendantes. Montrer que $U = X + Y$ et $V = X + Z$ sont indépendantes si et seulement si X est certaine.

Corollaire 13.33. **Variance d'une somme de v.a.d indépendantes**

Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes, admettant un moment d'ordre 2. On a alors:

$$V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k).$$

Exercice 13.20.

Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes admettant toutes une même variance σ^2. Déterminer la variance de la variable aléatoire

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Remarque 13.23.

- ✗ Dans l'analogie décrite ci-dessus, ce théorème correspond au théorème de Pythagore.
- ✗ Ce résultat permet de retrouver facilement la variance d'une loi binomiale.

On a enfin le résultat suivant, qui provient de l'inégalité de Cauchy-Schwarz.

Proposition 13.34.
Soient X, Y deux variables aléatoires admettant un moment d'ordre 2. Alors XY admet une espérance, et l'on a
$$E(XY)^2 \leq E\left(X^2\right) E\left(Y^2\right) \iff |E(XY)| \leq \sigma(X)\sigma(Y).$$

Preuve. Résultat admis (il suffit de constater que $\phi : (X,Y) \mapsto E(XY)$ est bilinéaire, symétrique et positive, et que cela suffit pour faire la preuve de l'inégalité de Cauchy-Schwarz). □

9 Série génératrice d'une variable aléatoire entière

Dans cette section, toutes les variables aléatoires considérées seront définies sur $(\Omega, \mathcal{T}, P)$, et supposées *entières* (c'est-à-dire à valeurs dans $\mathbb{N}$).

9.1 Définitions et résultats élémentaires

Définition 13.18. **Série génératrice**
Soit X une variable aléatoire entière. On appelle **série génératrice** de X (ou fonction génératrice de X), et on note G_X la fonction qui à t associe la somme de la série entière $\sum_{n \geq 0} P(X = n)t^n$.

Proposition 13.35.
Soit X une variable aléatoire entière. Le rayon de convergence de sa série génératrice G_X est supérieur à 1, et on a
$$\forall t \in [-1, 1],\ G_X(t) = \sum_{n=0}^{+\infty} P(X = n)t^n.$$

Le théorème de transfert permet alors immédiatement d'énoncer le résultat suivant.

Corollaire 13.36.
Soit X une variable aléatoire entière. On a $\forall t \in [-1, 1],\ G_X(t) = E\left(t^X\right)$.

Théorème 13.37.
Deux variables aléatoires entières ont même loi si et seulement si elles ont même série génératrice.

Théorème 13.38.
Soit X une variable aléatoire entière, de série génératrice G_X. Alors X admet une espérance si et seulement si G_X est dérivable en 1, et lorsque c'est le cas $E(X) = G'_X(1)$.

Citons le résultat suivant, qui n'est pas au programme, mais il faut être capable de retrouver la formule donnant la variance en fonction de la série génératrice.

Théorème 13.39.
Soit X une variable aléatoire entière, de série génératrice G_X. Alors X admet un moment d'ordre 2 si et seulement si G_X est deux fois dérivable en 1, et lorsque c'est le cas $V(X) = G''_X(1) + G'_X(1) - \left(G'_X(1)\right)^2$.

Théorème 13.40. **Somme de v.a. indépendantes et séries génératrices**
Soit X, Y deux variables aléatoires entières. On note R_X (resp. R_Y) le rayon de convergence de G_X (resp. G_Y). Si X, Y sont indépendantes, alors G_{X+Y} a un rayon de convergence supérieur à $R = \min(R_X, R_Y)$, et on a $\forall t \in]-R, R[,\ G_{X+Y}(t) = G_X(t) \cdot G_Y(t)$.

9.2 Cas des lois usuelles

Théorème 13.41. **Séries génératrices des lois usuelles**

Soit $n \in \mathbb{N}^*$ et $p \in]0,1[$.

i. Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$, alors $G_X(t) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} t^k$ et le rayon de convergence est infini.

ii. Si $X \sim \mathcal{B}(p)$, alors $G_X(t) = 1 - p + pt$ et le rayon de convergence est infini.

iii. Si $X \sim \mathcal{B}(n,p)$, alors $G_X(t) = (1 - p + pt)^n$ et le rayon de convergence est infini.

iv. Si $X \sim \mathcal{G}(p)$, alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$ et le rayon de convergence est $\dfrac{1}{1-p}$.

v. Si $X \sim \mathcal{P}(\lambda)$, alors $G_X(t) = e^{\lambda(t-1)}$ et le rayon de convergence est infini.

Exercice 13.21.

Retrouver, à l'aide du **Théorème 13.40** le résultat sur la loi d'une somme de variables aléatoires indépendantes suivant une loi de Poisson.

Exercice 13.22.

