Chapitre ALEA.15.

Couples aléatoires

Résumé & Plan

L'objectif de ce chapitre est d'étudier les couples aléatoires, *i.e.* les couples de deux variables aléatoires, principalement discrètes d'une part et à densité d'autre part. Nous verrons alors comment calculer la loi de sommes de variables aléatoires indépendantes.

Un être humain possède environ 150 000 cheveux (en tout cas, moins d'un million). Comme la ville de Paris compte 2,141 millions d'habitants, d'après le principe des tiroirs de Dirichlet*, il existe au moins deux personnes à Paris qui ont exactement le même nombre de cheveux.*

— **Le saviez-vous ?**

Cadre

Dans tout le chapitre, et même lorsque cela n'est pas précisé, on se fixe un espace probabilisé $(\Omega, \mathscr{T}, \mathbf{P})$.

1. COUPLES ALÉATOIRES

1.1. Généralités

Définition ALEA.15.1 | Couple aléatoire

On appelle *couple aléatoire* toute application $Z : \Omega \longrightarrow \mathbf{R}^2$ de la forme

$$Z : \Omega \longrightarrow (X(\omega), Y(\omega)),$$

où X, Y sont deux variables aléatoires réelles. On note $Z = (X, Y)$ cette application.
On dit que :

- Z est un *couple aléatoire discret* si X, Y sont discrètes,
- Z est un *couple aléatoire à densité* si X, Y sont à densité.

Notez bien que la notion de couple aléatoire dépend, comme pour la notion de variable aléatoire, de la tribu sous-jacente $\mathscr{T}$.

Proposition ALEA.15.1 | À propos de l'univers-image

Soit $Z = (X, Y)$ un couple aléatoire. Alors

$$Z(\Omega) = \{(X(\omega), Y(\omega)),\ \omega \in \Omega\} \subset \left\{(X(\omega_1), Y(\omega_2)),\ (\omega_1, \omega_2) \in \Omega^2 \times \Omega\right\} = X(\Omega) \times Y(\Omega).$$

Voyons deux exemples dans le cas discret puis un exemple à densité.

Exemple 1 — Notons $\Omega = \{a, b, c\}$, et X, Y deux variables aléatoires telles que $X(a) = X(b) = 1, X(c) = 2$ et $Y(a) = Y(c) = 3, Y(b) = 1$. On note $Z = (X, Y)$. Déterminer $X(\Omega)$, $X(\Omega) \times Y(\Omega)$ et $Z(\Omega)$. Le couple Z est un couple aléatoire discret.

Exemple 2 — On lance deux pièces. On appelle X le nombre de « pile » obtenus et on pose $Y = 1$ si on obtient au moins une fois face, et 0 sinon. Les quantités X et Y sont des variables aléatoires et le couple (X, Y) forme un vecteur aléatoire. De manière formelle, on considère

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

muni de la probabilité uniforme si l'expérience se déroule avec des pièces équilibrées, et X et Y sont définies par

$$X(PP) = 2, \quad Y(PP) = 0, \quad X(PF) = 1, \quad Y(PF) = 1,$$

$$X(FP) = 1, \quad Y(FP) = 1, \quad X(FF) = 0, \quad Y(FF) = 1.$$

Le vecteur (X, Y) est lui défini par

$$(X,Y)(PP) = (X(PP), \quad Y(PP)) = (2,0), \quad (X,Y)(PF) = (1,1),$$

$$(X,Y)(FP) = (1,1), \quad (X,Y)(FF) = (0,1).$$

Exemple 3 — ***à densité*** On observe deux bactéries et on s'intéresse à leurs durées de vie X, Y qui suivent une loi exponentielle de paramètre $\lambda > 0$, supposées indépendantes. Alors (X, Y) est un couple aléatoire à densité, mais c'est le cas aussi de $(X, \max(X, Y))$ puisque nous avons montré dans le Chapter ALEA.14 que $\max(X, Y)$ est à densité.

1.2. Fonction de répartition & Loi

Définition ALEA.15.2 | Fonction de répartition

On appelle *fonction de répartition* du couple aléatoire $Z = (X, Y)$ l'application $F_Z : \mathbf{R}^2 \longrightarrow [0,1]$ donnée, pour $(x, y) \in \mathbf{R}^2$, par

$$F_Z(x, y) = \mathbf{P}(X \leqslant x, Y \leqslant y).$$

La fonction de répartition d'un couple aléatoire est donc une fonction de deux variables.

Définition/Proposition ALEA.15.1 | Loi

On appelle *loi* du couple aléatoire $Z = (X, Y)$, la fonction $\mathbf{P}_Z$ qui à un pavé $I \times J \subset \mathbf{R}^2$ (I, J sont des intervalles) associe $\mathbf{P}_Z(I \times J)$ définie par :

$$\mathbf{P}_Z(I \times J) = \mathbf{P}(X \in I, Y \in J).$$

Proposition ALEA.15.2 | La loi ne dit rien sur l'égalité

Soient Z, Z' deux couples aléatoires.

1. $Z = Z' \quad \Longrightarrow \quad \mathbf{P}_Z = \mathbf{P}_{Z'}$.
2. La réciproque est fausse : si $\mathbf{P}_Z = \mathbf{P}_{Z'}$ alors Z n'est pas forcément égale à Z'.

Preuve

1. L'hypothèse nous dit que $X = X', Y = Y'$. Soient I, J deux intervalles réels. Alors

$$\begin{aligned} \mathbf{P}_Z(I \times J) &= \mathbf{P}\,(Z \in I \times J) \\ &= \mathbf{P}\,(X \in I, Y \in J) \\ &= \mathbf{P}\,(X' \in I, Y' \in J) = \mathbf{P}\,(Z' \in I \times J) \quad X = X', Y = Y' \\ &= \mathbf{P}_{Z'}(I \times J). \end{aligned}$$

2. **est faux.** On lance une pièce équilibrée. On note X la variable aléatoire égale à 1 si on obtient pile, égale à 0 si on obtient face. On note Y la variable aléatoire égale à 1 si on obtient face, égale à 0 si on obtient pile. Alors (X, Y) et (Y, X) ont même loi et pourtant $(X, Y) \neq (Y, X)$ en tant qu'applications.

Lorsque l'univers est discret, il n'existe que des couples aléatoires discrets.

Proposition ALEA.15.3 | Vecteur aléatoire défini sur un univers au plus dénombrable

Soit $Z = (X, Y)$ un couple aléatoire avec Ω au plus dénombrable. Alors :

Z est un couple aléatoire discret.

Preuve Puisque Ω est dénombrable, d'après le Chapter ALEA.13, X et Y sont discrètes. Donc Z est un couple aléatoire discret.

Proposition ALEA.15.4 | Structure d'espace vectoriel, opérations

1. L'ensemble des couples aléatoires réels muni de l'addition et de la multiplication externe des applications est un $\mathbf{R}$-espace vectoriel, *i.e.* : si Z_1, Z_2 sont deux couples aléatoires et $(\lambda, \mu) \in \mathbf{R}^2$, alors :

$$\lambda Z_1 + \mu Z_2 \quad \text{est un vecteur aléatoire réel.}$$

2. Si $Z_1 = (X_1, Y_1), Z_2 = (X_2, Y_2)$ sont deux couples aléatoires, alors :

$$\langle Z_1 | Z_2 \rangle = X_1 X_2 + Y_1 Y_2$$

est une variable aléatoire réelle.

Attention

Dans la proposition précédente, il s'agit d'additions dans $\mathbf{R}^2$, donc de couples.

Preuve

- $\lambda Z_1 + \mu Z_2 = (\lambda X_1 + \mu X_2, \lambda Y_1 + \mu Y_2)$ et $\lambda X_1 + \mu X_2, \lambda Y_1 + \mu Y_2$ sont des variables aléatoires réelles d'après le Chapter ALEA.12.
- Car l'ensemble des variables aléatoires réelles est stable par somme et produit.

2. COUPLES ALÉATOIRES DISCRETS

Dans cette section nous ne considèrerons que des couples aléatoires discrets du type $Z = (X, Y)$. On rappelle que cela signifie que $X(\Omega), Y(\Omega)$ sont des ensembles au plus dénombrables. Du fait de la présence de plusieurs variables aléatoires, de nouvelles notions de loi apparaissent, en plus de la notion classique : l'ancienne notion a son analogue pour les couples : la «loi conjointe».

Cadre

On se fixe donc dans toute la suite (X, Y) un couple aléatoire discret.

2.1. Système complet associé

Comme pour les variables aléatoires discrètes, précisions une notation avant de poursuivre.

Notation — Somme double infinie sur un ensemble dénombrable

Considérons E, F deux ensembles au plus dénombrable et $f : E \times F \longrightarrow \mathbf{R}$, on souhaite donner un sens à la quantité $\displaystyle\sum_{(x,y)\in E\times F} f(x,y)$.

- Si E, F sont finis, alors $E = \{x_n,\ n \in [\![0, N_1]\!]\}$, $F = \{y_n,\ n \in [\![0, N_2]\!]\}$ avec $N_1, N_2 \in \mathbf{N}$. Alors on pose :

$$\sum_{(x,y)\in E\times F} f(x,y) \underset{\textbf{(défi.)}}{=} \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} f(x_i, y_j) = \sum_{j=1}^{N_2}\sum_{i=1}^{N_1} f(x_i, y_j).$$

  C'est une somme double finie de première année, rien de plus à dire puisqu'elle ne dépend pas de la numérotation des éléments de E, F.
- Si E, F sont dénombrables, alors $E = \{x_n,\ n \in \mathbf{N}\}$, $F = \{y_n,\ n \in \mathbf{N}\}$. Alors, <u>**si**</u> les hypothèses de Fubini sont vérifiées :
  - pour tout $m \in \mathbf{N}$, $\left(\sum f(x_n, y_m)\right)_{n\geqslant 0}$ converge absolument, et
  - si $\left(\sum \sum_{n=0}^{\infty} f(x_n, y_m)\right)_{m\geqslant 0}$ converge absolument, alors

  on pose :

$$\sum_{(x,y)\in E\times F} f(x,y) \underset{\textbf{(défi.)}}{=} \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} f(x_n, y_m) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} f(x_n, y_m).$$

  En effet, on sait depuis le Chapter ANA.10, que les sommes infinies précédentes ne dépendront pas de la numérotation des éléments de E et F, et qu'en plus les hypothèses du théorème de Fubini seront vérifiées donc que l'on peut échanger l'ordre des sommes infinies.

Définition/Proposition ALEA.15.2 | Système d'évènements associé à un couple aléatoire discret

Si $Z = (X, Y)$ est un couple aléatoire discret, alors :

- $\{X = x, Y = y\}_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un système complet d'évènements.
- En particulier, il est quasi-complet :
  - $\displaystyle\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbf{P}(X = x, Y = y) = 1$,
  - $\{(X, Y) = (x, y), (X, Y) = (x', y')\} = \emptyset, \quad \forall (x, y) \neq (x', y')$.

On l'appelle le *système complet associé à* (X, Y).

Preuve Ne nous fatiguons par trop et utilisons le résultat analogue vu dans le Chapter ALEA.13.

2.2. Loi marginale, conjointe, conditionnelle

Rappelons que nous avons défini la loi de (X, Y) comme l'application qui à tout tout pavé réel $I \times J$ associe

$$\mathbf{P}_X(I \times J) = \mathbf{P}(X \in I, Y \in J).$$

Comment simplifier cette définition dans le cas de couples aléatoires discrets ? C'est la même démarche que pour les variables aléatoires discrètes.

$$\{X \in I, Y \in J\} = \{(X, Y) \in I \times J, (X, Y) \in (X, Y)(\Omega)\} \quad \text{donc,}$$

$$\mathbf{P}(X \in I, Y \in J) = \mathbf{P}\left((X, Y) \in I \times J, (X, Y) \in (X, Y)(\Omega)\right).$$

Or,

$$\{(X, Y) \in I \times J, (X, Y) \in (X, Y)(\Omega)\} = \biguplus_{\substack{(x,y)\in(X,Y)(\Omega)\\ x\in I, y\in J}} \{X = x, Y = y\},$$

$$\Longrightarrow \quad \mathbf{P}\left((X, Y) \in I \times J\right) = \sum_{\substack{(x,y)\in(X,Y)(\Omega)\\ x\in I, y\in J}} \mathbf{P}\left(X = x, Y = y\right).$$

Ainsi, pour obtenir la loi, il suffit de connaître :

- d'une part l'univers-image $(X, Y)(\Omega)$,
- et d'autre part tous les $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in (X, Y)(\Omega)$.

Ce constat nous mène tout droit à la définition suivante.

Définition ALEA.15.3 | Loi conjointe. Lois marginales.

Soit $Z = (X, Y)$ un couple aléatoire discret.

- **(Loi conjointe)** On appelle *loi conjointe* du couple aléatoire discret (X, Y) ou *fonction de masse* la fonction encore notée $\mathbf{P}_{(X,Y)}$ et définie par :

$$\mathbf{P}_{(X,Y)} : (x, y) \in (X, Y)(\Omega) \longmapsto \mathbf{P}\left(X = x, Y = y\right).$$

- **(Déterminer la loi conjointe)** *Déterminer la loi conjointe* d'un couple aléatoire discret c'est calculer $(X, Y)(\Omega)$ et $\mathbf{P}\left(X = x, Y = y\right)$ pour tout $x \in X(\Omega), y \in Y(\Omega)$.
- **(Marginales)**
  - La loi $\mathbf{P}_X : x \in X(\Omega) \longmapsto \mathbf{P}(X = x)$ de X est appelée *première loi marginale.*
  - La loi $\mathbf{P}_Y : y \in Y(\Omega) \longmapsto \mathbf{P}(Y = y)$ de Y est appelée *seconde loi marginale.*
- **(Avoir même loi que)** Soit $Z = (X', Y')$ un couple aléatoire discret définie sur $(\Omega, \mathscr{T}, \mathbf{P})$. On dit que $Z = Z'$ *ont même loi* si :

$$\forall (a, b) \in Z(\Omega) \cup Z'(\Omega)^{1}, \quad \mathbf{P}\left(Z = (a, b)\right) = \mathbf{P}\left(Z' = (a, b)\right).$$

[1] la plupart du temps, dans des contextes d'expériences aléatoires, les ensembles $Z(\Omega), Z'(\Omega)$ seront égaux

Notation

Lorsque $Z = (X, Y), Z' = (X', Y')$ ont même loi, on note $(X, Y) \sim (X', Y')$ ou $Z \sim Z'$.

Méthode Répondre à la question « déterminer la loi conjointe du vecteur aléatoire X, Y »

1. Commencer par déterminer son support $(X, Y)(\Omega)$ s'il n'est pas déjà donné.
2. Calculer les $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$. Si $X(\Omega), Y(\Omega)$ sont finis, il n'y a donc qu'un nombre fini de probabilités à déterminer, on peut les présenter sous forme d'un tableau comme ceci :

$Y = \ell$ / $X = k$	k_1	k_2	$\dots$
ℓ_1	$\mathbf{P}(Y = \ell_1, X = k_1)$	$\mathbf{P}(Y = \ell_1, X = k_2)$	$\dots$
$\vdots$			$\vdots$

Exemple 4 — ***Min / Max et lancer de dés*** On lance deux dés. On note D_1, D_2 les valeurs obtenues et on pose $X = \min(D_1, D_2)$, $Y = \max(D_1, D_2)$. Déterminer la loi conjointe de (X, Y).

Exemple 5 — On considère une urne avec quatre boules numérotées de 1 à 4 et les tirages de deux boules successifs avec remise. On note X le numéro de la boule au 1er tirage et Y le numéro de la boule au 2ème tirage. On pose enfin $M = \max(X, Y)$ et $Z = (X, M)$. Déterminer la loi conjointe de Z.

Exemple 6 — *Égalité de variables aléatoires discrètes* Soient X, Y deux variables aléatoires discrètes indépendantes de loi $\mathscr{G}(p)$ avec $p \in]0,1[$. Exprimer $\mathbf{P}(X=Y)$[2] à l'aide de la loi conjointe de (X, Y), puis calculer la somme.

[2] Au passage, les calculs faits ici prouvent que $\{X = Y\}$ est un évènement : on l'a écrit comme une réunion dénombrable d'évènements.

Généralisons les calculs faits dans le dernier exemple, dans une méthode.

Méthode Déterminer des probabilités d'évènements du type $\{X=Y\}$, …, $\{X_1 = \ldots = X_n\}$

Par exemple, si X, Y sont deux variables aléatoires discrètes, on souhaite calculer $\mathbf{P}(X=Y)$.

1. Introduire le système complet associé à X (ou Y) dans l'évènement $\{X=Y\}$:
$$\{X=Y\} = \bigcup_{k\in X(\Omega)} \{X=k, Y=k\}.$$
2. On passe ensuite aux probabilités :
$$\mathbf{P}(X=Y) = \sum_{k\in X(\Omega)} \mathbf{P}(X=k, Y=k).$$
En cas d'indépendance, on écrit alors
$$\mathbf{P}(X=Y) = \sum_{k\in X(\Omega)} \mathbf{P}(X=k)\,\mathbf{P}(Y=k).$$

De manière générale, pour $\mathbf{P}(X_1 = \ldots = X_n)$, on introduit le système complet d'évènements associé $n-1$ variables aléatoires parmi les n.

Nous avons défini la notion de loi conditionnelle d'une variable aléatoire discrète X sachant un évènement non négligeable A comme l'application
$$x \in X(\Omega) \longmapsto \mathbf{P}(X = x | A).$$

Nous pouvons en particulier regarder la loi conditionnelle de X sachant l'évènement $\{Y = y\}$ pour $y \in Y(\Omega)$ (ou l'inverse) si (X, Y) un couple aléatoire discret.

