CHAPITRE XV: VARIABLES ALÉATOIRES DISCRÈTES

I) Généralités

I-1) Variable aléatoire

Définition : Variable aléatoire discrète

Si Ω est un univers muni d'une tribu $\mathscr{A}$, une variable aléatoire discrète est une application X définie sur Ω telle que :

- $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$ est un ensemble fini ou dénombrable
- Si $x \in X(\Omega)$ alors $X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\} = (X = x)$ est un événement c'est à dire un élément de $\mathscr{A}$

Remarques :
- L'univers Ω n'est en général pas explicité.
- Les notions d'image et d'image réciproque sont celles associées à une application quelconque

- $X(\Omega) = \left\{X(\omega) \mid \omega \in \Omega\right\}$ est l'image de l'application X. Pour une variable aléatoire discrète X, puisque $X(\Omega)$ est au plus dénombrable, on écrira : $\boxed{X(\Omega) = \left\{x_n \mid n \in \mathbb{N}\right\}}$

  (éventuellement la suite $(x_n)_{n\in\mathbb{N}}$ ne contient qu'un nombre fini de valeurs distinctes lorsque $X(\Omega)$ est un ensemble fini)
- Si $A \subset X(\Omega)$, $X^{-1}(A) = \left\{\omega \in \Omega \mid X(\omega) \in A\right\}$ est l'image réciproque de A par X

  (qui existe même si X n'est pas bijective...Ne pas confondre avec l'application réciproque X^{-1} qui n'existe que si X est bijective !)

  Dans le cadre probabiliste, on utilise plutôt les notations

  $\boxed{(X \in A) \text{ ou } \{X \in A\}}$ pour l'événement à la place de $X^{-1}(A)$ pour $A \subset X(\Omega)$

  $\boxed{(X = x) \text{ ou } \{X = x\}}$ pour l'événement $X^{-1}\left(\{x\}\right)$ pour $x \in X(\Omega)$

  $\boxed{(X \geqslant x)}$ pour l'événement $X^{-1}\left([x, +\infty[\right)$ si X est à valeurs réelles (càd $X(\Omega) \subset \mathbb{R}$)

Il est peu probable qu'un sujet vous demande de justifier qu'une application est une variable aléatoire : ce sera admis ou même pas évoqué par le sujet.

Remarques importantes : Si X est une variable aléatoire discrète sur l'espace probabilisable $(\Omega, \mathscr{A})$ avec
$X(\Omega) = \left\{x_n \mid n \in \mathbb{N}\right\}$ où les $(x_n)_{n\in\mathbb{N}}$ sont 2 à 2 distincts
alors on associe naturellement à X le système complet d'événements $\Big((X = x_n)\Big)_{n\in\mathbb{N}}$.

Il s'agit de vérifier que :

- les événements $\Big((X = x_n)\Big)_{n\in\mathbb{N}}$ sont 2 à 2 incompatibles :
  Pour $n \neq m$, $\omega \in (X = x_n) \cap (X = x_m) \Rightarrow X(\omega) = x_n = x_m$ ce qui est absurde puisque $x_n \neq x_m$
- $\displaystyle\bigcup_{n\in\mathbb{N}} (X = x_n) = \Omega$: l'inclusion $\subset$ est triviale par définition de $(X = x_n)$ qui est bien une partie de Ω
  Si $\omega \in \Omega$ alors $X(\omega) \in X(\Omega) = \left\{x_n \mid n \in \mathbb{N}\right\}$ donc $\exists n \in \mathbb{N}, X(\omega) = x_n \Leftrightarrow \exists n \in \mathbb{N},\ \omega \in (X = x_n) \Leftrightarrow \omega \in \displaystyle\bigcup_{n\in\mathbb{N}} (X = x_n)$

I-2) Loi d'une variable aléatoire discrète

Définition : Loi d'une variable aléatoire

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,

on appelle loi de probabilité (ou simplement la loi de X) l'application $\left[\begin{array}{rcl} \mathrm{P}_{\mathrm{X}} : \mathrm{X}(\Omega) & \rightarrow & [0,1] \\ x & \mapsto & \mathrm{P}(\mathrm{X}=x) \end{array}\right]$

Lorsque $\mathrm{X}(\Omega) = \{x_n \mid n \in \mathbb{N}\}$, la loi de X consiste donc à donner une suite $(p_n)_{n\in\mathbb{N}}$ telle que : $\forall n \in \mathbb{N},\ p_n = \mathrm{P}(\mathrm{X}=x_n)$

Si X et Y sont deux variables aléatoires qui suivent la même loi, on notera $\mathrm{X} \sim \mathrm{Y}$

Notons que si $\mathrm{X}(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ où les $(x_n)_{n\in\mathbb{N}}$ sont 2 à 2 distincts alors

$$\sum_{n=0}^{+\infty} p_n = \sum_{n\in\mathbb{N}} \mathrm{P}(\mathrm{X}=x_n) =_{\sigma\text{-additivité}} \mathrm{P}\left(\bigcup_{n\in\mathbb{N}} (\mathrm{X}=x_n)\right) = \mathrm{P}\big(\mathrm{X} \in \mathrm{X}(\Omega)\big) = \mathrm{P}(\Omega) = 1$$

Proposition : Variable aléatoire $f(\mathrm{X})$

Soient X et Y des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,
alors, pour toute fonction f définie sur $\mathrm{X}(\Omega)$, $f(\mathrm{X})$ et $f(\mathrm{Y})$ sont encore des variables aléatoires.