On lance un dé deux fois de suite. X_1 et X_2 sont les numéros des faces obtenues lors, respectivement, du premier et du second lancer. $S = X_1 + X_2$ est la somme de ces deux numéros.

1. Trouver la fonction génératrice de S.

2. Est-il possible de truquer le dé pour que S suive une loi uniforme sur $\{2, \dots, 12\}$?

10 Quelques résultats asymptotiques

10.1 Approximation de la loi binomiale par une loi de Poisson

Le résultat suivant n'est pas au programme, mais permet de justifier que la loi de Poisson est appelée loi des évènements rares.

Proposition 13.42. **Convergence en loi de la binomiale vers une Poisson**

Soit $(p_n)_{n\in\mathbb{N}}$ une suite de réels compris dans $]0,1[$ et $\lambda > 0$ tels que $\lim\limits_{n\to+\infty} np_n = \lambda$.

Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires telles que $X_n \sim \mathcal{B}(n, p_n)$ pour tout $n \in \mathbb{N}$, alors

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

On dit que la suite (X_n) **converge en loi** vers une loi de Poisson. On note

$$X_n \xrightarrow[n\to+\infty]{\mathcal{L}} Z, \quad Z \sim \mathcal{P}(\lambda).$$

Remarque 13.24.

✗ L'intérêt de ce théorème réside dans le fait que le calcul numérique d'une loi de Poisson est plus simple que celui d'une loi binomiale.

✗ On peut par exemple utiliser ce résultat si $p_n = \dfrac{\lambda}{n}$.

✗ Dans une situation concrète, si n est grand et p est petit, on peut utiliser ce résultat si $np = \lambda$ reste petit par rapport à n. En pratique le résultat est exploitable si $n > 20, p \leq \dfrac{1}{10}$ et $np \leq 5$.

✗ Soit une urne contenant une proportion $p \in]0,1[$ de boules blanches. Si p est proche de 0 et n est assez grand, le théorème ci-dessus laisse penser que le nombre de boules blanches obtenues après n tirages suit approximativement une loi de Poisson de paramètre $\lambda = np$. C'est pour cette raison que la loi de Poisson est appelée la *loi des évènements rares*.

✗ On pourra retenir le résultat sous la forme suivante : $p_n \underset{n\to+\infty}{\sim} \dfrac{\lambda}{n} \Longrightarrow \mathcal{B}(n, p_n) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{P}(\lambda)$.

On illustre ce résultat avec `Python`.

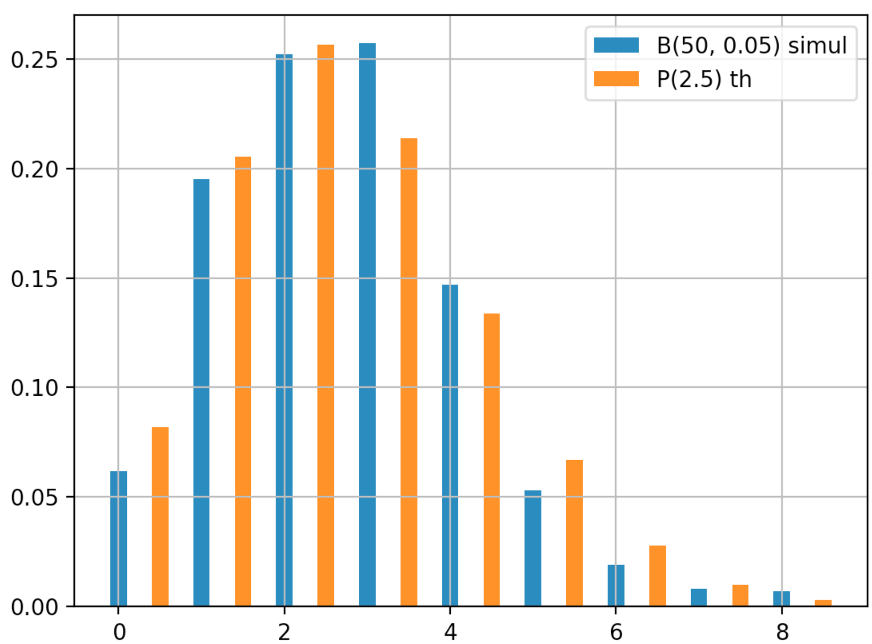
```
import numpy as np
import numpy.random as rd
import matplotlib.pyplot as plt

X=[rd.binomial(50, 0.05) for k in range(1000)] #echantillon taille 1000
m=max(X)
freq=np.zeros(m+1)
for k in range(len(X)):
    freq[X[k]]+=1
freq=freq/1000

def distr_poisson_th(lam,n):
    y=np.zeros(n+1)
    y[0]=np.exp(-lam)
    for k in range(1, n+1):
        y[k]=y[k-1]*lam/k
    return y

N=[k for k in range(m+1)]
Z=distr_poisson_th(2.5, m)

plt.grid()
plt.bar(N, freq, width=0.2, label='B(50, 0.05) simul')
plt.bar([k+0.5 for k in range(m+1)], Z, width=0.2, label='P(2.5) th')
plt.legend()
plt.show()
```



En bleu, les hauteur des bâtons correspondent aux *fréquences* d'apparition de chaque valeur lors de la simulation de 1000 variables aléatoires de loi $\mathcal{B}(50, 0.05)$.
En orange, les hauteurs *théoriques* de la distribution d'une loi de Poisson de paramètre 2.5.