Définition ALEA.15.4 | Loi conditionnelle

1. Soit $y \in Y(\Omega)$ tel que $\mathbf{P}\left(Y = y\right) \neq 0$. Alors on appelle *loi conditionnelle de* X *sachant* $\{Y = y\}$ l'application

$$\mathbf{P}(. \mid Y = y) \left| \begin{array}{ccc} X(\Omega) & \longrightarrow & [0,1], \\ x & \longmapsto & \mathbf{P}\left(X = x \middle| Y = y\right). \end{array} \right.$$

2. Soit $x \in X(\Omega)$ tel que $\mathbf{P}(X = x) \neq 0$. Alors on appelle *loi conditionnelle de* Y *sachant* $\{X = x\}$ l'application

$$\mathbf{P}(. \mid X = x) \left| \begin{array}{ccc} Y(\Omega) & \longrightarrow & [0,1], \\ y & \longmapsto & \mathbf{P}\left(Y = y \middle| X = x\right). \end{array} \right.$$

Liens entre les notions de loi. Les formules des probabilités totales et celle de Bayes permettent de faire le lien entre ces différentes notions de lois, voyons comment. Il n'y a donc, dans ce paragraphe, rien de neuf à apprendre.

Proposition ALEA.15.5 | Loi conjointe → Loi marginale

Les lois marginales de (X, Y) sont données, en fonction de la loi conjointe, par les formules suivantes :

$$\forall x \in X(\Omega), \quad \mathbf{P}_X(x) = \mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X = x, Y = y),$$
$$\forall y \in Y(\Omega), \quad \mathbf{P}_Y(y) = \mathbf{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x, Y = y).$$

Preuve *(Point clef — **Formule des probabilités totales**)*

Attention

La réciproque est fausse. En effet, si X et Y sont telles que $X(\Omega) = Y(\Omega) = \{0,1\}$, on peut définir deux lois de couple différentes[3] par :

$$\forall (x,y) \in \{0,1\}^2, \quad \mathbf{P}(X = x, Y = y) = \frac{1}{4}.$$

On constate que $X, Y \hookrightarrow \mathscr{U}\{0,1\}$.

Et si

$$\forall (x,y) \in \{0,1\}^2, \quad \mathbf{P}(X = x, Y = y) = \begin{cases} \frac{1}{2} & \text{si } (x,y) = (0,0) \text{ ou } (1,1), \\ 0 & \text{sinon,} \end{cases}$$

on constate là encore que $X, Y \hookrightarrow \mathscr{U}\{0,1\}$.

Remarque 2.1 — Si la loi conjointe est donnée sous forme d'un tableau, la formule se comprend de la manière suivante : pour calculer une loi marginale, il suffit de sommer les probabilités sur les lignes ou les colonnes du tableau.

[3]leur existence est garantie par le Théorème ALEA.15.1 puisque $\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4} = 1$ et $\frac{1}{2}+\frac{1}{2} = 1$

Proposition ALEA.15.6 | Lois maringales + indépendance → Loi conjointe

Si X et Y sont indépendantes, alors

$$\forall (x,y) \in (X,Y)(\Omega), \quad \mathbf{P}_Z(x,y) = \mathbf{P}_X(x) \cdot \mathbf{P}_Y(y).$$

Preuve

Exemple 7 — ***Cas fini*** Soit (X,Y) un couple aléatoire discret dont la loi marginale est donnée par le tableau ci-après :

X/Y	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{4}$

Déterminer la les lois marginales de (X,Y).

Exemple 8 — ***Cas « semi- » dénombrable*** Soit $p \in]0,1[$. Soient X et Y deux variables aléatoires à valeurs dans $\mathbf{N}$ telles que :

$$\forall (x,y)\in\mathbf{N}^2,\quad \mathbf{P}(X=x\cap Y=y)=\begin{cases} e^{-p}-p+pe^{-p} & \text{si } (x,y)=(0,0),\\ p-pe^{-p} & \text{si } (x,y)=(0,1),\\ pe^{-p} & \text{si } (x,y)=(1,1),\\ \frac{p^n}{n!}e^{-p} & \text{si } (x,y)=(n,0),\ \text{et } n\geq 2,\\ 0 & \text{sinon.}\end{cases}$$

Déterminer les lois marginales de (X,Y), on montrera que $X\hookrightarrow\mathscr{P}(p)$ et que $Y\hookrightarrow\mathscr{B}(p)$.

Proposition ALEA.15.7 | Lien loi conditionnelle / marginale / conjointe

Soit (X,Y) un couple aléatoire discret. Alors pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$:

$$\mathbf{P}(X=x,Y=y)=\mathbf{P}(X=x|Y=y)\,\mathbf{P}(Y=y)=\mathbf{P}(Y=y|X=x)\,\mathbf{P}(X=x),$$

et pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$:

$$\mathbf{P}(X=x)=\sum_{y\in Y(\Omega)}\mathbf{P}(X=x|Y=y)\,\mathbf{P}(Y=y),$$
$$\mathbf{P}(Y=y)=\sum_{x\in X(\Omega)}\mathbf{P}(Y=x|X=x)\,\mathbf{P}(X=x).$$

Remarque 2.2 — Rappel d'une convention On rappelle la convention ci-après déjà précisée dans le Chapter ALEA.12 : pour $(x,y)\in X(\Omega)\times Y(\Omega)$,

$$\mathbf{P}(X=x|Y=y)\,\mathbf{P}(Y=y)=0,\quad \text{si } \mathbf{P}(Y=y)=0.$$

Preuve *(Point clef — **Formule des probabilités totales, définition d'une probabilité conditionnelle**)*

Exemple 9 — *Urnes choisies aléatoirement* On considère $n+1$ urnes numérotées $0, \ldots, n$. On suppose que l'urne $i \in [\![0\,,\,n]\!]$ contient i boules noires et $n+1-i$ boules blanches. On procède à l'expérience suivante : on choisit une urne suivant une loi $\mathscr{B}\left(n, \frac{1}{2}\right)$ puis on effectue le tirage d'une boule. On note N le numéro de l'urne choisie à la première étape, et X la variable qui vaut 1 si on a obtenu une boule noire, et 0 sinon.

- Déterminer la loi conditionnelle de X sachant $\{N = i\}$ pour tout $i \in [\![0\,,\,n]\!]$.

- Déterminer la loi de X. Quelle est la probabilité que la boule tirée soit noire ?

- Déterminer la loi conjointe du couple (N, X).

Exemple 10 — *Pile ou Face jusqu'à deux piles* Dans une succession de piles ou faces pour laquelle la probabilité d'obtenir pile est $p \in]0,1[$ et d'obtenir face est $q = 1-p$, on note X le rang d'apparition du premier pile et Y celui du second.
Déterminer la loi conjointe de (X, Y) puis vérifier que le système quasi-complet d'évènements associé à (X, Y) en est bien un. *On a*

$$\boxed{(X,Y)(\Omega) = \left\{(k,\ell) \in \mathbf{N}^\star \times \mathbf{N}^\star,\ \ell > k\right\}}.$$

Nous avons clairement $X \hookrightarrow \mathscr{G}(p)$ *et par ailleurs, la loi conditionnelle de* Y *sachant* $\{X = k\}$ *pour* $k \in \mathbf{N}^\star$ *est donnée par l'expression suivante, puisque les lancers sont indépendants :*

$$\forall \ell \in [\![k+1\,,\,\infty]\!], \quad \mathbf{P}(Y = \ell \,|\, X = k) = pq^{\ell-k-1}.$$

On obtient alors pour tout couple $(k,\ell) \in (\mathbf{N}^\star)^2$ *tel que* $\ell > k$ *:*

$$\mathbf{P}(Y = \ell, X = k) = \mathbf{P}(Y = \ell \,|\, X = k)\,\mathbf{P}(X = k) = pq^{\ell-k-1} \times pq^{k-1} = p^2 q^{\ell-2}.$$

On vérifie ensuite que la somme des ces probabilités vaut bien un :

$$\sum_{k=1}^{\infty} \sum_{\ell=k+1}^{\infty} \mathbf{P}(Y = \ell, X = k) = \sum_{k=1}^{\infty} \sum_{\ell=k+1}^{\infty} p^2 q^{\ell-2} = \left(\frac{p}{q}\right)^2 \sum_{k=1}^{\infty} \frac{q^{k+1}}{1-q} = \frac{p}{q^2} q^2 \frac{1}{1-q} = \boxed{1}.$$

Exemple 11 — *Conditionnement Poisson / binomiale* Soient $n \in \mathbf{N}^\star$, $p \in]0,1[$, X une variable aléatoire telle que $X \hookrightarrow \mathscr{P}(\lambda)$, et Y une variable aléatoire telle que la loi conditionnelle sachant $\{X = n\}$ est une loi $\mathscr{B}(n,p)$. Déterminer la loi de Y.

Propriétés des couples aléatoires discrets. Précisons, en plus de celles mentionnées dans la Proposition ALEA.15.4, quelles sont les propriétés vérifiées par les couples aléatoires discrets.

Proposition ALEA.15.8 | Structure d'espace vectoriel, opérations sur les variables aléatoires réelles discrètes

1. L'ensemble des couples aléatoires discrets, muni de l'addition et de la multiplication de réels, est un $\mathbf{R}$-espace vectoriel, *i.e.* si Z et Z' sont deux couples aléatoires discrets, $(\lambda, \mu) \in \mathbf{R}^2$, alors :

$$\lambda Z + \mu Z'$$

est un couple aléatoire discret.
2. **(Produit scalaire)** Si $Z = (X, Y), Z' = (X', Y')$ sont deux couples discrets, alors :

$$\langle (X, Y) | (X', Y') \rangle = XX' + YY'$$

est une variable aléatoire discrète.
3. **(Image d'un couple aléatoire par une application)** Si (X, Y) est un couple aléatoire discret et $f : (X, Y)(\Omega) \longrightarrow \mathbf{R}$, alors

$$f(\mathrm{X},\mathrm{Y}) \left| \begin{array}{ccc} \Omega & \longrightarrow & \mathbf{R} \\ \omega & \longmapsto & f(\mathrm{X}(\omega),\mathrm{Y}(\omega)) \end{array} \right.$$ est une variable aléatoire réelle discrète.

Son univers-image est

$$f(\mathrm{X},\mathrm{Y})(\Omega) = f((\mathrm{X},\mathrm{Y})(\Omega))$$

et sa loi est donnée par :

$$\forall z \in f(\mathrm{X},\mathrm{Y})(\Omega), \quad \mathbf{P}\left(f(\mathrm{X},\mathrm{Y}) = z\right) = \sum_{\substack{(x,y)\in \mathrm{X}(\Omega)\times \mathrm{Y}(\Omega) \\ z=f(x,y)}} \mathbf{P}\left(\mathrm{X}=x, \mathrm{Y}=y\right).$$ [H.P]

Dans la dernière formule, on utilise donc la loi conjointe de (X,Y) pour déterminer celle de $f(\mathrm{X},\mathrm{Y})$.

Preuve

1.

2. Même preuve qu'en **1**.

3.

Existence d'un couple aléatoire discret de loi conjointe fixée. On généralise sans difficulté le théorème d'existence de variable aléatoire discrète de loi fixée, vu dans le Chapter ALEA.13.

Théorème ALEA.15.1 | Existence d'un espace probabilisé associé à une suite double de somme un

Soit $\mathbf{X} = \left\{(x_i, y_j),\ i, j \in \mathbf{N}\right\}$ $\left(\textit{resp.}\ \left\{(x_i, y_j),\ i \in [\![0\,,\mathrm{N}_1]\!], j \in [\![0\,,\mathrm{N}_2]\!]\right\} \text{ avec } \mathrm{N}_1, \mathrm{N}_2 \in \mathbf{N}\right)$ une partie au plus dénombrable de $\mathbf{R}^2$, et $(p_{i,j})_{(i,j)\in\mathbf{N}^2}$ $\left(\textit{resp.}\ (p_{i,j})_{(i,j)\in[\![1\,,\mathrm{N}_1]\!]\times[\![1\,,\mathrm{N}_2]\!]}\right)$ une famille de réels telle que :

$$\sum_{i,j\in\mathbf{N}} p_{i,j} = 1 \quad \left(\textit{resp.} \sum_{\substack{0\leqslant i\leqslant \mathrm{N}_1 \\ 0\leqslant j\leqslant \mathrm{N}_2}} p_{i,j} = 1\right) \quad \text{et} \quad p_{i,j} \geqslant 0.$$

Alors il existe un espace probabilisé $(\Omega, \mathscr{T}, \mathbf{P})$ et un couple aléatoire discret (X,Y) tel que :

- $(\mathrm{X},\mathrm{Y})(\Omega) = \mathbf{X}$.
- Pour tout (i,j), $\quad \mathbf{P}(\mathrm{X} = x_i, \mathrm{Y} = y_j) = p_{i,j}$.

La preuve est identique au cas des variables aléatoires discrètes.

Preuve

- On cherche un exemple, rien qu'un, d'espace probabilisé $(\Omega, \mathscr{T}, \mathbf{P})$ et de couple aléatoire qui répondent au problème posé. Posons alors :

$$(\Omega, \mathscr{T}) = (\mathbf{X}, \mathscr{P}(\mathbf{X})),$$
$$(\mathrm{X},\mathrm{Y}) = \mathrm{Id}_{\Omega\times\Omega} = \mathrm{Id}_{\mathbf{X}\times\mathbf{X}},$$

puis enfin pour $\mathbf{P}$ l'unique probabilité telle que : $\mathbf{P}\left(\left\{(x_i, y_j)\right\}\right) = p_{i,j}$ pour tout $i, j \in \mathbf{N}$ (ou $[\![1\,,\mathrm{N}_1]\!] \times [\![1\,,\mathrm{N}_2]\!]$) — *i.e.* celle définie dans la Section 2.4 du Chapter ALEA.12. On vérifie sans difficulté que ce choix convient.

Exemple 12 — Soit $a \in \mathbf{R}$. Déterminer a pour qu'il existe un couple aléatoire discret (X,Y) tel que $\quad \mathbf{P}\left(\mathrm{X} = i, \mathrm{Y} = j\right) = \frac{a}{2^{i+j}}$ pour tout $(i,j) \in (\mathbf{N}^\star)^2$.

- *Soit $i \in \mathbf{N}^\star$, alors étudions la convergence de* $\left(\sum \frac{1}{2^{i+j}}\right)_{j\geqslant 1}$. *Puisque* $|1/2| < 1$, *nous avons*

$$\sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = \frac{1}{2^i}\sum_{j=1}^{\infty} \frac{1}{2^j} = \frac{1}{2^i}\frac{1}{2}\frac{1}{1-1/2} = \frac{1}{2^i}.$$

- *De-même* $\sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2}\frac{1}{1-\frac{1}{2}} = 1.$
  Les hypothèses du théorème de Fubini *sont donc vérifiées, et nous avons*

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = 1.$$

Donc l'unique valeur de a telle que

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \boldsymbol{P}(X = i, Y = j) = 1$$

est $\boxed{a = 1}$. *De plus, pour tout* $(i, j) \in (\boldsymbol{N}^\star)^2$,

$$\frac{a}{2^{i+j}} \geqslant 0$$

d'où l'existence d'un couple aléatoire (X, Y) *de loi conjointe donnée précédemment.*

2.3. Espérance, Covariance

2.3.1. Espérance & Transfert

Définition ALEA.15.5 | Espérance d'un couple aléatoire discret

Soit $Z = (X, Y)$ un couple aléatoire discret.

- **(Admettre une espérance)** On dit que $Z = (X, Y)$ *admet une espérance* si X et Y admettent une espérance.
- Si X, Y admettent une espérance, alors on appelle *espérance de* (X, Y) le vecteur

$$\mathbf{E}(Z) = (\mathbf{E}(X), \mathbf{E}(Y)) \in \mathbf{R}^2.$$

L'espérance d'un vecteur aléatoire est donc aussi un vecteur : c'est le couple formé par les espérances.

Formule de transfert pour les couples aléatoires discrets. L'objectif est ici d'obtenir une formule pour calculer des espérances de fonctions de variables aléatoires $f(X, Y)$ sans avoir à trouver la loi. Nous allons voir également que la formule de transfert pour les couples aléatoires va nous permettre de démontrer certaines propriétés sur les variables aléatoires réelles discrètes admises jusque là. Notamment, en l'appliquant à $f : (x, y) \longmapsto \lambda x + \mu x$ pour $(\lambda, \mu) \in \mathbf{R}^2$, nous pourrons montrer la linéarité de l'espérance.

Théorème ALEA.15.2 | Transfert pour les couples aléatoires discrets

Soit (X, Y) un couple aléatoire discret et $g : (X, Y)(\Omega) \to \mathbf{R}$. Alors :

$$g(X, Y) \text{ possède une espérance} \iff \left(\sum g(x, y)\mathbf{P}(X = x, Y = y)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)} \text{ converge \textbf{absolument}}.$$

Dans ce cas, nous avons :

$$\begin{aligned}\mathbf{E}(g(X, Y)) &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} g(x, y) \times \mathbf{P}(X = x, Y = y)\\ &= \sum_{y\in Y(\Omega)}\sum_{x\in X(\Omega)} g(x, y) \times \mathbf{P}(X = x, Y = y).\end{aligned}$$

Preuve La permutation des sommes est justifiée par la convergence absolue et le Théorème ANA.10.5. Nous admettons le reste.

On peut donc enfin montrer la linéarité de l'espérance, dans le cas discret.

Corollaire ALEA.15.1 | Linéarité de l'espérance

Soient X, Y deux variables aléatoires discrètes admettant une espérance, et $\lambda, \mu \in \mathbf{R}$. Alors $\lambda X + \mu Y$ admet une espérance et :

$$\mathbf{E}\left(\lambda X + \mu Y\right) = \lambda \mathbf{E}\left(X\right) + \mu \mathbf{E}\left(Y\right).$$

Preuve

Corollaire ALEA.15.2 | Espérance d'un produit

- Si $X_1, \ldots, X_n$ est une collection de n variables aléatoires discrètes **indépendantes** admettant une espérance. Alors :

$$\mathbf{E}\left(X_1 \times \ldots \times X_n\right) = \mathbf{E}\left(X_1\right) \times \ldots \times \mathbf{E}\left(X_n\right).$$

- Soient X, Y deux variables aléatoires indépendantes discrètes ayant une espérance. Alors :

$$\mathbf{E}\left(X \cdot Y\right) = \mathbf{E}\left(X\right) \cdot \mathbf{E}\left(Y\right).$$

Attention **La réciproque est fausse**

Soit $X \hookrightarrow \mathscr{U}(\{-1, 0, 1\})$ et $Y = X^2$. Alors $\mathbf{E}\left(X\right) = (-1) \times 31 + 0 \times 13 + 1 \times 13 = 0$ et de-même, avec le théorème de transfert, $\mathbf{E}(XY) = \mathbf{E}(X^3) = 0$. Ainsi $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y) = 0$ et pourtant $Y = X^2$ n'est pas indépendante de X.

