Chapitre 36

Variables aléatoires (cas Ω fini)

Résumé -

La seconde grande idée du calcul des probabilités (après le conditionnement) est la notion de variables aléatoires. Il existe plusieurs façons de comprendre ce genre de variables. La manière rigoureuse est celle d'une application réciproque ou d'une classe d'équivalence puisqu'elle conduit à une partition de Ω, dont un système de représentant est $X(\Omega)$. Mais ce n'est pas la manière intuitive du physicien ou du probabiliste qui voit la ***variable aléatoire comme un nombre potentiel****. Ce nombre est multiple et prend telle ou telle valeur en fonction des événements de l'univers (on tient compte alors de la probabilité de réalisation).*
C'est typiquement le résultat chiffré d'une expérience de physique, dont la réalisation une seconde fois montre une certaine variabilité...
La question qui se pose, une fois que l'on maîtrise bien cette notion de variables aléatoires est celle de l'addition, multiplication (et autres opérations...) de variables aléatoires. Il faut dans ce cas, toujours commencer par étudier le couple de variables aléatoires dont on veut faire l'addition!

Sommaire

1. Problèmes

? Problème 158 - Variable aléatoire

Une variable peut prendre plusieurs valeurs, potentiellement, chacune avec une probabilité (a priori, ou non) fixé. Cette variable particulière comme par exemple la note à un devoir : au hasard elle est comprise entre 0 et 20 s'appelle une variable aléatoire.
On peut alors associer à chaque valeur, une probabilité, mais plus exactement un événement de l'univers. Donc les variables aléatoires partitionnent l'univers. On peut donc créer une relation d'équivalence $\asymp$: $\omega \asymp \omega'$, si pour ces deux issues de l'univers Ω, la variable aléatoire prend la même valeur.
Mais on a vu que dans ce cas là, une autre stratégie consiste à associer à la relation $\asymp$, une fonction f sur Ω, telle que : $\omega \asymp \omega' \Longleftrightarrow f(\omega) = f(\omega')$.
Quelle définition (mathématique, formelle) de variable aléatoire est la plus naturelle et pratique pour définir un nombre potentiel ?

? Problème 159 - Lois de probabilité fréquemment rencontrées

A une variable aléatoire, on associe une distribution de probabilité.
Quelles sont les distributions les plus courantes ? A quels types d'expériences aléatoires sont-elles associées ?

? Problème 160 - Espérance, variance...

Souvent, on ne peut se contenter d'une distribution de probabilité pour exprimer le résultat numérique d'une expérience. Il faut pouvoir résumer cette variable aléatoire ?
Qu'est-ce que l'espérance mathématique d'une variable aléatoire ? Pourquoi résume-t-elle d'une certaine façon la valeur de l'expérience ? Quelles sont les autres nombres à ajouter ?

? Problème 161 - Citation de Poincaré

Que pensez-vous de la citation de Henri Poincaré (Calcul des probabilités) :
« Vous me demandez de vous prédire les phénomènes qui vont se produire. Si, par malheur, je connaissais les lois de ces phénomènes, je ne pourrais y arriver que par des calculs inextricables et je devrais renoncer à vous répondre ; mais, comme j'ai la chance de les ignorer, je vais vous répondre tout de suite. Et, ce qu'il y a de plus extraordinaire, c'est que ma réponse sera juste. »

? Problème 162 - Deux variables aléatoires

Si l'expérience aléatoire conduit à la construction de deux (ou plus) variables aléatoires. Comment faire pour les étudier ensemble ?
On peut avoir à les considérer comme une base d'autres résultats possibles (d'autres variables aléatoires).

? Problème 163 - Produit scalaire

Ou bien, la question se pose parfois de mesurer la corrélation entre ces deux variables aléatoires. Sont-elles liées ? Par exemple : y a-t-il un lien entre fumer et développer un cancer ? Entre ne pas porter un masque et être touché par la Covid ?
On crée en probabilité-statistique un objet appelé coefficient de corrélation entre deux variables aléatoires : $\rho(X,Y)$.
Il mesure si X et Y sont proches : $\rho(X,Y)=1$, indépendants (ou quasiment) : $\rho(X,Y)=0$ ou sont plutôt opposées : $\rho(X,Y)=-1$.
Cela nous rappelle le produit scalaire (vu en physique et plus tard en maths) $\vec{x}\cdot\vec{y}$ qui mesure d'une certaine façon le lien entre $\vec{x}$ et $\vec{y}$ (modulo leurs normes). Y a-t-il un rapport entre la corrélation et la notion de produit scalaire.

2. Variable aléatoire

2.1. Quelques définitions

Soit Ω un univers fini lié à une expérience aléatoire.

Définition - Variable aléatoire
On appelle variable aléatoire (v.a.) sur Ω toute application $X:\Omega\to E$, où E est un ensemble. Dans le cas où $E\subset\mathbb{R}$, on parle de variable aléatoire réelle.
$X(\Omega)$ désigne donc l'ensemble image, c'est-à-dire les valeurs que prend l'application X. Cet ensemble est ici fini (car Ω fini), on dit alors que X est une v.a. discrète finie.

Pour aller plus loin - Variable alétoire sur un espace de probabilité infini
Lorsque la tribu définie sur Ω pour $\mathbf{P}$ *est plus compliqué que $\mathscr{P}(\Omega)$; pour créer une variable aléatoire réelle, il faut exiger en fait que pour tout I intervalle de $\mathbb{R}$, $X^{-1}(I)=\{\omega\in\Omega\mid X(\omega)\in I\}$ est un élément de la tribu.*
Cela, afin de pouvoir écrire et calculer : $\mathbf{P}(X^{-1}(I))$, *pour tout I raisonnable*

La stabilité linéaires des applications permet d'affirmer :

Proposition - Stabilité linéaire
Soient X,Y deux variables aléatoires réelles sur Ω, $\lambda\in\mathbb{R}$. Alors $X+Y$, XY et λX sont des variables aléatoires sur Ω.

Proposition - Composition
Soit X une v.a. sur Ω et g une application définie sur $X(\Omega)$, à valeurs dans un ensemble E', alors $g\circ X$ est une v.a. sur Ω, notée $g(X)$:

$$g(X):\begin{array}{ll}\Omega & \to E' \\ \omega & \mapsto g(X(\omega))\end{array}$$

Démonstration

Définition - variable aléatoire constante ou certaine
Si X est une application constante sur Ω, on dit que c'est une variable aléatoire constante ou certaine.