Si X et Y ont la même loi alors $f(\mathrm{X})$ et $f(\mathrm{Y})$ ont la même loi soit : $\mathrm{X} \sim \mathrm{Y} \Rightarrow f(\mathrm{X}) \sim f(\mathrm{Y})$

EXEMPLE N° 1

- On lance un dé cubique non truqué trois fois.
  Si X est le nombre de fois où le résultat est as ou deux et Y le nombre de résultats plus grand que cinq ,

  les variables aléatoires X **et** Y **ont la même loi mais ne sont pas égales !**

  D'abord $\mathrm{X}(\Omega) = \mathrm{Y}(\Omega) =$
  Pour déterminer la loi de X et Y, on reconnaît dans cette expérience aléatoire

  Ainsi, X et Y suivent donnée par :

- Déterminer la loi de la variable $\mathrm{U} = |\mathrm{X}-2|$

- On sait que Z est une variable aléatoire à valeurs dans $\mathbb{N}$ telle que : $\forall n \in \mathbb{N},\ 3\mathrm{P}(\mathrm{Z}=n+2) = 4\mathrm{P}(\mathrm{Z}=n+1) - \mathrm{P}(\mathrm{Z}=n)$
  Déterminer la loi de Z et donner la probabilité de l'événement (Z est pair)

II) Lois usuelles

Pour chacune des lois à connaître du programme de PT, il s'agit de :

- connaître la loi autrement dit, si X est une variable aléatoire réalisant cette loi :

$$X(\Omega) = ? \quad \text{et} \quad \forall x \in X(\Omega),\ P(X = x) = ?$$

- connaître aussi des schémas d'expériences aléatoires réalisant cette loi.

II-1) Loi uniforme sur un ensemble fini

Définition : Loi uniforme sur $[\![1, n]\!]$ où $n \in \mathbb{N}^*$

La variable aléatoire X suit une loi uniforme sur $[\![1, n]\!]$ ce qu'on note $X \sim \mathscr{U}\big([\![1, n]\!]\big)$ lorsque

$$X(\Omega) = [\![1, n]\!] \quad \text{et} \quad \forall k \in [\![1, n]\!],\ P(X = k) = \frac{1}{n}$$

Schémas type de réalisation :

-
-

II-2) Loi de Bernoulli

Définition : Loi de Bernoulli de paramètre p

La variable aléatoire X suit une loi de Bernoulli de paramètre $p \in]0, 1[$ ce qu'on note $X \sim \mathscr{B}(p)$ lorsque

$$X(\Omega) = \{0, 1\} \quad \text{et} \quad P(X = 1) = p = 1 - P(X = 0)$$

Plus généralement, cette loi apparaît dans les expériences qui n'ont que deux issues :
l'une qu'on appelle succès associé à 1 et l'autre l'échec associé à 0, le paramètre p est la probabilité de succès.

Schémas type de réalisation :

-
-
-
-

II-3) Loi binomiale

Définition : Loi binomiale de paramètre n et p

La variable aléatoire X suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0, 1[$ noté $X \sim \mathscr{B}(n, p)$ lorsque

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

Cette loi apparaît dans les expériences où on dénombre
le nombre de succès dans une répétition de n expériences de Bernoulli identiques et indépendantes de paramètre p.

On peut ainsi écrire $\boxed{X = X_1 + X_2 + \cdots + X_n \quad \text{où } (X_i)_{1 \leqslant i \leqslant n} \text{ mutuellement indépendantes de même loi } \mathscr{B}(p)}$

où X_i est le résultat de la ième expérience de Bernoulli dans la construction du schéma binomial

Schémas type de réalisation :

-
-
-

II-4) *Loi géométrique*

Définition : Loi géométrique de paramètre p

La variable aléatoire X suit une loi géométrique de paramètre $p \in]0,1[$ ce qu'on note $X \sim \mathscr{G}(p)$ lorsque

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

Cette loi apparaît dans les expériences où on recherche
le numéro du premier succès dans une répétition d'expériences de Bernoulli identiques et indépendantes de paramètre p.

On peut ainsi écrire

$X = \min\{k \in \mathbb{N}^* \mid X_k = 1\}$ où $(X_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoire indépendante et de même loi $\mathscr{B}(p)$

Schémas type de réalisation :

-
-
-

II-5) *Loi de Poisson*

Définition : Loi de Poisson de paramètre $\lambda > 0$

La variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$ ce qu'on note $X \sim \mathscr{P}(\lambda)$ lorsque

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

Schémas type de réalisation : Pas de modèle simple de réalisation de cette loi.

• Un bon exemple est celui du compteur Geiger pour la radioactivité : sur une plage donnée (1 unité de temps), il compte le nombre de désintégration radioactive. Ce décompte est une réalisation d'une loi de Poisson.
Unité de temps après unité de temps, on obtient des réalisations successives de la loi de Poisson.

• On dispose d'un mécanisme (imaginaire) qui décompte, par une unité de temps, le nombre de goutte de pluie qui tombe sur une petit plaque. Ce décompte est une réalisation d'une loi de Poisson.

La loi de Poisson est appelée **la loi des événements rares** : elle "correspond à une loi limite" de binomiale avec un très grand nombre n de répétitions de probabilité p petite de sorte que $np = \lambda$ tend à être fixe. Il s'agit donc du nombre de succès lors d'un grand nombre de répétitions d'événements ayant une probabilité faible d'intervenir.

Le théorème démontrant **la loi des événements rares** n'est plus au programme

III) Couple et suites de variables aléatoires discrètes

Il s'agit ici de généraliser le vocabulaire des couples de variables aléatoires vu dans le cas des univers finis au cas des variables aléatoires discrètes.

III-1) Loi conjointe, lois marginales et lois conditionnelles

Définitions : Loi conjointe, lois marginales et lois conditionnelles pour un couple de variables aléatoires discrètes