Preuve Nous démontrons le résultat pour $n = 2$: il s'étend alors automatiquement pour n quelconque par récurrence, notamment car $X_1 \times \ldots \times X_{n-1} \perp\!\!\!\perp X_n$.

Exemple 13 — *Cas fini* On considère une urne contenant n boules numérotées de 1 à n. On tire successivement deux boules avec remise. On note X le numéro de la 1ère boule, Y le numéro de la 2ème. Soit $Z = \max(X, Y)$. Quelle est l'espérance de Z ?

$$\begin{aligned}
\mathbf{E}(Z) &= \sum_{1 \leqslant i,j \leqslant n} \max(i,j) \mathbf{P}\left(\left(X = x_i\right) \cap \left(Y = y_j\right)\right) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \max(i,j) \\
&= \frac{1}{n^2} \sum_{i=1}^{n-1} \left[\sum_{j=1}^{i} \max(i,j) + \sum_{j=i+1}^{n} \max(i,j)\right] + \frac{1}{n^2} n^2 \\
&= \frac{1}{n^2} \left(\sum_{i=1}^{n-1} \left[\sum_{j=1}^{i} i + \sum_{j=i+1}^{n} j\right] + n^2\right) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \left[i^2 + \frac{n(n+1)}{2} - \frac{i(i+1)}{2}\right] \\
&= \frac{1}{2n^2} \sum_{i=1}^{n} \left[n(n+1) + i^2 - i\right] \\
&= \frac{1}{2n^2} \left(n^2(n+1) + \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}\right) \\
&= \frac{(n+1)(4n-1)}{6n}.
\end{aligned}$$

on sort de le terme $i = n$ + CHASLES

n^2 est le terme $i = n$ de la somme

Exemple 14 — *Cas «semi-»dénombrable* Soient X, Y deux variables aléatoires indépendantes telles que $X \hookrightarrow \mathscr{G}(p)$, avec $p \in]0,1[$ et $Y \hookrightarrow \mathscr{B}(p)$. Déterminer si $Z = \min(X, Y+1)$ admet une espérance, la calculer le cas échéant. *Il s'agit de calculer $\boldsymbol{E}(g(X,Y))$ si $g : (x,y) \in \boldsymbol{R}^2 \longmapsto \min(x, y+1)$. Étudions donc la convergence absolue de*

$$\left(\sum \min(k,\ell+1)\boldsymbol{P}(X=k, Y=\ell)\right)_{k\geqslant 1, \ell\in\{0,1\}}.$$

Soit $\ell \in \{0,1\}$. Il s'agit donc d'étudier la convergence absolue de la série en k, pour ces deux valeurs de ℓ.

- *Supposons que $\ell = 0$. Soit $k \geqslant 1$. Alors par indépendance*

  $$\min(k,\ell+1)\boldsymbol{P}(X=k, Y=\ell) = q\min(k,1)\boldsymbol{P}(X=k) = q\boldsymbol{P}(X=k)$$

  car $k \geqslant 1$. Comme $\sum_{k=1}^{\infty} \boldsymbol{P}(X=k) = 1$, nous déduisons la convergence souhaitée.
- *Supposons que $\ell = 1$. Alors par indépendance*

  $$\min(k,\ell+1)\boldsymbol{P}(X=k, Y=\ell) = p\min(k,2)\left(pq^{k-1}\right) = p^2\min(k,2)q^{k-1}.$$

  Comme $|q| < 1$, on a alors

  $$\sum_{k=1}^{\infty} p^2\min(k,2)q^{k-1} = p^2 + p^2\sum_{k=2}^{\infty} kq^{k-1} = p^2 + p^2\left(\frac{1}{(1-q)^2} - 1\right) = 1$$

  car $1-q = p$. D'où la convergence de la série double, et

  $$\boldsymbol{E}(Z) = \boldsymbol{E}(\min(X, Y+1)) = \boxed{q+1}.$$

2.3.2. Covariance

La covariance de deux variables est un objet qui mesure intuitivement la dépendance d'une des deux variables aléatoires par rapport à l'autre. Un cas limite étant celui de variables aléatoires indépendantes, qui correspondra à une covariance nulle. Voyons la définition.

Définition/Proposition ALEA.15.3 | Covariance

Soit (X,Y) un couple aléatoire discret, tel que X, Y admettent un un moment d'ordre deux. On appelle *covariance de* X *et* Y le réel noté $\mathbf{Cov}(X,Y)$ défini par :

$$\mathbf{Cov}(X,Y) = \mathbf{E}\left([X - \mathbf{E}(X)][Y - \mathbf{E}(Y)]\right).$$

Les variables aléatoires X et Y sont dites *non corrélées* si :

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

Preuve *(de l'existence de la covariance)* Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. En appliquant le théorème de transfert à $g(X, Y)$ avec $g : (x, y) \in \mathbf{R}^2 \longmapsto (x - \mathbf{E}(X))(y - \mathbf{E}(Y))$, on voit qu'il suffit de vérifier la convergence de la série double

$$\left(\sum (x - \mathbf{E}(X))(y - \mathbf{E}(Y))\mathbf{P}(X = x, Y = y)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)},$$

Or,

$$\begin{aligned}&\left|(x - \mathbf{E}(X))(y - \mathbf{E}(Y))\right|\mathbf{P}(X = x, Y = y)\\ &\leqslant \frac{1}{2}(x - \mathbf{E}(X))^2\mathbf{P}(X = x, Y = y) + \frac{1}{2}(y - \mathbf{E}(Y))^2\mathbf{P}(X = x, Y = y).^{a}\end{aligned}$$

- Pour le premier terme, constatons que d'après la formule des probabilités totales, nous avons

$$\sum_{y\in Y(\Omega)} (x - \mathbf{E}(X))^2\mathbf{P}(X = x, Y = y) = (x - \mathbf{E}(X))^2\mathbf{P}(X = x),$$

puis comme X admet une variance,

$$\sum_{x\in X(\Omega)} (x - \mathbf{E}(X))^2\mathbf{P}(X = x) = \mathbf{Var}(X).$$

D'où la convergence de

$$\left(\sum \frac{1}{2}(x - \mathbf{E}(X))^2\mathbf{P}(X = x, Y = y)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)}.$$

- La série

$$\left(\sum \frac{1}{2}(y - \mathbf{E}(Y))^2\mathbf{P}(X = x, Y = y)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)}.$$

converge aussi par symétrie des rôles entre x, y.

Donc d'après le théorème de comparaison pour les séries, la série double initiale converge.

On retiendra aussi de la démonstration l'écriture en série double de la covariance, qui est résumée par la proposition ci-après.

Proposition ALEA.15.9 | Covariance en somme double

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un moment d'ordre deux, *i.e.*

$$\left(\sum x^2\mathbf{P}(X = x)\right)_{x\in X(\Omega)}, \quad \left(\sum y^2\mathbf{P}(Y = y)\right)_{y\in Y(\Omega)} \quad \text{convergent.}$$

Alors :
$$\begin{aligned}\mathbf{Cov}(X, Y) &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} (x - \mathbf{E}(X))(y - \mathbf{E}(Y)) \times \mathbf{P}(X = x, Y = y)\\ &= \sum_{y\in Y(\Omega)}\sum_{x\in X(\Omega)} (x - \mathbf{E}(X))(y - \mathbf{E}(Y)) \times \mathbf{P}(X = x, Y = y).\end{aligned}$$

Proposition ALEA.15.10 | Propriétés de la covariance

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un un moment d'ordre deux.

1. **(Lien covariance/variance)** $\mathbf{Cov}(X, X) = \mathbf{Var}(X)$.
2. **(Formule de König-Huygens)** $\mathbf{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$.
3. **(Symétrie)** $\mathbf{Cov}(X, Y) = \mathbf{Cov}(Y, X)$.
4. **(Constante)** $\mathbf{Cov}(X, c) = \mathbf{Cov}(c, X) = 0$ pour toute constante $c \in \mathbf{R}$.
5. **(Covariance et somme)** Les applications $\mathbf{Cov}(., Y) : Z \longmapsto \mathbf{Cov}(Z, Y)$ et $\mathbf{Cov}(X, .) : Z \longmapsto \mathbf{Cov}(X, Z)$ sont des applications linéaires sur l'espace vectoriel des variables aléatoires admettant un moment d'ordre deux, *i.e.* si $(\lambda, \mu) \in \mathbf{R}^2$ et X, Y, Z sont trois variables aléatoires admettant un moment d'ordre deux, alors

$$\begin{aligned}\mathbf{Cov}(\lambda X + \mu Y, Z) &= \lambda\mathbf{Cov}(X, Z) + \mu\mathbf{Cov}(Y, Z),\\ \mathbf{Cov}(Z, \lambda X + \mu Y) &= \lambda\mathbf{Cov}(Z, X) + \mu\mathbf{Cov}(Z, Y).\end{aligned}$$

6. **(Variance d'une somme (cas $n = 2$))**

$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) + 2\mathbf{Cov}(X, Y).$$

En particulier, si X et Y sont non corrélées, nous avons :

$$\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y).$$

[a] On rappelle $|ab| \leqslant \frac{a^2}{2} + \frac{b^2}{2}$ pour tout a, b réels.

Preuve

1.

2.

3.

4.

5.

6.

La propriété de bilinéarité doit être comprise comme une règle de développement classique sur les nombres réels du collège. Ainsi, si je décide de noter « · » l'opérateur de covariance, l'assertion[4]

$$\mathbf{Cov}\left(Z, \lambda X + \mu Y\right) = \lambda \mathbf{Cov}\,(Z, X) + \mu \mathbf{Cov}\,(Z, Y) \quad \text{devient simplement}$$
$$Z \cdot (\lambda X + \mu Y) = \lambda Z \cdot X + \mu Z \cdot Y.$$

Exemple 15 — Soient X, Y deux variables aléatoires admettant un moment d'ordre deux. Développer $\mathbf{Cov}\,(X - Y, X + Y)$ et $\mathbf{Cov}\,(2X + 3Y, X - 4Y)$.

[4] Très simple à retenir donc.

L'expression type König-Huygens de la covariance nous permet d'établir facilement le corolaire qui suit.

Corollaire ALEA.15.3 | Indépendantes $\Longrightarrow$ non-corrélées

Soient X, Y deux variables aléatoires indépendantes, alors X, Y sont non-corrélées.

Preuve D'après le Corollaire ALEA.15.2, $\mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = \mathbf{Cov}(X,Y) = 0$, c'est justement dire que les variables aléatoires X, Y sont non-corrélées.

Attention

La non corrélation n'implique pas l'indépendance.] Soit $X \hookrightarrow \mathscr{U}(\{-1,0,1\})$ et $Y = X^2$. Alors on a déjà montré que $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y) = 0$ donc $\mathbf{Cov}(X,Y) = 0$, et pourtant $Y = X^2$ n'est pas indépendante de X.

Méthode Développement d'une variance de somme

Soit la quantité $\mathbf{Var}(X+Y)$ avec X, Y deux variables aléatoires discrètes ayant une variance.

1. Écrire la quantité en fonction de la covariance : $\mathbf{Var}(X+Y) = \mathbf{Cov}(X+Y, X+Y)$.
2. Développer en utilisant la bilinéarité de la covariance.

Exemple 16 — Développer $\mathbf{Var}(X-2Y+3Z)$ lorsque X, Y, Z sont trois variables aléatoires discrètes ayant une variance.

Proposition ALEA.15.11 | Variance d'une somme, cas général

Soient $X, \dots, X_n$ une famille de n variables aléatoires possédant un moment d'ordre deux. Alors :

$$\mathbf{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbf{Var}(X_i) + 2 \sum_{\substack{i<j \\ (i,j)\in[\![1,n]\!]^2}} \mathbf{Cov}\left(X_i, X_j\right)$$
$$= \sum_{i=1}^{n} \mathbf{Var}(X_i) + \sum_{\substack{i\neq j \\ (i,j)\in[\![1,n]\!]^2}} \mathbf{Cov}\left(X_i, X_j\right). \quad \text{[H.P]}$$

En particulier, si elles sont **deux à deux non corrélées** : $\mathbf{Cov}\left(X_i, X_j\right) = 0$ pour tout $(i,j) \in \{1, \dots, n\}^2$, $i \neq j$, ou mieux deux à deux indépendantes, alors :

$$\mathbf{Var}(X + \dots + X_n) = \mathbf{Var}(X) + \dots + \mathbf{Var}(X_n).$$

Cette formule n'est pas au programme, elle est cependant très classique. Il faut donc en connaître la démonstration mais pas l'énoncé par coeur.

Preuve

Coefficient de corrélation. Un nouvel objet peut être défini à l'aide de la covariance : il s'agit du coefficient de corrélation, qui est un outil de détection de relations affines entre variables aléatoires.

Définition/Proposition ALEA.15.4 | Inégalité de Cauchy-Schwarz/ Coefficient de corrélation.

Soit (X, Y) un couple aléatoire discret, tel que X, Y admettent un un moment d'ordre deux. Alors :

1. **(Inégalité de Cauchy-Schwarz)**

$$|\mathbf{Cov}(X, Y)| \leqslant \sigma_X \cdot \sigma_Y.$$

Si X, Y sont d'écart-type non nul. On appelle *coefficient de corrélation entre* X *et* Y la quantité

$$\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\sigma_X . \sigma_Y} \in [-1, 1].$$

2. **(Cas d'égalité)** $\rho(X, Y) = \pm 1 \iff \exists a, b \in \mathbf{R}, \quad \mathbf{P}(Y = aX + b) = 1.$

Remarque 2.3 — Interprétation du coefficient de corrélation Ainsi, pour savoir si une variable aléatoire dépend de manière affine de l'autre (avec probabilité 1), il suffit de calculer leur coefficient de corrélation.

- s'il est proche de 1 ou -1, on envisage une relation affine entre les deux variables aléatoires.
- En revanche, plus il est proche de zéro, plus l'existence d'une telle relation devient hautement improbable.

Preuve

1. *Soit* $P : t \in \boldsymbol{R} \longrightarrow \mathbf{V}(tX + Y) = \boldsymbol{Cov}(tX + Y, tX + Y)$, *cette fonction est bien définie puisque* X, Y *admettent une variance donc* $tX + Y$ *aussi pour tout* t*. Alors, par bilinéarité de la covariance :*

$$\forall t \in \mathbf{R}, \quad P(t) = \mathbf{V}(X) t^2 + 2t \boldsymbol{Cov}(X, Y) + \mathbf{V}(Y).$$

Comme $\boldsymbol{Var}(X) = \sigma_X^2 \neq 0$ *par hypothèse,* P *est une fonction polynomiale de degré 2, positive (par définition et car la variance est positive), ainsi son discriminant est négatif, soit*

$$4\boldsymbol{Cov}(X, Y)^2 - 4\mathbf{V}(X)\mathbf{V}(Y) \leq 0 \iff \boldsymbol{Cov}(X, Y)^2 \leq \mathbf{V}(X)\mathbf{V}(Y).$$

Il reste à composer par la fonction $\sqrt{\cdot}$*, qui est bien croissante, de chaque côté et de l'inégalité. Puisque* $\sigma_X \sigma_Y > 0$ *par hypothèse, il reste ensuite à diviser de chaque côté par* $\sigma_X \sigma_Y$*. On obtient bien*

$$|\rho(X, Y)| \leq 1.$$

2.

3. SOMMES DE VARIABLES ALÉATOIRES INDÉPENDANTES & CONVOLUTION

Nous allons voir dans cette partie comment analyser la loi de sommes de variables aléatoires indépendantes, discrètes ou à densité. Dans le cas discret, nous verrons que la formule fait intervenir la loi conjointe du couple. Dans le cas à densité, nous aurons une formule. On applique ensuite ces résultats aux lois usuelles étudiées.

3.1. Cas discret

Proposition ALEA.15.12 | Produit de convolution discret

- Soit (X, Y) un couple aléatoire discret. Alors $X+Y$ est une variable aléatoire réelle discrète et pour tout $k \in (X+Y)(\Omega)$, on a :

$$\mathbf{P}\left(X+Y=k\right) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X=\ell, Y=k-\ell), \text{ ou de manière équivalente :}$$

$$\mathbf{P}_{X+Y}(k) = \sum_{\ell \in X(\Omega)} \mathbf{P}_{(X,Y)}\left(\ell, k-\ell\right).$$

- En particulier, si X et Y sont **indépendantes**, alors : pour tout $k \in (X+Y)(\Omega)$, on a :

$$\mathbf{P}\left(X+Y=k\right) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X=\ell) \cdot \mathbf{P}(Y=k-\ell), \text{ ou de manière équivalente :}$$

$$\mathbf{P}_{X+Y}(k) = \sum_{\ell \in X(\Omega)} \mathbf{P}_X\left(\ell\right) \cdot \mathbf{P}_Y\left(k-\ell\right).$$

Là encore, cette formule n'est pas à retenir par coeur : mémoriser surtout qu'il s'agit simplement d'une formule des probabilités totales où l'on introduit le système complet d'évènements associé à l'une ou l'autre des variables aléatoires. On retiendra que :

- la loi de la somme s'exprime donc en fonction de la loi conjointe du couple.
- En cas d'indépendance, la loi de la somme s'exprime donc en fonction des lois marginales.