10.2 Loi (faible) des grands nombres

Proposition 13.43. **Inégalité de Markov**

Soit X une variable aléatoire à valeurs dans $\mathbb{R}_+$ admettant une espérance. Alors :

$$\forall a > 0,\ P(X \geq a) \leq \frac{E(X)}{a}.$$

Proposition 13.44. **Inégalité de Bienaymé-Tchebychev**

Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors :

$$\forall \varepsilon > 0,\ P\left(|X - E(X)| \geq \varepsilon\right) \leq \frac{V(X)}{\varepsilon^2}.$$

Preuve. Il suffit d'appliquer l'inégalité de Markov à la variable aléatoire $(X - E(X))^2$. □

Remarque 13.25.

- ✗ Cette inégalité majore la probabilité que X s'éloigne d'au moins ε de sa moyenne.
- ✗ Cette probabilité est d'autant plus petite que ε est grand : plus on s'éloigne de la moyenne moins on a de chance de trouver X.
- ✗ Cette probabilité est d'autant plus petite que $V(X)$ est petit : une variance faible correspond à une variable aléatoire resserrée autour de sa moyenne.
- ✗ Enfin, dans le cadre d'une expérimentation où l'on mesure plusieurs fois de suite un même phénomène naturel X, cette inégalité majore la probabilité que l'on ait commis une erreur de mesure supérieure à un seuil ε donné (on considère en général un seuil critique).

Définition 13.19. **Moyenne empirique**

Soit $(X_n)_{n\in\mathbb{N}}$ une famille de variables aléatoires indépendantes suivant la même loi. Pour tout $n \in \mathbb{N}$, on appelle **moyenne empirique** de $X_1, \ldots, X_n$ la variable aléatoire $\overline{X}_n$ définie par

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Exemple 13.10.

On lance n fois une pièce de monnaie qui renvoie *Pile* avec probabilité p. Le résultat du k-ième lancer correspond à une variable aléatoire X_k, qui prend la valeur 1 si l'on a obtenu *Pile* et 0 sinon. Les X_k sont indépendantes et suivent la même loi. Dans ce cas, $\overline{X}_n$ est le nombre moyen de *Pile* obtenus. La *loi faible des grands nombres* établira que ce nombre moyen tend vers p lorsque n tend vers $+\infty$, ce qui permet d'ailleurs d'estimer p par observations si jamais le paramètre est inconnu.

Théorème 13.45. **Loi (faible) des grands nombres**

Soit $(X_n)_{n\in\mathbb{N}}$ une famille de variables aléatoires indépendantes suivant la même loi, admettant une espérance μ et une variance σ^2.
Alors, pour tout $n \in \mathbb{N}$ et tout réel $\varepsilon > 0$ on a

$$P\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \quad \text{et} \quad \lim_{n\to+\infty} P\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) = 0.$$

Remarque 13.26.

- ✗ Ce résultat établit que, pour tout $\varepsilon > 0$, il devient improbable que $\overline{X}_n \notin]\mu - \varepsilon, \mu + \varepsilon[$ lorsque $n \to +\infty$ (mais pas que cela est impossible). On a de plus un majorant de cette probabilité.
- ✗ Ce résultat est d'une grande importance théorique et pratique. Il indique par exemple qu'en mesurant n fois un phénomène physique, on obtient une moyenne empirique $\overline{X}_n$ qui tend vers la moyenne théorique μ lorsque $n \to +\infty$.
- ✗ Ce résultat permet également d'estimer la valeur moyenne d'une variable dans une population à l'aide d'un sondage.

Exemple 13.11.

Soit $\varepsilon = \dfrac{1}{10}$, et supposons $V(X_1) \leq 1$. Pour $n > 10^4$, on a

$$P\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{1}{100}.$$

On dit que

$$\left]\overline{X}_n - \frac{1}{10}, \overline{X}_n + \frac{1}{10}\right[$$

est un *intervalle de confiance de μ au seuil de confiance de 99% (ou au seuil de risque de 1%).*

$\overline{X}_n$ est appelé un **estimateur** de μ, dont l'écart à μ sera donc inférieur à $\dfrac{1}{10}$ dans 99% des cas.