Si X et Y sont deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$

- le couple (X, Y) est une variable aléatoire discrète à valeurs dans $\mathrm{X}(\Omega) \times \mathrm{Y}(\Omega)$
- on définit **la loi conjointe** du couple (X, Y) par la donnée de
  - l'ensemble $\mathrm{X}(\Omega) \times \mathrm{Y}(\Omega)$ qui contient les valeurs prises par le couple (X, Y)
  - la valeur des probabilités : $\forall (x, y) \in \mathrm{X}(\Omega) \times \mathrm{Y}(\Omega),\ \mathrm{P}\big((\mathrm{X}, \mathrm{Y}) = (x, y)\big) = \mathrm{P}(\mathrm{X} = x, \mathrm{Y} = y) = \mathrm{P}\Big((\mathrm{X} = x) \cap (\mathrm{Y} = y)\Big)$
- **la loi marginale** de X (resp. de Y) est la loi de la variable aléatoire X (resp.Y)
  <u>On peut déterminer les lois marginales à partir de la loi conjointe</u> :
  $\forall x \in \mathrm{X}(\Omega),\ \mathrm{P}(\mathrm{X} = x) = \sum\limits_{y \in \mathrm{Y}(\Omega)} \mathrm{P}(\mathrm{X} = x, \mathrm{Y} = y) \qquad \Big(\text{resp.} \quad \forall y \in \mathrm{Y}(\Omega),\ \mathrm{P}(\mathrm{Y} = y) = \sum\limits_{x \in \mathrm{X}(\Omega)} \mathrm{P}(\mathrm{X} = x, \mathrm{Y} = y)\Big)$
  <u>On ne peut pas, en général, déterminer la loi conjointe à partir des lois marginales</u> (sauf si X et Y indépendantes)
- Pour $y \in \mathrm{Y}(\Omega)$ avec $\mathrm{P}(\mathrm{Y} = y) > 0$,
  la loi conditionnelle de X **sachant** $(\mathrm{Y} = y)$ est la probabilité $\left[\begin{array}{rcl} \mathrm{P}_{(\mathrm{Y}=y)} : \mathrm{X}(\Omega) & \to & [0, 1] \\ x & \mapsto & \mathrm{P}(\mathrm{X} = x | \mathrm{Y} = y) \end{array}\right]$
  On définit de manière analogue la loi conditionnelle de Y sachant $(\mathrm{X} = x)$ pour $x \in \mathrm{X}(\Omega)$ avec $\mathrm{P}(\mathrm{X} = x) > 0$.
  <u>Si on connaît les lois conditionnelles, on peut retrouver la loi conjointe</u> : (formule des probabilités composées)
  $$\forall (x, y) \in \mathrm{X}(\Omega) \times \mathrm{Y}(\Omega), \quad \mathrm{P}(\mathrm{X} = x, \mathrm{Y} = y) = \mathrm{P}(\mathrm{Y} = y) \times \mathrm{P}_{(\mathrm{Y}=y)}(\mathrm{X} = x) = \mathrm{P}(\mathrm{X} = x) \times \mathrm{P}_{(\mathrm{X}=x)}(\mathrm{Y} = y)$$

Lorsque X et Y sont des variables aléatoires prenant un nombre fini de valeurs, on représente en général cette loi à l'aide d'un tableau à double entrée : Si $\mathrm{X}(\Omega) = \{x_1, \dots, x_n\}$ et $\mathrm{Y}(\Omega) = \{y_1, \dots, y_p\}$ alors

$(\mathrm{X} = x_i) / (\mathrm{Y} = y_j)$	y_1	y_2	$\dots$	y_j	$\dots$	y_p	$\mathrm{P}(\mathrm{X} = x_i)$
x_1							
x_2							
$\vdots$							
x_i				$\mathrm{P}(\mathrm{X} = x_i, \mathrm{Y} = y_j)$			
$\vdots$							
x_n							
$\mathrm{P}(\mathrm{Y} = y_j)$							1

Lorsque $\mathrm{X}(\Omega)$ et $\mathrm{Y}(\Omega)$, on peut "théoriquement" imaginer ce tableau sans pour autant pourvoir le représenter.

EXEMPLE N° 2 On dispose de 6 urnes contenant chacune 6 boules indiscernables au toucher telles que l'urne n°i contient i boules blanches et les autres sont noires.
On lance un dé équilibré qui fixe le numéro d'un urne et on tire des boules dans cette urne avec remise jusqu'à obtenir une boule blanche.
On appelle N la variable aléatoire correspondant au numéro de l'urne et X celle donnant le nombre de tirages nécessaires pour obtenir la boule blanche. Déterminer la loi du couple (N, X).

EXEMPLE N° 3 **(D'après oral maths 2)**

Soient $a \in \mathbb{R}$ et (X, Y) un couple de variables aléatoires à valeurs dans $\mathbb{N}^2$ dont la loi est donnée par :

$$\forall j \in \mathbb{N}, \forall k \in \mathbb{N}, \quad P(X = j, Y = k) = a\frac{j+k}{3^{j+k}}$$

1. Pour $k \in \mathbb{N}$ fixé, justifier la convergence et calculer la somme S_k de la série $\sum_{j \geqslant 0} \frac{j+k}{3^{j+k}}$

2. Déterminer alors la valeur de a

3. Déterminer les lois marginales de X et Y.

III-2) *Variables aléatoires indépendantes*

Définition et propositions : Couple de variables aléatoires indépendantes

Si X et Y sont deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathscr{A}, P)$,

X et Y sont indépendantes, notées $X \perp\!\!\!\perp Y$ lorsque $\forall A \subset X(\Omega), \forall B \subset Y(\Omega), \quad P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$

On a : $X \perp\!\!\!\perp Y \Leftrightarrow \forall (x, y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants

soit : $P(X = x, Y = y) = P(X = x)P(Y = y)$

SUITE EXEMPLE N° 3 Les variables X et Y sont-elles indépendantes ? Calculer $P(X = Y)$

Définition et proposition : Variables aléatoires (mutuellement) indépendantes

• Si $X_1, X_2, \ldots, X_n$ sont des variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathscr{A}, P)$,

on dit que $X_1, X_2, \ldots, X_n$ sont (mutuellement) indépendantes si, pour $(x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$,

les événements $\big((X_i = x_i)\big)_{i \in [\![1,n]\!]}$ sont (mutuellement) indépendants : $\forall J \subset [\![1, n]\!],\ P\left(\bigcap_{j \in J}(X_j = x_j)\right) = \prod_{j \in J} P(X_j = x_j)$

En particulier :

- des variables mutuellement indépendantes sont donc aussi 2 à 2 indépendantes

(mais l'indépendance 2 à 2 n'entraîne pas l'indépendance mutuelle)

- des variables mutuellement indépendantes vérifient $P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$

(mais cette seule égalité ne suffit pas à prouver l'indépendance mutuelle)

On admet que :

Proposition :

Si $X_1, X_2, \ldots, X_n$ sont mutuellement indépendantes alors

$\forall i \in [\![1, n]\!], \forall A_i \subset X_i(\Omega)$, les événements $\big(X_i \in A_i\big)_{1 \leqslant i \leqslant n}$ sont mutuellement indépendants

Définition : Suite de variables aléatoires indépendantes, suite de variables aléatoire i.i.d

• Les variables aléatoires d'une suite $(X_n)_{n \in \mathbb{N}}$ de variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ sont indépendantes lorsque, pour tout $n \in \mathbb{N}$, les variables aléatoires $(X_k)_{k \in [\![0,n]\!]}$ sont mutuellement indépendantes.

