Probabilités

I Axiomatique des probabilités (Kolmogorov)

I.1 Tribus ou σ-algèbre

Définition

Soit Ω un ensemble (appelé l'univers). Une tribu (ou σ-algèbre) sur Ω est un sous-ensemble $\mathcal{A} \subset \mathcal{P}(\Omega)$ ayant les propriétés suivantes :

i. $\Omega \in \mathcal{A}$

ii. $E \in \mathcal{A} \Longrightarrow \overline{E} \in \mathcal{A}$

iii. Si $(E_n)_{n\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathcal{A}$, alors leur réunion appartient à $\mathcal{A}$

→ stable par réunion dénombrable, et passage au complémentaire.

Vocabulaire.

- On note généralement un tribu $\mathcal{A}$ ou $\mathcal{T}$.
- Les éléments de $\mathcal{A}$ sont appelés 'événements'.
- $(\Omega, \mathcal{A})$ s'appelle un espace probabilisable.
- Dans l'axiome (iii), on peut indexer par n'importe quelle ensemble au plus dénombrable.

Explication du formalisme dans le jeu de pile ou face : On souhaite modéliser une expérience aléatoire, par exemple, tirer indéfiniment à pile ou face (en notant 0 pour pile et 1 pour face).

-Un 'aléa' est l'un des résultats possibles de cette expérience : c'est à dire ici :

-L'univers est l'ensemble des résultats possibles, c'est à dire qu'ici $\Omega = \{0,1\}^{\mathbb{N}}$.

Généralement, l'univers est trop complexe pour être décrit.
- Les événements sont certains sous-ensemble de résultat. Par exemple, $B =$ [le premier tirage est pile] est un événement. C'est un sous ensemble de Ω. ici :

I.2 Exemples de tribus

Il existes deux tribus dites triviales :
- $\mathcal{A} = \{\emptyset, \Omega\}$
- $\mathcal{A} = \mathcal{P}(\Omega)$

Lorsque Ω est fini, on utilise systématiquement la deuxième.
Le résultat suivant est immédiat :

Proposition (intersection)

L'intersection quelconque de tribus est une tribu.

Ceci justifie la définition ci-après (hors programme)

Définition

Soit B une partie de $\mathcal{P}(\Omega)$. On note $\mathcal{T}(B)$ la tribu engendrée par B, définie ainsi :

$$\boxed{\mathcal{T}(B) = \bigcap_{\substack{\mathcal{A} \text{ tribu} \\ \mathcal{A} \supset B}} \mathcal{A}}$$

C'est la plus petite tribu qui contient B

Exemple. Si par exemple, si on prend un sous ensemble Ω_1 de Ω. Quelle est la plus petite tribu contenant Ω_1 ?

I.3 Propriété des événements et vocabulaire

- Ω s'appelle l'événement certain.
- $\emptyset$ s'appelle l'événement impossible.
- Si E est un événement, $\overline{E}$ s'appelle l'événement contraire de E.
- Deux événement sont dits incompatibles lorsque leur intersection est vide.
- On dit qu'un famille finie ou dénombrable $(E_n)_{n\in I}$ est un système complet d'événements lorsque $\bigcup\limits_{n\in I} E_n = \Omega$ et que les E_n sont deux à deux incompatibles.

I.4 Manipulation d'événements

De la définition on déduit immédiatement que

Proposition

Si $(E_n)_{n\in I}$ une famille finie dénombrable d'événements. Alors :

$$\boxed{\bigcap_{n\in I} E_n}$$

et

$$\boxed{\bigcup_{n\in I} E_n}$$

sont aussi des événements.

Remarque. De façon informelle, toute succession finie ou dénombrable d'opérations sur les événements, à base de $\cup$ et de $\cap$, définit un événement.

Manipulation de quantificateurs :
Le vocabulaire des événements transforme en intersections et réunions les assertions quantifiées, en vertu de la constatation suivante :

"les quantificateur $\forall$ sont associés à des intersections, les quantificateur $\exists$ sont associés à des réunions."

Voici des exemples de cette situation.

Exemple. Dans un jeu de pile ou face, on note $E_n =$ [le nième tirage est pile]
Montrer que les ensembles suivants sont des événements :

$F =$ [tous les tirages pairs sont des pile]
$G =$ [exactement un des tirage est pile]
$H =$ [si on tire une fois pile, alors tous les tirages suivants sont aussi des pile]

L'exemple qui suit est plus subtil :

Exemple. limites supérieures et inférieures d'une famille d'événements
Soit $\mathcal{A}$ une tribu, et soit $(E_n)_{n\in\mathbb{N}}$ une suite d'événements de E. On note $S =$ [une infinité de E_n sont réalisés] .
Alors S est un événement (appelé la limite supérieure des E_n).
En effet, si on note

$$B_n = \bigcup_{k\geqslant n} E_k$$

Alors

$$S = \bigcap_{n\in\mathbb{N}} B_n = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geqslant n} E_k$$

Noter que toutes les intersections et réunions faites portent sur des familles dénombrables d'ensembles.

De même si on note $I =$ [tous les E_n sont réalisés à partir d'un certain rang], alors I est un événement(la limite inférieure de (E_n)) : prendre le temps d'écrire cet événement avec les symboles intersection et réunion.

I.5 Indicatrice

Si A est un événement , on définit $\mathbb{1}_A$ la fonction indicatrice de A :

$$\boxed{\mathbb{1}_A : \begin{array}{|lcl} \Omega & \longrightarrow & \{0,1\} \\ \omega & \longmapsto & \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{sinon} \end{cases} \end{array}}$$

Remarquons que si A et B sont deux événements, alors $\boxed{\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A}$, que $\boxed{\mathbb{1}_{A\cap B} = \mathbb{1}_A \times \mathbb{1}_B}$ et que $\boxed{1 - \mathbb{1}_{A\cup B} = (1-\mathbb{1}_A)\times(1-\mathbb{1}_B)}$, ce qui rendra très utile les indicatrices pour les calculs.

On démontrera de plus que $\boxed{E(\mathbb{1}_A) = \mathbb{P}(A)}$.

II Probabilité sur un espace probabilisable

II.1 Définition

Définition

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. Une probabilité (ou loi de probabilité) sur $(\Omega, \mathcal{A})$ est une application $\mathbb{P} : \mathcal{A} \longrightarrow [0,1]$ vérifiant :

i. $\mathbb{P}(\Omega) = 1$
ii. Si $(E_n)_{n\in\mathbb{N}}$ est une famille d'éléments incompatibles, alors :

$$\boxed{\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(E_n)}$$

$(\Omega, \mathcal{A}, \mathbb{P})$ est appelé espace probabilisé

Remarque.
— En prenant $E_0 = \Omega$ et $E_n = \emptyset$ pour $n > 0$, on obtient $\mathbb{P}(\emptyset) = 0$.
— L'axiome (ii) s'appelle **l'axiome de σ-additivité.** Il est aussi vrai pour une famille $(E_n)_{n\in J}$ avec $J \subset \mathbb{N}$ fini : il suffit de la compléter avec $E_n = \emptyset$ pour $n \in \mathbb{N}\backslash J$. Il est également vrai pour toute famille au plus dénombrable :
— Si $(E_i)_{i\in I}$ est une famille dénombrable d'événements incompatibles, on déduit de (ii) que : la famille $(\mathbb{P}(E_i))_{i\in I}$ est sommable et

$$\sum_{i\in I} \mathbb{P}(E_i) = \mathbb{P}\left(\bigcup_{i\in I} E_i\right)$$

II.2 Propriétés finies

Proposition

On se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et on considère A, B et $A_1, ..., A_n$ des événements.

i. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
ii. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
iii. Inégalité de Boole. Pour toute famille fine d'événements on a :

$$\boxed{\mathbb{P}\left(\bigcup_{n=0}^{N} A_n\right) \leqslant \sum_{n=0}^{N} \mathbb{P}(A_n)}$$

iv. La formule du crible :

$$\boxed{\mathbb{P}(A_1 \cup \cdots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{i_1<\cdots<i_k} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k})}$$

Démonstration. $\begin{cases} A \cup \bar{A} = \Omega \\ \text{union disjointe} \end{cases} \Rightarrow \mathbb{P}(A) + \mathbb{P}(\bar{A}) = 1 = \mathbb{P}(\Omega).$
$\Rightarrow \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A).$

Boole finie : Par récurrence.
$n=2$: $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2) \leq \mathbb{P}(E_1) + \mathbb{P}(E_2).$
On suppose vrai pour n.
$\mathbb{P}(E_1 \cup \dots \cup E_{n+1}) \underset{\text{cas } n=2}{\leq} \mathbb{P}(E_1 \cup \dots \cup E_n) + \mathbb{P}(E_{n+1})$
$\underset{H.R}{\leq} \mathbb{P}(E_1) + \dots + \mathbb{P}(E_{n+1}).$

II.3 Continuité monotone

Proposition (Continuité croissante)

Soit $(A_n)_{n\in\mathbb{N}}$ une famille croissante pour l'inclusion d'événements. On pose $B=\bigcup_{n\in\mathbb{N}} A_n$. Alors :

$$\mathbb{P}(B)=\lim_{n\to+\infty}\mathbb{P}(A_n)$$

Cette propriétés est d'une importance capitale pour la suite, et contrairement à l'intuition elle n'est pas du tout évidente.

preuve : on note : $B_0=A_0$; $B_1=A_1\setminus A_0$ ... $B_n=A_n\setminus A_{n-1}$

Par construction : (i) B_n disjoints (ii) $A_n=B_0\cup B_1\cup\dots\cup B_n$.