Méthode Déterminer la loi de la somme de deux variables aléatoires discrètes indépendantes

1. Introduire le système complet associé à X où Y dans l'évènement $\{X+Y=k\}$ pour tout $k \in (X+Y)(\Omega)$.
2. Utiliser la formule des probabilités totales pour avoir

$$\mathbf{P}\left(X+Y=k\right) = \sum_{\ell \in X(\Omega)} \mathbf{P}(X=\ell) \cdot \mathbf{P}(Y=k-\ell).$$

3. Pour poursuivre le calcul, regarder pour quels $\ell \in X(\Omega)$, on a :

$$\mathbf{P}(X=\ell) \cdot \mathbf{P}(Y=k-\ell) \neq 0.$$

Puis conclure en finissant le calcul de la somme.

Preuve

Exemple 17 — *Cas fini* Soient X, Y deux variables aléatoires indépendantes suivant une loi $\mathscr{B}(p)$, avec $p \in]0,1[$. Déterminer la loi de $X+Y$, que remarquer ?

Exemple 18 — *Cas dénombrable* Soient X, Y deux variables aléatoires indépendantes suivant une loi $\mathscr{G}(p)$, avec $p \in]0,1[$.

1. Déterminer la loi de $X+Y$. Étudier l'existence d'une espérance, la calculer le cas échéant.

2. Pour $n \geqslant 2$ fixé, déterminer la loi conditionnelle de X sachant $\{X+Y=n\}$.

On peut enfin démontrer certains résultats du Chapter ALEA.13 qui ont été admis.

Théorème ALEA.15.3 | Stabilité des lois binomiale et de Poisson

Soient $X, \dots, X_d$ des variables aléatoires réelles **indépendantes**, $d \in \mathbf{N}^\star$. Alors :

1. si $X, \dots, X_d$ sont d variables aléatoires réelles mutuellement indépendantes de lois $\mathscr{B}(n_1, p), \dots, \mathscr{B}(n_d, p)$, alors
$$X + \dots + X_d \hookrightarrow \mathscr{B}(n_1 + \dots + n_d, p).$$
2. Si $X, \dots, X_d$ sont d variables aléatoires réelles mutuellement indépendantes de loi $\mathscr{P}(\lambda_1), \dots, \mathscr{P}(\lambda_d)$, avec $\lambda_1, \dots, \lambda_d \in \mathbf{R}^+$ alors :
$$X + \dots + X_d \hookrightarrow \mathscr{P}(\lambda_1 + \dots + \lambda_d).$$

La propriété pour la binomiale se comprend aisément puisque si l'on réalise n_1 puis n_2 *etc.* puis n_d expériences de Bernoulli identiques indépendantes et qu'on somme les nombres de succès obtenus, cela revient à effectuer $n_1 + \dots + n_d$ cette même expérience et compter les succès.

Attention

Pour la loi binomiale, on ne peut en général pas sommer par rapport au paramètre p, et même en cas d'indépendance.

Preuve Faisons les preuves pour $d = 2$, le résultat général s'en déduit alors par récurrence évident sur d.

1. Rappelons la formule de Vandermonde vue dans le TD de dénombrement : pour tout $k \in \mathbf{Z}$ tel que $k \leqslant n_1, n_2$, nous avons
$$\binom{n_1+n_2}{k} = \sum_{\ell=0}^{k} \binom{n_1}{\ell}\binom{n_2}{k-\ell}.$$

2.

Exemple 19 — Soient X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ et μ.

1. Donner la loi de $X+Y$. Quelle est son espérance ?

2. Pour $n \geqslant 0$ fixé, déterminer la loi conditionnelle de X sachant $\{X+Y=n\}$.

3.2. Cas à densité

Le théorème suivant est fondamental et sera admis. L'idée est la suivante : on connaît une formule pour la densité de la somme de deux variables aléatoires indépendantes à densité.

Théorème ALEA.15.4 | Produit de convolution à densité

- Soient X et Y deux variables aléatoires **indépendantes**[5] de densités f_X et f_Y. Alors l'application notée $f_X \star g_X$ définie par :

$$\forall z \in \mathbf{R}, \quad (f_X \star f_Y)\left(\boxed{z}\right) = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} f_X(x) f_Y\left(\boxed{z} - x\right) \mathrm{d}x & \text{en cas de convergence,} \\ 0 & \text{sinon.} \end{cases}$$

est une est une densité de $X+Y$. Autrement dit,

$$f_{X+Y} = f_X \star f_Y.$$

- Plus généralement, si $X_1, \ldots, X_n$ sont à densité **indépendantes**, alors $X_1 + \ldots + X_n$ est également à densité.

Notez l'analogie avec la convolution discrète : le symbole somme est remplacé par une intégrale, les lois sont remplacées par les densités.

Preuve Admis.

Attention

L'hypothèse d'indépendance est cruciale. En effet, soit X est une variable aléatoire à densité. Alors X et $-X$ ne sont pas indépendantes, et

$$X + (-X) = 0 \quad \text{n'est même pas une variable aléatoire à densité (car discrète).}$$

[5] Il existe aussi une formule dans le cas non-indépendant, elle fait intervenir la notion de « densité conjointe » du couple, mais qui n'est pas au programme.

Remarque 3.1 —

- Comme cela est précisé dans le programme officiel, la condition « $f_X \star f_Y$ est définie et continue sur $\mathbf{R}$ sauf éventuellement en un nombre fini de points » est souvent admise, il convient alors uniquement de justifier que $X \perp\!\!\!\perp Y$ — l'indépendance est quant à elle attendue dans votre rédaction !
- Il est facile de voir que, dès que l'une des deux densités au moins est bornée, l'intégrale $f_X \star f_Y$ converge. Et c'est bien le cas de toutes les lois usuelles vues dans ce chapitre.
- On a, par changement de variable :

$$\int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,\mathrm{d}y.$$

Méthode Déterminer une densité de la somme de deux variables aléatoires indépendantes à densité

1. Rappeler l'hypothèse d'indépendance puis écrire l'expression de la convolée des densités :

$$\forall z \in \mathbf{R}, \quad (f_X \star f_Y)(z) = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x & \text{en cas de convergence,} \\ 0 & \text{sinon.} \end{cases}$$

2. Pour poursuivre le calcul, regarder pour quels $x \in \mathbf{R}$, on a :

$$f_X(x) f_Y(z-x) \neq 0.$$

Puis conclure en réduisant l'intervalle d'intégration.

Exemple 20 — Calculer une densité de $X+Y$ lorsque X et Y sont deux variables aléatoires indépendantes qui suivent les lois indiquées.

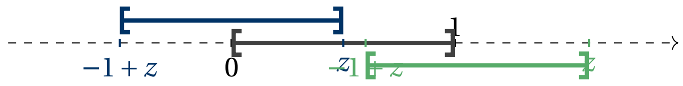
1. $X, Y \hookrightarrow \mathscr{U}[0,1]$. *Puisque X, Y sont indépendantes, alors une densité de f_{X+Y} est $f_X \star f_Y$. Ainsi, pour tout $z \in \mathbf{R}$,*

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}z,$$

or : $\quad f_X(x) f_Y(z-x) \neq 0 \quad \Longleftrightarrow \quad 0 \leqslant x \leqslant 1,\ 0 \leqslant z-x \leqslant 1$, *i.e. :*

$$0 \leqslant x \leqslant 1,\ -1+z \leqslant x \leqslant z.$$

Dessinons ces deux intervalles en x.



Plusieurs cas de figure se présentent alors.

- *si $z \leqslant 0$ ou $-1+z \geqslant 1$ alors on intègre la fonction nulle et $f_{X+Y}(z) = 0$.*
- *Si $0 < z \leqslant 1$, alors $f_{X+Y}(z) = \displaystyle\int_0^z 1\,\mathrm{d}t = z$.*
- *Si $0 < -1+z \leqslant 1$, alors $f_{X+Y}(z) = \displaystyle\int_{-1+z}^{1} 1\,\mathrm{d}t = 2-z$.*

On déduit alors une densité de $X+Y$:

$$\boxed{f_{X+Y}(z) = \begin{cases} 0 & \text{si } z \leqslant 0, \text{ ou } z \geqslant 2, \\ z & \text{si } 0 < z \leqslant 1, \\ 2-z & \text{si } 1 < z < 2. \end{cases}}$$

2. $X \hookrightarrow \mathscr{U}[-1,1], Y \hookrightarrow \mathscr{U}[0,1]$.

3. $X \hookrightarrow \mathscr{E}(1), Y \hookrightarrow \mathscr{E}(2)$.

On peut enfin démontrer le résultat de stabilité de la loir normale du Chapter ALEA.14 qui avait été admis et dont on rappelle l'énoncé.

Proposition ALEA.15.13 | Somme de lois normales indépendantes

Soient $X_1, \dots, X_n$ une collection de n variables aléatoires indépendantes telles que :

$$\forall i \in [\![1, n]\!], \quad X_i \hookrightarrow \mathcal{N}(\mu_i, \sigma_i^2), \quad \mu_i \in \mathbf{R}, \quad \sigma_i > 0.$$

- Alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right) = \mathcal{N}\left(\mathbf{E}\left(\sum_{i=1}^n X_i\right), \mathbf{Var}\left(\sum_{i=1}^n X_i\right)\right).$$

- Et plus généralement : pour toute famille $(\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n \setminus \{(0, \dots, 0)\}$, on a :

$$\sum_{i=1}^n \lambda_i X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n \lambda_i \cdot \mu_i, \sum_{i=1}^n \lambda_i^2 \cdot \sigma_i^2\right)$$
$$= \mathcal{N}\left(\mathbf{E}\left(\sum_{i=1}^n \lambda_i \cdot X_i\right), \mathbf{Var}\left(\sum_{i=1}^n \lambda_i \cdot X_i\right)\right).$$

Preuve Nous allons nous contenter de montrer le cas $n = 2, \mu_1 = \mu_2, \sigma_1^2 = \sigma_2^2 = 1$ et $\lambda_1 = \lambda_2 = 1$, *i.e.* montrons que $X_1 + X_2 \hookrightarrow \mathcal{N}(0, 2)$.

- Commençons par justifier que $-\frac{1}{2}t^2 - \frac{1}{2}(x-t)^2 = -\frac{1}{4}x^2 - \frac{1}{2}\left(\frac{t-x/2}{1/\sqrt{2}}\right)^2$ pour tout $t, x \in \mathbf{R}$.

- On peut alors déduire que $X_1 + X_2 \hookrightarrow \mathcal{N}(0, 2)$.

Le cas général se démontre en utilisant le produit de convolution mais les calculs sont plus compliqués.

Corollaire ALEA.15.4 | Application rigolote

Soient X et Y deux variables aléatoires **indépendantes** à densité. Alors :

$$\mathbf{P}(X = Y) = 0.$$

Preuve

⋆ ⋆ ⋆ **Fin du chapitre** ⋆ ⋆ ⋆

4. EXERCICES

4.1. Calculs déterministes

Exercice ALEA.15.1 | **Sommes doubles finies** Soit $n \in \mathbf{N}^\star$. Calculer les sommes suivantes :

- $\sum_{i=0}^{n} \sum_{j=1}^{n} (2i+j)$,
- $\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2^{i+j}$,
- $\sum_{i=1}^{n} \sum_{j=i}^{n} \frac{i}{j}$,
- $\sum_{1 \leqslant i,j \leqslant n} |i-j|$,
- $\sum_{1 \leqslant i<j \leqslant n} (i+j)$.

Exercice ALEA.15.2 | **Séries doubles** Étudier la convergence des séries doubles ci-après. Calculer leur somme le cas échéant.

$$\left(\sum \frac{ij^2}{3^{i+j}}\right)_{i,j \geqslant 0}, \quad \left(\sum \frac{i^2-1}{i!j!}\right)_{i,j \geqslant 0}.$$

Exercice ALEA.15.3 | **Intégration de fonctions constantes par morceaux** Calculer les intégrales suivantes. *On admettra leur existence*

1. Pour tout $x \in \mathbf{R}$, $\int_{-\infty}^{2x} f$ et $\int_{-3x}^{3x} f$ où : $f : t \longmapsto \begin{cases} \frac{1}{4} & \text{si } t \in [-1,3] \\ 0 & \text{sinon.} \end{cases}$
2. Pour tout $x \in \mathbf{R}$, $\int_{-\infty}^{\infty} g(t)g(x-t)\,\mathrm{d}t$ où : $g : t \longmapsto \begin{cases} \mathrm{e}^{-t} & \text{si } t \geqslant 0, \\ 0 & \text{sinon.} \end{cases}$

4.2. Généralités

Exercice ALEA.15.4 | **Coefficient de corrélation de combinaisons linéaires** Soient X et Y deux variables aléatoires réelles admettant une variance non nulle, et $a, b, c, d \in \mathbf{R}$ tels que $a \neq 0, c \neq 0$. Montrer que :

$$\rho(a\mathrm{X}+b, c\mathrm{Y}+d) = \operatorname{sgn}(ac)\rho(\mathrm{X},\mathrm{Y}), \quad \text{où} \quad \operatorname{sgn}(x) = \begin{cases} 1 & \text{si } x \geqslant 0, \\ -1 & \text{si } x < 0. \end{cases}$$

Solution (exercice ALEA.15.4)

Tout d'abord, précisons que l'on suppose $a \neq 0$, $c \neq 0$ pour que $\sigma_{a\mathrm{X}+b} \neq 0$ et $\sigma_{c\mathrm{X}+d} \neq 0$: en effet,

$$\sigma^2_{a\mathrm{X}+b} = \mathbf{Var}\left(a\mathrm{X}+b\right) = a^2\mathbf{Var}\left(\mathrm{X}\right) \neq 0$$

dès que $a \neq 0$, de-même pour $\sigma_{c\mathrm{X}+d}$.

Avec les notations de l'énoncé, nous avons

$$\rho(a\mathrm{X}+b, c\mathrm{Y}+d) = \frac{\mathbf{Cov}\left(a\mathrm{X}+b, c\mathrm{Y}+d\right)}{\sqrt{\mathbf{Var}\left(a\mathrm{X}+b\right)}\sqrt{\mathbf{Var}\left(c\mathrm{Y}+d\right)}},$$

or d'après les propriétés classiques sur la variance,

$$\mathbf{Var}\left(a\mathrm{X}+b\right) = a^2\mathbf{Var}\left(\mathrm{X}\right) = a^2\sigma^2_{\mathrm{X}}, \quad \mathbf{Var}\left(c\mathrm{Y}+d\right) = c^2\mathbf{Var}\left(\mathrm{Y}\right) = c^2\sigma^2_{\mathrm{Y}}.$$

Puis, par bilinéarité de la covariance :

$$\mathbf{Cov}\left(a\mathrm{X}+b, c\mathrm{Y}+d\right) = ac\mathbf{Cov}\left(\mathrm{X},\mathrm{Y}\right) + a\mathbf{Cov}\left(\mathrm{X},d\right) + c\mathbf{Cov}\left(b,\mathrm{Y}\right) + \mathbf{Cov}\left(b,d\right).$$

Or la covriance d'une variable aléatoire avec n'importe quelle constante est nulle, donc on trouve finalement

$$\mathbf{Cov}\left(a\mathrm{X}+b, c\mathrm{Y}+d\right) = ac\mathbf{Cov}\left(\mathrm{X},\mathrm{Y}\right).$$

Ainsi,

$$\rho(aX+b, cY+d) = \frac{ac\mathbf{Cov}(X,Y)}{|a|\,|c|\,\sigma_X\sigma_Y} = \frac{ac}{|a|\,c}\rho(X,Y).$$

Or, $\frac{ac}{|a|c} = \operatorname{sgn}(ac)$, on obtient alors la formule

$$\boxed{\rho(aX+b, cY+d) = \operatorname{sgn}(ac)\rho(X,Y)}.$$

Exercice ALEA.15.5 | **Équivalence entre indépendance et non-corrélation pour les couples de Bernoulli** Soient X et Y deux variables aléatoires de Bernoulli de paramètre $p \in [0,1]$.

1. Montrer que

$$|\mathbf{Cov}(X,Y)| \leqslant \frac{1}{4}.$$

2. Montrer que X, Y sont indépendantes, si et seulement si $\mathbf{Cov}(X,Y) = 0$. *Propriété également symptomatique des vecteurs aléatoires de lois normales, appelés « vecteurs gaussiens » : l'indépendance est équivalente à la non-corrélation deux à deux.*

Solution (exercice ALEA.15.5)

1. D'après le cours, $\mathbf{Var}(X) = p(1-p) = \mathbf{Var}(Y)$. Or,

$$|\rho(X,Y)| = \left|\frac{\mathbf{Cov}(X,Y)}{p(1-p)}\right| \leqslant 1,$$

dès lors, nous obtenons

$$|\mathbf{Cov}(X,Y)| \leqslant p(1-p).$$

Or, par étude de la fonction $f : x \longmapsto x(1-x)$ sur $[0,1]$, nous déduisons que pour tout $x \in [0,1]$,

$$x(1-x) \leqslant \frac{1}{4}.$$

Ainsi, $\boxed{|\mathbf{Cov}(X,Y)| \leqslant \frac{1}{4}}$.