• Une suite $(X_n)_{n \in \mathbb{N}^*}$ de variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ est i.i.d

(independant and identically distribued)

lorsque toutes les variables ont la même loi et que les variables de la suite $(X_n)_{n \in \mathbb{N}^*}$ sont indépendantes.

Exemple : Un jeu de pile ou face avec une pièce truqué (qui ne s'arrête jamais) est une suite $(X_n)_{n \in \mathbb{N}*}$ de variables aléatoires i.i.d. de loi de Bernoulli de paramètre p.

Proposition : Fonctions de variables indépendantes

• Si X et Y sont indépendantes et si f et g des applications définies respectivement sur $X(\Omega)$ et $Y(\Omega)$ alors

$f(X)$ et $g(Y)$ sont des variables aléatoires discrètes également indépendantes

ainsi : $X \perp\!\!\!\perp Y \Rightarrow f(X) \perp\!\!\!\perp g(Y)$

• Si $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes et $(f_n)_{n \in \mathbb{N}^*}$ une suite de fonctions où f_n est définie sur $X_n(\Omega)$ alors $\big(f_n(X_n)\big)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes.

Proposition : Lemme des coalitions

Si $(X_k)_{k\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes alors

• Pour $(m,n)\in\mathbb{N}^2$ avec $1\leqslant m<n$ alors les variables aléatoires $f(X_1,\dots,X_m)$ et $g(X_{m+1},\dots,X_n)$ sont indépendantes pour toutes fonctions f et g définies sur les espaces adéquats

• Pour $(m_1,m_2,\dots,m_n)\in\mathbb{N}^n$ avec $1\leqslant m_1<m_2<\dots<m_n$ avec $n\in\mathbb{N}^*$ alors les variables aléatoires $f_1(X_1,\dots,X_{m_1}),\ f_2(X_{m_1+1},\dots,X_{m_2}),\ \dots\ ,\ f_n(X_{m_{n-1}+1},\dots,X_{m_n})$ sont indépendantes pour toutes fonctions $f_1,f_2,\dots,f_{m_n}$ définies sur les espaces adéquats

EXEMPLE N° 4 On considère X, Y et Z trois variables aléatoires indépendantes et de même loi uniforme sur $[\![1,n]\!]$.

1. Donner la loi du couple (X,Y)
2. Démontrer que : $P(X+Y=k)=\dfrac{k-1}{n^2}$ si $k\in[\![2,n+1]\!]$ et $P(X+Y=k)=\dfrac{2n-k+1}{n^2}$ si $k\in[\![n+1,2n]\!]$
3. Démontrer que $P(X+Y=Z)=\dfrac{n-1}{2n^2}$
4. Quelle est la loi de $T=n+1-Z$. Que dire des variables $X+Y$ et T ? En déduire, sans calcul, $P(X+Y+Z=n+1)$.

IV) Espérance, variance et covariance

IV-1) Variable aléatoire discrète d'espérance finie

En PTSI, l'espérance d'une variable aléatoire réelle X avec $X(\Omega) = \{x_1, \dots, x_n\}$ a été définie par :

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

Si on veut généraliser la définition lorsque $X(\Omega)$ est un ensemble dénombrable, il y a deux difficultés :

- la série $\sum\limits_{x\in X(\Omega)} x P(X = x)$ converge-t-elle ?
- l'ordre dans lequel on somme les termes a-t-il une influence ? Attention, cette question n'est pas triviale.

  Considérons par exemple la série harmonique alternée $\sum\limits_{n\geqslant 1} \dfrac{(-1)^n}{n}$ (étudié dans le chapitre série) :

  Nous avions démontré que $\sum\limits_{n=1}^{+\infty} \dfrac{(-1)^n}{n} = -\ln 2$ soit $-1 + \dfrac{1}{2} - \dfrac{1}{3} + \dfrac{1}{4} + \cdots - \dfrac{1}{2k-1} + \dfrac{1}{2k} + \cdots = -\ln 2$

  Pourtant :

$$\begin{aligned}\left(-1+\frac{1}{2}+\frac{1}{4}\right)+\left(-\frac{1}{3}+\frac{1}{6}+\frac{1}{8}\right)+\cdots+\left(-\frac{1}{2k-1}+\frac{1}{4k-2}+\frac{1}{4k}\right)+\cdots &= \left(-\frac{1}{2}+\frac{1}{4}\right)+\left(-\frac{1}{6}+\frac{1}{8}\right)+\cdots+\left(-\frac{1}{4k-2}+\frac{1}{4k}\right)+\dots \\ &= \frac{1}{2}\left(-1+\frac{1}{2}\right)+\frac{1}{2}\left(-\frac{1}{3}+\frac{1}{4}\right)+\cdots+\frac{1}{2}\left(-\frac{1}{2k-1}+\frac{1}{2k}\right)+\dots \\ &= \frac{1}{2}\times -\ln 2\end{aligned}$$

  Lorsqu'on somme une série, l'ordre de sommation a donc de l'importance !

  Nous admettrons que l'absolue convergence de la série conduit à assurer que la somme de dépend de l'ordre de sommation des termes. Cela nous amène donc à la définition suivante pour l'espérance d'une variable aléatoire discrète :

Définition : Variable aléatoire réelle discrète d'espérance finie et calcul de son espérance

Si X est une variable aléatoire réelle discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$ avec

$$X(\Omega) = \left\{x_n \mid n \in \mathbb{N}\right\} \text{ où les } (x_n)_{n\in\mathbb{N}} \text{ sont des réels 2 à 2 distincts}$$

X est d'espérance finie si la série $\sum\limits_{n\geqslant 0} x_n P(X = x_n)$ est <u>absolument convergente</u> (et pas seulement convergente...)