(iii) $\bigcup_{n=0}^{\infty} A_n=\bigcup_{n=0}^{\infty} B_n$. $\mathbb{P}(A_n)=\sum_{k=0}^{n}\mathbb{P}(B_k)\xrightarrow[n\to\infty]{}\sum_{k=0}^{\infty}\mathbb{P}(B_k)$ (conv de la série).

$=\mathbb{P}\left(\bigcup_{n=0}^{\infty} B_n\right)$ par σ-additivité

$=\mathbb{P}\left(\bigcup_{n=0}^{\infty} A_n\right)$.

Proposition

Si (A_n) n'est pas croissante, on a tout de même :

$$\lim_{n\to+\infty}\mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right)=\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right)$$

Proposition (Continuité décroissante)

Soit (A_n) une suite décroissante d'événements. Alors :

$$\lim_{n\to+\infty}\mathbb{P}(A_n)=\mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

Proposition

Si (A_n) n'est pas décroissante, on a tout de même :

$$\lim_{n\to+\infty}\mathbb{P}\left(\bigcap_{k=0}^{n} A_k\right)=\mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

II.4 Autres propriétés dénombrables

Proposition (Inégalité de Boole)

Pour toute famille $(A_n)_{n\in\mathbb{N}}\in\mathcal{A}^{\mathbb{N}}$,

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right)\leqslant\sum_{n=0}^{\infty}\mathbb{P}(A_n)\in[0,+\infty]$$

preuve : $\underbrace{\mathbb{P}\left(\bigcup_{n=0}^{N} A_n\right)}_{u_N}\leq\underbrace{\sum_{n=0}^{N}\mathbb{P}(A_n)}_{v_N}$ (Boole partiel).

continuité croissante : $u_N\longrightarrow\mathbb{P}\left(\bigcup_{n=0}^{\infty} A_n\right)$.

th de somme série : $v_N\xrightarrow[N\to\infty]{}\sum_{n=0}^{\infty}\mathbb{P}(A_n)$.

Proposition (décomposition sur un système complet)

Soit (A_n) un système complet d'événements. Alors pour tout $B\in\mathcal{A}$:

$$\mathbb{P}(B)=\sum_{n=0}^{+\infty}\mathbb{P}(B\cap A_n)$$

(i) $A_i\cap A_j=\emptyset$ si $i\neq j$
(ii) $\bigcup A_n=\Omega$

dém : on pose : $E_n=E\cap A_n$. $E=\bigcup_{n=0}^{\infty}E_n$ et les E_n sont disjoints.

Soit $x\in E$. comme $x\in\Omega$, et $\Omega=\bigcup_{n=0}^{\infty}A_n$; $\exists n, x\in A_n$. Alors, $x\in E\cap A_n=E_n$. $\to E\subset\bigcup E_n$. Si $x\in E_n\Rightarrow x\in E$. $\to\mathbb{P}(E)=\sum\mathbb{P}(E\cap A_n)$ (σ-additivité).

II.5 Événements presque sûrs, événements négligeables

dans cette section, on introduit de nouveaux objets qui n'existaient pas dans les probabilités finies et qui vont jouer un grand rôle.

Définition

Soit $(\Omega,\mathcal{A},\mathbb{P})$ un espace probabilisé et soit E un événement. On dit que E est négligeable lorsque $\mathbb{P}(E)=0$ et que E est presque sûr lorsque $\mathbb{P}(E)=1$.

Proposition

— Toute réunion finie ou dénombrable d'événements négligeables est négligeable.
— Toute intersection finie ou dénombrable d'événements presque sûrs est presque sûre.

→ A_n négligeables : $\mathbb{P}\left(\bigcup_{n=0}^{\infty}A_n\right)\leq\sum_{n=0}^{\infty}\mathbb{P}(A_n)=0$.

A_n presque sûrs : $\mathbb{P}\left(\bigcap_{n=0}^{\infty}A_n\right)=1-\mathbb{P}\left(\overline{\bigcap_{n=0}^{\infty}A_n}\right)$

$=1-\underbrace{\mathbb{P}\left(\bigcup_{n=0}^{\infty}\overline{A_n}\right)}_{=0}$.

Exemple. On joue à pile ou face une infinité de fois avec une pièce équilibrée.
Démontrer que les événements suivants sont négligeables ou presque surs.

F = [on tire au moins une fois pile]
G = [on tire un nombre fini une fois pile]
H = [on tire un nombre infini de pile et un nombre infini de face]

II.6 Systèmes quasi complet d'événements

Une famille (A_n) est un système quasi complet d'événements, si les A_i sont disjoints et $\cup A_i$ est presque sur.
Si l'on dispose d'un tel système alors la formule : pour tout $B \in \mathcal{A}$:

$$\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B \cap A_n)$$

reste valide. (les événements négligeables ne comptent pas pour les probabilités)

II.7 Exemples de lois de probabilités

Loi uniforme sur un univers fini

Lorsque Ω est fini, on prend $\mathcal{A} = \mathcal{P}(\Omega)$. Pour un événement (sous ensemble) E, on pose :

$$\boxed{\mathbb{P}(E) = \frac{|E|}{|\Omega|}}$$

Alors :

Proposition

$\mathbb{P}$ est une loi probabilité appelée loi uniforme. Tous les singletons sont équiprobables.

Exemple : loi uniforme sur un dé, sur un jeu de carte.

Les lois uniformes sont les plus naturelles, mais pas les seules (et heureusement). Le résultat suivant donne la forme générale des lois de probabilité sur un univers fini.

Loi quelconque sur un univers fini

Proposition

Soit $\Omega = \{\omega_1, .., \omega_n\}$. Soit $(p_1, ..., p_n) \in \mathbb{R}_+^n$ tel que $p_1 + \cdots + p_n = 1$. Il existe une unique probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ telle que $\mathbb{P}(\{\omega_i\}) = p_i$ pour tout $i \in [\![1, n]\!]$.

Exemple : On joue trois fois à pile ou face. On compte le nombre de pile obtenu : ceci fournit sur l'ensemble $\{0, 1, 2, 3\}$ une loi de probabilité qui n'est pas uniforme.

Supposons maintenant l'univers infini, la situation est différente. on est en particulier confronté à une difficulté : **il n'existe pas de loi uniforme.**

Probabilité discrète sur un univers dénombrable

Dans ce paragraphe, $\Omega = \{\omega_n,\ n \in \mathbb{N}\}$ est un univers dénombrable. On le munit de la tribu $\mathcal{A} = \mathcal{P}(\Omega)$ (qui elle n'est pas dénombrable). Le théorème ci dessous caractérise les lois de probabilités sur $(\omega, \mathcal{A})$.

Théorème

- Soit $(p_n)_{n \in \mathbb{N}}$ une famille de réels positifs sommable et de somme 1. Alors il existe une unique loi de probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ telle que pour tout n on ait $\mathbb{P}(\{\omega_n\}) = p_n$.
- Réciproquement, si $\mathbb{P}$ est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, alors $\sum_{i=0}^{+\infty} \mathbb{P}(\{\omega_i\}) = 1$.

Vocabulaire : la famille $(p_n)_{n \in \mathbb{N}}$ est aussi quelquefois indexée $(p_\omega)_{\omega \in \Omega}$ elle est appelée la **distribution de probabilité discrète,** ou le germe de probabilité.