Définition - Notations

Soit $X : \Omega \to E$ une variable aléatoire sur Ω.
Alors, pour $A \subset E$, $X^{-1}(A)$ est un événement (car $X^{-1}(A) \in \mathscr{P}(\Omega)$) noté

$$(X \in A) \text{ ou } \{X \in A\} \quad (= X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}).$$

Dans le cas d'une variable aléatoire réelle, pour $x \in \mathbb{R}$, on note :

$$(X \leqslant x) = X^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid X(\omega \leqslant x\},$$
$$(X < x) = X^{-1}(]-\infty, x[),$$
$$(X = x) = X^{-1}(\{x\})$$

de même pour $(X \geqslant x), (X > x), (a \leqslant X \leqslant b)\ldots$

Heuristique - Deux points de vue sur X^{-1}

De manière générale, ce que l'on a c'est une famille d'événement paramétrée par une variable réelle et pour laquelle on cherche des probabilités de réalisation.
Par exemple, on s'intéresse à l'évolution de la température (moyenne ou sur un point du globe) sur 10 ans.
On note H, la variable qui indique cette évolution. On sait que $H \in \mathbb{R}$ (même si toutes les valeurs réelles ne sont pas réalistes).
Ce que l'on cherche alors est : la probabilité d'une hausse supérieure à 2°

$$\mathbf{P}(H \geqslant 2)$$

Il faut donc que l'ensemble $H^{-1}([2, +\infty[)$ soit un événement mesurable en probabilité.
Donc ce qui nous intéresse, dans la pratique, c'est plutôt $H^{-1}(I)$, comme élément de Ω et dont on cherche une probabilité.
On fera bien attention à la manière de lire les expressions mathématiques du type $\mathbf{P}(H = 0)$
On a alors deux points de vue

- H est une application et H^{-1} est une application réciproque, en règle générale non bijective, donc une application de l'ensemble des parties de E (ou de $\mathbb{R}$ pour une var) sur l'ensemble des parties de Ω (ou une tribu de Ω).
  C'est le point de vue choisi dans le cours.
- H^{-1} fait la partition de Ω en classes d'équivalence.

  $$\omega \mathscr{R}_H \omega' \iff H(\omega) = H(\omega')$$

  De ce point de vue, le théorème suivant sur le système complet d'événements est trivial.

2.2. Loi de probabilité

On se place désormais sur un espace probabilisé fini $(\Omega, \mathbf{P})$.

Définition - Loi (de probabilité) d'une variable aléatoire

Soit X une v.a. sur Ω. Alors l'application

$$\mathbf{P}_X : \begin{array}{ll} X(\Omega) & \to [0,1] \\ x & \mapsto \mathbf{P}(X = x) \end{array}$$

s'appelle la loi (de probabilité) de X.

Remarque - Changement d'espace

En fait X transforme l'espace de probabilité $(\Omega, \mathbf{P})$ en l'espace $(X(\Omega), \mathbf{P}_X)$.
A proprement parlé, il faudrait ajouter les tribus pour parler d'espace de probabilité. Mais ici Ω est fini, donc, cette omission n'est pas grave.

Savoir faire - Définir la loi d'une variable aléatoire

Définir la loi de probabilité d'une v.a. X finie, c'est donc donner $X(\Omega)$ ainsi que les probabilités $\mathbf{P}(X=x)$ pour $x \in X(\Omega)$.

Remarque - Notation

Deux petites remarques :

— On peut noter f_X à la place de $\mathbf{P}_X$ pour éviter toute confusion avec une probabilité conditionnelle.

— On note usuellement $\mathbf{P}(X \in A)$ la probabilité de l'événement $(X \in A)$ et non $\mathbf{P}\Big((X \in A)\Big)$ ou $\mathbf{P}(\{X \in A\})$.

Définition - Relation d'équivalence sur les variables aléatoires

Soient X et Y deux variables aléatoires définies sur un même univers Ω.
On dit que X suit la même loi que Y, et on note $X \sim Y$, si $\mathbf{P}_X = \mathbf{P}_Y$.
Il s'agit d'une relation d'équivalence sur les variables aléatoires.

Attention - Les variables X et Y peuvent être différentes!

Pour une pièce bien équilibrée si X indique le nombre de pile pour n lancers et Y le nombre de face.
Alors $X \sim Y$, alors que $Y = n - X$ (et donc $X \neq Y$).

Démonstration

Exercice

Montrer que si $X \sim Y$, alors pour tout $f : \mathbb{R} \to \mathbb{R}$, $f(X) \sim f(Y)$.
La réciproque est-elle vraie ?

Proposition - Variable aléatoire et système complet d'événements

Soit X une variable aléatoire définie sur Ω.
Alors la famille $\Big((X = x)\Big)_{x \in X(\Omega)}$ est un système complet d'événements, appelé système complet d'événements associé à X. En particulier,

$$\sum_{x \in X(\Omega)} \mathbf{P}(X = x) = 1.$$

Démonstration

Proposition - Existence d'une v.a.
Soient E un ensemble fini, $E = \{x_1, \dots, x_n\}$ et $p_1, \dots, p_n$ des réels positifs tels que $\sum_{i=1}^{n} p_i = 1$.
Alors, si Ω est un ensemble fini tel que $\operatorname{Card}\Omega \geqslant n$,
il existe une probabilité $\mathbf{P}$ sur Ω et une v.a. X définie sur Ω vérifiant :
$\forall i \in [\![1, n]\!], \mathbf{P}(X = x_i) = p_i$

Démonstration

Proposition - Loi d'une fonction de X
Soient X v.a. sur Ω et g une fonction définie sur $X(\Omega)$.
Alors la loi de probabilité de $Y = g(X)$ est donnée par $Y(\Omega) = g(X(\Omega))$ et

$$\forall y \in Y(\Omega), \mathbf{P}(Y = y) = \sum_{x \in g^{-1}(\{y\})} \mathbf{P}(X = x) = \sum_{x \mid g(x) = y} \mathbf{P}(X = x)$$

Démonstration

Définition - Fonction de répartition
Soit X une v.a.r sur Ω. L'application

$$F_X : \begin{array}{ll} \mathbb{R} & \to [0,1] \\ x & \mapsto \mathbf{P}(X \leqslant x) \end{array}$$

s'appelle la fonction de répartition de X ; si $X(\Omega) = \{x_1, x_2, \dots, x_n\}$, alors

$$\forall x \in \mathbb{R}, F_X(x) = \sum_{i \mid x_i \leqslant x} \mathbf{P}(X = x_i).$$

Exemple - Fonction de répartition du max

Exercice

On lance deux dés à six faces parfaitement équilibrés. Soit X la variable aléatoire égale à la somme des points obtenus. Donner la loi de X, sa fonction de répartition ainsi que la loi de $Y = |X-7|$.

2.3. Lois usuelles

$(\Omega, \mathbf{P})$ désigne un espace probabilisé fini.

Loi uniforme

Définition - Loi uniforme
On dit qu'une v.a. X suit une loi uniforme sur $X(\Omega)$ si

$$\forall x \in X(\Omega), \mathbf{P}(X=x) = \frac{1}{\text{Card}\,(X(\Omega))}.$$

Dans le cas particulier où $X(\Omega) = [\![1,n]\!]$, on note $X \hookrightarrow \mathscr{U}_n$ et on a

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

Remarque - Modèle pour la loi uniforme $\mathscr{U}_n$
On choisit un objet au hasard parmi n objets numérotés de 1 à n. X désigne le numéro de l'objet tiré. (ex : lancer d'un dé non truqué)

Loi de Bernoulli

Définition - Loi de Bernoulli
Soit $p \in [0,1]$. On dit qu'une v.a.r. X suit une loi de Bernoulli de paramètre p si $X(\Omega) = \{0,1\}$ et $\mathbf{P}(X=1) = p$, et donc $\mathbf{P}(X=0) = 1-p$.
On note $X \hookrightarrow \mathscr{B}(p)$.