2. Montrons que X, Y sont indépendantes, si et seulement si $\mathbf{Cov}(X,Y) = 0$.
$\boxed{\Longrightarrow}$ Conséquence directe du cours : deux variables indépendantes sont non-corrélées.
$\boxed{\Longleftarrow}$ Supposons que $\mathbf{Cov}(X,Y) = 0$. D'après le théorème de transfert, applicable car on travaille sur des univers-images finis, nous avons

$$\mathbf{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = 0+0+0+1\times 1\mathbf{P}(X=1, Y=1) - p^2,$$

donc l'hypothèse donne

$$\mathbf{P}(X=1, Y=1) - p^2 = 0 \quad \Longleftrightarrow \quad \mathbf{P}(X=1, Y=1) = \mathbf{P}(X=1)\mathbf{P}(Y=1).$$

Maintenant, il faut magouiller avec les évènements pour obtenir le reste de la définition d'indépendance.

$$\mathbf{P}(X=1, Y=0) + \mathbf{P}(X=1, Y=1) = \mathbf{P}(X=1) = p,$$

donc on obtient

$$\mathbf{P}(X=1, Y=0) = p - p^2 = p(1-p) = \mathbf{P}(X=1)\mathbf{P}(Y=0),$$

on obtient par le même procédé

$$\mathbf{P}(X=0, Y=1) = p - p^2 = (1-p)p = \mathbf{P}(X=0)\mathbf{P}(Y=1),$$

puis enfin

$$\begin{aligned}\mathbf{P}(X=0, Y=0) &= \mathbf{P}(X=0) - \mathbf{P}(X=0, Y=1) = 1 - p - (1-p)p\\ &= (1-p)(1-p) = \mathbf{P}(X=0)\mathbf{P}(Y=0).\end{aligned}$$

4.3. Étude de lois

Exercice ALEA.15.6 | Soient X et Y deux variable aléatoire prenant leurs valeurs dans $\{0,1,2\}$. On suppose que la loi du couple (X,Y) est donnée par le tableau :

Y \ X	0	1	2
0	p	$\frac{p}{2}$	$\frac{p}{4}$
1	$2p$	p	$\frac{p}{2}$
2	$4p$	$2p$	p

1. Déterminer une condition sur p pour que ce tableau représente effectivement la loi conjointe d'un couple aléatoire.
2. Déterminer alors les lois marginales de ce couple.
3. Calculer le coefficient de corrélation entre X et Y.
4. Les variables X et Y sont-elles indépendantes ?

Solution (exercice ALEA.15.6)

1. D'après le cours, le tableau représente la loi conjointe d'un couple (X,Y) si et seulement si tous les termes du tableau sont positifs, et si leur somme vaut 1. Et en effet, leur somme est

$$3p+8p+p+\frac{p}{4}=\frac{49p}{4}.$$

Donc la seule valeur possible de p est $\frac{4}{49}$. Par ailleurs pour cette valeur de p tous les termes du tableau sont bien positifs. Donc $\boxed{p=\frac{4}{49}}$.

2. D'après le tableau, on a $\boxed{X(\Omega)=Y(\Omega)=\{0,1,2\}}$. D'après la formule des probabilités totales, on a :

$$\mathbf{P}(X=0)=\sum_{\ell=0}^{2}\mathbf{P}(X=0,Y=\ell)=\boxed{7p=\frac{4}{7}},$$
$$\mathbf{P}(X=1)=\sum_{\ell=0}^{2}\mathbf{P}(X=1,Y=\ell)=\boxed{\frac{7p}{2}=\frac{2}{7}},$$
$$\mathbf{P}(X=2)=\sum_{\ell=0}^{2}\mathbf{P}(X=2,Y=\ell)=\boxed{\frac{7p}{4}=\frac{1}{7}}.$$

En faisant de-même pour Y, on trouve

$$\boxed{\mathbf{P}(Y=0)=\frac{7p}{4}=\frac{1}{7},\quad \mathbf{P}(Y=1)=\frac{7p}{2}=\frac{2}{7},\quad \mathbf{P}(Y=2)=7p=\frac{4}{7}}.$$

3. On commence par calculer la covariance, *i.e.* $\mathbf{Cov}(X,Y)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y)$, qui existe puisque X, Y sont des variables aléatoires finies. Tout d'abord,

$$\mathbf{E}(X)=\frac{14}{49}+\frac{2}{7}=\frac{4}{7},$$
$$\mathbf{E}(Y)=\frac{14}{49}+\frac{8}{7}=\frac{10}{7}.$$

D'après le théorème de transfert, on a :

$$\mathbf{E}(XY)=0+p+2\frac{p}{2}+2(2p)+4p$$
$$=10p=\frac{40}{49}.$$

Donc

$$\mathbf{Cov}(X,Y)=\frac{40}{49}-\frac{4}{7}\frac{10}{7}=0.$$

Donc $\boxed{\rho(X,Y)=\frac{\mathbf{Cov}(X,Y)}{\sigma_X\sigma_Y}=0.}$

4. Oui, vérifier tous les cas possibles.

Exercice ALEA.15.7 | Loi multinomiale – D'après concours G2E Soient p_1, p_2 et p_3 des réels strictement positifs. Pour tout $n \in \mathbf{N}$, on considère $D_n = \{(i,j) \in \mathbf{N}^2 \ : \ i+j \leq n\}$ et la fonction de deux variable f définie sur D_n par

$$f(i,j) = \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j}.$$

1. Déterminer une condition nécessaire et suffisante d'existence de (X,Y) de loi conjointe f.
2. Déterminer les lois marginales des variables aléatoires X et Y, ainsi que leur espérance et variance.

Solution (exercice ALEA.15.7)

1. Il s'agit de calculer la somme double. On a

$$\begin{aligned}\sum_{(i,j)\in D_n} f(i,j) &= \sum_{i=0}^{n}\sum_{j=0}^{n-i} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j} \\ &= \sum_{i=0}^{n} \frac{n!p_1^i}{i!} \sum_{j=0}^{n-i} \frac{1}{j!(n-i-j)!} p_2^j p_3^{n-i-j} \\ &= \sum_{i=0}^{n} \frac{n!p_1^i}{i!(n-i)!} \sum_{j=0}^{n-i} \binom{n-i}{j} p_2^j p_3^{n-i-j} && \textit{binôme} \\ &= \sum_{i=0}^{n} \frac{n!p_1^i}{i!(n-i)!} (p_2+p_3)^{n-i} && \textit{binôme} \\ &= (p_1+p_2+p_3)^n.\end{aligned}$$

Donc

$$\sum_{(i,j)\in D_n} f(i,j) = 1 \iff (p_1+p_2+p_3)^n = 1 \iff p_1+p_2+p_3 = 1.$$

Par ailleurs, tous les $f(i,j)$, $(i,j) \in D_n$ sont bien entendu positifs, donc f définit une loi conjointe si et seulement si

$$\boxed{p_1+p_2+p_3 = 1}.$$

2. On a $\boxed{X(\Omega) = [\![0\,,\,n]\!] = Y(\Omega).}$ De plus, soit $i \in [\![0\,,\,n]\!]$, alors d'après la formule des probabilités totales :

$$\begin{aligned}\mathbf{P}(X=i) &= \sum_{j=0}^{n} f(i,j) \\ &= \sum_{j=0}^{n-i} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j} \\ &= \frac{n!}{i!(n-i)!} p_1^i \sum_{j=0}^{n-i} \frac{(n-i)!}{j!(n-i-j)!} p_2^j p_3^{n-i-j} && \textit{binôme} \\ &= \binom{n}{i} p_1^i (p_2+p_3)^{n-i} && p_2+p_3 = 1-p_1 \\ &= \binom{n}{i} p_1^i (1-p_1)^{n-i}.\end{aligned}$$

Donc $\boxed{X \hookrightarrow \mathscr{B}(n,p_1)}$. De-même, en échangeant les rôles de i et j, on trouve $\boxed{Y \hookrightarrow \mathscr{B}(n,p_2).}$ Donc

$$\boxed{\mathbf{E}(X) = np_1,\ \mathbf{Var}(X) = np_1(1-p_1),\ \mathbf{E}(Y) = np_2,\ \mathbf{Var}(Y) = np_2(1-p_2).}$$

Exercice ALEA.15.8 | Soient X et Y deux variables aléatoires à valeurs dans $\mathbf{N}^\star$ telles que :

$$\mathbf{P}(X=i, Y=j) = \frac{a}{2^{i+j}} \quad \text{avec } (i,j) \in (\mathbf{N}^\star)^2 \text{ et } a \in \mathbf{R}.$$

1. Calculer a pour que la famille $(\mathbf{P}(X=i, Y=j))_{i,j}$ soit bien une loi conjointe d'un couple aléatoire.

2. Déterminer les lois marginales de X et Y.
3. Les variables aléatoires X et Y sont-elles indépendantes ?

Solution (exercice ALEA.15.8)

1. — Soit $i \in \mathbf{N}^\star$, alors étudions la convergence de $\left(\sum \frac{1}{2^{i+j}}\right)_{j\geqslant 1}$. Puisque $|1/2| < 1$, nous avons
$$\sum_{j=1}^{\infty} \frac{1}{2^{i+j}} = \frac{1}{2^i}\sum_{j=1}^{\infty}\frac{1}{2^j} = \frac{1}{2^i}\frac{1}{2}\frac{1}{1-1/2} = \frac{1}{2^i}.$$
— De-même $\sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2}\frac{1}{1-\frac{1}{2}} = 1$.
Les hypothèses du théorème de Fubini sont donc vérifiées, et nous avons
$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{1}{2^{i+j}} = 1.$$
Donc l'unique valeur de a telle que
$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\mathbf{P}(X = i, Y = j) = 1$$
est $\boxed{a = 1}$. De plus, pour tout $(i,j) \in (\mathbf{N}^\star)^2$,
$$\frac{a}{2^{i+j}} \geqslant 0$$
d'où l'existence d'un couple aléatoire (X, Y) de loi conjointe donnée précédemment.
2. Nous avons clairement $\boxed{X(\Omega) = Y(\Omega) = \mathbf{N}^\star}$. De plus, pour tout $i \geqslant 1$, nous avons d'après le cours :
$$\mathbf{P}(X = i) = \sum_{j=1}^{\infty}\mathbf{P}(X = i, Y = j) = \frac{1}{2^i}\sum_{j=1}^{\infty}\frac{1}{2^j} = \frac{1}{2^i} = \frac{1}{2}\frac{1}{2^{i-1}},$$
de-même, par symétrique des indices, on a
$$\mathbf{P}(Y = j) = \frac{1}{2^j} = \frac{1}{2}\frac{1}{2^{j-1}}.$$
On reconnaît ainsi des lois géométriques : $\boxed{X, Y \hookrightarrow \mathscr{G}(1/2).}$
3. Les variables aléatoires X, Y sont indépendantes puisque pour tout $(i,j) \in (\mathbf{N}^\star)^2$,
$$\mathbf{P}(X = i)\,\mathbf{P}(Y = j) = \frac{1}{2^i}\frac{1}{2^j} = \frac{1}{2^{i+j}} = \mathbf{P}(X = i, Y = j)$$
donc $\boxed{X \perp\!\!\!\perp Y.}$

Exercice ALEA.15.9 | Une urne contient 3 jetons numérotés 1, 2 et 3, indiscernables au toucher. On effectue une suite de tirages d'un jeton de cette urne, en replaçant à chaque fois le jeton obtenu avant le tirage suivant.

1. On note Y le nombre de tirage nécessaires pour obtenir, pour la première fois, deux numéros différents. Déterminer la loi de Y et son espérance.
2. On note Z le nombre de tirages nécessaires pour obtenir, pour la première fois, les trois numéros. Déterminer la loi du couple (Z, Y).
3. Déterminer la loi de Z et son espérance.

Solution (exercice ALEA.15.9)

1. La loi en question ressemble à une géométrique, sachant le résultat du premier tirage. Notons T_i l'évènement «le jeton i sort au tirage 1» pour tout $i \in [\![1, 3]\!]$. Alors sachant l'un de ces évènements, Y compte le temps d'attente de l'un des deux autres jetons, de probabilité $\frac{2}{3}$. Donc $\boxed{Y(\Omega) = [2, \infty[\, \cap\, \mathbf{N}}$, et de plus, pour tout $k \geqslant 2$,
$$\begin{aligned}\mathbf{P}(Y = k) &= \sum_{i=1}^{3}\mathbf{P}(Y = k | T_i)\,\mathbf{P}(T_i)\\ &= \sum_{i=1}^{3}\frac{2}{3}\left(\frac{1}{3}\right)^{(k-1)-1} \times \frac{1}{3}\\ &= \boxed{\frac{2}{3}\left(\frac{1}{3}\right)^{k-2}}.\end{aligned}$$

Il faut bien faire attention à la dernière étape au décalage de 1 dans la puissance : c'est seulement une fois le premier lancer réalisé que l'on commence à attendre un succès. Afin de minimiser nos efforts, on peut constater que $Y-1 \hookrightarrow \mathscr{G}\left(\frac{2}{3}\right)$ puisque $(Y-1)(\Omega) = \mathbf{N}^\star$ et que pour tout $k \in \mathbf{N}^\star$, $\mathbf{P}\left(Y-1=k\right) = \mathbf{P}\left(Y=k+1\right) = \frac{2}{3}\left(\frac{1}{3}\right)^{k-1}$. Donc $Y-1$ admet une espérance et

$$\mathbf{E}\left(Y-1\right) = \frac{3}{2} \quad \Longleftrightarrow \quad \boxed{\mathbf{E}\left(Y\right) = \frac{5}{2}}.$$

2. On note Z le nombre de tirages nécessaires pour obtenir, pour la première fois, les trois numéros. Nécessairement, il faut plus de temps pour obtenir les 3 numéros que les 2. Ainsi

$$\boxed{(Z,Y)(\Omega) = \left\{(k,\ell) \in ([2,\infty[\,\cap \mathbf{N})^2,\ k>\ell\right\}}.$$

Soient $k,\ell \geqslant 2$ de sorte que $k>\ell$. Alors si on a obtenu deux numéros distincts au tirage ℓ, on attendra un dernier succès, de probabilité $\frac{1}{3}$, donc nous aurons une loi géométrique à nouveau.

$$\begin{aligned}\mathbf{P}\left(Z=k, Y=\ell\right) &= \mathbf{P}(Z=k \mid Y=\ell)\,\mathbf{P}\left(Y=\ell\right)\\ &= \frac{1}{3}\left(\frac{2}{3}\right)^{k-1-\ell} \times \frac{2}{3}\left(\frac{1}{3}\right)^{\ell-2}\\ &= \boxed{\left(\frac{1}{3}\right)^{\ell-1} \cdot \left(\frac{2}{3}\right)^{k-\ell}}.\end{aligned}$$

3. On a $\boxed{Z(\Omega) = [3,\infty[\,\cap \mathbf{N}}$ puisqu'il faut attendre au moins trois lancers pour avoir tous les numéros. Soit donc $k \geqslant 3$, alors

$$\begin{aligned}\mathbf{P}\left(Z=k\right) &= \sum_{\ell=2}^{k-1}\left(\frac{1}{3}\right)^{\ell-1} \cdot \left(\frac{2}{3}\right)^{k-\ell}\\ &= 3\left(\frac{2}{3}\right)^k \sum_{\ell=2}^{k-1}\left(\frac{1}{2}\right)^{\ell}\\ &= 3\left(\frac{2}{3}\right)^k \left(\frac{1}{2}\right)^2 \frac{1-\left(\frac{1}{2}\right)^{k-2}}{\frac{1}{2}}\\ &= \boxed{\frac{3}{2}\left(\frac{2}{3}\right)^k\left(1-\left(\frac{1}{2}\right)^{k-2}\right)}.\end{aligned}$$

4.4. Somme de variables aléatoires & Stabilité de lois

4.4.1. Cas discret

Exercice ALEA.15.10 | On fait deux biopsies à un patient. Dans la première n cellules sont étudiées et on désigne par X le nombre de cellules malignes. Dans la seconde m cellules sont étudiées et on note Y le nombre de cellules malignes. La probabilité qu'une cellule soit maligne est notée p.

1. Par quelle loi peut-on modéliser les variables aléatoires X et Y?
2. Que représente $X+Y$? Déterminer la loi de $X+Y$.
3. Le laborantin a mélangé par inadvertance les deux éprouvettes. Quelle est alors la loi conditionnelle de X sachant $\{X+Y=k\}$, $k \in [\![0\,,\,n+m]\!]$?

Solution (exercice ALEA.15.10)

1. La variable aléatoire X compte le nombre de succès (avoir une cellule maligne, de probabilité p) dans une répétition de n expériences de BERNOUILLI (considérer une cellule et analyser si elle est maligne ou pas) indépendantes. Donc $\boxed{X \hookrightarrow \mathscr{B}(n,p)}$, de-même $\boxed{Y \hookrightarrow \mathscr{B}(m,p)}$.
2. La variable aléatoire X + Y compte alors le nombre total de cellules malignes sur les deux prélèvements, et par indépendance, d'après le cours nous avons $\boxed{X+Y \hookrightarrow \mathscr{B}(n+m,p)}$.
3. Le laborantin a mélangé par inadvertance les deux éprouvettes. Soit $\ell \in [\![0\,,\,n]\!]$, alors

$$\begin{aligned}
\mathbf{P}(X=\ell | X+Y=k) &= \frac{\mathbf{P}\left(X=\ell, X+Y=k\right)}{\mathbf{P}\left(X+Y=k\right)} \\
&= \frac{\mathbf{P}\left(X=\ell, Y=k-\ell\right)}{\mathbf{P}\left(X+Y=k\right)} \\
&= \frac{\mathbf{P}\left(X=\ell\right)\mathbf{P}\left(Y=k-\ell\right)}{\mathbf{P}\left(X+Y=k\right)} \qquad \textit{indépendance} \\
&= \frac{\binom{n}{k}p^k(1-p)^{n-k}\cdot\binom{m}{k-\ell}p^{k-\ell}(1-p)^{m-k+\ell}}{\binom{n+m}{k}p^k(1-p)^{n+m-k}} \\
&= \frac{\binom{n}{k}\cdot\binom{m}{k-\ell}}{\binom{n+m}{k}}. \qquad \textit{simplifications}
\end{aligned}$$

On reconnaît alors une loi $\boxed{\mathscr{H}\left(n+m, k, \frac{n}{n+m}\right).}$

Exercice ALEA.15.11 | **Sommer exceptionnellement en le paramètre p** Une urne contient des jetons numérotés de 1 à $k \geqslant 2$. Chaque jeton $j \in [\![1\,,\,k]\!]$ est en proportion $p_j \in]0,1[$. On effectue $n \geqslant 1$ tirages avec remise.