Démonstration. $\boxed{\Leftarrow}$ $\Omega = \bigcup_{i \in \mathbb{N}} \{\omega_i\}$ est une union disjointe dénombrable donc $1 = \mathbb{P}(\Omega) = \sum_{i=0}^{+\infty} \mathbb{P}(\{\omega_i\})$.
$\boxed{\Rightarrow}$ On note $p_\omega = \mathbb{P}(\{\omega\})$. On a alors, par σ-additivité :

$$\forall A \in \mathcal{P}(\Omega),\ \mathbb{P}(A) = \sum_{\omega \in A} p_\omega$$

Ceci détermine la fonction $\mathbb{P}$. On vérifie ensuite que $\mathbb{P}$ définie ainsi est bien une probabilité sur Ω. L'axiome $\mathbb{P}(\Omega) = 1$ vient du fait que la famille a pour somme 1. L'axiome de σadditivité est une conséquence immédiate du théorème de Fubini positif.

Exemple. Existe-t-il une loi de probabilité sur $\mathbb{N}^*$ telle que $\mathbb{P}(\{n\})$ soit proportionnel à $\frac{1}{n}$? et à $\frac{1}{n!}$?

Sur un univers indénombrable

Sur les univers non dénombrables, **l'existence de tribus ayant des propriétés requises sera toujours admise.** C'est par exemple le cas du jeu de pile ou face infini examiné précédemment. En effet, dans ce cas, L'univers est l'ensemble des suites infinies de $0, 1$. On peut démontrer que cet ensemble n'est pas dénombrable.

III Conditionnement–Indépendance

Dans tout cette section, on se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

III.1 Conditionnement

Définition

Soient A et B deux éléments de $\mathcal{A}$, avec B non négligeable. On appelle probabilité de A conditionnellement à B, que l'on note $\mathbb{P}(A|B)$ ou $\mathbb{P}_B(A)$, la quantité :

$$\boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}}$$

Proposition

$\mathbb{P}_B : \mathcal{A} \longrightarrow [0,1]$ est une loi de probabilité sur $(\Omega, \mathcal{A})$, appelée loi conditionnelle relativement à B.

III.2 Les trois formules des probabilités conditionnelles

La formule des probabilités totales

Proposition

Soit $(A_n)_{n\in\mathbb{N}}$ un système complet d'événements, fini ou dénombrable. On suppose que pour tout $n \in \mathbb{N}$, $\mathbb{P}(A_n) \neq 0$. Alors, pour tout $B \in \mathcal{A}$:

$$\boxed{\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}_{A_n}(B)\mathbb{P}(A_n)}$$

Remarque. S'il y a des A_n négligeables dans la suite, alors $\mathbb{P}_{A_n}(B)$ n'est pas défini, mais $\mathbb{P}(A_n \cap B) = 0$. On considère donc que $\mathbb{P}_{A_n}(B)\mathbb{P}(A_n) = 0$ ce qui permet d'écrire la formule.

utilisation : On utilisera cette formule lorsque apparaissent des disjonctions de cas.

Exemple1
On lance un dé jusqu'à l'obtention d'un 6. Quelle est la probabilité que tous les dés tirés soient pairs ?

Exemple 2 On suppose que les naissance Fille Garçons sont équiprobables et que la probabilité de l'événement $A_n =$ [avoir n enfants] est de $p_n = p(1-p)^n$. Calculer la probabilité de l'événement $B =$ [avoir deux garçons].

On trouve $P(B) = \frac{2p(1-p)^2}{(1+p)^3}$

La formule des probabilités composées

Proposition

Soient $A_1, ..., A_n$ des événements tels que $\mathbb{P}(A_1 \cap \cdots A_n) \neq 0$. Alors :

$$\boxed{\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}_{A_1}(A_2)\mathbb{P}_{A_1\cap A_2}(A_3)\cdots\mathbb{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n)}$$

Utilisation : cette formule est utile lorsqu'il y a des choix successifs

Exemple. Soit (a_n) une famille d'entiers. On part d'une urne contenant $n_1 = 1$ boules noires et $b_1 = 1$ boule blanche. On répète l'expérience suivante : Au $i^{\text{ième}}$ tirage, on tire une boule, si elle est noire on s'arrête, sinon on rajoute 1 boule blanche. On note B_k l'événement : le jeu s'arrête au $k$ième tirage. .

i. Calculer la probabilité de B_k
ii. Calculer la probabilité que le jeu termine
iii. Même question si on rajoute 2^i boules blanches après le $i$eme tirage.

La formule de Bayes

La formule de Bayes a pour but d'inverser le conditionnement. Elle est peu utile dans le programme de MP.

Proposition

Soit $(B_n)_{n\in\mathbb{N}}$ un système complet d'événements, soit A ~~un événement, alors :~~

$$\boxed{\mathbb{P}_A(B_n) = \frac{\mathbb{P}_{B_n}(A)\mathbb{P}(B_n)}{\sum_{k=0}^{+\infty} \mathbb{P}_{B_k}(A)\mathbb{P}(B_k)}} = \frac{P(A|B_n)P(B_n)}{P(A)}$$

Exemple. Une maladie touche en moyenne 1 personne sur 10000. Un test permettant de déceler cette maladie existe. Celui-ci est positif pour 99% des individus malades, et pour seulement 0,1% des individus sains.

i. Un patient fait le test, celui-ci est positif. Quelle est la probabilité qu'il soit malade ? (on trouve quasiment 1/10
ii. Pour confirmer le premier résultat, le patient refait le test, qui est encore positif. Quelle est la nouvelle probabilité qu'il soit malade ? (cela revient à considérer l'événement A' consistant en 2 tests positifs consécutifs. on trouve quasiment 1 cette fois).

III.3 Indépendance

Cas de deux événements

Définition

Soit $(A, B) \in \mathcal{A}^2$. On dit que A et B sont indépendants lorsque $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Notation : on note parfois $A \perp\!\!\!\perp B$ pour signifier que les événements A et B sont indépendants.

Remarque. -Si $\mathbb{P}(B) \neq 0$, alors A et B sont indépendants si et seulement si $\mathbb{P}(A|B) = \mathbb{P}(A)$.

- si A est presque sur, alors il est indépendant de tout événement.
-Si A est indépendant de lui même alors il est presque sur ou négligeable.
-Si A et B sont indépendants, alors A et $\overline{B}$ aussi. → preuve p.1.

Exemple
On tire 2 dés. On considère les événements suivants :
— $A =$ [le premier dé est pair]
— $B =$ [le second dé est pair]
— $C =$ [la somme est paire]
Etudier leur indépendance deux à deux.

Cas général

Définition

Soit (A_n) une famille dénombrable d'événements. On dit que les A_n sont mutuellement indépendants lorsque :

$$\forall k \in \mathbb{N}^*,\ \forall (i_1, ..., i_k) \in \mathbb{N}^k,\ \mathbb{P}\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j})$$

Proposition

Si les A_n sont mutuellement indépendants, alors ils sont indépendants deux à deux. La réciproque est fausse.

Exemple. On reprend l'exemple précédent
— $A =$ [le premier dé est pair]
— $B =$ [le second dé est pair]
— $C =$ [la somme est paire]

Vérifier que ces trois événements ne sont pas mutuellement indépendants.

Il est intéressant de noter qu'on a une analogie avec l'indépendance linéaire . Ce qui se produit ici est que l'événement C est obtenu à partir de A et B par intersections reunions et complémentaires, autrement dit, il est dans la tribu engendré par A et B.

Propriétés

Proposition

Passage au complémentaire :
Soit (A_n) une famille d'événements, et soit (B_n) telle que pour tout $n \in \mathbb{N}$, ($B_n = A_n$ ou $B_n = \overline{A_n}$). Alors les A_n sont mutuellement indépendants si et seulement si les B_n sont mutuellement indépendants.

Coalitions finies :
Soit I fini, soit $(A_i)_{i\in I}$ une famille d'événements mutuellement indépendants, et soit $J_1 \cup \cdots \cup J_r$ une partition de I. On pose $B_k = \bigcap_{i \in J_k} A_i$. Alors $(B_1, ..., B_r)$ est une famille d'événements mutuellement indépendants.

Intersections infinies.
Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'événements mutuellement indépendants et I, J une partition de $\mathbb{N}$. On note $B_1 = \bigcap_I A_i$ et $B_2 = \bigcap_J A_i$. Alors B_1 et B_2 sont indépendants.

Démonstration.

Par passage au complémentaire, les résultats sont vrais pour les réunions. En fait les résultats ci-dessus sont un cas particulier du théorème suivant qui bien que hors programme peut être utilisé :

Proposition (Lemme des coalitions, HP)

Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'événements mutuellement indépendants. On considère deux sous-familles disjointes $(A_n)_{n\in I_1}$ et $(A_n)_{n\in I_2}$. On note $\mathcal{A}_1$ et $\mathcal{A}_2$ les tribus engendrées par ces familles, respectivement. Alors si $B_1 \in \mathcal{A}_1$ et $B_2 \in \mathcal{A}_2$, B_1 et B_2 sont indépendants.