Dans ce cas, on appelle espérance de X, noté E(X), le réel $E(X) = \sum\limits_{n=0}^{+\infty} x_n P(X = x_n)$

En particulier, si $X(\Omega)$ est un ensemble fini, alors X est toujours d'espérance finie (la série est en fait une somme finie) et on retrouve la définition de l'espérance vu en première année.

<u>Remarque</u> : Le critère d'absolue convergence assure non seulement la convergence mais il assure aussi que la somme ne dépend pas de l'ordre d'énumération.

EXEMPLE N° 5 1) $X \sim \mathscr{G}(p)$ (loi géométrique de paramètre $p \in]0,1[$). Vérifier que X est d'espérance finie qu'on déterminera.

On rappelle que : $X(\Omega) = \mathbb{N}^*$ et $\forall n \in \mathbb{N}^*$, $P(X = n) = p(1-p)^{n-1}$

• On doit vérifier que $\sum\limits_{n\geqslant 1} nP(X = n)$ CVA or c'est une série à termes $\geqslant 0$, donc la CVA est de la CV.

On remarque que : $\sum\limits_{n\geqslant 1} nP(X = n) = p\sum\limits_{n\geqslant 1} n(1-p)^{n-1} = p\sum\limits_{n\geqslant 1} nx^{n-1}$ avec $x = 1-p$ où $p \in]0,1[$

on reconnaît la série entière dérivée termes à termes de la série géométrique $\sum\limits_{n\geqslant 0} x^n$ donc il y a CVA pour $|x| < 1$.

Or $x = 1-p \in]0,1[$ aussi $\boxed{\text{X est d'espérance finie}}$

• Enfin : $\sum\limits_{n=1}^{+\infty} nx^{n-1} = \left(\sum\limits_{n=0}^{+\infty} x^n\right)' = \left(\dfrac{1}{1-x}\right)' = \dfrac{1}{(1-x)^2}$ donc $E(X) = p \times \dfrac{1}{(1-(1-p))^2} = \dfrac{p}{p^2} = \boxed{\dfrac{1}{p} = E(X)}$

2) $X \sim \mathscr{P}(\lambda)$ (loi de Poisson de paramètre $\lambda > 0$) Vérifier que X est d'espérance finie qu'on déterminera.

On rappelle que : $X(\Omega) = \mathbb{N}$ et $\forall n \in \mathbb{N},\ P(X = n) = e^{-\lambda}\dfrac{\lambda^n}{n!}$

• On doit vérifier que $\sum_{n\geqslant 0} nP(X = n)$ CVA or c'est une série à termes $\geqslant 0$, donc la CVA est de la CV.

On remarque que : $\sum_{n\geqslant 0} nP(X = n) = \lambda e^{-\lambda} \sum_{n\geqslant 1} n\dfrac{\lambda^{n-1}}{n!} = \lambda e^{-\lambda} \sum_{n\geqslant 1} \dfrac{\lambda^{n-1}}{(n-1)!} = \lambda e^{-\lambda} \sum_{k\geqslant 0} \dfrac{\lambda^k}{k!}$ et on reconnaît une série exponentielle donc il y a CVA pour $\lambda \in \mathbb{R}$ donc aussi pour $\lambda > 0$ et $\boxed{\text{X est d'espérance finie}}$

• Enfin : $\sum_{n\geqslant 0} nP(X = n) = \lambda e^{-\lambda}e^{\lambda}$ donc $\boxed{E(X) = \lambda}$

Propositions : Linéarité,positivité, croissance de l'espérance et autres

Soient X et Y des variables aléatoires réelles discrètes d'espérances finies sur $(\Omega, \mathscr{A}, P)$

- Linéarité de l'espérance (admis) : Pour $(\lambda, \mu) \in \mathbb{R}^2$, $\lambda X + \mu Y$ est d'espérance finie et : $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$
- Positivité de l'espérance : Si $X \geqslant 0$ alors $E(X) \geqslant 0$

  Si $X \geqslant 0$ alors $X(\Omega) = \{x_k \mid k \in \mathbb{N}\} \subset [0, +\infty[$ et ainsi
  $E(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n)$ est la somme d'une série convergence à termes positifs donc $E(X) \geqslant 0$
- Croissance de l'espérance : Si $X \leqslant Y$ alors $E(X) \leqslant E(Y)$

  $Y - X$ est une variable aléatoire discrète d'espérance finie (par la linéarité) et positive donc, par positivité :
  $E(Y - X) = E(Y) - E(X) \geqslant 0$
- Si $|X| \leqslant Y$ et si Y est d'espérance finie alors X est d'espérance finie. (admis)
- Si $X \geqslant 0$ et d'espérance nulle alors $(X = 0)$ est presque sûr c'est à dire $P(X = 0) = 1$

  Notons $X(\Omega) = \{0\} \cup \{x_k \mid k \in \mathbb{N}^*\}$ où $x_k > 0$ alors : $E(X) = \sum_{k=1}^{+\infty} x_k P(X = x_k) = 0 \Rightarrow \forall k \in \mathbb{N}^*,\ x_k P(X = x_k) = 0$ (somme nulle de positifs) et donc, puisque $x_k \neq 0$: $\forall k \in \mathbb{N}^*,\ P(X = x_k) = 0$
  En utilisant le système complet d'événements $\{(X = x_n)\}_{n\in\mathbb{N}}$ (où $x_0 = 0$) alors $\sum_{n=0}^{+\infty} P(X = x_n) = 1 \Leftrightarrow P(X = 0) = 1$

Définition et proposition : Variable aléatoire centrée

Une variable aléatoire réelle discrète d'espérance finie sur un espace $(\Omega, \mathscr{A}, P)$ est centrée lorsque $E(X) = 0$.