1. Déterminer la loi de N_i définie comme le nombre de jetons numéros i tirés.
2. Soient i, j deux entiers de $[\![1\,,\,k]\!]$.
   - **2.1)** Déterminer la loi de $N_i + N_j$ si $i \neq j$, son espérance, sa variance.
   - **2.2)** Calculer $\mathbf{Cov}\left(N_i, N_j\right)$.

Solution (exercice ALEA.15.11)

1. On répète n fois de façon indépendante et identique la même expérience de BERNOULLI qui consiste à tirer un jeton dans l'urne. Le succès est de tirer un jeton numéroté i et sa probabilité est p_i. Comme N_i compte le nombre de succès, on a $\boxed{N_i \hookrightarrow \mathscr{B}(n,p_i)}$.
2. Il s'agit de traiter globalement la variable $N_i + N_j$:
   - **2.1)** De nouveau comme dans la question précédente, on répète n fois de façon indépendante et identique la même expérience de BERNOULLI qui consiste à tirer un jeton dans l'urne. Mais cette fois, le succès est de tirer un jeton numéroté i ou j et sa probabilité est donc $p_i + p_j$ (proportion de jetons portant le numéro i ou j).
     Comme $N_i + N_j$ compte le nombre de succès, on a $\boxed{N_i + N_j \hookrightarrow \mathscr{B}(n, p_i + p_j)}$
     D'où $\mathbf{E}\left(N_i + N_j\right) = n(p_i + p_j)$ et $V(N_i + N_j) = n(p_i + p_j)(1-(p_i + p_j))$.
   - **2.2)** Or on sait aussi que $\mathbf{Var}\left(N_i + N_j\right) = \mathbf{Var}\left(N_i\right) + \mathbf{Var}\left(N_j\right) + 2\mathbf{Cov}\left(N_i, N_j\right)$. On en déduit

$$\begin{aligned}
\mathbf{Cov}\left(N_i, N_j\right) &= \frac{1}{2}\left[\mathbf{Var}\left(N_i+N_j\right) - \mathbf{Var}\left(N_i\right) + \mathbf{Var}\left(N_j\right)\right] \\
&= \frac{1}{2}\left[n(p_i+p_j)(1-p_i-p_j) - np_i(1-p_i) - np_j(1-p_j)\right] \\
&= \frac{1}{2}\left[np_i + np_j - np_i^2 - np_j^2 - 2np_ip_j - np_i + np_i^2 - np_j + np_j^2\right] \\
&= -np_ip_j
\end{aligned}$$

En conclusion : $\boxed{\mathbf{Cov}\left(N_i, N_j\right) = -np_ip_j.}$

Exercice ALEA.15.12 | Soient $p \in]0,1[$ et $q = 1-p$, ainsi que X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbf{N}$ telle que :

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X = k) = \mathbf{P}(Y = k) = pq^k.$$

On définit alors $U = \max(X, Y)$ et $V = \min(X, Y)$.

1. Donner la loi conjointe de (U, V), puis les marginales de U et V.
2. Les variables aléatoires U, V sont-elles indépendantes ?

Solution (exercice ALEA.15.12)

1. Nous avons $\boxed{(U,V)(\Omega) = \left\{(k,\ell) \in \mathbf{N}^2,\ k \geqslant \ell\right\}}$. Soit donc $(k,\ell) \in (U,V)(\Omega)$, alors

$$\mathbf{P}(U = k, V = \ell) = \mathbf{P}(\{X = k, Y = \ell\} \cup \{X = \ell, Y = k\}).$$

Les deux évènements apparaissant dans la réunion sont disjoints uniquement si $k \neq \ell$.

- Donc si $k \neq \ell$, nous avons $\mathbf{P}(U = k, V = \ell) = \mathbf{P}(X = k, Y = \ell) + \mathbf{P}(X = \ell, Y = k) = 2(pq^k)(pq^\ell) = 2p^2q^{k+\ell}$ puisque $X \perp\!\!\!\perp Y$ et que X, Y ont même loi.
- Si $k = \ell$, alors $\mathbf{P}(U = k, V = k) = \mathbf{P}(X = k, Y = k) = (pq^k)^2$.

Donc :

$$\boxed{\forall (k,\ell) \in (U,V)(\Omega), \quad \mathbf{P}(U = k, V = \ell) = \begin{cases} 2p^2q^{k+\ell} & \text{si } \ell < k, \\ p^2q^{2k} & \text{si } k = \ell. \end{cases}}$$

Passons à présent aux lois marginales.

- $\boxed{U(\Omega) = V(\Omega) = \mathbf{N}.}$
- Soit $k \in U(\Omega)$, alors d'après le cours,

$$\mathbf{P}(U = k) = \sum_{\ell=0}^{k} \mathbf{P}(U = k, V = \ell).$$

Faisons deux cas. Si $k \geqslant 1$:

$$\begin{aligned} \mathbf{P}(U = k) &= \sum_{\ell=0}^{k-1} \left(2p^2q^{k+\ell}\right) + (pq^k)^2 \\ &= 2p^2q^k \frac{1-q^k}{p} + (pq^k)^2. \end{aligned}$$

Par ailleurs, si $k = 0$:

$$\mathbf{P}(U = 0) = \mathbf{P}(X = 0)\,\mathbf{P}(Y = 0) = p^2.$$

Donc :

$$\mathbf{P}(U = k) = \begin{cases} 2pq^k + p(p-2)q^{2k} & \text{si } k \neq 0, \\ p^2 & \text{si } k = 0. \end{cases} = \boxed{2pq^k + p(p-2)q^{2k}}.$$

De-même, pour tout $\ell \in \mathbf{N}$,

$$\mathbf{P}(V = \ell) = \sum_{k=\ell}^{\infty} \mathbf{P}(U = k, V = \ell) = \sum_{k=\ell+1}^{\infty} \left(2p^2q^{k+\ell}\right) + (pq^\ell)^2.$$

Ainsi,

$$\mathbf{P}(V = \ell) = 2p^2q^\ell q^{\ell+1}\frac{1}{p} + p^2q^{2\ell} = \boxed{2pq^{2\ell+1} + p^2q^{2\ell}}.$$

2. $\boxed{\text{Non}}$, car

$$\mathbf{P}(U = 0)\,\mathbf{P}(V = 0) = p^2(2pq + p^2) = p^3(2q + p) = p^3(2 - p),$$

et

$$\mathbf{P}(U = 0, V = 0) = p^2,$$

mais $p^2 = p^3(2-p) \iff 1 = p(2-p) \iff p = 1$ (seule solution dans $[0,1]$), ce qui est exclu.

Exercice ALEA.15.13 | **Loi binomiale négative** Une équipe de géologues travaille en zone sismique et installe un sismomètre au sommet d'un volcan. Chaque jour, si ce sismomètre détecte une onde sismique, ce dernier envoie une alerte par satellite à un camp de base situé au pied du volcan. On note X_1 la variable aléatoire réelle correspondant au nombre de jours sans alerte avant une première alerte, par exemple si cette première alerte survient au troisième jour après l'installation du sismomètre alors X_1 prend la valeur 2.

On admet que ce sismomètre envoie au plus une alerte par jour, et que la probabilité que survienne une alerte est constante et égale à p, indépendante des résultats obtenus les jours précédents.

1. **1.1)** Déterminer l'ensemble des valeurs prises par X_1 noté $X_1(\Omega)$, et pour tout $j \in X_1(\Omega)$, calculer $\mathbf{P}(X_1 = j)$.
   1.2) En déduire que $1+X_1$ suit la loi géométrique de paramètre p ce qu'on notera dorénavant $X_1 \hookrightarrow \mathscr{G}^\star(p)$.
2. Ce sismomètre étant extrêmement sensible, les géologues décident de mener des analyses complémentaires à partir d'une seconde alerte. Aussi on note X_2 la variable aléatoire réelle correspondant au nombre de jours sans alerte avant cette seconde alerte, depuis l'installation du sismomètre. Par exemple, si une première alerte survient au troisième jour après l'installation de ce sismomètre et une seconde au septième jour alors X_2 prend la valeur 5
   2.1) Déterminer $(X_1, X_2)(\Omega)$.
   2.2) Soit $j \in \mathbf{N}$. Déterminer la loi conditionnelle de X_2 sachant $\{X_1 = j\}$.
   2.3) En déduire que :
$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X_2 = k) = (k+1)q^k p^2.$$
3. On généralise l'étude précédente et pour tout $n \in \mathbf{N}^*$, on note X_n la variable aléatoire réelle correspondant au nombre de jours sans alerte avant une n-ième alerte.
   3.1) Démontrer par récurrence sur $k \in \mathbf{N}$ que :
$$\forall k \in \mathbf{N}, \quad \binom{n+k}{k} = \sum_{j=0}^{k} \binom{n+j-1}{j}.$$
   3.2) Démontrer par récurrence sur $n \in \mathbf{N}^*$ que :
$$\forall k \in \mathbf{N}, \quad \mathbf{P}(X_n = k) = \binom{n+k-1}{k} q^k p^n.$$

Solution (exercice ALEA.15.13)

1. **1.1)** La première alerte peut survenir dès le premier jours, donc $\boxed{X_1(\Omega) = \mathbf{N}}$. Soit $j \in X_1(\Omega)$, alors $\{X_1 = j\}$ signifie que nous avons eu j jours sans alerte puis une alerte au jour $j+1$, donc :
$$\boxed{\mathbf{P}(X_1 = j) = q^j p}.$$
   1.2) On a clairement $(X_1+1)(\Omega) = \mathbf{N}^\star$, donc pour tout $\ell \geqslant 1$, on a
$$\mathbf{P}(X_1 + 1 = \ell) = \mathbf{P}(X_1 = \ell - 1) = q^{\ell-1}p,$$
   donc $\boxed{1+X_1 \hookrightarrow \mathscr{G}(p).}$
2. **2.1)** Nécessairement $X_2 \geqslant X_1$, et on peut éventuellement avoir égalité si deux alertes surviennent l'une après l'autre. Donc
$$\boxed{(X_1, X_2)(\Omega) = \left\{(j,k) \in \mathbf{N}^2, j \leqslant k\right\}}.$$
   2.2) Soit $j \in \mathbf{N}$. Soit $k \geqslant j$, alors
$$\boxed{\mathbf{P}(X_2 = k | X_1 = j) = q^{k-j}p},$$
   qui correspond à $k-j$ jours sans alerte puis une alerte.

2.3) On a $\boxed{X_2(\Omega)=\mathbf{N}}$, le cas zéro étant obtenu lorsque deux alertes consécutives ont lieu les deux premiers jours. D'après la formule des probabilités totales, on déduit alors :

$$\begin{aligned}\mathbf{P}(X_2=k)&=\sum_{j=0}^{\infty}\mathbf{P}(X_2=k,X_1=j)\\&=\sum_{j=0}^{k}\mathbf{P}(X_2=k|X_1=j)\,\mathbf{P}(X_1=j)\\&=\sum_{j=0}^{k}q^{k-j}pq^jp\\&=\boxed{\mathbf{P}(X_2=k)=(k+1)q^kp^2.}\end{aligned}$$

3. On généralise l'étude précédente et pour tout $n\in\mathbf{N}^*$, on note X_n la variable aléatoire réelle correspondant au nombre de jours sans alerte avant une n-ième alerte.

3.1) Posons H_k : $\binom{n+k}{k}=\sum_{j=0}^{k}\binom{n+j-1}{j}$ pour tout $k\in\mathbf{N}$.

- **Initialisation.** $\sum_{j=0}^{0}\binom{n+j-1}{j}=\binom{n-1}{0}=1=\binom{n}{0}$ donc H_0 est vraie.
- **Hérédité.** Soit k fixé tel que H_k soit vraie. Alors :

$$\begin{aligned}\sum_{j=0}^{k+1}\binom{n+j-1}{j}&=\sum_{j=0}^{k}\binom{n+j-1}{j}+\binom{n+k}{k+1}\\&=\binom{n+k}{k}+\binom{n+k}{k+1}.\end{aligned}$$

De plus, la formule de Pascal donne $\binom{n+k}{k}+\binom{n+k}{k+1}=\binom{n+k+1}{k+1}$. Donc H_{k+1} est vraie. Donc par principe de récurrence :

$$\boxed{\forall k\in\mathbf{N},\quad\binom{n+k}{k}=\sum_{j=0}^{k}\binom{n+j-1}{j}}.$$

3.2) Démontrons par récurrence sur $n\in\mathbf{N}^*$ que :

$$\forall k\in\mathbf{N},\quad\mathbf{P}(X_n=k)=\binom{n+k-1}{k}q^kp^n.$$

Posons, pour $n\geqslant1$: H_n « $\mathbf{P}(X_n=k)=\binom{n+k-1}{k}q^kp^n$ pour tout $k\in\mathbf{N}$. »

- **Initialisation.** Pour $n=1$, nous avons bien

$$\mathbf{P}(X_1=k)=\binom{k-1}{k}q^k=q^kp,$$

car cela a déjà été prouvé dans une question précédente.

- **Hérédité.** Supposons la propriété prouvée au rang n, c'est-à-dire :

$$\forall k\in\mathbf{N},\quad\mathbf{P}(X_n=k)=\binom{n+k-1}{k}q^kp^n.$$

Alors cherchons la loi de X_{n+1}. Soit donc $k\in\mathbf{N}$, nous avons encore d'après la formule des probabilités totales :

$$\begin{aligned}\mathbf{P}(X_{n+1}=k)&=\sum_{j=0}^{\infty}\mathbf{P}(X_{n+1}=k|X_n=j)\,\mathbf{P}(X_n=j)\\&=\sum_{j=0}^{k}q^{k-j}p\mathbf{P}(X_n=j)\\&=\sum_{j=0}^{k}q^{k-j}p\binom{n+j-1}{j}q^jp^n\\&=\binom{n+k}{k}q^kp^{n+1}.\end{aligned}$$

la loi condi. se calcule comme dans une question précédente, $k-j$ sans alerte puis une alerte

hypothèse de récurrence

question précédente

Donc par principe de récurrence, nous avons pour tout $n\in\mathbf{N}^\star$:

$$\boxed{\mathbf{P}(X_n=k)=\binom{n+k-1}{k}q^kp^n.}$$

Exercice ALEA.15.14 | Soient X,Y,Z trois variables aléatoires indépendantes suivant toutes la loi de Poisson $\mathscr{P}(\lambda)$ avec $\lambda>0$. Montrer que $\frac{X+Y}{1+Z}$ admet une espérance, et la calculer.

Solution (exercice ALEA.15.14)

On peut voir $\frac{X+Y}{1+Z}$ comme une fonction de trois variables en X,Y,Z mais nous ne disposons pas de théorème de transfert pour ce cadre. En revanche, notons $T = X+Y$ et $g : (x,y) \in \mathbf{R}^2 \longmapsto \frac{x}{1+y}$. Alors puisque $X \perp\!\!\!\perp Y$, $T \hookrightarrow \mathscr{P}(2\lambda)$. Et $\frac{X+Y}{1+Z} = h(T,Z)$. Donc $\frac{X+Y}{1+Z}$ admet une espérance si et seulement si

$$\sum_{k,\ell=0}^{\infty} \frac{k}{1+\ell}\mathbf{P}\left(T=k, Z=\ell\right) \quad \text{converge absolument.}$$

Ou encore par indépendance et positivité :

$$\sum_{k,\ell=0}^{\infty} \frac{k}{1+\ell}\mathbf{P}\left(T=k\right)\mathbf{P}\left(Z=\ell\right).$$

Fixons $\ell \in \mathbf{N}$ et étudions

$$\sum_{k=0}^{\infty} \frac{k}{1+\ell}\frac{e^{-2\lambda}(2\lambda)^k}{k!}\mathbf{P}\left(Z=\ell\right).$$

Soit $N \in \mathbf{N}$, alors

$$\begin{aligned}\sum_{k=0}^{N} \frac{k}{1+\ell}\frac{e^{-2\lambda}(2\lambda)^k}{k!}\mathbf{P}\left(Z=\ell\right) &= \frac{\mathbf{P}\left(Z=\ell\right)}{1+\ell}\sum_{k=0}^{N} k\frac{e^{-2\lambda}(2\lambda)^k}{k!} \\ &= \frac{\mathbf{P}\left(Z=\ell\right)}{1+\ell}e^{-2\lambda}(2\lambda)\sum_{k=1}^{N}\frac{(2\lambda)^{k-1}}{(k-1)!} \\ &\xrightarrow{N\longrightarrow\infty} \frac{\mathbf{P}\left(Z=\ell\right)}{1+\ell}e^{-2\lambda}(2\lambda)e^{2\lambda} \\ &= (2\lambda)\frac{e^{-\lambda}\lambda^{\ell}}{(1+\ell)!} \\ &= 2e^{-\lambda}\frac{\lambda^{\ell+1}}{(1+\ell)!}.\end{aligned}$$

Maintenant, on somme en l'autre variable. On sait qu'elle converge aussi d'après le cours sur la série exponentielle, et de plus par changement d'indice :

$$\sum_{\ell=0}^{\infty} 2e^{-\lambda}\frac{\lambda^{\ell+1}}{(1+\ell)!} = 2e^{-\lambda}\left(e^{\lambda}-1\right).$$

Ainsi, la série double étudiée est absolument convergente et par théorème de transfert, $\frac{X+Y}{1+Z}$ admet une espérance égale à :

$$\boxed{\mathbf{E}\left(\frac{X+Y}{1+Z}\right) = 2e^{-\lambda}\left(1-e^{-\lambda}\right)}.$$

Exercice ALEA.15.15 | On considère un concessionnaire de voitures. Soit X la variable aléatoire égale au nombre de voitures vendues par jour, on suppose que X suit une loi de Poisson de paramètre $\lambda > 0$. On considère qu'un acheteur de voiture fait un prêt pour se procurer sa voiture avec une probabilité $p \in]0,1[$. Soit Y la variable aléatoire égale au nombre d'acheteurs par jour de voitures qui font un prêt.