Cela signifie que tout événement construit "à partir des $(A_n)_{n\in I_1}$" est indépendant de tout événement construit "à partir des $(A_n)_{n\in I_2}$"

III.4 Indépendance dans le jeu de pile ou face

On modélise le pile ou face par l'univers $\Omega = \{0, 1\}^{\mathbb{N}}$. Nous aurons besoin du résultat suivant :

Proposition (admise)

Soit $p \in [0, 1]$. On peut munir Ω d'une tribu $\mathcal{A}$ contenant les événements $A_k =$ [le $n^{\text{ième}}$ tirage est pile] et d'une probabilité $\mathbb{P}$ telle que :
i. $\forall n \in \mathbb{N}$, $\mathbb{P}(A_n) = p$
ii. $(A_n)_{n\in\mathbb{N}}$ est une famille d'événements mutuellement indépendants.

IV Annexe : Rappels sur les cardinaux usuels

Dans les univers finis munis de la loi uniforme, les problèmes de probabilité se ramènent à des problèmes de dénombrements.
Il faut donc connaître les cardinaux de référence.

Ensemble	Cardinal
$\mathfrak{S}(E)$	$\lvert E\rvert!$
F^E	$\lvert F\rvert^{\lvert E\rvert}$
$E_1 \times \cdots \times E_n$	$\lvert E_1\rvert \times \cdots \times \lvert E_n\rvert$
$\bigsqcup_{i\in I} A_i$ disjointe	$\sum_{i\in I} \lvert A_i\rvert$
$\bigcup_{i\in I} A_i$	$\sum_{k=1}^{n}(-1)^{k-1}\sum_{i_1<\cdots<i_k} \lvert (A_{i_1}\cap\cdots\cap A_{i_k}\rvert$
$\{f : [\![1,p]\!] \longrightarrow [\![1,n]\!]$ injective$\}$	$\frac{n!}{(n-p)!}$
$\{f : [\![1,p]\!] \longrightarrow [\![1,n]\!]$ strictement croissante$\}$	$\binom{n}{p}$
$\{f : [\![1,p]\!] \longrightarrow [\![1,n]\!]$ croissante$\}$	$\binom{n+p-1}{p}$
$\{(a_1,\dots,a_p) \in (\mathbb{N}^*)^p,\ a_1+\cdots+a_p = n\}$	$\binom{n-1}{p-1}$
$\{(a_1,\dots,a_p) \in \mathbb{N}^p,\ a_1+\cdots+a_p = n\}$	$\binom{n+p-1}{p-1}$

$P(E) \Rightarrow 2^{|E|}$

n
→ p-uplets ordonnés parmi n éléments
→ pas d'ordre.

On peut également connaître et savoir redémontrer les deux formules suivantes :

Van der Monde :

$$\sum_{p+q=k}\binom{a}{p}\binom{b}{q} = \binom{a+b}{k}$$

Multinôme à p termes :

$$\sum_{k_1+\ldots+k_p=n}\frac{n!}{k_1!\ldots k_p!}z_1^{k_1}\ldots z_p^{k_p} = (z_1+\ldots+z_p)^n$$

Voici quelques exemples de problèmes de dénombrement utilisant les formules du tableau :

Exemple. Utilisation de la formule de Poincaré :

Dénombrer les dérangements de S_n c'est à dire les permutations sans point fixe.

Exemple. Utilisation d'une bijection judicieuse :

Démontrer la formule de la dernière ligne du tableau ci-dessus en utilisant celle de l'avant dernière ligne.

Combien y a-t-il de façons de placer p prospectus dans n boîtes aux lettres ?

Combien y a t'il de chemins reliant le point $(0,0)$ au point (p,q) utilisant des déplacements d'un case vers le haut ou vers la droite ?

Exemple. Tirages sans remise (jeu de carte par exemple, utilisation de coefficients binomiaux.)

Déterminer le nombre de mains de 5 cartes d'un jeu de 52 qui contiennent :

une paire exactement (1098240)
deux paires exactement (123552)
un brelan (54912)
une suite (5 cartes consecutives pas forcément de la meme couleur, pas toutes de la même couleur) (9180)
un full (3744)
un carré (624)
une quinte flush suite à la même couleur (36)

Le problème des anagrammes :
On dispose d'un jeu de p lettres, dans lequel la lettre numéro i apparait m_i fois. Combien de mots différents pouvons-nous former en utilisant toutes les lettres ?

Exemple. Tirages avec remise. (problèmes d'urnes)

Les successions de Laplace : on se donne $n+1$ urnes numérotées de 0 à n. L'urne U_k contient k boules noires et $n-k$ blanches.
On choisit aléatoirement U_k une urne dans laquelle sont faits tous les tirages avec remise. On note B_m l'événement les m premiers tirages sont blancs et E_k l'événement l'urne tirée est U_kDéterminer les limites quand m tend vers l'infini de $P(B_m|B_{m-1})$ et $P(E_k|B_m)$.

Exemple. Utilisation de récurrences.

Déterminer le nombres de suites de longueur n ne contenant que des 0 et des 1 mais ne contenant deux 1 successifs.

Exemple. Utilisation de série génératrice :

On note p_n le nombre de permutations$f$ qui alternantes c'est à dire tel que la suite $(f(i))_{1\le i\le n}$ n'ait pas trois termes consécutifs rangés dans le même sens.
Montrer que $2p_{n+1} = \sum_0^n \binom{k}{n} p_k p_{n-k}$
Montrer que la série $\sum \frac{p_n}{n!}x^n$ a pour somme $\tan x + \frac{1}{\cos x}$ et en déduire une méthode de calcul de p_n

Variables aléatoires discrètes

I Généralités – Définitions et propriétés

I.1 Définition générale d'une variable aléatoire

Définition

Soient $(\Omega, \mathcal{A})$ et $(\Omega', \mathcal{A}')$ deux espaces probabilisables et $X : \Omega \longrightarrow \Omega'$. On dit que X est une variable aléatoire si l'image réciproque de tout événement par X est un événement. Autrement dit :

$$X \text{ est une variable aléatoire } \iff \forall B \in \mathcal{A}',\ X^{-1}(B) \in \mathcal{A}$$

Notation : Dans le cadre des variables aléatoires, l'événement $X^{-1}(B)$ est toujours noté $[X \in B]$.

Nous traduisons cette définition dans le cas particulier du programme :

I.2 Variables aléatoires discrètes

Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, soit E un ensemble, soit $X : \Omega \longrightarrow E$ une application. On dit que X est une variable aléatoire discrète lorsque :

i. $X(\Omega)$ est fini ou dénombrable

ii. $\forall y \in X(\Omega)$, $X^{-1}(\{y\})$ est un événement

Notation : l'événement $X^{-1}(\{y\})$ est toujours noté $[X = y]$. Cette définition coincide avec la précédente, en effet on a :

Proposition

Soit $X : \Omega \longrightarrow E$ une application telle que $X(\Omega)$ est fini ou dénombrable. Alors les deux propositions suivantes sont équivalentes :

i. X est une variable aléatoire discrète

ii. $\forall B \subset E$, l'ensemble $[X \in B]$ est un événement

Démonstration.
$\boxed{\Leftarrow}$ Il suffit de prendre $B = \{y\}$ où $y \in X(\Omega)$.
$\boxed{\Rightarrow}$ $[X \in B] = [X \in B \cap X(\Omega)] = \cup_{b \in B \cap X(\Omega)} [X = b]$. Or $[B \cap X(\Omega)]$ est dénombrable. $[X \in B]$ est une union dénombrable d'événements, c'est donc un événement.

Remarque. Si Ω est fini, toutes les applications définies sur Ω sont des variables aléatoires. Ce résultat est toujours valable si Ω est dénombrable et muni de la tribu $\mathcal{A} = \mathcal{P}(\Omega)$.

I.3 Premiers exemples

Indicatrices :

Soit $A \subset \Omega$ un événement. $X = \mathbb{1}_A$ est une variable aléatoire discrète.

Technique de démonstration sur un exemple

On joue à pile ou face indéfiniment. On note X l'indice du premier pile, avec éventuellement $X = +\infty$. Montrer que X est une variable aléatoire réelle discrète.

I.4 Propriétés

Proposition

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ une variable aléatoire discrète. Soit $f : E \longrightarrow F$ une application quelconque. Alors $f \circ X$ est une variable aléatoire discrète.