Remarque - Modèle pour la loi de Bernoulli $\mathscr{B}(p)$
Lors d'une expérience aléatoire effectuée une seule fois et ayant deux issues possibles, on note X la v.a. qui vaut 1 si l'on a l'une des issues (succès) et 0 si l'on a l'autre (échec).
(ex : on lance une pièce de monnaie, le succès désignant le fait d'obtenir "pile".)
X est alors la fonction indicatrice de l'événement A "avoir un succès", usuellement notée $\mathbb{1}_A$, définie, pour tout $\omega \in \Omega$ par $\mathbb{1}_A(\omega) = 1$ si $\omega \in A$, $\mathbb{1}_A(\omega) = 0$ si $\omega \notin A$ et on a

$$\mathbb{1}_A \hookrightarrow \mathscr{B}(\mathbf{P}(A)).$$

Savoir faire - Exploitation d'indicatrice (1)
Très souvent, on associera à l'événement $A \subset \Omega$, la variable aléatoire $X = \mathbb{1}_A : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$.
Dans ce cas $X \hookrightarrow \mathscr{B}(p)$ avec $p = \mathbf{P}(A)$

Loi binomiale

Définition - Loi binomiale
Soient $n \in \mathbb{N}^*$ et $p \in [0,1]$. On dit qu'une v.a.r. suit une loi binomiale de paramètres n et p si $X(\Omega) = [\![0,n]\!]$ et

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

On note $X \hookrightarrow \mathscr{B}(n,p)$.

Remarque - Modèle pour la loi binomiale $\mathscr{B}(n,p)$
On considère une urne avec deux catégories de boules : des blanches en proportion p et des non blanches en proportion $q = 1-p$. On tire n boules **avec remise** de cette urne. X désigne la v.a. égale au nombre de boules blanches obtenues dans ce tirage.
Plus généralement, c'est la loi de la variable aléatoire X égale au nombre de succès lors de la répétition de n expériences de Bernoulli indépendantes, la probabilité d'un succès étant p.

3. Couples de variables aléatoires

Soit $(\Omega, \mathbf{P})$ un espace probabilisé fini. On a déjà rencontré cette situation dans le cas du lancer de dés.

Définition - Couple de variables aléatoires
Soient X et Y deux variables aléatoires sur Ω, à valeurs respectivement dans les ensembles E et F. L'application

$$Z : \begin{array}{ll} \Omega & \to E \times F \\ \omega & \mapsto (X(\omega), Y(\omega)) \end{array}$$

est appelée couple de variables aléatoires, c'est donc une v.a. à valeurs dans $E \times F$. On note $Z = (X,Y)$.
On a $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$.
Si X et Y sont des v.a.r. on parle de couple de v.a.r.

Définition - Loi conjointe, lois marginales
On appelle loi conjointe de X et Y la loi du couple $Z = (X, Y)$, c'est-à-dire l'application

$$\mathbf{P}_{(X,Y)} : \begin{array}{ll} Z(\Omega) & \longrightarrow [0,1] \\ (x,y) & \longmapsto \mathbf{P}\Big((X,Y) = (x,y)\Big) = \mathbf{P}\Big((X = x) \cap (Y = y)\Big) \end{array}$$

On appelle lois marginales du couple (X, Y) les deux lois de probabilité, respectivement de X et Y.
On a $\mathbf{P}_X : x \mapsto \sum\limits_{y \in Y(\Omega)} \mathbf{P}\big((X,Y) = (x,y)\big) = \mathbf{P}(X = x)$ (et de même pour P_Y).

Attention - Lien entre les deux lois
A partir de la loi conjonte d'un couple de v.a. on peut déterminer les lois marginales. La réciproque est fausse.

Définition - Loi conditionnelle
Soit (X, Y) un couple de variables aléatoires.
Pour $x \in X(\Omega)$ fixé tel que $\mathbf{P}(X = x) \neq 0$, on appelle loi conditionnelle de Y sachant $(X = x)$ la loi de Y pour la probabilité $\mathbf{P}_{(X=x)}$, c'est-à-dire l'application :

$$\begin{array}{ll} Y(\Omega) & \longrightarrow [0,1] \\ y & \longmapsto \mathbf{P}(Y = y | X = x) \end{array}$$

De même on peut définir, pour $y \in Y(\Omega)$ tel que $\mathbf{P}(Y = y) \neq 0$, la loi conditionnelle de X sachant $(Y = y)$.

Exercice
On lance trois fois de suite une pièce équilibrée. Soient X_1 la v.a. qui vaut 1 si le premier jet donne "pile" et 0 sinon, et X_2 la v.a. égale au nombre de "face" obtenu.
Déterminer la loi conjointe ainsi que les lois marginales du couple (X_1, X_2).
Déterminer la loi conditionnelle de X_1 sachant $(X_2 = 1)$.

Exercice
Un sac contient 4 boules numérotées de 1 à 4. On tire deux boules avec remise et on note X_1 et X_2 les numéros obtenus. On pose $X = X_1$ et $Y = \max(X_1, X_2)$.
Déterminer la loi conjointe ainsi que les lois marginales du couple (X_1, X_2).
Déterminer la loi conjointe ainsi que les lois marginales du couple (X, Y).
Déterminer la loi conditionnelle de Y sachant $(X = 2)$

Proposition - Caractérisation de loi de couple de v.a.
Soient E et F deux ensembles.

$$\left\{\Big((x_i, y_j), p_{ij}\Big) \in (E \times F) \times \mathbb{R} \,\Big|\, 1 \leqslant i \leqslant r, 1 \leqslant j \leqslant s\right\}$$

représente la loi de probabilité d'un couple de v.a. si et seulement si

$$\forall (i,j) \in [\![1,r]\!] \times [\![1,s]\!], p_{ij} \geqslant 0 \text{ et } \sum_{i=1}^{r} \sum_{j=1}^{s} p_{ij} = 1.$$

Démonstration

Les définitions précédentes peuvent se généraliser à des n-uplets de v.a.

Définition - Vecteur aléatoire
Soient $n \in \mathbb{N}^*$, $E_1, \ldots, E_n$ des ensembles et $X_1, \ldots, X_n$ n v.a. sur Ω. La variable aléatoire à valeurs dans $E_1 \times \cdots \times E_n$

$$X : \begin{array}{ll} \Omega & \to E \times \cdots \times E_n \\ \omega & \mapsto (X_1(\omega), \ldots, X_n(\omega)) \end{array}$$

est appelée n-uplet de variables aléatoires. On note $X = (X_1, \ldots, X_n)$.
On a $X(\Omega) \subset X_1(\Omega) \times \cdots \times X_n(\Omega)$.
Si les X_i sont des v.a.r. on parle de n-uplet de v.a.r. ou de vecteur aléatoire à valeur dans $\mathbb{R}^n$.
La loi conjointe de $X_1, \ldots, X_n$ est la loi du n-uplet $X = (X_1, \ldots, X_n)$ Les lois marginales du n-uplet $X = (X_1, \ldots, X_n)$ sont les n lois de probabilité des v.a. $X_1, \ldots, X_n$, elles peuvent se déduire de la loi conjointe.

4. Indépendance

4.1. Indépendance de deux variables aléatoires

Définition - Indépendance de deux v.a.
Deux variables aléatoires discrètes X et Y sur l'espace probabilisable $(\Omega, \mathbf{P})$ sont dites indépendantes si, pour tout couple (x, y) de $X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants, c'est-à-dire si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \mathbf{P}(\{X = x\} \cap \{Y = y\}) = \mathbf{P}(X = x)\mathbf{P}(Y = y).$$

On notera (en MPSI3) : $X \perp\!\!\!\perp Y$ pour indiquer que X et Y sont indépendantes

C'est une relation symétrique, elle n'est ni réflexive, ni transitive.

Proposition - Caractérisation d'indépendance
Deux v.a. X et Y sur $(\Omega, \mathbf{P})$ sont indépendantes si et seulement si,
pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants.