1. Déterminer la loi conjointe de (X,Y).
2. Déterminer la loi de Y, puis donner $\mathbf{E}(Y)$ et $\mathbf{Var}(Y)$.
3. On pose $Z = X - Y$. Quelle est la loi de Z ? Les variables aléatoires Y,Z sont-elles indépendantes ? Calculer $\mathbf{Cov}(X,Y)$.

Solution (exercice ALEA.15.15)

1. Il y a forcément moins d'acheteurs faisant un prêt que d'acheteurs. Donc

$$\boxed{(X,Y)(\Omega) = \left\{(k,\ell) \in \mathbf{N}^2,\ k \geqslant \ell\right\}}.$$

Soit donc $(k,\ell) \in \mathbf{N}^2$ vérifiant $k \geqslant \ell$. Ici, d'après l'énoncé, nous connaissons la loi du nombre d'acheteurs faisant un prêt conditionnellement à $\{X = k\}$, qui est une $\mathscr{B}(k,p)$: en effet, il s'agit de compter le nombre de succès (contraction d'un prêt) dans une répétition de k expériences de Bernouilli indépendantes. Donc

$$\mathbf{P}(Y=\ell|X=k) = \binom{k}{\ell}p^{\ell}(1-p)^{k-\ell}.$$

Donc, d'après la définition d'une probabilité conditionnelle :

$$\begin{aligned}\mathbf{P}(Y=\ell, X=k) &= \mathbf{P}(Y=\ell \mid X=k)\,\mathbf{P}(X=k)\\ &= \binom{k}{\ell} p^{\ell}(1-p)^{k-\ell} e^{-\lambda}\frac{\lambda^k}{k!}\\ &= \boxed{\frac{e^{-\lambda}}{\ell!(k-\ell)!}(\lambda p)^{\ell}(\lambda(1-p))^{k-\ell}.}\end{aligned}$$

2. On a $\boxed{Y(\Omega)=\mathbf{N}}$ et d'après la formule des probabilités totales, nous déduisons pour tout $\ell \in \mathbf{N}$

$$\begin{aligned}\mathbf{P}(Y=\ell) &= \sum_{k=0}^{\infty}\mathbf{P}(Y=\ell, X=k)\\ &= \sum_{k=\ell}^{\infty}\frac{e^{-\lambda}}{\ell!(k-\ell)!}(\lambda p)^{\ell}(\lambda(1-p))^{k-\ell}\\ &= e^{-\lambda}\frac{(\lambda p)^{\ell}}{\ell!}\sum_{k=\ell}^{\infty}\frac{(\lambda(1-p))^{k-\ell}}{(k-\ell)!}\\ &= e^{-\lambda}\frac{(\lambda p)^{\ell}}{\ell!}\sum_{k=0}^{\infty}\frac{(\lambda(1-p))^{k}}{k!} \quad \textit{changement d'indice}\\ &= e^{-\lambda}\frac{(\lambda p)^{\ell}}{\ell!}e^{-\lambda(1-p)}\\ &= e^{-\lambda p}\frac{(\lambda p)^{\ell}}{\ell!}.\end{aligned}$$

Ainsi, $\boxed{Y \hookrightarrow \mathscr{P}(\lambda p)}$.

3. On pose $Z = X - Y$. Il s'agit dans notre expérience du nombre d'acheteurs n'ayant pas contracté de prêt. On peut donc utiliser par analogie les résultats précédents en remplaçant p par $1-p$ (ce qui revient dans le compte de succès à inverser succès et échec). Ainsi, $\boxed{Z \hookrightarrow \mathscr{P}(\lambda(1-p)).}$

Les variables aléatoires $\boxed{\text{Y, Z ne sont clairement pas indépendantes}}$, car il y a une «balance» entre les acheteurs empruntant et les autres. Pour le voir, on peut par exemple écrire :

$$\mathbf{P}(Z=1, Y=0) = \mathbf{P}(X=1, Y=0) = e^{-\lambda}\lambda(1-p),$$

alors que

$$\mathbf{P}(Z=1)\,\mathbf{P}(Y=0) = e^{-\lambda(1-p)}e^{-\lambda p}.$$

La covariance existe puisque X, Y admettent une variance. On a au choix deux méthodes, la première est à privilégier.

- **(Méthode 1 : utiliser la loi de** Z**)** On a par bilinéarité de la covariance :

$$\begin{aligned}\mathbf{Var}(X-Y) &= \mathbf{Cov}(X-Y, X-Y)\\ &= \mathbf{Var}(X) - \mathbf{Cov}(X,Y) - \mathbf{Cov}(Y,X) + \mathbf{Var}(Y)\\ &= \mathbf{Var}(X) + \mathbf{Var}(Y) - 2\mathbf{Cov}(X,Y).\end{aligned}$$

Ainsi,

$$\begin{aligned}\mathbf{Cov}(X,Y) &= \frac{1}{2}(\mathbf{Var}(X) + \mathbf{Var}(Y) - \mathbf{Var}(Z))\\ &= \frac{1}{2}(\lambda + \lambda p - \lambda(1-p))\\ &= \boxed{\lambda p.}\end{aligned}$$

- **(Méthode 2 : König-Huygens)** On peut sinon revenir à König-Huygens, mais les calculs sont beaucoup plus longs. On $\mathbf{E}(X) = \lambda, \mathbf{E}(Y) = \lambda p$. Et

$$\begin{aligned}
\mathbf{E}(XY) &= \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} k\ell\,\mathbf{P}(X=k, Y=\ell)\\
&= \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} k\ell \frac{e^{-\lambda}}{\ell!(k-\ell)!}(\lambda p)^{\ell}(\lambda(1-p))^{k-\ell}\\
&= e^{-\lambda}\sum_{k=0}^{\infty} k \sum_{\ell=1}^{k}\left(\frac{1}{(\ell-1)!(k-\ell)!}(\lambda p)^{\ell}(\lambda(1-p))^{k-\ell}\right)\\
&= \lambda p e^{-\lambda}\sum_{k=1}^{\infty}\frac{k}{(k-1)!}\sum_{\ell=1}^{k}\left(\binom{k-1}{\ell-1}(\lambda p)^{\ell-1}(\lambda(1-p))^{(k-1)-(\ell-1)}\right) && \text{changement d'indice}\\
&= \lambda p e^{-\lambda}\sum_{k=1}^{\infty}\frac{k}{(k-1)!}\sum_{\ell=0}^{k-1}\left(\binom{k-1}{\ell}(\lambda p)^{\ell}(\lambda(1-p))^{(k-1)-\ell}\right) && \text{somme interne binôme}\\
&= \lambda p e^{-\lambda}\sum_{k=1}^{\infty}\frac{k}{(k-1)!}(\lambda p+\lambda(1-p))^{k-1}\\
&= \lambda p e^{-\lambda}\sum_{k=1}^{\infty}\frac{k-1+1}{(k-1)!}\lambda^{k-1}\\
&= \lambda p e^{-\lambda}\left(\lambda\sum_{k=2}^{\infty}\frac{1}{(k-2)!}\lambda^{k-2}+\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\lambda^{k-1}\right) && \text{changement d'indice + série exp.}\\
&= \lambda p e^{\lambda}\left(\lambda e^{-\lambda}+e^{\lambda}\right)\\
&= \lambda p(\lambda+1).
\end{aligned}$$

Donc par la formule de König-Huygens :

$$\boxed{\mathbf{Cov}(X,Y) = \lambda p(\lambda+1) - \lambda^2 p = \lambda p}.$$

Exercice ALEA.15.16 | **Produits de lois Rademacher** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes telles que $\mathbf{P}(X_n=-1)=p$ et $\mathbf{P}(X_n=1)=1-p$ avec $p\in\,]0,1[$. On pose $Z_n=\prod_{i=1}^n X_i$.

1. Calculer l'espérance de Z_n.
2. Déduire de la question précédente la loi de Z_n.
3. Donner une condition sur p pour que Z_1 et Z_2 soient indépendantes.

Solution (exercice ALEA.15.16)

1. Par mutuelle indépendance des X_i, on a $\mathbf{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbf{E}(X_i) \iff \mathbf{E}(Z_n) = (\mathbf{E}(X_1))^n$. Or $\mathbf{E}(X_1) = \mathbf{P}(X_1=1) - \mathbf{P}(X_1=-1) = 1-p-p = 1-2p$, d'où $\boxed{\mathbf{E}(Z_n) = (1-2p)^n.}$
2. On a deux égalités : $\mathbf{P}(Z_n=1) - \mathbf{P}(Z_n=-1) = \mathbf{E}(Z_n)$ et $\mathbf{P}(Z_n=1) + \mathbf{P}(Z_n=-1) = 1$, d'où un système d'équations que l'on peut résoudre : $\begin{cases}\mathbf{P}(Z_n=1)-\mathbf{P}(Z_n=-1)=(1-2p)^n\\ \mathbf{P}(Z_n=1)+\mathbf{P}(Z_n=-1)=1\end{cases} \Leftrightarrow \begin{cases}2\mathbf{P}(Z_n=1)=(1-2p)^n+1\\ 2\mathbf{P}(Z_n=-1)=1-(1-2p)^n\end{cases} \Leftrightarrow \begin{cases}\mathbf{P}(Z_n=1)=\frac{1+(1-2p)^n}{2}\\ \mathbf{P}(Z_n=-1)=\frac{1-(1-2p)^n}{2}\end{cases}$.
   Ainsi, $\boxed{Z_n(\Omega)=\{-1,1\} \text{ et } \mathbf{P}(Z_n=1)=\frac{1+(1-2p)^n}{2}, \mathbf{P}(Z_n=-1)=\frac{1-(1-2p)^n}{2}.}$
3. On peut commencer par calculer sous quelle condition la covariance est nulle. Nous avons.

$$\begin{aligned}
\mathbf{Cov}(Z_1,Z_2) &= \mathbf{Cov}(X_1,X_1X_2)\\
&= \mathbf{E}(X_1^2X_2) - \mathbf{E}(X_1)\mathbf{E}(X_1X_2) && \text{indépendance et même loi pour les } X_i\\
&= \mathbf{E}(X_1^2)\mathbf{E}(X_2) - \mathbf{E}(X_1)^3 && X_1^2 = 1\\
&= \mathbf{E}(X_2) - \mathbf{E}(X_1)^3 = 1-2p-(1-2p)^3\\
&= (1-2p)4p(1-p).
\end{aligned}$$

Or, comme $p\in\,]0,1[$, Z_1, Z_2 sont donc non corrélées uniquement pour $p=\frac{1}{2}$. Inversement, montrons que pour $p=\frac{1}{2}$, les variables aléatoires Z_1, Z_2 sont bien indépendantes. On calcule.

$$\mathbf{P}(Z_1=1,Z_2=1)=\mathbf{P}(X_1=1,X_1X_2=1)=\mathbf{P}(X_1=1,X_2=1)=\mathbf{P}(X_1=1)\mathbf{P}(X_2=1)=\frac{1}{4}$$

$$\begin{aligned}\mathbf{P}(Z_1=1)\mathbf{P}(Z_2=1)&=\mathbf{P}(X_1=1)\mathbf{P}(X_1X_2=1)\\&=\frac{1}{2}\left(\frac{1}{4}+\frac{1}{4}\right)\\&=\frac{1}{4}.\end{aligned}$$

$\{X_1X_2=1\} = \{X_1=1,X_2=1\}\uplus\{X_1=-1,X_2=-1\}$

On passe aux autres vérifications.

$$\mathbf{P}(Z_1=-1,Z_2=1)=\mathbf{P}(X_1=-1,X_1X_2=1)=\mathbf{P}(X_1=-1,X_2=-1)=\frac{1}{4}$$

$$\begin{aligned}\mathbf{P}(Z_1=-1)\mathbf{P}(Z_2=1)&=\mathbf{P}(X_1=-1)\mathbf{P}(X_1X_2=1)\\&=\frac{1}{2}\left(\frac{1}{4}+\frac{1}{4}\right)\\&=\frac{1}{4}.\end{aligned}$$

$\{X_1X_2=1\} = \{X_1=1,X_2=1\}\uplus\{X_1=-1,X_2=-1\}$

$$\mathbf{P}(Z_1=1,Z_2=-1)=\mathbf{P}(X_1=1,X_1X_2=-1)=\mathbf{P}(X_1=1,X_2=-1)=\frac{1}{4}$$

$$\begin{aligned}\mathbf{P}(Z_1=1)\mathbf{P}(Z_2=-1)&=\mathbf{P}(X_1=1)\mathbf{P}(X_1X_2=-1)\\&=\frac{1}{2}\left(\frac{1}{4}+\frac{1}{4}\right)\\&=\frac{1}{4}.\end{aligned}$$

$\{X_1X_2=-1\} = \{X_1=1,X_2=-1\}\uplus\{X_1=-1,X_2=1\}$

Et enfin :

$$\mathbf{P}(Z_1=-1,Z_2=-1)=\mathbf{P}(X_1=-1,X_2=1)=\frac{1}{4}$$

$$\begin{aligned}\mathbf{P}(Z_1=-1)\mathbf{P}(Z_2=-1)&=\mathbf{P}(X_1=-1)\mathbf{P}(X_1X_2=-1)\\&=\frac{1}{2}\left(\frac{1}{4}+\frac{1}{4}\right)\\&=\frac{1}{4}.\end{aligned}$$

$\{X_1X_2=-1\} = \{X_1=1,X_2=-1\}\uplus\{X_1=-1,X_2=1\}$

Donc $\boxed{Z_1, Z_2 \text{ sont indépendantes si et seulement si } p=\frac{1}{2}.}$

Exercice ALEA.15.17 | Quand l'indépendance n'est pas loin Soit (X_i) une suite de variables de BERNOULLI de même paramètre p, indépendantes. Pour tout $i\in\mathbf{N}$, on pose $Y_i=X_iX_{i+1}$.

1. Quelle est la loi de Y_i pour tout $i\in\mathbf{N}$?
2. Soit $n\in\mathbf{N}$, on pose : $\quad S_n=\sum_{i=1}^{n}Y_i$.
   - **2.1)** Démontrer que :
   $$\mathbf{Var}(S_n)=\sum_{i=1}^{n}\mathbf{Var}(Y_i)+2\sum_{1\leqslant i<j\leqslant n}\mathbf{Cov}\left(Y_i,Y_j\right).$$
   - **2.2)** Déterminer $\mathbf{E}(S_n)$ et $\mathbf{Var}(S_n)$.

Solution (exercice ALEA.15.17)

1. On voit que le support de Y_i est $\{0,1\}$ pour tout i. De plus, $\mathbf{P}(Y_i=1)=\mathbf{P}(X_i=1)\mathbf{P}(X_{i+1}=1)=p^2$ par indépendance. Ainsi, $\mathbf{P}(Y_i=0)=1-p^2$ et : $\boxed{Y_i\hookrightarrow\mathscr{B}(p^2).}$
2. L'espérance est linéaire, on en déduit immédiatement que : $\boxed{\mathbf{E}(S_n)=np^2}$.
   En revanche, étant donné que les variables aléatoires Y_i ne sont pas indépendantes, la variance se développe avec des termes de covariance :
   $$\begin{aligned}\mathbf{Var}(S_n)&=\sum_{i=1}^{n}\mathbf{Var}(Y_i)+2\sum_{1\leqslant i<j\leqslant n}\mathbf{Cov}\left(Y_i,Y_j\right)\\&=np^2(1-p^2)+2\sum_{1\leqslant i<j\leqslant n}\mathbf{Cov}\left(Y_i,Y_j\right).\end{aligned}$$
   Intéressons-nous aux termes de covariance. On remarque que dès que $j>i+1$, alors les variables $Y_i=X_iX_{i+1}$ et $Y_j=X_jX_{j+1}$ sont indépendantes (elles n'ont aucun facteur X_ℓ commun, et on peut alors appliquer le lemme des coalitions).
   Ainsi,
   $$\sum_{1\leqslant i<j\leqslant n}\mathbf{Cov}\left(Y_i,Y_j\right)=\sum_{1\leqslant i\leqslant n}\mathbf{Cov}(Y_i,Y_{i+1})=\sum_{1\leqslant i\leqslant n}\left(\mathbf{E}(Y_iY_{i+1})-p^2\right).$$

Et $\mathbf{E}(Y_iY_{i+1}) = 1 \times \mathbf{P}(Y_iY_{i+1} = 1) = \mathbf{P}(Y_i = 1, Y_{i+1} = 1) = \mathbf{P}(X_iX_{i+1} = 1, X_{i+1}X_{i+2} = 1) = \mathbf{P}(X_i = 1, X_{i+1} = 1, X_{i+2} = 1) = p^3$.
D'où finalement : $\boxed{\mathbf{Var}(S_n) = np^2(1-p^2) + 2(n-1)p^3(1-p).}$

4.4.2. Cas continu

Exercice ALEA.15.18 | On considère U et V deux variables indépendantes suivant la loi uniformes sur $[0,1]$

1. Déterminer la loi de $X = -\ln(U)$ et de $Y = -\ln(V)$.
2. Déterminer la loi de $Z = X + Y$.
3. Déterminer la loi de $T = e^Z$, et en déduire la loi de $\frac{1}{UV}$.

Exercice ALEA.15.19 | **Autour de la loi Gamma** Soit $n \in \mathbf{N}^*$. On pose

$$\Gamma(n) = \int_0^{+\infty} t^{n-1}e^{-t}\,dt.$$

1. Montrer par récurrence que pour tout $n \in \mathbf{N}^*$, $\Gamma(n)$ est convergente.
2. Calculer $\Gamma(n)$ en fonction de n.
3. On pose $f_n(t) = \begin{cases} \frac{1}{\Gamma(n)} t^{n-1}e^{-t} & \text{si } t \geq 0 \\ 0 & \text{sinon.} \end{cases}$ Montrer que f_n est une densité.
4. Soit X_n une variable aléatoire de densité f_{n+1}. Montrer que pour tout $x > 0$, si Y_x suit une loi de Poisson de paramètre x, on a :

$$\mathbf{P}(X_n \leq x) = 1 - \mathbf{P}(Y_x \leq n).$$

Indication : On pourra chercher à dériver cette inégalité.