11 Sélection d'exercices

Exercice 13.23. **Simulation de lois usuelles à partir de la loi $\mathcal{U}(\{0,1\})$**

1. Écrire, en `Python`, une fonction `Binomiale(n,p)` qui renvoie une simulation de la loi $\mathcal{B}(n,p)$, **sans utiliser** la commande `binomial(n,p)`.
2. Écrire, en `Python`, une fonction `Geometrique(p)` qui renvoie une simulation de la loi $\mathcal{G}(p)$, **sans utiliser** la commande `geometric(p)`.

Exercice 13.24. **Variables indicatrices**

Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé et A et B deux évènements.

On appelle *variable indicatrice* de l'évènement A la variable aléatoire notée $\mathbb{1}_A$ définie par

$$\forall \omega \in \Omega, \qquad \mathbb{1}_A(\omega) = \begin{cases} 1, & \text{si } \omega \in A \\ 0, & \text{si } \omega \notin A \end{cases}$$

1. Reconnaître la loi de $\mathbb{1}_A$. Préciser son espérance et sa variance.
2. Montrer que $\mathbb{1}_A \times \mathbb{1}_B = \mathbb{1}_{A\cap B}$.
3. Calculer alors $\text{Cov}(\mathbb{1}_A, \mathbb{1}_B)$. Vérifier que $\text{Cov}(\mathbb{1}_A, \mathbb{1}_B) \leq P(A \cap B)$.

Exercice 13.25.

Soit $n \geqslant 3$ un entier. n personnes jettent chacune une pièce équilibrée, et une personne gagne si elle a obtenu le contraire de toutes les autres.

1. Quelle est la probabilité qu'une partie donnée ait un gagnant ?
2. Combien de parties faut-il jouer en moyenne pour obtenir un gagnant ?

Exercice 13.26.

Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ. On définit Y par $Y = X^2$ si X est pair non nul, et $Y = 0$ sinon. Déterminer la loi de Y et son espérance.

Exercice 13.27.

Une urne contient des boules vertes et des boules blanches, indiscernables au toucher. La proportion de boules vertes est p, $0 < p < 1$; la proportion de boules blanches est $1 - p$.
On effectue une suite de tirages successifs d'une boule avec remise. (Toute boule tirée de l'urne y est remise avant de procéder au tirage suivant.)

On note N_V la variable aléatoire égale au nombre de tirages nécessaires pour obtenir la première boule verte, et N_B la variable aléatoire égale au nombre de tirages nécessaires pour obtenir la première boule blanche.

1. Quelles sont les lois des variables aléatoires N_V et N_B ? Rappeler la valeur de leurs espérances.

On définit la variable aléatoire X de sorte que que $X = i$ si les i premières boules tirées sont blanches et la $(i+1)-$ème est verte, ou si les i premières boules tirées sont vertes et la $(i+1)-$ème est blanche.

2. Déterminer la loi de la variable aléatoire X.

3. Montrer que la variable aléatoire X admet une espérance et que
$$E(X)=\frac{p}{1-p}+\frac{1-p}{p}.$$

4. Montrer que $E(X)$ est minimale lorsque $p=\frac{1}{2}$, et calculer cette valeur minimale.

Exercice 13.28. Guichets

Une gare dispose de deux guichets. Trois clients notés C_1, C_2, C_3 arrivent en même temps. Les clients C_1 et C_2 se font servir tandis que le client C_3 attend puis effectue son opération dès que l'un des deux guichets se libère.

On définit X_1, X_2, X_3 les variables aléatoires égales à la durée de l'opération des clients C_1, C_2,. C_3 respectivement. Ces durées sont mesurées en minutes et arrondies à l'unité supérieure ou égale. On suppose que les variables aléatoires X_1, X_2, X_3 suivent la loi géométrique de paramètre p, $p\in\,]0;1[$ et qu'elles sont indépendantes. On note $q=1-p$.

On note A l'événement, : "C_3 termine en dernier son opération". On se propose de calculer la probabilité de A. On remarque que l'événement A est égal à l'événement
$$\left(\left(\min\left(X_1,X_2\right)+X_3\right)>\max\left(X_1,X_2\right)\right).$$

1. Rappeler la loi de X_1 ainsi que son espérance $E\left(X_1\right)$ et sa variance $V\left(X_1\right)$. On définit la variable aléatoire Δ par $\Delta=\left|X_1-X_2\right|$.