Proposition (vecteur aléatoire)

On se donne, pour $i \in [\![1, n]\!]$, $X_i : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E_i$ une variable aléatoire discrète. Posons :

$$X : \left| \begin{array}{lcl} (\Omega, \mathcal{A}, \mathbb{P}) & \longrightarrow & E_1 \times \cdots \times E_n \\ \omega & \longmapsto & (X_1(\omega), \ldots, X_n(\omega)) \end{array} \right.$$

Alors X est une variable aléatoire discrète.

Démonstration. $X(\Omega)$ est dénombrable, et comme chaque X_i est un variable aléatoire, $[X = (x_1, \ldots, x_n)]$ est l'intersection des $[X_i = x_i]$, c'est donc un événement.

Proposition

Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, si f est une fonction quelconque, alors $Z = f(X, Y)$ est une variable aléatoire discrète.

En particulier, $X + Y$, $X \times Y$, λX....... sont des variables aléatoires discrètes.

Ces propositions prouvent que la notion de variable aléatoire est très souple. Voici un exemple un peu plus sophistiqué pour s'en rendre compte.

Si X et Y sont deux variables aléatoires discrètes à valeurs réelles, posons :

$$Z = \begin{cases} XY \text{ si } e^X \geqslant Y \\ X - 1 \text{ sinon} \end{cases}$$

Alors Z est une variable aléatoire (en effet, on peut par exemple écrire $Z = XY \mathbb{1}_{e^X \geqslant Y} + (X-1) \mathbb{1}_{e^X < Y}$).

Dans la pratique, toutes les applications que nous rencontrerons seront des variables aléatoires.

I.5 Loi d'une variable aléatoire

Définition

Soit $X = (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ une variable aléatoire discrète. Alors en posant, pour $B \subset X(\Omega)$, $\mathbb{P}_X(B) = \mathbb{P}(\{X \in B\})$, on définit une loi de probabilité sur $X(\Omega)$ muni de la tribu $\mathcal{P}(X(\Omega))$. $\mathbb{P}_X$ s'appelle la loi de X.

Démonstration.

Proposition

La loi d'une variable aléatoire discrète est entièrement déterminée par la famille des réels $p_x = \mathbb{P}(X = x)$ pour x décrivant $X(\Omega)$ (les p_x constituent la distribution de probabilités)

Dans la pratique, pour déterminer la loi d'une variable aléatoire :

-On commence par identifier l'ensemble $X(\omega)$
-On détermine pour tout $x \in X(\omega)$ la probabilité $\mathbb{P}(X = x)$.

Exemple. On joue à pile ou face. On note p la probabilité que la pièce tombe sur pile, et $q = 1 - p$.
On note X la variable aléatoire égale au rang du premier pile. Alors :

$X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$,
pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = pq^{n-1}$
$\mathbb{P}(X = +\infty) = 0$.

La loi de X est enitèrement déterminée.

Variables identiquement distribuées
Deux variables aléatoires différentes peuvent avoir la même loi. Par exemple, pour $1 \leq i \leq 6$ notons X_i la variable aléatoire indicatrice du résultat d'un lancer de dé. La loi de X_i est donnée par :

$P(X_i = 1) = 1/6$ et $P(X_i = 0) = 5/6$ $\forall i$

Ainsi toutes les variables X_i ont la même loi, mais ces variables sont tout à fait différentes.

Notation : Lorsque deux variables X et Y ont la même loi, on notera $X \sim Y$.

Des variables de même loi sont dites identiquement distribuées.

II Espérance d'une variable aléatoire réelle

II.1 Définition

Définition (cas des variables aléatoires réelles discrètes positives)

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}_+$ une variable aléatoire réelle discrète positive. On pose dans $[0, +\infty]$:

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

$E(X)$ s'appelle l'espérance de X

Remarque.
$X(\Omega)$ est dénombrable, donc l'espérance de X est la somme de la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$. Il y a alors deux possibilités :
— ou bien la famille est sommable, et alors l'espérance et bien définie et appartient à $\mathbb{R}_+$.
— ou bien elle n'est pas sommable, et on pose alors par convention $E(X) = +\infty$.

Qualitativement : l'espérance doit être vue comme la valeur moyenne des valeurs de X (pondérées par leur probabilité d'apparition).

Pour définir l'espérance dans le cas général, on fait comme pour les familles sommables :

Définition (cas des variables aléatoires réelles discrètes)

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}$ une variable aléatoire réelle discrète.
On dit que X possède une espérance si la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable c'est à dire si

$$\boxed{\sum_{x \in X(\Omega)} |x|\mathbb{P}(X = x) < +\infty}$$

ou encore si $E(|X|) < +\infty$

Si c'est le cas on pose

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

Comme dans le cas des familles sommables, il faut faire attention à la place des valeurs absolues dans cette définition.

Exemple.
— Espérance des constantes. $\rightarrow X(\Omega) = \{c\} \rightarrow E(X) = c\,\mathbb{P}(X = c)$.
— Espérance des indicatrices. Si A est un événement, on a
$$\boxed{E(\mathbb{1}_A) = \mathbb{P}(A)}$$
— Espérance des VA presque surement nulles.
— Espérance lors d'un lancer de dé.

Variables identiquement distribuées :
L'espérance d'une variable ne dépend que de sa loi. Si $X \sim Y$ on a donc $E(X) = E(Y)$.

Un exemple

Temps d'attente du premier succès dans un pile ou face. On tire à pile ou face (avec probabilité p pour obtenir pile). On note X la variable aléatoire égale au premier instant ou l'on obtient pile. Montrer que X est presque surement finie, d'espérance finie, et calculer son espérance.

II.2 Le théorème de transfert

Théorème

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ une variable aléatoire discrète. Soit $f : E \longrightarrow \mathbb{R}$. Alors la variable aléatoire réelle discrète $f(X)$ possède une espérance si et seulement si $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas :

$$\boxed{E(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X = x)}$$

Démonstration. Formel, sous réserve de justification — justifie Sommabilité par paquets

$\to S = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X=x) = \sum_{y \in f(X(\Omega))} \left(\sum_{\substack{f(x)=y \\ x \in X(\Omega)}} f(x) \mathbb{P}(X=x) \right)$

$= \sum_y y \sum_{\substack{x \in X(\Omega) \\ (f(x)=y)}} \mathbb{P}(X=x)$ Mais, $\sum \mathbb{P}(X=x) = \mathbb{P}(f(x)=y)$

$\to \sum_x f(x)\mathbb{P}(X=x) = \sum_y y\mathbb{P}(f(X)=y) = E(f(X))$

II.3 Linéarité de l'espérance

Théorème

Soient X et Y deux variables aléatoires discrètes admettant chacune une espérance. Alors $X+Y$ admet également une espérance, et $E(X+Y) = E(X) + E(Y)$.

Démonstration. Soient $X : \Omega \longrightarrow \mathbb{R}$ et $Y : \Omega \longrightarrow \mathbb{R}$. Posons $f : (u,v) \in \mathbb{R}^2 \longmapsto u + v$. Alors, en appliquant le théorème de transfert, $f(X,Y)$ possède une espérance si et seulement si $((x+y)\mathbb{P}((X,Y) = (x,y)))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable. On note $\alpha_{x,y} = |x+y|\mathbb{P}((X,Y) = (x,y))$.

Remarquons que $\alpha_{x,y} \leqslant (|x| + |y|)\mathbb{P}(X = x, Y = y)$. Ainsi :

$$\begin{aligned}
&\sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} |x+y|\mathbb{P}(X = x, Y = y) \\
&\leqslant \sum_{x \in X(\Omega)} |x| \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y) + \sum_{y \in Y(\Omega)} |y| \sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y) \\
&= \sum_{x \in X(\Omega)} |x|\mathbb{P}(X = x) + \sum_{y \in Y(\Omega)} |y|\mathbb{P}(Y = y) \\
&= E(|X|) + E(|Y|)
\end{aligned}$$

La famille est donc bien sommable, $X+Y$ possède une espérance, et en reprenant le même calcul (sans les valeurs absolues), qui est validé par la sommabilité, on montre que $E(X+Y) = E(X) + E(Y)$.

Exemple

Temps d'attente du deuxième succès dans un pile ou face. On tire à pile ou face (avec probabilité p pour obtenir pile). On note Z la variable aléatoire égale au second instant ou l'on obtient pile. Déterminer l'espérance de cette variable aléatoire , soit directement, soit en utilisant la linéarité de l'espérance.

$\to \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}) = \{X,\ \text{tq } E(|X|) < \infty\}$

Corollaire

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. L'ensemble $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ des variables aléatoires réelles discrètes est un $\mathbb{R}$-espace vectoriel. De plus, l'application espérance est une forme linéaire sur cet espace.