Si X est une variable aléatoire discrète d'espérance finie alors $X - E(X)$ est centrée

Considérons la variable aléatoire $X_p : \omega \longmapsto 1$ (pas très aléatoire car toujours constante à a valeur 1)
Alors $X_p(\Omega) = \{1\}$ est finie donc X_p est d'espérance finie avec $E(X_p) = 1 \times P(X_p = 1) = 1 \times 1 = 1$
Dés lors, si X est une variable aléatoire quelconque : $X - E(X) = X + \lambda X_p$ où $\lambda = -E(X)$
Par linéarité : $E\big(X - E(X)\big) = E(X) + \lambda E(X_p) = E(X) - E(X) \times 1 = 0$ donc $X - E(X)$ est une variable aléatoire centrée

Théorème : Théorème de transfert (Admis)

Si X est une variable aléatoire réelle discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$
et si f est une application à valeurs réelles définie sur $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ alors

$$f(X) \text{ est une variable aléatoire discrète d'espérance finie} \Leftrightarrow \sum_{n\geqslant 0} f(x_n)P(X = x_n) \text{ converge absolument}$$

et on a alors : $E\big(f(X)\big) = \sum_{n=0}^{+\infty} f(x_n)P(X = x_n)$

SUITE EXEMPLE N° 5 Calculer $E\big(X(X-1)\big)$ lorsque : a) $X \sim \mathscr{G}(p)$ b) $X \sim \mathscr{P}(\lambda)$

• On utilise la formule de transfert : $\displaystyle\sum_{n=1}^{+\infty} k(k-1)p(1-p)^{k-1} = p(1-p)\sum_{n=1}^{+\infty} k(k-1)p^{k-2}$

On reconnaît la dérivée terme à terme seconde de $\dfrac{1}{1-x}$ pour $x = p$ avec $|x| < 1$:

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \frac{1}{(1-x)^2} = \sum_{n=0}^{+\infty} nx^{n-1}, \quad \frac{2}{(1-x)^3} = \sum_{n=0}^{+\infty} n(n-1)x^{n-2}$$

La série CVA donc $E\big(X(X-1)\big) < +\infty$ et : $E\big(X(X-1)\big) = p(1-p) \times \dfrac{2}{\big(1-(1-p)\big)^3} = \dfrac{2(1-p)}{p^2}$

• On utilise la formule de transfert : $\displaystyle\sum_{n=0}^{+\infty} k(k-1)\frac{\lambda^k}{k!} = \lambda^2 e^{-\lambda}\sum_{n=0}^{+\infty} k(k-1)\frac{\lambda^{k-2}}{k!}$

On reconnaît la dérivée terme à terme seconde de $e^x = \displaystyle\sum_{k=0}^{+\infty}\frac{x^k}{k!}$ pour $x = \lambda \in \mathbb{R}$ qui est aussi de somme e^x

La série CVA donc $E\big(X(X-1)\big) < +\infty$ et : $E\big(X(X-1)\big) = \lambda^2 e^{-\lambda}e^{\lambda} = \lambda^2$

IV-2) Variables aléatoires discrètes de variance finie

Proposition et définitions : Variance et écart-type

Soit X est une variable aléatoire discrète sur un espace $(\Omega, \mathscr{A}, P)$.

Si X^2 est d'espérance finie alors X est aussi d'espérance finie et on définit alors

- la variance de X notée $V(X)$ par : $V(X) = E\big((X-E(X))^2\big) = E(X^2) - E(X)^2$
- l'écart-type de X noté $\sigma(X)$ par : $\sigma(X) = \sqrt{V(X)}$

On remarque qu'alors : $\boxed{\text{Pour } a \text{ et } b \text{ des réels, } V(aX+b) = a^2V(X)}$ (La variance n'est pas linéaire !)

Notons $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$ où les x_k sont des réels 2 à 2 distincts.
On sait que $\displaystyle\sum_{n\geqslant 0} x_n^2 P(X = x_n)$ converge. Prouvons que $\displaystyle\sum_{n\geqslant 0} |x_n| P(X = x_n)$ converge.

On remarque que : $\forall x \in \mathbb{R},\ |x| \leqslant 1 + x^2$ En effet : soit $|x| \leqslant 1 \leqslant 1 + x^2$ soit $|x| > 1$ mais alors $|x| \leqslant x^2 \leqslant 1 + x^2$

Aussi : $\forall n \in \mathbb{N},\quad |x_n|P(X = x_n) \leqslant (1 + x_n^2)P(X = x_n) = P(X = x_n) + x_n^2P(X = x_n)$
Or : $\displaystyle\sum_{n\geqslant 0} P(X = x_n) = P(\Omega) = 1$ converge et $\displaystyle\sum_{n\geqslant 0} x_n^2P(X = x_n)$ converge donc $\displaystyle\sum_{n\geqslant 0}(1 + x_n^2)P(X = x_n)$ converge.

Par majoration, $\displaystyle\sum_{n\geqslant 0} x_nP(X = x_n)$ converge bien absolument. Ainsi : $\boxed{X^2 \text{ d'espérance finie} \Rightarrow X \text{ d'espérance finie.}}$

Puisque X^2, X et $X_p = 1$ sont d'espérances finies,
$(X-E(X))^2 = X^2 - 2E(X)X + (E(X))^2 \times 1 = X^2 - 2\lambda X + \lambda^2 X_p$ est aussi d'espérance finie par linéarité avec $\lambda = E(X) \in \mathbb{R}$
et : $V(X) = E\big((X-E(X))^2\big) = E(X^2) - 2\lambda E(X) + \lambda^2 \times 1 = E(X^2) - 2E(X) \times E(X) + (E(X))^2 = E(X^2) - (E(X))^2$

Enfin $(aX+b)^2 = a^2X^2 + 2abX + b^2 \times 1$ sera aussi d'espérance finie par linéarité donc on peut définir $V(aX+b)$:
$V(aX+b) = E\big((aX+b)^2\big) - \big(E(aX+b)\big)^2 = a^2E(X^2) + 2abE(X) + b^2 \times 1 - \big(aE(X)+b\big)^2 = a^2\big(E(X^2) - (E(X))^2\big) = a^2V(X)$

Si $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ alors la formule de transfert donne :
$V(X) = \displaystyle\sum_{n=0}^{+\infty}(x_n - E(X))^2 P(X = x_n)$ donc la variance mesure la dispersion de X autour de sa moyenne $E(X)$
En particulier : $V(X) = 0 \Rightarrow P(X = E(X)) = 1$ autrement dit la variable aléatoire X est constante presque sûrement.