2. Calculer la probabilité $\mathrm{P}\left(\Delta=0\right)$.

3. Soit n un entier naturel non nul.

a. Justifier que
$$\mathrm{P}\left(X_1-X_2=n\right)=\sum_{k=1}^{+\infty}\mathrm{P}\left(X_2=k\right)\mathrm{P}\left(X_1=n+k\right).$$

b. En déduire que
$$\mathrm{P}\left(\Delta=n\right)=2\frac{pq^n}{1+q}.$$

4. **a.** Montrer que Δ admet une espérance $E\left(\Delta\right)$ et la calculer.

b. Montrer que
$$E\left(\left(X_1-X_2\right)^2\right)=2V\left(X_1\right).$$
En déduire que Δ admet une variance $V\left(\Delta\right)$ et la calculer.

5. Montrer que l'événement A est égal à l'événement $\left(X_3>\Delta\right)$.

6. **a.** En déduire que
$$\mathrm{P}\left(A\right)=\sum_{k=0}^{+\infty}\mathrm{P}\left(\Delta=k\right)\mathrm{P}\left(X_3>k\right).$$

b. Exprimer $\mathrm{P}\left(A\right)$ à l'aide de p et q.

Exercice 13.29.

Soient $\lambda,a>0$ tels que $a\neq1$. On suppose avoir un couple (X,Y) de variables aléatoires tels que :
$$\forall(i,j)\in\mathbb{N}^2,\ P(X=i,Y=j)=\begin{cases}\dfrac{\lambda a^i}{j!} & \text{si } 0\leqslant i\leqslant j\\ 0 & \text{sinon}\end{cases}.$$

1. Déterminer la loi de Y et calculer λ pour qu'il s'agisse bien d'une loi de couple.

2. Les variables X et Y sont-elles indépendantes ?

Exercice 13.30.

On lance une pièce équilibrée plusieurs fois de suite. X est le rang pour lequel on obtient pour la deuxième fois Pile.

On tire alors une boule dans une urne contenant $X-1$ boules numérotées de 1 à $X-1$, et on appelle Y le numéro de la boule obtenue.

1. X et Y sont-elles indépendantes ?
2. Déterminer les variances de X et Y, ainsi que leur covariance.

Exercice 13.31. **L'ascenseur**

Un immeuble de p étages est équipé d'un ascenseur. N personnes montent dans l'ascenseur au rez de chaussée et descendent chacune à un étage au hasard et de façon indépendante.
Soit X le nombre d'arrêts de l'ascenseur.
On note X_i la variable valant 1 si l'ascenseur s'arrête à l'étage i et 0 sinon et on note E_k la variable aléatoire qui prend la valeur de l'étage où descend la $k-$ième personne.

1. Pour tout $k \in [\![1; N]\!]$, quelle est la loi de E_k?
2. En utilisant les variables aléatoires E_k, déterminer $P(X_i = 0)$.
3. Calculer $E(X)$.
4. Déterminer pour i différent de j : $P(X_i = 0 \cap X_j = 0)$.
En déduire $P(X_i X_j = 0)$.
5. Déduire de la question précédente $\mathrm{Cov}(X_i, X_j)$ puis calculer alors $V(X)$.

Exercice 13.32. **Lois sans mémoire**

On s'intéresse dans cet exercice à une modélisation particulière de la durée de vie d'un composant électronique ou électrique (par exemple d'une ampoule ou d'un pixel d'écran).

Soit X est une variable aléatoire discrète à valeurs entières, on dit que X suit une loi **sans mémoire** si,

✗ X est à valeurs positives ;
✗ pour tout $k \in \mathbb{N}$ et pour tout $j \in \mathbb{N}$, $P(X > k) > 0$ et

$$P_{[X>k]}(X > k + j) = P(X > j).$$

On peut interpréter la définition comme suit; si le composant a déjà survécu un temps k, la probabilité qu'il survive un temps j supplémentaire est la même qu'au tout début de son utilisation. Tout se passe comme si le composant avait *oublié* qu'il avait déjà vécu un temps t.

1. Soit X une variable aléatoire qui suit une loi sans mémoire. Montrer que $P(X > 0) = 1$ puis que $P(X = 0) = 0$.

On pourra ainsi considérer dans la suite que $X(\Omega) \subset]0, +\infty[$ *si* X *suit une loi sans mémoire.*

L'objet de cette partie est de démontrer le théorème suivant.

Théorème. **Lois sans mémoires discrètes**

Soit X une variable aléatoire discrète. Les deux propriétés qui suivent sont équivalentes :
i. $X \sim \mathcal{G}(p)$ où $p \in]0, 1[$.
ii. X est discrète à valeurs entières et X suit une loi sans mémoire.

2. Soit $p \in]0, 1[$ et soit $X \sim \mathcal{G}(p)$. On pose $q = 1 - p$.

a. Montrer que, pour tout $k \in \mathbb{N}$, $P(X > k) = q^k$.
b. Soit $k \in \mathbb{N}$ et soit $j \in \mathbb{N}$. Montrer que

$$P_{[X>k]}(X > k + j) = P(X > j).$$

c. Conclure.