De plus, la formule de transfert nous donne que $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ si et seulement si $|X| \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition

i. L'espérance est une forme linéaire positive : si X est positive, alors $E(X) \geqslant 0$. $x \geqslant y$, $E(x) \geqslant E(y)$.

ii. $|E(X)| \leqslant E(|X|)$ pour toute VA ayant une espérance.

iii. Si X est positive d'espérance nulle, alors X est presque sûrement nulle.

Démonstration.

Théorème de comparaison

Soit X et Y deux variables aléatoires réelles discrètes, avec Y positive. On suppose que $|X| \leqslant Y$ et que Y admet une espérance. Alors X admet une espérance, et $E(|X|) \leqslant E(Y)$.

voir feuille 1.

Proposition (variables bornées)

Toutes les variables aléatoires bornées discrètes possèdent une espérance, et si $|X|$ est majorée par M, alors $|E(X)| \leqslant M$.

preuve $\quad$ si $|X| \leq M$ $\quad$ (M = cste)

$E(|X|) \leq E(M) = M.$

II.4 Moments d'ordre p

Définition

Soit X une variable aléatoire réelle discrète. On dit que X admet un moment d'ordre p (pour $p \in \mathbb{N}^*$) lorsque X^p possède une espérance. Alors, $E(X^p)$ s'appelle le moment d'ordre p de X.

Proposition

Si X possède un moment d'ordre p, alors pour tout $q \in [\![1, p]\!]$, $E(|X|^q)$ existe.

Démonstration.

II.5 Moments d'ordre 2, l'espace $\mathcal{L}^2$, Cauchy-Schwarz

L'espace $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ $\longrightarrow$ précisions p. 1

Définition

On pose $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P}) = \{X,\ E(|X|^2) < +\infty\}$. Cet ensemble est un espace vectoriel.

Démonstration. Remarquons simplement que $|X+Y|^2 \leqslant 2(|X|^2 + |Y|^2)$.

Inégalité de Cauchy-Schwarz

Proposition

Considérons l'application ϕ suivante :

$$\phi : \left| \begin{array}{ccc} \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})^2 & \longrightarrow & \mathbb{R} \\ (X, Y) & \longmapsto & E(XY) \end{array} \right.$$

ϕ est une forme bilinéaire symétrique positive.

Théorème (Cauchy-Schwarz)

Soient X et Y des variables aléatoires réelles discrètes ayant un moment d'ordre 2. Alors

$$\boxed{E(XY)^2 \leqslant E(X^2)E(Y^2)}$$

Démonstration. Identique au cas des produits scalaires.

Exemples d'utilisation de l'inégalité de Cauchy Schwarz

L'inégalité de Cauchy Schwarz est encore plus importante pour les variables aléatoires que pour les fonctions. Il faut y penser dès que le carré des variables aléatoires intervient. En voici des exemples.

- Montrer que pour toute VA ayant un moment d'ordre 2 on a $E(X^2) \leqslant E(X)^2$. $\in$

$Y = 1$

- (plus subtil) Soit X une vard ayant un moment d'ordre 2 et X^+ sa partie positive. Montrer que X^+ a une espérance et que : $\quad \rightarrow \boxed{X^+ = \max(X, 0)}$.

$$\boxed{E(X^+) \leqslant \sqrt{\mathbb{P}(X \geqslant 0)E(X^2)}}$$

$\rightarrow$ faire sortir l'indicatrice.

$\hookrightarrow \mathbb{1}_{X \geq 0}$ $\quad \rightarrow$ VA de la fct $x \geq 0$.

$\hookrightarrow X^+ = X\mathbb{1}_{X\geq 0}$. $\rightarrow \mathbb{E}(X^+) = E(X\mathbb{1}_{X\geq 0}) \leq \sqrt{E(X^2)\, E(\mathbb{1}_{X\geq 0}^2)}$

Or, $\mathbb{1}_{X\geq 0}^2 = \mathbb{1}(X \geq 0)$ et $E(\mathbb{1}_{X\geq 0}) = P(X \geq 0)$ $\rightarrow$ OK

II.6 Variance

Définition

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. On pose :

$$\begin{aligned} V(X) &= E\left((X - E(X))^2\right) \\ &= E(X^2) - E(X)^2 \quad \text{(Formule de König-Huygens)} \end{aligned}$$

$V(X) \geq 0$

La variance est un indicateur de dispersion par rapport à la moyenne comme l'explique la proposition ci-dessous.

Proposition

- Si a est une constante, $V(X + a) = V(X)$.
- $V(X) = 0 \iff X$ presque sûrement constante.

De plus on a la propriété algébriques suivante :

Proposition

La variance est homogène de degré 2 :

$$\boxed{V(aX) = a^2V(X)}$$

C'est la raison pour laquelle, sa racine carrée est intéressante :

Écart-type

Définition

Soit X une vard ayant une variance. On appelle l'écart-type de X la quantité $\boxed{\sigma(X) = \sqrt{V(X)}}$.

Cet indicateur est essentiel en statistiques (on l'appelle aussi l'écart quadratique moyen)

Variables aléatoires centrées réduites

On peut toujours renormaliser une variable aléatoire en faisant des transformation affines (translation pour changer la moyenne, homothétie pour changer l'écart type) :

Définition

- On dit que $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ est centrée lorsque $E(X) = 0$.
- On dit que $X \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ est réduite lorsque $V(X) = 1$.

Proposition

Soit $X \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ une vard non presque sûrement constante. On pose :

$$\boxed{Y = \frac{X - E(X)}{\sigma(X)}}$$

Alors Y est centrée réduite.

Voir p. 1

II.7 Covariance

Définition

Soient X et Y deux variables aléatoires réelles discrètes ayant un moment d'ordre 2. On pose :

$$\boxed{\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)}$$

Proposition

La covariance est une forme bilinéaire symétrique positive sur l'ensemble des variables admettant un moment d'ordre 2 de $(\Omega, \mathcal{A}, \mathbb{P})$. Elle vérifie donc Cauchy Schwarz.

Définition

Soit X et Y deux variables admettant un moment d'ordre 2. On définit $\rho(X, Y)$, le coefficient de corrélation de (X, Y) :

$$\boxed{\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}}$$

Par Cauchy-Schwarz, le coefficient de corrélation est inférieur à 1 en module. Cet indicateur mesure si les variables sont corrélées : plus il est proche de 1 (ou -1). plus les variables sont corrélées positivement (ou négativement). Lorsqu'il est nul on dit que les VA X et Y sont **non corrélées**.

III Propriété spécifique des vard à valeurs entières

III.1 Une formule pour l'espérance

Proposition

Soit X une vard à valeurs dans $\mathbb{N}$. Alors :

$$\boxed{E(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n)}$$

preuve : dans $[0, +\infty]$, $\sum_{n=1} \mathbb{P}(X \geqslant n) = \sum \left(\sum \mathbb{P}(X = \ell) \right)$ par σ-additivité

$= \sum_{1 \leq n \leq \ell} \mathbb{P}(X = \ell)$ ([illegible]) $= \sum \left(\sum_{n=1} \mathbb{P}(X = \ell) \right) = \sum_{\ell=1} \ell \, \mathbb{P}(X = \ell)$.

Exemple : Temps d'attente du premier succès et du deuxième succès dans un pile ou face. On tire à pile ou face (avec probabilité p pour obtenir pile). On note X la variable aléatoire égale au premier instant ou l'on obtient pile. Déterminer l'espérance de cette variable aléatoire en utilisant la formule ci-dessus. (à comparer avec la méthode déjà employée).

III.2 La fonction génératrice

Définition

Soit X une vard à valeurs dans $\mathbb{N}$. Alors la série entière $\sum_{n \geqslant 0} \mathbb{P}(X = n)t^n$ converge normalement sur $[0, 1]$ (elle a donc a un rayon de convergence au moins égal à 1). La fonction :

$$\boxed{G_X : t \in [0, 1] \longmapsto \sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n}$$

s'appelle fonction (ou série) génératrice de X.

Démonstration. $R \geqslant 1$ car CVN sur $[0, 1]$

→ Dem : $\forall t \in [0, 1]$, $0 \leq P(X = n)t^n \leq P(X = n)$. Or, $\sum P(X = n)$ CV. Donc, $\sum P(X = n)t^n$ CV par comp.

L'intérêt de cette série entière est qu'elle contient toutes les informations sur la loi de la vard X.

Exemple : le lancer de dé

Soit X la variable aléatoire égale au résultat d'un lancer de dé à 6 faces. La fonction génératrice de X est :

III.3 Propriétés remarquables

Proposition

-Pour tout $t \in [0,1]$, $G_X(t) = E(t^X)$.