Solution (exercice ALEA.15.19)

1. Montrer la convergence par récurrence.

■ **Initialisation.** Soit $A > 0$, alors $\int_0^A 1e^{-t}\,dt = 1 - e^{-A} \xrightarrow{A\to\infty} 1$. Donc $\Gamma(1)$ converge et vers 1.

■ **Hérédité.** Supposons que $\Gamma(n)$ converge, alors la fonction $A \longmapsto \int_0^A t^{n-1}e^{-t}\,dt$ admet une limite quand $A \longrightarrow \infty$. Alors, puisque les fonctions $t \longmapsto t^n$, $t \longmapsto -e^{-t}$ sont de classe $\mathscr{C}^1$, on obtient par intégration par parties :

$$\int_0^A t^n e^{-t}\,dt = -n\int_0^A t^{n-1}\left(-e^{-t}\right)dt + \left[-t^n e^{-t}\right]_0^A.$$

Mais $\left[-t^ne^{-t}\right]_0^A \xrightarrow{A\to\infty} 0$ par croissances comparées, et par hypothèse de récurrence, $-n\int_0^A t^{n-1}\left(-e^{-t}\right)dt \xrightarrow{A\to\infty} n\Gamma(n)$. Donc $\Gamma(n+1)$ converge.
Finalement, par principe de récurrence, on a montré que $\boxed{\Gamma(n) \text{ converge pour tout } n \in \mathbf{N}^\star}$ et $\boxed{\Gamma(n) = (n-1)\Gamma(n-1).}$

2. Dans la question précédente, nous avons montré que

$$\forall n \in \mathbf{N}^\star, \quad \Gamma(n+1) = n\Gamma(n).$$

On établit alors facilement par récurrence que $\Gamma(n) = (n-1)!\Gamma(1)$. Mais $\Gamma(1) = 1$ comme nous l'avons montré dans la question précédente. Donc

$$\boxed{\forall n \in \mathbf{N}^\star, \quad \Gamma(n) = (n-1)!}.$$

3. La fonction f_n est continue sauf peut-être en zéro, et f_n est positive puisque Γ apparaît comme l'intégrale d'une fonction positive pour tout n donc est positive. Et enfin, d'après le résultat des questions précédentes, $\int_0^1 t^{n-1}e^{-t}\,dt$ converge et vaut $\Gamma(n)$. Donc l'intégrale totale vaut par linéarité de l'intégrale $\frac{\Gamma(n)}{\Gamma(n)} = 1$. Donc $\boxed{f_n \text{ est bien une densité de probabilité.}}$
4. L'inégalité dérivée est la suivante, par linéarité de la dérivation et puisque X_n est à densité f_{n+1} :

$$\forall x > 0, \quad f_n(x) = -\sum_{k=0}^n \frac{d\left(e^{-x}\frac{x^k}{k!}\right)}{dx}.$$

Soit en utilisant la dérivée d'un produit :

$$\forall x > 0, \quad f_n(x) = -\sum_{k=0}^n e^{-x}\left(\frac{kx^{k-1}}{k!} - \frac{x^k}{k!}\right) = -e^{-x}\left(\sum_{k=1}^n\left(\frac{x^{k-1}}{(k-1)!} - \frac{x^k}{k!}\right) + (0-1)\right).$$

Donc par téléscopage et puisque $\Gamma(n+1) = n!$, on déduit l'égalité suivante à démontrer :

$$\forall x > 0, \quad f_n(x) = \mathrm{e}^{-x}\frac{x^n}{\Gamma(n+1)}.$$

On retrouve bien l'expression de f_n sur $\mathbf{R}^{+\star}$. Il ne reste plus qu'à primitiver pour revenir à l'égalité initiale, on déduit l'existence d'une constante $C \in \mathbf{R}$ de sorte que :

$$\forall x > 0, \quad \mathbf{P}(X_n \leqslant x) = 1 - \mathbf{P}(Y_x \leqslant n) + C.$$

Faisons $x \xrightarrow{x\longrightarrow 0}$: on obtient par continuité de la fonction de répartition,

$$\mathbf{P}(X_n \leqslant 0) = 0 = 1 - \lim_{x\longrightarrow 0}\mathbf{P}(Y_x \leqslant n) + C.$$

Mais

$$\mathbf{P}(Y_x \leqslant n) = \sum_{k=0}^{n}\mathrm{e}^{-x}\frac{x^k}{k!},$$

donc

$$\lim_{x\longrightarrow 0}\mathbf{P}(Y_x \leqslant n) = 1 + 0.$$

On trouve alors $C = 0$. La formule est donc établie :

$$\boxed{\forall x > 0, \quad \mathbf{P}(X_n \leq x) = 1 - \mathbf{P}(Y_x \leq n)}.$$

Exercice ALEA.15.20 | **Loi du χ^2 de degré 2** Soient X_1, X_2 deux variables aléatoires mutuellement indépendantes, suivant toutes la loi normale centrée réduite. On note

$$Z = \frac{1}{2}(X_1^2 + X_2^2).$$

1. Montrer que $\frac{1}{2}X_1^2$ suit une loi usuelle que l'on déterminera.
2. En déduire une densité de Z. *Le résultat fera intervenir une intégrale qu'on ne cherchera pas à calculer.*

Solution (exercice ALEA.15.20)

1. Commençons par déterminer la loi de X_1^2. Soit $x \in \mathbf{R}^+$, alors

$$\begin{aligned}\mathbf{P}\left(\frac{1}{2}X_1^2 \leqslant x\right) &= \mathbf{P}\left(|X_1| \leqslant \sqrt{2x}\right)\\ &= \Phi\left(\sqrt{2x}\right) - \Phi\left(-\sqrt{2x}\right)\\ &= 2\Phi\left(\sqrt{2x}\right) - 1.\end{aligned}$$

Et donc :

$$F_{X_1^2/2}(x) = \begin{cases} 2\Phi\left(\sqrt{2x}\right) - 1 & \text{si } x > 0,\\ 0 & \text{sinon.}\end{cases}$$

Donc comme Φ est continue, on a $2\Phi\left(\sqrt{2x}\right) - 1 \xrightarrow{x\longrightarrow 0} 2\Phi(0) - 1 = 1 - 1 = 0$, et Φ est $\mathscr{C}^1$ donc par composition, $F_{X_1^2/2}$ est continue et $\mathscr{C}^1$ sauf peut-être en zéro. Ainsi, $\frac{1}{2}X_1^2$ est à densité dont une densité est

$$f_{X_1^2/2} : x \longmapsto \begin{cases} = 2\varphi(\sqrt{2x})\frac{\mathrm{d}\left(\sqrt{2x}\right)}{\mathrm{d}x} & \text{si } x > 0,\\ 0 & \text{sinon.}\end{cases}.$$

Ce qui livre en utilisant l'expression de la densité de la $\mathscr{N}(0,1)$:

$$\boxed{f_{X_1^2/2} : x \longmapsto \begin{cases} \frac{1}{\sqrt{\pi x}}\mathrm{e}^{-x} & \text{si } x > 0,\\ 0 & \text{sinon.}\end{cases}}$$

2. Pour simplifier on notera $f = f_{X_1^2/2} = f_{X_2^2}$ puisque $\frac{X_1^2}{2}$ et $\frac{X_2^2}{2}$ ont même loi. Par lemme des coalitions, $\frac{1}{2}X_1^2 \perp\!\!\!\perp \frac{1}{2}X_2^2$. Donc Z admet une densité, donnée par

$$\forall z \in \mathbf{R}, \quad f_Z(z) = \int_{-\infty}^{\infty} f(x)f(z-x)\,\mathrm{d}x.$$

Par ailleurs, pour tout x réel, $f(x)f(z-x) \neq 0$ si et seulement si $x > 0$ et $z - x > 0$, soit encore si et seulement si : $x > 0, x < z$. Deux cas se présentent alors. Si $z < 0$, on a $f_Z(z) = 0$, reste ensuite las cas $z \geqslant 0$:

$$\begin{aligned} f_Z(z) &= \int_0^z \frac{1}{\sqrt{\pi x}} e^{-x} \frac{1}{\sqrt{\pi(z-x)}} e^{-(z-x)} \, dx \\ &= \frac{e^{-z}}{\pi} \int_0^z \frac{1}{\sqrt{x(x-z)}}. \end{aligned}$$

En résumé :

$$\boxed{\forall z \in \mathbf{R}, \quad f_Z(z) = \begin{cases} \frac{e^{-z}}{\pi} \int_0^z \frac{1}{\sqrt{x(x-z)}} & \text{si } z > 0, \\ 0 & \text{sinon.} \end{cases}}$$

Exercice ALEA.15.21 | **Matrice aléatoire uniforme** Soient X, Y deux variables aléatoires réelles à densité de loi $\mathscr{U}([0,1])$ et indépendantes. On note F_X et F_Y les fonctions de répartitions associées.

1. Montrer que $Z = X^2 - Y$ admet la fonction h ci-dessous pour densité :

$$\forall x \in \mathbf{R}, \quad h(x) = \begin{cases} \sqrt{x+1} & \text{si } -1 \leqslant x < 0, \\ 1 - \sqrt{x} & \text{si } 0 \leqslant x \leqslant 1, \\ 0 & \text{sinon.} \end{cases}$$

2. Déterminer la probabilité pour que la matrice aléatoire $M = \begin{pmatrix} 0 & -1 \\ Y & 2X \end{pmatrix}$ soit diagonalisable dans $\mathfrak{M}_2(\mathbf{R})$.

Solution (exercice ALEA.15.21)

1. On a par stabilité de la loi uniforme, que $-Y \hookrightarrow \mathscr{U}[-1,0]$. Cherchons à présent une densité de X^2. On a clairement $\mathbf{P}(X^2[0,1]) = 1$. Soit donc $x \in [0,1]$,

$$\begin{aligned} \mathbf{P}(X^2 \leqslant x) &= \mathbf{P}\left(-\sqrt{x} \leqslant X \leqslant \sqrt{x}\right) \\ &= F_X(\sqrt{x}) - F_X(-\sqrt{x}) \\ &= F_X(\sqrt{x}) = \sqrt{x}. \end{aligned}$$

car F_X est nulle sur $\mathbf{R}^-$

En résumé,

$$\forall x \in \mathbf{R}, \quad F_{X^2}(x) = \begin{cases} 1 & \text{si } x > 1, \\ \sqrt{x} & \text{si } x \in [0,1], \\ 0 & \text{sinon.} \end{cases}$$

Or $\sqrt{x} \xrightarrow{x \to 0} 0$, $\sqrt{x} \xrightarrow{x \to 1} 1$, la fonction F_{X^2} est continue. De plus, elle est $\mathscr{C}^1$ sauf peut-être en 0 et 1. Donc X^2 est à densité donnée par

$$\boxed{\forall x \in \mathbf{R}, \quad f_{X^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{si } x \in]0,1], \\ 0 & \text{sinon.} \end{cases}}$$

Puisque $X \perp\!\!\!\perp Y$ on a par coalitions : $X^2 \perp\!\!\!\perp -Y$. Donc $X^2 - Y$ est à densité, et une densité est

$$\forall z \in \mathbf{R}, \quad f_Z(z) = \int_{-\infty}^{\infty} f_{X^2}(x) f_{-Y}(z-x) \, dx.$$

Par ailleurs,

$$f_{X^2}(x) f_{-Y}(z-x) \neq 0 \iff x \in [0,1], z - x \in [-1,0],$$

ou de manière équivalente $0 \leqslant x \leqslant 1, z \leqslant x \leqslant z+1$. Plusieurs cas se présentent alors.

$$\forall z \in \mathbf{R}, \quad f_Z(z) = \begin{cases} \int_0^{z+1} \frac{1}{2\sqrt{x}} \, dx & \text{si } z+1 \in [0,1] \\ \int_z^1 \frac{1}{2\sqrt{x}} \, dx & \text{si } z \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$

Ce qui livre après calculs des intégrales :

$$\boxed{\forall z \in \mathbf{R}, \quad f_Z(z) = h(z) = \begin{cases} \sqrt{z+1} & \text{si } -1 \leqslant z < 0, \\ 1-\sqrt{z} & \text{si } 0 \leqslant z \leqslant 1, \\ 0 & \text{sinon.} \end{cases}}$$

2. Commençons par déterminer les valeurs propres de $M = \begin{pmatrix} 0 & -1 \\ Y & 2X \end{pmatrix}$. Soit $\lambda \in \mathbf{R}$, alors :

$$\det \begin{pmatrix} -\lambda & -1 \\ Y & 2X-\lambda \end{pmatrix} = -\lambda(2X-\lambda) + Y = \lambda^2 - 2X\lambda + Y.$$

Donc notons $\Delta = 4X^2 - 4Y = 4(X^2 - Y)$. Alors on a l'égalité d'évènements :

$$\{M \text{ diagonalisable sur } \mathfrak{M}_{2,2}(\mathbf{R})\} = \{\Delta > 0\},$$

puisque si $\Delta < 0$ elle n'a même pas de valeur propre réelle, et si $\Delta = 0$ elle aurait une unique valeur propre et donc serait de la forme XI_2 ce qui est exclut. Mais

$$\mathbf{P}(\Delta > 0) = \mathbf{P}\left(X^2 - Y > 0\right) = \int_0^1 \left(1-\sqrt{z}\right) = \left[z - \frac{z^{3/2}}{3/2}\right]_0^1 = \boxed{\frac{1}{3}}.$$

Exercice ALEA.15.22 | On considère deux variables aléatoires réelles X et Y définies sur le même espace probabilisé, indépendantes, et de même loi $\mathscr{E}(1)$.

1. Soit $t > 0$. Montrer que $Y - tX$ admet pour densité h définie par :

$$\forall x \in \mathbf{R}, \quad h(x) = \begin{cases} \frac{e^{-x}}{t+1} \text{ si } x > 0, \\ \frac{e^{x/t}}{t+1} \text{ si } x \leq 0. \end{cases}$$

2. En déduire une densité de $Z = \frac{Y}{X}$.

Solution (exercice ALEA.15.22)

1. Soit $t > 0$. Alors $\mathbf{P}(-tX \leqslant 0) = 1$, donc pour tout $x \in \mathbf{R}^{-\star}$,

$$\begin{aligned} \mathbf{P}(-tX \leqslant x) = \mathbf{P}\left(X \geqslant -\frac{x}{t}\right) &= 1 - F_X\left(-\frac{x}{t}\right) \\ &= \begin{cases} 1 - \left(1 - e^{\frac{x}{t}}\right) & \text{si } -\frac{x}{t} \geqslant 0, \\ 1 - 0 & \text{sinon,} \end{cases} \\ &= \begin{cases} e^{\frac{x}{t}} & \text{si } x \leqslant 0, \\ 1 & \text{sinon.} \end{cases} \end{aligned}$$

Puisque $e^{\frac{x}{t}} \xrightarrow{x \longrightarrow 0} 1$, la fonction F_{-tX} est continue par composition, de plus elle est $\mathscr{C}^1$ sauf peut-être en zéro. Donc $-tX$ est à densité et une densité est donnée par :

$$\boxed{\forall x \in \mathbf{R}, \quad f_{-tX}(x) = \begin{cases} \frac{1}{t} e^{\frac{x}{t}} & \text{si } x \leqslant 0, \\ 0 & \text{sinon.} \end{cases}}$$

Par lemme des coalitions, $Y \perp\!\!\!\perp -tX$. Donc $Y - tX$ est à densité et admet pour densité h définie par :

$$\forall x \in \mathbf{R}, \quad h(x) = \int_{-\infty}^{\infty} f_Y(z) f_{-tX}(x-z)\, dz.$$

Mais $f_Y(z) f_{-tX}(x-z) \neq 0$ si et seulement si $z \geqslant 0, x - z \leqslant 0$ soit $z \geqslant 0, x \leqslant z$. Deux cas se présentent alors.

$$\begin{aligned} h(x) &= \begin{cases} \int_x^{\infty} e^{-z} \frac{1}{t} e^{\frac{x-z}{t}}\, dt \text{ si } x > 0, \\ \int_0^{\infty} e^{-z} \frac{1}{t} e^{\frac{x-z}{t}}\, dz \text{ si } x \leq 0. \end{cases} \\ &= \boxed{\begin{cases} \frac{e^{-x}}{t+1} \text{ si } x > 0, \\ \frac{e^{x/t}}{t+1} \text{ si } x \leq 0. \end{cases}} \end{aligned}$$

2. Calculons les fonctions de répartition des deux variables aléatoires. Soit $t \in \mathbf{R}^{+\star}$, alors puisque X prend des valeurs positives, on a :

$$\begin{aligned}
\mathbf{P}(\mathrm{Z} \leqslant t) &= \mathbf{P}(\mathrm{Y} \leqslant t\mathrm{X}) \\
&= \mathbf{P}(\mathrm{Y} - t\mathrm{X} \leqslant 0) \\
&= \int_{-\infty}^{0} h(x)\,\mathrm{d}x = \int_{-\infty}^{0} \frac{\mathrm{e}^{x/t}}{t+1}\,\mathrm{d}x \\
&= \frac{t}{t+1}\left[\mathrm{e}^{x/t}\right]_{-\infty}^{0} \\
&= \frac{t}{t+1}.
\end{aligned}$$

Par ailleurs, pour $t \leqslant 0$, on a évidemment $\mathbf{P}(\mathrm{Z} \leqslant t) = 0$ puisque Z prend ses valeurs dans $\mathbf{R}^{+}$. De plus, puisque $\frac{t}{t+1} \xrightarrow{t \longrightarrow 0} 0$, la fonction F_{Z} est donc continue sur $\mathbf{R}$ et $\mathscr{C}^1$ sauf éventuellement en zéro. Donc Z est à densité, et une densité est donnée par

$$\boxed{f_{\mathrm{Z}} : t \longmapsto \begin{cases} \frac{1}{(t+1)^2} & \text{si } t \geqslant 0, \\ 0 & \text{sinon.} \end{cases}}$$