3. Soit X une variable aléatoire discrète à valeurs entières et qui suit une loi sans mémoire. On pose $q = P(X > 1)$ et, pour tout $k \in \mathbb{N}$, $u_k = P(X > k)$.

a. Justifier : $X(\Omega) \subset \mathbb{N}^*$.
b. Justifier : $q > 0$.

c. Montrer que, pour tout $k \in \mathbb{N}$, $u_{k+1} = qu_k$. En déduire une expression simple de u_k pour tout $k \in \mathbb{N}$.
d. Montrer que, pour tout $k \in \mathbb{N}^*$, $P(X = k) = u_{k-1} - u_k$.
e. Montrer que $q < 1$. On pourra faire un raisonnement par l'absurde.
f. Conclure.

Exercice 13.33. — Qui conditionne qui

On considère, dans cet exercice, deux variables aléatoires X et Y (définies sur le même espace probabilisé qu'on ne précise pas ici) à valeurs dans $\mathbb{N}$.
On suppose, pour tout l'exercice, que, pour tout $n \in \mathbb{N}$, $P(X = n) > 0$ et que la loi conditionnelle de Y sachant $[X = n]$ est la loi uniforme $\mathcal{U}(\llbracket 0, n \rrbracket)$.

1. Établir que, pour tout $k \in \mathbb{N}$, $P(Y = k) = \displaystyle\sum_{n=k}^{+\infty} P_{[X=n]}(Y = k)P(X = n)$.

2. Montrer, sans déterminer explicitement leurs lois, que $X - Y$ et Y suivent la même loi.

3. Montrer que, pour tout $k \in \mathbb{N}$, on a $P(Y = k) = \displaystyle\sum_{n=k}^{+\infty} \frac{1}{n+1} P(X = n)$.

4. On suppose, dans cette question que $Z = Y + 1$ suit une loi géométrique de paramètre $p \in]0; 1[$.

a. Calculer, pour $k \in \mathbb{N}$, $P(Y = k) - P(Y = k + 1)$.
b. En déduire la loi de X.
c. Vérifier que X admet bien une espérance et la calculer.
d. Montrer que $X - Y$ et Y sont indépendantes.
e. En déduire la valeur de $\text{cov}(X, Y)$.

Exercice 13.34. — Classique, facile et incontournable

On effectue une succession infinie de lancers indépendants d'une pièce donnant pile avec probabilité p. On note $q = 1 - p$.
On dit que la première série est de longueur $n > 1$ si les n premiers lancers ont amené le même côté de la pièce et le $(n + 1)$-ième l'autre côté. De même la deuxième série commence au lancer suivant la fin de la première série et se termine au lancer précédant un changement de côté.

Soient X_1 et X_2 les variables aléatoires égales aux longueurs de la première et deuxième série.

1. Déterminer la loi de X_1.

2. Montrer que X_1 admet une espérance et que $E(X_1) = \dfrac{p}{q} + \dfrac{q}{p}$.

3. Déterminer la loi du couple (X_1, X_2).

4. En déduire la loi de X_2.

5. Montrer que X_2 admet une espérance et que $E(X_2) = 2$.

6. Montrer que $E(X_1X_2)$ existe et que $E(X_1X_2) = \dfrac{1}{q} + \dfrac{1}{p}$.

7. On suppose que $p = 1/2$. Montrer que X_1 et X_2 sont indépendantes.

8. On suppose que $p \neq 1/2$.
a. En considérant $P([X_1 = 1] \cap [X_2 = 1])$ montrer que X_1 et X_2 ne sont pas indépendantes.
b. Vérifier que $\text{cov}(X_1, X_2) = 4 - \dfrac{1}{pq}$ puis en déduire une nouvelle preuve que X_1 et X_2 ne sont pas indépendantes.

Exercice 13.35.

On considère une suite (X_n) de variables de Bernoulli, $X_n \hookrightarrow \mathcal{B}(p_n)$, un réel $\lambda \in]0; 1[$ et on suppose que

$$P_{[X_n=1]}(X_{n+1} = 1) = p_n, \qquad P_{[X_n=0]}(X_{n+1} = 1) = \lambda p_n.$$

1. Montrer que
$$p_{n+1} = (1-\lambda)p_n^2 + \lambda p_n.$$
2. Calculer $\mathrm{Cov}(X_n, X_{n+1})$ en fonction de p_{n+1} et p_n, puis en fonction de p_n et λ.
3. Les variables X_n et X_{n+1} sont-elles indépendantes?