- G_X est continue sur $[0,1]$ fermé, en particulier au point 1. → Abel Radial.

-Si $G_X = G_Y$, alors X et Y suivent la même loi.

Démonstration.

III.4 Application à l'espérance et à la variance

→ précisés p.1

Proposition

Il est équivalent de dire :

i. $E(X) < +\infty$

ii. G_X est dérivable en 1

Dans ce cas, $\boxed{G'_X(1) = E(X)}$ (cette formule doit être connue)

Il est équivalent de dire :

i. $V(X) < +\infty$

ii. G_X est deux fois dérivable en 1

Dans ce cas, $\boxed{G''_X(1) = E(X(X-1))}$ (cette formule permet de retrouver la variance sans difficulté)

Démonstration.

Proposition (Cas général)

X possède un moment d'ordre p si et seulement si G_X est p fois dérivable en 1

La preuve se fait de la même façon.

IV Indépendance des variables aléatoires

IV.1 Définitions

Cas de deux variables aléatoires discrètes

Définition

Soient $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ et $Y : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow F$. On dit que X et Y sont indépendants si :

$$\forall x \in X(\Omega), \forall y \in Y(\Omega), [X = x] \text{ et } [Y = y] \text{ sont indépendants}$$

Cas général

Définition

Soit $(X_i)_{i \in I}$ une famille de vad. On dit qu'elles sont indépendantes lorsque $\forall n \in \mathbb{N}^*$, $\forall (i_1, \dots, i_n) \in I^n$, $\forall (a_{i_1}, \dots, a_{i_n})$, les événements $[X_{i_1} = a_{i_1}], \dots, [X_{i_n} = a_{i_n}]$ sont mutuellement indépendants.

Dans les expériences telles que les tirages de pile ou face ou de dés, les variables aléatoires associées aux différents tirages seront systématiquement supposées indépendantes.

Exemple de référence :

Dans une suite de pile ou face, on note A_n la variable aléatoire égale au numero du tirage ou l'on obtient le $n$ième pile. Montrer que les variables aléatoires A_n sont indépendantes.

IV.2 Espérance et variance

En plus des propriétés usuelles de l'espérance et de la variance on a des résultats remarquables en cas d'indépendance. Les voici :

Proposition

Si X et Y sont indépendantes et ont une espérance :

$$\boxed{E(XY) = E(X)E(Y)}$$

Démonstration.

Sous réserve

Corollaire

SI X et Y sont indépendantes et ont une variance alors :

a) $\boxed{V(X+Y) = V(X) + V(Y)}$

b) X et Y sont non corrélées : $\text{cov}(X,Y) = \rho(X,Y) = 0$ (réciproque fausse)

IV.3 Indépendance et séries génératrices

Proposition

Soient X et Y des vard à valeurs dans $\mathbb{N}$. On suppose que X et Y sont indépendantes. Alors

$$\boxed{G_{X+Y} = G_X \cdot G_Y}$$

IV.4 Lemme des coalitions

Le résultat suivant est très utile pour décider si des variables sont indépendantes.

Fonctions de variables indépendantes

Théorème

Soient X et Y deux vad indépendantes. Soient f et g deux fonctions définies sur $X(\Omega)$ et $Y(\Omega)$ réciproquement. Alors $f(X)$ et $g(Y)$ sont indépendantes.

et plus généralement :

Le lemme des coalitions

Théorème

Soient $(X_1, ..., X_n)$ une famille finie de variables aléatoires indépendantes. Soit $p \in [\![1,n]\!]$. Soient Y et Z des variables aléatoires de la forme $Y = f(X_1, ..., X_p)$ et $Z = g(X_{p+1}, ..., X_n)$. Alors Y et Z sont indépendantes.

Démonstration.

V Lois usuelles

V.1 La loi uniforme

Définition

On dit qu'une variable aléatoire X suit la loi uniforme $U(n)$ lorsque :

i. $X(\Omega) = [\![1,n]\!]$

ii. $\forall k \in [\![1,n]\!]$, $\mathbb{P}(X=k) = 1/n$

Proposition

— $G_X(t) = \frac{1}{n}(t + \cdots + t^n)$

— $E(X) = \frac{n+1}{2}$

— $V(X) = \frac{n^2-1}{12}$

l'exemple de référence est le lancer de dé.

Remarque. On peut également considérer la loi uniforme $U([\![a,b]\!])$, en opérant une translation de $a-1$ (qui va affecter la moyenne mais pas la variance).

Notation : à la place de l'expression "X suit la loi uniforme $U(n)$" on écrit de façon simplifiée $X \hookrightarrow U(n)$. On fera de même pour toutes les lois usuelles.

V.2 La loi de Bernoulli

Définition

Soit $p \in [0,1]$ et $q = 1-p$. On dit que $X \hookrightarrow B(p)$ lorsque :

i. $X(\Omega) = \{0,1\}$

ii. $\mathbb{P}(X=1) = p$

Proposition

— $G_X(t) = pt + q$

— $E(X) = p$

— $V(X) = pq$

La loi de Bernoulli modélise les tirages de pile ou face, ou plus généralement, tous les tirages ayant deux issues. Elle est très importante.

Variante de la loi de Bernoulli . La loi de **Rademacher** est donnée par :

i. $X(\Omega) = \{-1,1\}$

ii. $\mathbb{P}(X=1) = \frac{1}{2}$

Cette loi sert à modéliser les marches aléatoires symétriques (on diminue ou augmente de 1 avec équiprobabilité). Cette loi n'a pas de série génératrice car elle n'est pas à valeur dans $\mathbb{N}$, mais elle a espérance et variance bien entendu.

V.3 La loi binomiale $B(n,p)$

Proposition

Soit $(X_1,\dots,X_n)$ un n−uplet de variables indépendantes suivant la loi $B(p)$.
Alors la variable aléatoire $X = X_1 + \cdots + X_n$ vérifie :

i. $X(\Omega) = [\![0,n]\!]$

ii. $\mathbb{P}(X=k) = \binom{n}{k} p^k q^{n-k}$ pour $k \in [\![1,n]\!]$ avec $q = 1-p$.

On dit que la variable X suit une loi binomiale de paramètres (n,p) :

$$\boxed{X \hookrightarrow B(n,p)}$$

Démonstration. On conclut par théorème d'unicité.

La loi $B(n,p)$ modélise le nombre de succès dans la répétition de n expériences de Bernoulli indépendantes, de paramètre p.

Proposition

— $G_X(t) = (pt+q)^n$
— $E(X) = np$
— $V(X) = npq$

Additivité des variables binomiales indépendantes

Proposition

Soit (X,Y) un couple de variables indépendantes avec $X \hookrightarrow B(n,p)$ et $Y \hookrightarrow B(m,p)$. Alors :

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

V.4 La loi géométrique

Définition

On dit qu'une variable aléatoire X définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ suit $\mathcal{G}(p)$, la loi géométrique de paramètre p lorsque :

i. $X(\Omega) = \mathbb{N}^*$

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

Elle modélise le premier succès dans la répétition d'expériences de Bernoulli indépendantes de paramètre p.

Proposition

— $G_X(t) = \frac{pt}{1-qt}$
— $E(X) = \frac{1}{p}$
— $V(X) = \frac{q}{p^2}$

Proposition (loi sans mémoire)

La loi géométrique est caractérisée par la propriété suivante (on dit qu'elle est sans mémoire)
Pour tout n, pour tout k,

$$\mathbb{P}(X > n+k | X > n) = \mathbb{P}(X > k)$$

Réciproquement, si X est non presque sûrement constante avec $X(\Omega) \subset \mathbb{N}^*$, alors il existe p tel que $X \hookrightarrow \mathcal{G}(p)$.

V.5 La loi de Poisson

Définition

Soit $\lambda \in \mathbb{R}_+^*$, on dit que la variable aléatoire X suit une loi de Poisson de paramètre λ ($X \hookrightarrow \mathcal{P}(\lambda)$) lorsque :

i. $X(\Omega) = \mathbb{N}$

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

Les lois de Poisson modélisent généralement les files d'attentes, ou le nombre de succès dans un laps de temps fixe.