Démonstration

Théorème - de Coalition
Soient X et Y deux variables aléatoires sur $(\Omega, \mathbf{P})$, indépendantes.
Alors, pour toutes fonctions f et g, les v.a. (si elles sont bien définies) $f(X)$ et $g(Y)$ sont indépendantes.

Démonstration

4.2. Indépendance de plusieurs variables aléatoires

Définition - Indépendance mutuelle
Soient n variables aléatoires $X_1, \ldots, X_n$ définies sur (Ω, P).
On dit que $X_1, \ldots, X_n$ sont mutuellement indépendantes (ou indépendantes) si pour tout $(x_1, x_2, \ldots, x_n) \in X_1(\Omega) \times X_2(\Omega) \times \ldots \times X_n(\Omega)$, on a

$$\mathbf{P}(\bigcap_{i=1}^{n} \{X_i = x_i\}) = \prod_{i=1}^{n} \mathbf{P}(X_i = x_i)$$

c'est-à-dire si les événements $\Big((X_i = x_i)\Big)_{1 \leqslant i \leqslant n}$ sont mutuellement indépendants.

Proposition - Caractérisation de l'indépendance mutuelle
Soient n variables aléatoires $X_1, \ldots, X_n$ définies sur (Ω, P).
$X_1, \ldots, X_n$ sont mutuellement indépendantes si et seulement si pour tout $(A_1, \ldots, A_n) \subset X_1(\Omega) \times \cdots \times X_n(\Omega)$, les événements $(X_i \in A_i)$ sont mutuellement indépendants.

Démonstration

Proposition - Indépendance mutuelle implique l'indépendance deux à deux
n variables aléatoires $X_1,\dots,X_n$ sur $(\Omega,\mathbf{P})$ mutuellement indépendantes sont indépendantes deux à deux.

Démonstration

Théorème - Coalitions (n vari ables)
Soient $X_1,\dots,X_n$ n variables aléatoires définies sur (Ω,P) mutuellement indépendantes. Alors :

— pour toutes fonctions $g_1,\dots,g_n$, les v.a. (si elles sont bien définies) $g_1(X_1),\dots,g_n(X_n)$ sont mutuellement indépendantes;

— pour toutes fonctions φ et ψ, les deux variables aléatoires (si elles sont définies) $\varphi(X_1,X_2,\dots,X_p)$ et $\psi(X_{p+1},\dots,X_n)$ sont indépendantes;

— plus généralement si $Y_1,\dots,Y_k$ sont k variables aléatoires telles que Y_i soit fonction des X_j pour $j\in J_i$, avec les ensembles J_i deux à deux disjoints, alors $Y_1,\dots,Y_k$ sont mutuellement indépendantes.

On ne fait pas la première démonstration, il suffit d'apter le cas de deux variables.
On ne fait pas non plus la dernière démonstration, très pénible dans les notations.

Démonstration

Exemple - Une pièce qui n'en finit pas d'être tirée

4.3. Opération de variables aléatoires (indépendantes)

Si E est un ensemble muni d'une loi (E : espace vectoriel, par exemple), on peut faire agir la même loi sur des variables aléatoires X_i à valeurs dans E.

Théorème - Stabilité de la loi binomiale pour la somme
Soit $(\Omega, \mathbf{P})$ un espace de probabilité fini.

— Si X et Y sont deux v.a.r. indépendantes sur (Ω, P) telles que $X \hookrightarrow \mathscr{B}(n, p)$ et $Y \hookrightarrow \mathscr{B}(m, p)$, alors

$$X + Y \hookrightarrow \mathscr{B}(n + m, p).$$

— Si $X_1, \dots, X_n$ sont des v.a.r. mutuellement indépendantes sur (Ω, P) telles que pour tout i,
$X_i \hookrightarrow \mathscr{B}(n_i, p)$, alors

$$\sum_{i=1}^{n} X_i \hookrightarrow \mathscr{B}\Big(\sum_{i=1}^{n} n_i, p\Big).$$

La loi de Bernoulli étant un cas particulier de la loi binomiale ($n = 1$), on a

Corollaire - Addition de Bernoulli
Si $X_1, \dots, X_n$ sont des v.a.r. mutuellement indépendantes sur $(\Omega, \mathbf{P})$, toutes de loi $\mathscr{B}(p)$, alors

$$\sum_{i=1}^{n} X_i \hookrightarrow \mathscr{B}(n, p).$$

Attention - Le même coefficient p
On soulignera qu'il faut que ce soit la même probabilité p en paramètre aux lois additionnées

Démonstration

Savoir faire - Modélisation

Une suite finie $(X_i)_{1\leqslant i\leqslant n}$ de variables aléatoires indépendantes fournit un moyen de modéliser une succession de n épreuves dont les résultats sont indépendants, en particulier les répétitions indépendantes d'une même épreuve se modélisent par la donnée de n v.a. indépendantes **équidistribuées** (c'est-à-dire de même loi).
Une suite de n lancers de pile ou face aux résultats indépendants se modélise par la donnée de n v.a. indépendantes $X_1,\dots,X_n$ de loi $\mathscr{B}(p)$, X_i étant le résultat du i-ième lancer.

Pour aller plus loin - Comment redémontrer la formule de Vandermonde

Deux stratégies :

— *par combinatoire.*
On décompose un ensemble de $m+n$ éléments en deux sous ensembles : un de n et l'autre de m éléments.
Et on calcule de deux façons différentes le nombre de sous-ensembles à k éléments.

— *avec les polynômes.*
On développe, avec la formule du binôme de Newton, de deux façons différentes $(1+x)^n \times (1+x)^m$, puis on identifie le coefficient devant x^k.

Exercice

Soit $(X_1,\dots,X_n)$ un n-uplet de v.a.r. indépendantes de même loi $\mathscr{B}(100,\frac{1}{5})$.
On note N la variable aléatoire égale à Card $\{k\in[\![1,n]\!]\mid X_k=99\}$. Déterminer la loi de N.

Exercice

Une urne contient n boules indiscernables au toucher numérotées de 1 à n.
On prélève deux boules successivement et on note X_1 (resp. X_2) la v.a. égale au numéro de la première (resp. deuxième) boule tirée. On pose $Z=\max(X_1,X_2)$.
Déterminer la loi conjointe du couple (X_1,X_2), les lois marginales, ainsi que la loi de Z dans les deux cas suivants :

— le tirage se fait avec remise ;
— le tirage se fait sans remise.

Truc & Astuce pour le calcul - Etude du $\max(X_i)$, où les X_i sont indépendantes

On note $Z=\max(X_1,X_2,\dots X_n)$. En deux temps :

1. On constate que $Z\leqslant k \Longleftrightarrow \forall\ i\leqslant n, X_i\leqslant k$.
   Par indépendance : $\mathbf{P}(Z\leqslant k)=\prod_{i=1}^{n}\mathbf{P}(X_i\leqslant k)$
2. Revenir au cas $Z=k$:
   $[Z\leqslant k]=[Z=k]\cup[Z\leqslant k-1]$, événements incompatibles,
   Donc $\mathbf{P}(Z\leqslant k)=\mathbf{P}(Z=k)+\mathbf{P}(Z\leqslant k-1)$,
   et $\mathbf{P}(Z=k)=\mathbf{P}(Z\leqslant k)-\mathbf{P}(Z\leqslant k-1)$

Exercice

Retrouver la loi de Z dans l'exercice précédent, en exploitant le savoir-faire

4.4. Suite de variables aléatoires

Définition - Indépendance pour une suite de v.a.
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur $(\Omega,\mathbf{P})$.
$(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires mutuellement indépendantes si, pour tout $n\in\mathbb{N}^*$, les variables aléatoires $X_1,\dots,X_n$ sont mutuellement indépendantes.