Exercice 13.36. **Déterminant aléatoire**

Soit $n \geq 1$ un entier. On introduit l'ensemble $\mathcal{R}_n$ des matrices *aléatoires* $M \in \mathcal{M}_n(\mathbb{R})$ dont tous les coefficients sont des variables aléatoires indépendantes de même loi uniforme $\mathcal{U}(\{-1,1\})$:
$$M = (X_{i,j})_{1\leq i,j\leq n}, \qquad \forall (i,j) \in [\![1,n]\!]^2, \quad X_{i,j} \sim \mathcal{U}(\{-1,1\}).$$
On s'intéresse aux variables aléatoires réelles $\det(M)$, où $M \in \mathcal{R}_n$.

1. Déterminer, pour $n \in \mathbb{N}^*$, pour toute matrice $M \in \mathcal{R}_n$, $E(\det(M))$.
2. Montrer, par récurrence sur $n \in \mathbb{N}^*$, que $\forall M \in \mathcal{R}_n$, $V(\det(M)) = n!$.
3. Soit $M \in \mathcal{R}_n$. On pose $\Delta = \det(M)$. Soit $(a_n)_{n\geq 1}$ une suite de réels strictement positifs. Déterminer une condition suffisante sur a_n pour que
$$\lim_{n\to+\infty} P(\Delta > a_n) = 0.$$

Exercice 13.37.

Soit X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètre a. Soit $Z = X + 3Y$.

1. Calculer l'espérance et la variance de Z.
2. Déterminer la fonction génératrice de Z. Retrouver alors l'espérance et la variance de Z.

Exercice 13.38.

Soit $p \in \,]0,1[$, c un réel et X une variable aléatoire à valeurs dans $\mathbb{N}$ telle que, pour tout $n \in \mathbb{N}$ on a $P(X = n) = \dfrac{p^{n+1}}{(n+1)\,c}$.

1. Déterminer c.
2. Déterminer la fonction génératrice de X puis calculer son espérance et sa variance.

Exercice 13.39.

Soient X, Y deux variables aléatoires indépendantes, de même loi, et à valeurs dans $\mathbb{N}$. On suppose que $Z = X + Y + 1$ suit une loi géométrique de paramètre $p \in]0,1[$.

1. Déterminer l'espérance de X.
2. Calculer la fonction génératrice de X.
3. En déduire la loi de X.

Exercice 13.40.

Soit (X_n) une suite de v.a. (mutuellement) indépendantes avec $X_n \sim \mathcal{B}(p_n)$. On note $\overline{X}_n = (X_1 + ... + X_n)/n$. On veut montrer que, pour tout $\varepsilon > 0$,
$$\lim_{n\to+\infty} P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{i=1}^{n} p_i\right| \geq \varepsilon\right) = 0.$$

1. Pourquoi ne peut-on pas appliquer la loi faible des grands nombres?
2. Quelle est l'espérance de $\overline{X}_n$? Sa variance? Démontrer que $V(\overline{X}_n) \leq \dfrac{1}{n}$.
3. En déduire le résultat.

Exercice 13.41.

Une entreprise fabrique des jouets. Malgré les contrôles, on observe que 0,6 % en moyenne des jouets mis sur le marché sont défectueux.
On considère un lot de n jouets, et on note X le nombre de ceux qui sont défectueux dans ce lot.

1. Quelle est la loi de X ?
2. Dans cette question, on prend $n = 500$. Donner une approximation de la probabilité d'avoir, dans ce lot de 500 jouets, au plus deux jouets défectueux.

Exercice 13.42.

Soit $n \in \mathbb{N}^*$ X une variable aléatoire suivant la loi géométrique $\mathcal{G}\left(\frac{1}{n}\right)$.

1. Montrer que $P(X \geqslant n^2) \leqslant \frac{1}{n}$.

2. Montrer que $P(|X-n| \geqslant n) \leqslant 1-\frac{1}{n}$. En déduire que $P(X \geqslant 2n) \leqslant 1-\frac{1}{n}$.

Exercice 13.43. **Théorème de Bernoulli**

Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables indépendantes suivant toutes une loi de Bernoulli de même paramètre $p \in]0,1[$.

1. Montrer que pour $\varepsilon > 0$ on a $P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| \geqslant \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}$. Quelle interprétation peut-on en donner ?

2. En lançant 1000 fois une pièce, on a obtenu 480 fois Pile. Peut-on, avec un seuil de confiance de 90%, penser que la pièce est équilibrée ?

12 Bonus. Le *German Tank Problem*

Le contexte. Été 1943, les Alliés essaient de percer le bloc de l'Axe en créant un nouveau front via l'Italie. Ils rencontrent un nouveau type de char allemand, le bien nommé *Sonderkraftfahrzeug 171* plus connu des afficionados de machines de combat sous le nom de *Panther*.