SUITE EXEMPLE N° 5 Justifier que X est de variance finie qu'on déterminera lorsque : a) $X \sim \mathcal{G}(p)$ b) $X \sim \mathcal{P}(\lambda)$

$X^2 = X(X-1) + X$ or $E\big(X(X-1)\big) < +\infty$ et $E(X) < +\infty$ donc $E(X^2) < +\infty$ et X admet une variance
En outre : $V(X) = E(X^2) - E(X)^2 = E\big(X(X-1)\big) + E(X) - E(X)^2$ soit :
si $X \sim \mathcal{G}(p)$ où $E(X) = \dfrac{1}{p}$ et $E\big(X(X-1)\big) = \dfrac{2(1-p)}{p^2}$ d'où $V(X) = \dfrac{2(1-p)}{p^2} + \dfrac{p}{p^2} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$
si $X \sim \mathcal{P}(\lambda)$ où $E(X) = \lambda$ et $E\big(X(X-1)\big) = \lambda^2$ d'où $V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

Définition et proposition : Variable aléatoire centrée réduite

Une variable aléatoire réelle discrète admettant une variance sur un espace $(\Omega, \mathscr{A}, P)$ est réduite lorsque $V(X) = 1$.

Si X est une variable aléatoire discrète admettant une variance non nulle alors $\dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite

Si $V(X) \neq 0$ existe alors X est d'espérance finie et, par linéarité :
$E\left(\dfrac{X-E(X)}{\sigma(X)}\right) = \dfrac{1}{\sigma(X)}(E(X) - E(X)) = 0$ donc $\dfrac{X-E(X)}{\sigma(X)}$ est centrée
De plus $V\left(\dfrac{X-E(X)}{\sigma(X)}\right) =_{V(aX)=a^2V(X)} \dfrac{1}{\sigma(X)^2} V\big(X - E(X)\big) =_{V(X+b)=V(X)} \dfrac{1}{\sigma(X)^2} V(X) = 1$ donc $\dfrac{X-E(X)}{\sigma(X)}$ est réduite

Théorème : Inégalité de Cauchy-Schwarz

Si X^2 et Y^2 sont des variables aléatoires discrètes réelles d'espérance finie, alors XY est aussi d'espérance finie et :

$$E(XY)^2 \leqslant E(X^2)E(Y^2)$$

Le cas d'égalité est obtenue lorsque les variables X et Y sont proportionnelles presque sûrement.

$(|X| - |Y|)^2 \geqslant 0 \Leftrightarrow X^2 - 2|XY| + Y^2 \geqslant 0 \Leftrightarrow |XY| \leqslant \dfrac{1}{2}(X^2 + Y^2)$ or $\dfrac{1}{2}(X^2+Y^2)$ est d'espérance finie par linéarité
aussi XY est bien d'espérance finie et tous les termes $E(XY)$, $E(X^2)$ et $E(Y^2)$ existent.

$\forall t \in \mathbb{R}$, $(tX + Y)^2 = t^2X^2 + 2tXY + Y^2 \geqslant 0$ et c'est une variable d'espérance finie avec $E((tX+Y)^2) \geqslant 0$.
Par linéarité : $t^2E(X^2) + 2tE(XY) + E(Y^2) \geqslant 0$ pour tout t réel aussi le discriminant de cette expression en t doit rester négatif : $\Delta = 4E(XY)^2 - 4E(X^2)E(Y^2) \leqslant 0 \Leftrightarrow E(XY)^2 \leqslant E(X^2)E(Y^2)$
Le cas d'égalité induit l'existence d'une racine double t_0 soit : $E((t_0X+Y)^2) = 0$ où $(t_0X+Y)^2 \geqslant 0$
donc $P((t_0X+Y)^2 = 0) = 1 \Leftrightarrow P(t_0X + Y = 0) = 1$ autrement dit X et Y sont proportionnelles presque sûrement

SUITE EXEMPLE 1 La variable Z admet-elle une espérance ? une variance ? Si oui les calculer.

EXEMPLE N° 6 Soit X une variable aléatoire discrète sur l'espace probabilité $(\Omega, \mathscr{A}, P)$ de loi

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall k \in \mathbb{N}^*,\ P(X = k) = \frac{(2k+1)a}{k^2(k+1)^2} \text{ où } a > 0 \text{ est fixé}$$

1. On pose $u_k = \dfrac{1}{k^2}$ pour $k \in \mathbb{N}^*$. Calculer $u_k - u_{k+1}$ et en déduire la valeurs de a.
2. La variable aléatoire X admet-elle une espérance ? une variance ?
3. La variable aléatoire $Y = (X+1)^2$ admet-elle une espérance ?
4. Retrouver, par une autre méthode, que Y n'est pas d'espérance finie.

IV-3) Covariances et cas des variables aléatoires indépendantes

Définition : Covariance et variables aléatoires décorrélées

Si X et Y sont des variables aléatoires discrètes de variances finies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,

la covariance de X et Y, notée cov(X, Y), est le réel $\quad \operatorname{cov}(X,Y) = E\big((X-E(X))(Y-E(Y))\big) = E(XY) - E(X)E(Y)$

Lorsque $\operatorname{cov}(X,Y) = 0$, on dit que les deux variables aléatoires X et Y sont décorrélées.