5. Moments d'une variable aléatoire réelle finie

$(\Omega,\mathbf{P})$ désigne un espace probabilisé fini.

5.1. Espérance

Définition - Espérance. Loi centrée
Soit X une v.a.réelle, $X(\Omega)=\{x_1,\dots,x_n\}$.
On appelle **espérance** de X le réel noté $\mathbf{E}(X)$ défini par :

$$\mathbf{E}(X)=\sum_{k=1}^{n} x_k\mathbf{P}(X=x_k)=\sum_{x\in X(\Omega)} x\mathbf{P}(X=x)$$

X est dite **centrée** si $\mathbf{E}(X)=0$.

Proposition - Espérance des lois usuelles
Soit $(\Omega,\mathbf{P})$ un espace de probabilité fini
— si X est constante égale à a alors $\mathbf{E}(X)=a$;
— si $X\hookrightarrow\mathscr{B}(p)$ alors $\mathbf{E}(X)=p$;
— si $X\hookrightarrow\mathscr{B}(n,p)$ alors $\mathbf{E}(X)=np$.

Démonstration

Pour aller plus loin - $\displaystyle\sum\!\!\!\!\!\!\int_x$
Selon la nature de $X(\omega)$, le calcul de l'espérance sera une somme (potentiellement infinie) si $X(\Omega)$ est dénombrable ou une intégrale si $X(\Omega)$ à la puissance du continue.
Pour ce dernier cas (hors-programme des CPGE), on associe à la variable aléatoire une densité f_X de probabilité telle que $\mathbf{P}(X\leqslant x)=\int_{-\infty}^{x} f_(t)\mathrm{d}t$.
On a alors $\mathbf{E}(X)=\int_{-\infty}^{+\infty} x f_X(t)\mathrm{d}t$ (généralisée à tous les moments).
L'intégrale selon STIELJES *peut aider à unifier tous ces cas ; on pourrait la noter* $\displaystyle\sum\!\!\!\!\!\!\int_x$

Savoir faire - Exploitation d'indicatrice (2)
Si $A\subset\Omega$ est un événement, alors $\mathbb{1}_A\hookrightarrow\mathscr{B}(\mathbf{P}(A))$ et donc $\mathbf{E}(\mathbb{1}_A)=\mathbf{P}(A)$.
On exploitera cette relation associée à des propriétés essentielles de l'espérance (linéarité...)

Proposition - Formulation équivalente
Soit X une v.a.r. On a
$$\mathbf{E}(X)=\sum_{\omega\in\Omega} X(\omega)\mathbf{P}(\{\omega\}).$$

On peut aussi noter que $X = \sum_{i=1}^{p} x_i \mathbb{1}_{[X=x_i]}$.

Démonstration

Théorème - Formule de transfert
Soient X une v.a. définie sur Ω, à valeurs dans un ensemble E et $g : X(\Omega) = \{x_1, \dots, x_n\} \to \mathbb{R}$.
Alors l'espérance de la v.a.r. $Z = g(X)$ est donnée par la formule :

$$\mathbf{E}(Z) = \sum_{k=1}^{n} g(x_k)\mathbf{P}(X = x_k) = \sum_{x \in X(\Omega)} g(x)\mathbf{P}(X = x)$$

C'est-à-dire que l'on n'a pas besoin de connaître la loi de Z pour calculer son espérance, la loi de X suffit.

Démonstration

Corollaire - Application : pseudo-linéarité
Soient X une v.a.r., a et b deux réels. Alors $\mathbf{E}(aX + b) = a\mathbf{E}(X) + b$.
En particulier $X - \mathbf{E}(X)$ est une v.a. centrée (appelée v.a. centrée associée à X).

Démonstration

Corollaire - Espérance de couple
Soient (X, Y) un couple de v.a. définies sur Ω et $g : (X, Y)(\Omega) \to \mathbb{R}$.
Alors l'espérance de la v.a.r. $g(X, Y)$ est donnée par la formule :

$$\mathbf{E}(g(X, Y)) = \sum_{(x,y) \in (X,Y)(\Omega)} g(x, y)\mathbf{P}\big((X, Y) = (x, y)\big)$$

Démonstration

On peut généraliser la formule à une v.a. du type $g(X_1, \cdots, X_n)$.

Proposition - Propriétés
Si X et Y sont deux v.a.r. (plus généralement si $X_1, \cdots, X_n$ sont n v.a.r) définies sur un même espace probabilisé fini $(\Omega, \mathbf{P})$. Alors :

i) $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$ (linéarité de l'espérance)

ii) si $X \geqslant 0$ p.s. alors $\mathbf{E}(X) \geqslant 0$ (positivité de l'espérance)

iii) si $X \geqslant 0$ p.s. et $\mathbf{E}(X) = 0$ alors $X = 0$ p.s.

iv) si $X \leqslant Y$ p.s. alors $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$ (croissance de l'espérance)

v) si X, Y sont indépendantes, $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$

vi) $\mathbf{E}(X_1 + X_2 + \cdots + X_n) = \mathbf{E}(X_1) + \mathbf{E}(X_2) + \cdots + \mathbf{E}(X_n)$

vii) si $X_1, \cdots, X_n$ sont indépendantes, $\mathbf{E}(X_1 X_2 \cdots X_n) = \mathbf{E}(X_1)\mathbf{E}(X_2) \cdots \mathbf{E}(X_n)$

Pour aller plus loin - Lien avec l'intégration
Les formules suivantes rappellent en grande partie des résultat équivalent sur le calcul d'intégrale de fonctions positives.
Pour l'intégration : si $f \geqslant 0$ alors $\int_I f \geqslant 0$.
L'équivalence pour la nullité étaient assurée à condition que f soit continue (et positive).
Dans le cadre des var, la continuité n'a pas de sens, alors il est nécessaire de considérer des ***var nulle presque sûrement***

Démonstration

Remarque - Espérance d'une loi binomiale
On peut ainsi retrouver facilement l'espérance d'une v.a. binomiale : c'est une somme de n Bernoulli.
Exercice
Faites le calcul

Proposition - Inégalité de Markov
Toute v.a.r. positive sur Ω fini vérifie l'inégalité :

$$\forall a > 0, \mathbf{P}(X \geqslant a) \leqslant \frac{\mathbf{E}(X)}{a}$$

Savoir faire - Exploitation d'indicatrice (3)
Il y a un événement naturel à étudier ici $A = [X \geqslant a]$, puis on exploite la propriété de l'espérance de la variable $\mathbb{1}_A$.

Démonstration

Exercice
Soit X, une variable aléatoire à valeurs entières.
Montrer que $\mathbf{P}(X=0) \geqslant 1-\mathbf{E}(X)$

Pour aller plus loin - Minoration de la probabilité d'absence (ou majoration de présence)
On trouve pour $X(\Omega) \in \mathbb{N}$, $\mathbf{P}(X=0) \geqslant 1-\mathbf{E}(X)$, une minoration de la probabilité d'absence

Savoir faire - Composition avec exp **(pour avoir une va positive). Méthode de Chernoff**
Il arrive fréquemment qu'on compose avec exp la variable X. On parle de comparaison avec des vecteurs sous-gaussiens.
On a alors $\mathbf{P}(X \geqslant x) = \mathbf{P}(e^X \geqslant e^x) \leqslant \dfrac{\mathbf{E}(e^X)}{e^x}$.