Ce dernier est mieux équipé et plus performant que ceux rencontrés jusqu'alors. Il a été conçu en réponse à l'excellent *T-34* utilisé par les soviétiques sur le front de l'Est.
Sans rentrer dans les détails d'armement de cette subtile et sympathique machine, il peut percer les défenses et détruire la majorité des tank alliés.

Néanmoins, malgré sa puissance théorique, celui-ci ne peut avoir un réel impact sur l'issue de la guerre que si le nombre d'unités produites est suffisant. Il apparait crucial pour les Alliés de déterminer ou plutôt d'*estimer* combien de *Panther* étaient produits. La tâche fut confiée à [la] *Economic Warfare Division of the American Embassy in London*[1].

La modélisation. On suppose que l'ennemi produit une série de chars immatriculés par des entiers en commençant par 1. En plus de cela, quelle que soit la date de production du char, ses années de service, ou encore son numéro de série, la distribution des numéros d'immatriculation est considérée comme étant uniforme dès l'instant où on mène l'analyse.

Dans notre modélisation, les allemands disposent de N tanks numérotés de 1 à N. Les force alliées observent aléatoirement, uniformément et "avec remise" n numéros de séries $(X_1, ..., X_n)$ et cherchent à estimer le paramètre N.

On considère dans tout le problème un $n-$échantillon $(X_1, ..., X_n)$ de la loi $\mathcal{U}([\![1, N]\!])$ et une première idée serait de considérer la *moyenne* des valeurs observées, on commence donc par poser

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

1. Que vaut $E(\overline{X}_n)$. Il serait *pratique* qu'en moyenne, la variable aléatoire choisie pour estimer N renvoie N. Expliciter alors une variable aléatoire T_n, fonction du $n-$échantillon $(X_1, ..., X_n)$ telle que $E(T_n) = N$.
Un $n-$échantillon est une famille de n variables aléatoires indépendantes de même loi.

[1]Comme le raconte l'article *An Empirical Approach to Economic Intelligence in World War II*, R. Ruggles & H. Brodie, Journal of the American Statistical Association **42**-237 (1947), 72-91

2. Calculer $V(T_n)$ et montrer, à l'aide de l'inégalité de Bienaymé-Tchebychev, que

$$\forall \varepsilon > 0, \qquad \lim_{n\to+\infty} P\left(|T_n - N| > \varepsilon\right) = 0.$$

3. Ce résultat semble affirmer que l'*estimateur* T_n *converge* (dans un certain sens) vers N, c'est à dire que si n est assez grand (si on dispose de suffisamment de données) la valeur approchée de N obtenue avec la définition de T_n appliquée à l'observation est proche de N. En revanche, imaginons qu'on ait 5 données $X = [8, 322, 15, 135, 69]$, que vaut l'estimation obtenue avec T_5 correspondant à cette observation? Que cela motive-t-il ?

On introduit alors le nouvel *estimateur*, c'est à dire une nouvelle fonction du $n-$échantillon

$$M_n = \max(X_1, ..., X_n)$$

4. Calculer, pour tout $k \in \mathbb{N}$, $P(M_n \leq k)$.

5. Soit Y un v.a à valeurs dans $[\![1; N]\!]$. Montrer que $E(Y) = \displaystyle\sum_{k=0}^{N-1} P(Y > k)$.

6. Montrer alors que

$$E(M_n) = N - \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^n.$$

7. Vérifier que, pour tout $k \in [\![0; N-1]\!]$,

$$0 \leq \left(\frac{k}{N}\right)^n \leq N \int_{k/N}^{(k+1)/N} t^n \mathrm{d}t.$$

8. En déduire que

$$N - \frac{N}{n+1} \leq E(M_n) \leq N$$

puis que

$$\lim_{n\to+\infty} E(M_n) = N.$$

(On dit que l'estimateur M_n est *asymptotiquement sans biais.*)

Si l'estimateur M_n parait naturel, il a clairement un défaut; il sous-estime nécessairement N (puisqu'il renverra toujours une valeurs inférieure (ou égale) à N). On va donc essayer d'y apporter une légère *correction.*

Commençons par introduire le numéro du plus petit tank observé $m_n = \min(X_1, ..., X_n)$.

Comme N est inconnu, on ne connait pas l'écart entre N et M_n, mais il parait raisonnable de penser qu'il y a (en moyenne) autant de tanks *non observés* entre M_n et N qu'entre 1 et m_n. Entre le plus petit numéro observé et le tank avec le numéro de série 1, il y a $m_n - 1$ numéros de tank.
On pense alors à ajouter la correction

$$\tilde{M}_n = M_n + (m_n - 1).$$

9. En s'inspirant des calculs précédents pour M_n, déterminer $E(m_n)$ sous forme d'une somme qu'on ne cherchera pas à simplifier.

10. Montrer que $\tilde{M}_n$ vérifie maintenant $E(\tilde{M}_n) = N$.