Proposition

- $G_X(t) = e^{\lambda(t-1)}$
- $E(X) = V(X) = \lambda$

Additivité des variables de Poisson

Proposition

Soit (X, Y) un couple de variables indépendantes avec $X \hookrightarrow \mathcal{P}(\lambda)$ et $Y \hookrightarrow \mathcal{P}(\mu)$. Alors :

$$\boxed{X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)}$$

Proposition (Approximation de la loi binomiale pour les événements rares)

Soit (X_n) une suite de variables aléatoires suivant la loi $\mathcal{B}(n, p_n)$. Supposons que $np_n \xrightarrow[n\to+\infty]{} \lambda$.
Alors $\mathbb{P}(X_n = k) \xrightarrow[n\to+\infty]{} \mathbb{P}(X = k)$ avec $X \hookrightarrow \mathcal{P}(\lambda)$.

La loi de Poisson permet de faire une approximation de la loi binomiale pour les grandes valeurs de n et les petites valeurs de p avec le paramètre $\lambda = np$
Par exemple : On tire 520 fois une carte avec remise dans un jeu de 52. Quelle est la probabilité d'obtenir 10 fois l'as de pique?

Ici $n = 520; p = 1/52$ la probabilité cherchée est $\frac{(51)^{510}}{52^{520}}\binom{520}{10}$ qui vaut environ 0.126330 l'approximation de Poisson est $\frac{10^{10}}{10!}e^{-10}$. La formule de Stirling approxime ceci en $\frac{1}{\sqrt{20\pi}}$ qui vaut environ 0.126156

VI Vecteurs aléatoires, lois conjointes, lois marginales

VI.1 Loi d'un couple

Définition

Soient X, Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle loi conjointe des variables aléatoires la loi du couple (X, Y) donnée par $(\mathbb{P}(X = x, Y = y))_{x \in X(\Omega), y \in Y(\Omega)}$.
On appelle lois marginales les lois de X et de Y.

On peut ceci sous forme d'un tableau :

	x_1	...	x_i	...	x_n	$X(\Omega)$
y_1						
$\vdots$						
y_j			$P(X = x_i, Y = y_j)$			$P(Y = y_j)$
$\vdots$						
y_p						
$Y(\Omega)$			$P(X = x_i)$			1

En sommant sur les lignes on trouve la loi de Y, en sommant sur les colonnes on trouve la loi de X.
La somme de tous les éléments du tableau vaut 1.

Proposition

La loi conjointe détermine les lois marginales, mais la réciproque est fausse.

preuve Les événements $(Y = y)$ pour $y \in Y(\Omega)$ forment un système complet.
Donc : $\forall x \in X(\Omega), \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x | Y = y)\mathbb{P}(Y = y)$.
$= \sum_{y \in Y(\Omega)} \mathbb{P}((X = x) \cap (Y = y)) = \sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y))$.

VI.2 Caractérisation de l'indépendance

Proposition

Si X et Y sont indépendantes, les lois marginales déterminent les lois conjointes

Plus précisément le tableau ci dessus est entièrement déterminé par la dernière colonne et la dernière ligne, puisque $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$.

Ceci est faux si les variables ne sont pas indépendantes.

Cette définition s'étend sans difficulté :

Définition (cas de n variables)

Soit $(X_1, \ldots, X_n)$ un n-uplet de variables aléatoires discrètes. La loi conjointe est donnée par $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)$ pour $(x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$.

Dans ce cas aussi, la loi conjointe est déterminée par les lois marginales dans le cas des variables indépendantes, mais pas en général.

VI.3 Loi conditionnelle

Définition

Soit $A \in \mathcal{A}$. On appelle loi conditionnelle de X relativement à l'événement A la loi de probabilité $\mathbb{P}_X(\cdot|A)$ définie par :

$$\boxed{\mathbb{P}_X(B|A) = \mathbb{P}(X \in B|A)}$$

→ pas très important.

VI.4 Etude complète d'un exemple

Variante autour de l'exercice 98 du poly d'oral CCP.
Soit n un entier. On effectue n épreuves de Bernoulli de paramètre p. Soit X le nombre de succès. On effectue ensuite $n - X$ épreuves de Bernoulli. On note Y le nombre de nouveaux succès.
Déterminer :
a) La loi de X
b) La loi de Y conditionnellement à l'événement $X = i$
c) La loi conjointe du couple (X, Y).
d) si les variables X, Y sont indépendantes ou non.
e) La loi de Y.

VII Inégalité de Markov, de Bienaymé-Tchebychev, loi faible des grands nombres

VII.1 Inégalité de Markov

Proposition

Soit X une variable aléatoire à valeurs positives et ayant une espérance. Alors :

$$\boxed{\forall a \in \mathbb{R}_+, \mathbb{P}(X \geqslant a) \leqslant \frac{E(X)}{a}}$$

Démonstration. $X = X 1\!\!1_{X\geq a} + X 1\!\!1_{X<a}$.

$\to E(X) = E(X 1\!\!1_{X\geq a}) + E(X 1\!\!1_{X<a})$.

$X 1\!\!1_{X\geq a} \geq a 1\!\!1_{X\geq a} \to E(X 1\!\!1_{X\geq a}) \geq E(a 1\!\!1_{X\geq a}) = a\mathbb{P}(X\geq a)$

$E(X 1\!\!1_{X<a}) \geq 0 \to E(X) \geq a\mathbb{P}(X \geq a)$.

Remarque. On utilise souvent une amélioration de l'inégalité de Markov :

Proposition

Si f est une fonction strictement croissante et strictement positive telle que $f(X)$ est d'espérance finie alors :

$$\boxed{\mathbb{P}(X \geqslant a) \leqslant \frac{E(f(X))}{f(a)}}$$

VII.2 Inégalité de Bienaymé-Tchebychev

Proposition

Soit X une variable aléatoire ayant un moment d'ordre 2. Soit $\epsilon > 0$.

$$\boxed{\mathbb{P}(|X - E(X)| \geqslant \epsilon) \leqslant \frac{V(X)}{\epsilon^2}}$$

Démonstration.

Ces inégalités sont utilisées pour contrôler l'écart entre une variable aléatoire et sa moyenne. Nous en verrons des utilisations en exercice.
Par exemple : on tire 1000 fois à pile ou face (donc en moyenne 500 pile). Majorer la probabilité de faire au moins 550 pile par l'inégalité de Bienaymé Tchebychev.

L'application théorique la plus importante de ces inégalités est le résultat ci-dessous.

VII.3 Loi faible des grands nombres

On se donne une variable aléatoire X qui modélise une certaine expérience. On suppose l'existence d'une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires définies sur un espace probabilisé telle que :

- pour tout $n \in \mathbb{N}$, $X_n \sim X$
- les X_n sont mutuellement indépendantes.

On dit alors que (X_n) est une suite de variables aléatoires indépendantes identiquement distribuées (iid).

Théorème (Loi faibles des grands nombres)

Soit (X_n) une suite de variables aléatoires discrètes indépendantes identiquement distribuées suivant la loi de X. On suppose que X possède une variance. Alors :

$$\boxed{\forall \epsilon > 0,} \boxed{\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - E(X)\right| \geqslant \epsilon\right) \xrightarrow[n\to+\infty]{} 0}$$

Démonstration. $Y_n = \frac{X_1 + \dots + X_n}{n}$ alors comme $X_1, \dots, X_n$ ont une variance, Y_n possède une variance. De +, $E(Y_n) = E\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} \underbrace{E(X_k)}_{=E(X)}$

$= E(X)$ et d'après B.T $\to \mathbb{P}(|Y_n - E(Y_n)| \geq \varepsilon) \leq \frac{V(Y_n)}{\varepsilon^2}$.

$V(Y_n) = \frac{1}{n}V(X)$. Par suite : $V(Y_n) \xrightarrow[n\to\infty]{} 0$)

VII.4 Types de convergences(HP)

Soit (X_n) une suite de variables aléatoires réelles discrètes. On distingue plusieurs type de convergences de (X_n) vers une éventuelle variable aléatoire limite X :

- La convergence en loi :

$$\forall (a, b) \in \mathbb{R}^2, \mathbb{P}(X_n \in [a, b]) \xrightarrow[n\to+\infty]{} \mathbb{P}(X \in [a, b])$$

- La convergence en probabilité :

$$\forall \epsilon > 0, \mathbb{P}(|X_n - X| \geqslant \epsilon) \xrightarrow[n\to+\infty]{} 0$$

- La convergence presque sûre :

$$\mathbb{P}\left(X_n \xrightarrow[n\to+\infty]{} X\right) = 1$$

L'étude de ces différentes notion n'est pas au programme. Avec ce vocabulaire, la loi faible des grands nombres exprime que la suite de variables aléatoires $\frac{X_1+\cdots+X_n}{n}$ converge en probabilité vers sa moyenne. C'est un fait très intuitif.

A titre d'exercice, comment s'interprète l'approximation d'une loi binomiale par une loi de Poisson en termes de convergences de variables aléatoires ?