Exercice
On suppose que $X(\Omega) = \{-2,-1,0,1,2\}$, avec $p_2 = \mathbf{P}(X=2) = \mathbf{P}(X=-2)$ et $p_1 = \mathbf{P}(X=1) = \mathbf{P}(X=-1)$.
Majorer $\mathbf{P}(X \geqslant \epsilon)$.

5.2. Variance (d'une variable aléatoire)

Définition - Moment d'ordre r
Soit X une v.a.r. sur $(\Omega, \mathbf{P})$ fini. On appelle **moment d'ordre r** **(**$r \in \mathbb{N}$**)** de X le réel $m_r(X) = \mathbf{E}(X^r)$ et **moment centré d'ordre r** de X le réel $\mu_r(X) = \mathbf{E}[(X-\mathbf{E}(X))^r]$.

Savoir faire - Formulation calculatoire (transfert) - moment
Si X est centrée : $\mu_r(X) = \sum_{x \in X(\Omega)} x^r \mathbf{P}(X=x)$.

Définition - Variance et écart-type
Soit X une v.a.r. sur $(\Omega, \mathbf{P})$ fini. On appelle **variance** de X, notée $\mathbf{V}(X)$, le moment centré d'ordre 2 de X : $\mathbf{V}(X) = \mathbf{E}\Big((X-\mathbf{E}(X))^2\Big)$.
$\sigma(X) = \sqrt{\mathbf{V}(X)}$ est appelé **écart-type** de X.
Si $\sigma(X) = 1$ on dit que X est **réduite**.

Remarque - Positivité de la variance
On notera que $Y = \big(X - \mathbf{E}(X)\big)^2 \geqslant 0$, donc $\mathbf{V}(X) = \mathbf{E}(Y) \geqslant 0$.

Savoir faire - Formulation calculatoire (transfert) - variance
On a donc $\mathbf{V}(X) = \sum_{x \in X(\Omega)} (x-\mathbf{E}(X))^2 \mathbf{P}(X=x)$.

Proposition - Propriétés
Soit $(\Omega, \mathbf{P})$ un espace de probabilité fini. Alors

- **i)** $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$ formule de Huygens
- **ii)** $\mathbf{V}(aX+b) = a^2\mathbf{V}(X)$
- **iii)** $\sigma(aX+b) = |a|\sigma(X)$
- **iv)** si $\sigma(X) > 0$, $X^* = \frac{X-\mathbf{E}(X)}{\sigma(X)}$ est centrée réduite, on l'appelle **v.a. centrée réduite associée** à X
- **v)** $\mathbf{V}(X) = 0 \Leftrightarrow X$ est constante presque sûrement

Exemple - Variance de la loi uniforme sur $[[1, n]]$**.**

Démonstration

Proposition - Variance des lois usuelles
- si $X \hookrightarrow \mathscr{B}(p)$ alors $\mathbf{V}(X) = p(1-p)$;
- si $X \hookrightarrow \mathscr{B}(n,p)$ alors $\mathbf{V}(X) = np(1-p)$.

Démonstration

Remarque - A propos de la loi hypergéométrique
Il est bon de connaître et reconnaître la loi définie dans l'exercice suivante. C'est la loi hypergéométrique.
Il s'agit d'une sorte de répétition de loi de Bernoulli, avec dépendance cette fois-ci.
Notons enfin que dans l'exercice, on considère une succession de tirage sans remise. Le résultat est identique (même loi) si l'on considère plutôt un tirage d'une poignée de n boules.

Exercice

Une urne contient N boules, de deux catégories : des blanches en proportion p et des non blanches en proportion $q = 1-p$ ($Np \in \mathbb{N}$ désigne donc le nombre de boules blanches et $Nq \in \mathbb{N}$ celui de non blanches).
On tire successivement n boules de cette urne **sans remise** et on note X la v.a. égale au nombre de boules blanches obtenues.
Déterminer la loi de X, son espérance et sa variance (on pourra, pour cette dernière, commencer par calculer $\mathbf{E}(X(X-1))$).

Théorème - Inégalité de Bienaymé-Tchebychev
Soit X une v.a.r. sur $(\Omega, \mathbf{P})$ fini. Alors :

$$\forall\, \epsilon > 0,\ \mathbf{P}(|X - \mathbf{E}(X)| \geqslant \epsilon) \leqslant \frac{\mathbf{V}(X)}{\epsilon^2}$$

Pour aller plus loin - Majoration de la probabilité d'absence (ou minoration de présence)
On trouve pour $X(\Omega) \in \mathbb{N}$, $\mathbf{P}(X = 0) \leqslant \dfrac{\mathbf{V}(X)}{\mathbf{E}(X)^2}$, une majoration de la probabilité d'absence

Démonstration

Exercice
Soit X, une variable aléatoire à valeurs entières.
Montrer que $\mathbf{P}(X=0) \leqslant \dfrac{\mathbf{V}(X)}{\mathbf{E}(X)^2}$

Savoir faire - Majoration avec deux v.a.

Supposons que $(X_n) \to X$ et $(Y_n \to Y)$ (en probabilité).
1. Notons : $|x| < \frac{\epsilon}{2}$ et $|y| < \frac{\epsilon}{2} \Longrightarrow |x+y| \leqslant |x| + |y| < \epsilon$
par contraposée : $|x+y| \geqslant \epsilon \Longrightarrow |x| \geqslant \frac{\epsilon}{2}$ ou $|y| \geqslant \frac{\epsilon}{2}$.
2. En terme d'événements, on a donc l'inclusion :

$$\left[|(X_n+Y_n)-(X+Y)| \geqslant \epsilon\right] \subset \left([|X_n - X| \geqslant \frac{\epsilon}{2}]\right) \cup \left([|Y_n - Y| \geqslant \frac{\epsilon}{2}]\right)$$

Et en probabilité :

$$\mathbf{P}\left[|(X_n+Y_n)-(X+Y)| \geqslant \epsilon\right] \leqslant \mathbf{P}\left([|X_n - X| \geqslant \frac{\epsilon}{2}]\right) + \mathbf{P}\left([|Y_n - Y| \geqslant \frac{\epsilon}{2}]\right)$$

(Notons cette méthode s'applique également si $(Y_n)n$ (par exemple) n'est pas une variable aléatoire...)

Application - Variable aléatoire et suite numérique

5.3. Covariance (de deux variables aléatoires)

Définition - Covariance
Soient X, Y deux v.a.r sur (Ω, P) fini. On appelle **covariance** de X et Y le réel :

$$\mathbf{Cov}(X,Y) = \mathbf{E}\Big(\big(X - \mathbf{E}(X)\big)\big(Y - \mathbf{E}(Y)\big)\Big)$$

Exemple - Application

Proposition - Propriétés de la covariance
Soient X, X', Y, Y' des v.a.r. sur $(\Omega, \mathbf{P})$ fini et $a, b, c, d, \lambda \in \mathbb{R}$. On a

- **i)** $\mathbf{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$
- **ii)** $\mathbf{Cov}(aX+b, cY+d) = ac\mathbf{Cov}(X,Y)$
- **iii)** $\mathbf{Cov}(X,Y) = \mathbf{Cov}(Y,X)$
- **iv)** $\mathbf{Cov}(X,X) = \mathbf{V}(X)$
- **v)** $\mathbf{Cov}(X+X', Y) = \mathbf{Cov}(X,Y) + \mathbf{Cov}(X',Y)$ et $\mathbf{Cov}(\lambda X, Y) = \lambda\mathbf{Cov}(X,Y)$
  $\mathbf{Cov}(X, Y+Y') = \mathbf{Cov}(X,Y) + \mathbf{Cov}(X,Y')$ et $\mathbf{Cov}(X, \lambda Y) = \lambda\mathbf{Cov}(X,Y)$

Démonstration

Remarque - Cov comme un produit scalaire
La dernière propriété est répétée pour souligner le caractère de la bilinéarité de **Cov** sur l'espace $(\Omega, \mathbf{P})$.
Est-ce un produit scalaire?
On a également : $\mathbf{Cov}(X,X) = \mathbf{V}(X) \geqslant 0$.
En revanche, on n'a pas : $\mathbf{Cov}(X,X) = 0 \Rightarrow X = 0$. Seulement $X = c$ (une constante) p.s.
On peut alors changer l'égalité pour obtenir un produit scalaire : $X \doteq Y$, pour exprimer $\mathbf{V}(X-Y) = 0$.
C'est bien une relation d'équivalence. On préfère noter $X = Y + c$ p.s. (lire presque sûrement),
la valeur de c est alors $\mathbf{E}(X) - \mathbf{E}(Y)$.
C'est équivalent à : $\{\omega \mid X(\omega) - \mathbf{E}(X) \neq Y(\omega) - \mathbf{E}(Y)\} = [(X-Y) \neq \mathbf{E}(X-Y)]$ a une probabilité nulle.
Si les variables aléatoires X et Y ne sont pas indépendantes :

Pour aller plus loin* - $\mathbf{Cov}(X) = 0 \Longrightarrow X = 0$ *ps.
On retrouve le même problème que pour l'intégrale

Théorème - Lien variance-covariance
Pour des v.a.r. définies sur $(\Omega, \mathbf{P})$ fini, on a

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

$$\mathbf{V}(X_1 + \cdots + X_n) = \mathbf{V}(X_1) + \cdots + \mathbf{V}(X_n) + 2\sum_{i<j} \mathbf{Cov}(X_i, X_j) = \sum_{i,j} \mathbf{Cov}(X_i, X_j)$$

Démonstration

Théorème - Cas d'indépendance
Si X et Y sont deux v.a. **indépendantes** (plus généralement si $X_1, \cdots, X_n$ sont n v.a. **deux à deux indépendantes**) définies sur un même espace probabilisé fini. Alors :

i) $\mathbf{Cov}(X, Y) = 0$ (on dit que X et Y sont **non corrélées**)

ii) $\mathbf{V}(X+Y) = \mathbf{V}(X) + \mathbf{V}(Y)$

iii) $\mathbf{V}(X_1 + X_2 + \cdots + X_n) = \mathbf{V}(X_1) + \mathbf{V}(X_2) + \cdots + \mathbf{V}(X_n)$

D'une certaines façons, deux variables aléatoires non corrélées sont orthogonales pour le pseudo-produit scalaire **Cov** d'où la notation :

Définition - Variables non corrélées (notation)
Si X et Y sont deux variables aléatoires, non corrélées (i.e. $\mathbf{Cov}(X, Y) = 0$), on note $X \perp Y$.
On a donc $X \perp\!\!\!\perp Y \Longrightarrow X \perp Y$

Démonstration

Remarque - Variance d'une binomiale
On peut ainsi retrouver facilement la variance d'une v.a. binomiale.
Comme lors d'un exercice précédent, $\mathbf{V}(X) = \sum_{i=1}^{n} \mathbf{V}(X_i) = np(1-p)$.
Ce coup-ci l'indépendance est très importante.
La variance d'une hypergéométrique (addition de Bernoulli non indépendantes) est plus compliquée.

Attention - La réciproque est fausse.
Cela signifie que deux var peuvent avoir une covariance nulle (ou plus loin un coefficient de corrélation linéaire), sans être indépendantes.

Exercice
Soit X et Y deux variables aléatoires qui suivent une loi de Bernoulli et telles que la loi conjointe est donnée par le tableau :

$Y \backslash X$	0	1
0	a	b
1	c	d

1. Calculer la $\mathbf{Cov}(X,Y)$.
2. Montrer que (X,Y) sont indépendantes ssi la matrice est de rang 1.
   En déduire que si (X,Y) sont indépendantes, il existe $\lambda,\mu>0$ tel que $b=\lambda d$, $c=\lambda d$ et $a=\lambda\mu d$.
   Calculer $\mathbf{Cov}(X,Y)$
3. Montrer que dans ce cas $\mathbf{Cov}(X,Y)=0$ si et seulement si X et Y sont indépendantes.

Analyse - Inégalité de Cauchy-Schwarz

Définition - Coefficient de corrélation linéaire
Soient X et Y deux v.a. d'écart type non nul, (i.e. X,Y non constants presque sûrement).
On appelle **coefficient de corrélation linéaire** de X et Y le réel

$$\rho(X,Y)=\frac{\mathbf{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Proposition - Propriétés
$|\rho(aX+b,cY+d)|=|\rho(X,Y)|$.
On a toujours $|\rho(X,Y)|\leqslant 1$, c'est-à-dire $|\mathbf{Cov}(X,Y)|\leqslant\sigma(X)\sigma(Y)$, et $|\rho(X,Y)|=1$ si et seulement si il existe a et b réels tels que $Y=aX+b$ presque sûrement, c'est-à-dire tels que $\mathbf{P}(Y=aX+b)=1$.

Démonstration

Savoir faire - Tableau récapitulatif

On considère p un réel de l'intervalle $]0,1[$ et on pose $q = 1-p$.

nom	$X(\Omega)$	loi	espérance	variance
v.a constante (certaine)	$\{a\}$	$P(X=a)=1$	a	0
loi uniforme sur $\{1,2,\cdots n\}$ $\mathscr{U}_n$	$\{1,2,\cdots n\}$	$P(X=k)=\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
loi de Bernoulli de paramètre p $\mathscr{B}_p$ ou $\mathscr{B}(1,p)$	$\{0,1\}$	$P(X=0)=q$ $P(X=1)=p$	p	pq
loi binomiale de paramètres n,p $\mathscr{B}(n,p)$	$\{0,1,2,\cdots n\}$	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	np	npq

6. Bilan

Synthèse

⇝ Une variable aléatoire, issue d'une expérience, prend différentes valeurs numériques en fonction de l'état de l'univers. Il s'agit donc d'une fonction $X : \Omega \to \mathbb{R}$. La connaissance de l'univers est d'ailleurs souvent réduite à l'image donnée par différents $X_i(\Omega)$. Réciproquement, on peut partitionner Ω en fonction de $X^{-1}(k)$, où $k \in X(\Omega)$.

⇝ On associe alors à toute variable, une loi : la suite des valeurs $\mathbf{P}(X=x)$ où $x \in X(\Omega)$. Et comme X est à valeur numérique, on lui associe aussi une espérance $\mathbf{E}(X) = \sum_{x\in X(\omega)} x\mathbf{P}(X=x)$ qui résume, en une valeur toutes les valeurs prises par X. Ce résumé est évidemment pas bon, on lui associe alors $\mathbf{V}(X)$ qui évolue cette erreur.

⇝ Certaines variables aléatoires sont très fréquentes, on apprend donc à les reconnaitre : la variable peut suivre une loi uniforme, ou bien une loi de Bernoulli (comme la variable indicatrice) ou encore une loi binomiale qui compte le nombre de succès lorsqu'on répète n fois et indépendamment une même expérience élémentaire dont la probabilité de succès est p. D'autres lois sont classiques pour des univers dénombrables ou ayant la puissance du continu.

⇝ Plusieurs variables aléatoires peuvent qualifiées d'indépendantes si les événements générés par ces variables sont indépendants. Pour mesurer cela, il faut donc s'intéresser aux événements $(X_1 = x_1)\cap(X_2 = x_2)\dots$ que l'on étudie à partir de la variable aléatoire du couple (X_1,X_2) (ou n-uplet) : c'est l'événement $[(X_1,X_2) = (x_1,x_2)]$.
Si ces variables ne sont pas indépendantes, on peut mesurer la corrélation entre ces variables aléatoires.

⇝ Enfin, deux inégalités sont importantes pour contrôler des événements aléatoires, connaissant la variable aléatoire.
L'inégalité de Markov : si $X \geqslant 0$, $\mathbf{P}(X \geqslant a) \leqslant \frac{\mathbf{E}(X)}{a}$ donne une minoration sur la propriété d'absence d'une variable entière.
L'inégalité de Bienaymé-Tchébychev $\mathbf{P}(|X-a| \geqslant \epsilon) \leqslant \frac{\mathbf{V}(X)}{\epsilon^2}$ donne une majoration sur la propriété d'absence d'une variable entière.

Savoir-faire et Truc & Astuce du chapitre

- — Savoir-faire - Définir la loi d'une variable aléatoire
- — Savoir-faire - Exploitation d'indicatrice (1)
- — Savoir-faire - Modélisation
- — Truc & Astuce pour le calcul - Etude du $\max(X_i)$ où les X_i sont indépendantes
- — Savoir-faire - Exploitation d'indicatrice (2)
- — Savoir-faire - Exploitation d'indicatrice (3)
- — Savoir-faire - Composition avec exp (pour avoir une va positive). Méthode de Chernoff
- — Savoir-faire - Formulation calculatoire (transfert) - Moment
- — Savoir-faire - Formulation calculatoire (transfert) - Variance
- — Savior-faire - Majoration avec deux v.a.
- — Savoir-faire - Tableau récapitulatif

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$[X=a], [X\leqslant a]$	Respectivement les événements $X^{-1}(\{a\})$ et $X^{-1}(]-\infty,a])$	$\{\omega \mid X(\omega)=a\}\dots$	On raisonne ici par image de fonctions $X:\Omega\to R$ ou par classe d'équivalence. .
$\mathbb{1}_A$	Variable aléatoire indicatrice de l'événement A	$\mathbb{1}_A(x)=1$ ssi $x\in A$	$\mathbf{E}(\mathbb{1}_A)=\mathbf{P}(A)$. On s'en sert pour partitionner Ω .
$X\hookrightarrow \mathscr{U}(\llbracket 1,n\rrbracket)$ (ou $\mathscr{U}_n$	X (v.a.) suit la loi uniforme sur $\llbracket 1,n\rrbracket$	$X(\Omega)=\llbracket 1,n\rrbracket$, $\mathbf{P}(X=k)=\frac{1}{n}$	$\mathbf{E}(X)=\dfrac{n+1}{2}$ et $\mathbf{V}(X)=\dfrac{n^2-1}{12}$.
$X\hookrightarrow\mathscr{B}(p)$	X (v.a.) suit la loi de Bernoulli de paramètre p	$X(\Omega)=\{0,1\}$, $\mathbf{P}(X=1)=p$	$\mathbf{E}(X)=p$ et $\mathbf{V}(X)=p(1-p)$.
$X\hookrightarrow\mathscr{B}(n,p)$	X (v.a.) suit la loi binomiale de paramètres n et p	$X(\Omega)=\llbracket 0,n\rrbracket$, $\mathbf{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	$\mathbf{E}(X)=np$ et $\mathbf{V}(X)=np(1-p)$.
	Nombre de succès lors d' une répétition de n expériences aléatoires indépendantes	avec probabilité de succès pour chacune égale à p	
$X \perp\!\!\!\perp Y$	Les variables X et Y sont indépendantes	$\forall\ (x,y)\in X(\Omega)\times Y(\Omega)$, $\mathbf{P}([(X=x)\cap(Y=y)])=\mathbf{P}(X=x)\mathbf{P}(Y=y)$	Notation propre à la MPSI3
$X=Y$p.s.	Les variables X et Y sont égales presque sûrement	$\mathbf{P}\{\omega\in\Omega \mid X(\omega)\neq Y(\omega)\})=0$	
$\mathbf{E}(X)$	Espérance de X	$\mathbf{E}(X)=\sum_{x\in X(\Omega)} x\mathbf{P}(X=x)=\sum_{\omega\in\Omega}X(\omega)\mathbf{P}(\{\omega\})$	Meilleur résumé en un nombre d'une variable aléatoire(...).
$\mathbf{V}(X)$	Variance de X	$\mathbf{V}(X)=\mathbf{E}((X-\mathbf{E}(X))^2)=\sum_{x\in X(\Omega)}(x-\mathbf{E}(X))^2\mathbf{P}(X=x)$	$\mathbf{V}(X)\geqslant 0$.
$\sigma(X)$	Ecart-type de X	$\sigma(X)=\sqrt{\mathbf{V}(X)}$	Meilleur résumé de la dispersion (autour de sa valeur moyenne) d'une variable aléatoire(...).
$\mathrm{Cov}(X,Y)$	Covariance de X et Y.	$\mathrm{Cov}(X,Y)=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y)$ (on exploite la loi du couple)	$\mathrm{Cov}(X,X)=\mathbf{V}(X)$ et $\mathrm{Cov}(X,Y)\leqslant\sqrt{\mathbf{V}(X)}\sqrt{\mathbf{V}(Y)}$.
$\rho(X,Y)$	Corrélation de X et Y.	$\rho(X,Y)=\sqrt{\mathrm{Cov}(X,Y)}$	Entre -1 (cas $X=-Y$) et 1 (cas $X=Y$), elle mesure le commun entre X et Y.
$X\perp Y$	Les v.a. X et Y sont non corrélées	$\mathrm{Cov}(X,Y)=0$	$X\perp\!\!\!\perp Y\Rightarrow X\perp Y$

Retour sur les problèmes

158. Comment répondre mathématiquement à la question : qu'est-ce qui est le plus naturelle? Quelle que soit la réponse intime que se donne le lecteur, le plus important est qu'il soit capable de comprendre l'autre point de vue (fonction image/classe d'équivalence).

159. Cours. Sur un univers fini : lois uniformes, loi de Bernoulli (0 ou 1), loi binomiale, loi hypergéométrique...
160. L'espérance est le bon résumé d'une variable aléatoire, car le nombre m qui annule $\mathbf{E}(X-m)$ est $\mathbf{E}(X)$. Mais cela donne une définition auto-réferente. A éviter absolument !
161. Moins on en connait sur un phénomène, plus on peut exploiter le calcul de probabilités. Il est parfois plu simple.
Une autre confirmation de ce point de vue est la méthode de Monte-Carlo proposé par Van Neumann pour calculer l'intégrale d'une fonction compliquée. Pour calculer $\int_a^b f(t)\mathrm{d}t$, on tire au hasard des nombres x_i entre a et b, et on fait la moyenne des $f(x_i)\dots$
162. On étudie la loi du couple. Finalement, c'est comme si Ω était partitionner en une partiiton plus fine contenant celle induit par la variable X et celle induite par $Y\dots$
163. Voir le cours. On retrouve sauf le fait que le produit scalaire soit défini. Sauf à faire évoluer = en = p.s.