Remarque : $\operatorname{cov}(X,X) = V(X)$ et $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$

On rappelle que, pour $(x,y) \in \mathbb{R}^2$: $|xy| \leqslant \dfrac{x^2+y^2}{2}$ (identité remarquable $(|x|-|y|)^2 \geqslant 0$) Par croissance et linéarité : $E(|XY|) \leqslant \frac{1}{2}E(X^2) + \frac{1}{2}E(Y^2) < +\infty$ puisque X et Y admettent une variance aussi XY est d'espérance finie.

$$\boxed{\text{Ainsi : Si } E(X^2) < +\infty \text{ et } E(Y^2) < +\infty \text{ alors } E(XY) < +\infty}$$

Par suite, par linéarité : $(X-E(X))(Y-E(Y)) = XY - E(X)Y - E(Y)X + E(X)E(Y) \times 1$ a une espérance finie car XY, X, Y et 1 sont toutes d'espérances finies si X^2 et Y^2 d'espérances finies.

$$\boxed{\text{Ainsi : Si } E(X^2) < +\infty \text{ et } E(Y) < +\infty \text{ alors } \operatorname{cov}(X,Y) \text{ existe}}$$

Et on a : $\operatorname{cov}(X,Y) = E(XY) - E(X)\times E(Y) - E(Y)\times E(X) + E(X)E(Y)\times 1 = E(XY) - E(X)E(Y)$

Proposition : Variances d'une somme finie de variables aléatoires (éventuellement décorrélées)

- Si X et Y sont des variables aléatoires discrètes de variances finies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,

$$V(X+Y) = V(X) + V(Y) + 2\operatorname{cov}(X,Y)$$

Lorsque X et Y sont décorrélées : $V(X+Y) = V(X) + V(Y)$

- Si $X_1, X_2, \dots, X_n$ sont des variables aléatoires de variances finies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,

$$V(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) + 2\sum_{1\leqslant i<j\leqslant n} \operatorname{cov}(X_i, X_j)$$

Lorsque les variables aléatoires sont 2 à 2 décorrélées : $V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)$

$V(X+Y) = E\big((X+Y)^2\big) - (E(X)+E(Y))^2 = E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 = V(X) + V(Y) + 2\operatorname{cov}(X,Y)$

Puis récurrence simple avec pour l'hérédité :

$$\begin{aligned} V(X_1 + \cdots + X_n + X_{n+1}) &= V(X_1 + \cdots + X_n) + V(X_{n+1}) + 2\operatorname{cov}(X_1 + \cdots + X_n, X_{n+1}) \\ &=_{HR_n} V(X_1) + \cdots + V(X_n) + 2\sum_{1\leqslant i<j\leqslant n} \operatorname{cov}(X_i, X_j) + V(X_{n+1}) \\ &\quad +2\Big(E\big((X_1 + \cdots + X_n)X_{n+1}\big) - \big(E(X_1) + \cdots + E(X_n)\big)E(X_{n+1})\Big) \end{aligned}$$

or : $E\big((X_1 + \cdots + X_n)X_{n+1}\big) - \big(E(X_1) + \cdots + E(X_n)\big)E(X_{n+1}) = \displaystyle\sum_{i=1}^{n} \big(E(X_i X_{n+1}) - E(X_i)E(X_{n+1})\big) = \sum_{i=1}^{n} \operatorname{cov}(X_i, X_{n+1})$

qui correspond aux couples $(i, n+1)$ où $1 \leqslant i \leqslant n$ dans la somme $\displaystyle\sum_{1\leqslant i<j\leqslant n+1} \operatorname{cov}(X_i, X_j)$

Proposition : Espérance, variance et covariance dans le cas d'un couple de variables aléatoires indépendantes (ADMIS)

Soient X et Y sont des variables aléatoires discrètes de variances finies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,

Si X **et** Y **sont indépendantes alors** $E(XY) = E(X)E(Y) \Leftrightarrow \operatorname{cov}(X,Y) = 0 \Leftrightarrow V(X+Y) = V(X) + V(Y)$

soit : $X \perp\!\!\!\perp Y \Rightarrow E(XY) = E(X)E(Y)$, $\operatorname{cov}(X,Y) = 0$ et $V(X+Y) = V(X) + V(Y)$

Attention ! La réciproque est fausse.

Un contre-exemple : X suit une loi uniforme sur $\{-1, 0, 1\}$ et $Y = X^2$

$E(X) = \boxed{-1\times\frac{1}{3} + 0\times\frac{1}{3} + 1\times\frac{1}{3} = 0}$ $\quad E(Y) = \boxed{(-1)^2\times\frac{1}{3} + 0\times\frac{1}{3} + 1\times\frac{1}{3} = \frac{2}{3}}$ $\quad E(XY) = E(X^3) = E(X) = \boxed{0}$

Aussi : $E(XY) = E(X)E(Y) \Leftrightarrow \operatorname{cov}(X,Y) = 0$ pourtant X et Y ne sont pas indépendantes puisque

$P(X=0, Y=1) \neq P(X=0)P(Y=1)$ car $\boxed{P(X=0,Y=1) = P(\varnothing) = 0 \text{ et } P(X=0)P(Y=1) = \frac{1}{3}\times\frac{2}{3}}$

IV-4) Cas particulier des variables aléatoires à valeurs dans $\mathbb{N}$: utilisation des séries génératrices

Dans cette partie, on considère une variable aléatoire X à valeurs dans $\mathbb{N}$, ce qui est souvent le cas pour nous.
Le résultat suivant donne une autre façon d'obtenir l'espérance (un de vos camarades a eu cette démo à faire à l'oral!)

Proposition : Calcul de l'espérance d'une variable aléatoire à valeurs dans $\mathbb{N}$

Si X est une variable aléatoire à valeurs dans $\mathbb{N}$ d'espérance finie alors $\mathrm{E}(\mathrm{X}) = \sum_{n=1}^{+\infty} \mathrm{P}(\mathrm{X} \geqslant n)$

Pour $n \in \mathbb{N}$, on a : $(\mathrm{X} \geqslant n) = (\mathrm{X} \geqslant n+1) \cup (\mathrm{X} = n)$ et ces événements sont incompatibles donc :

$$\forall n \in \mathbb{N},\ \mathrm{P}(\mathrm{X} \geqslant n) = \mathrm{P}(\mathrm{X} \geqslant n+1) + \mathrm{P}(\mathrm{X} = n)$$

Mais alors, pour $\mathrm{N} \in \mathbb{N}$:

$$\sum_{n=0}^{\mathrm{N}} n\mathrm{P}(\mathrm{X}=n) = \sum_{n=0}^{\mathrm{N}} n\big(\mathrm{P}(\mathrm{X} \geqslant n) - \mathrm{P}(\mathrm{X} \geqslant n+1)\big) = \sum_{n=0}^{\mathrm{N}} \big(n\mathrm{P}(\mathrm{X} \geqslant n) - (n+1)\mathrm{P}(\mathrm{X} \geqslant n+1) + \mathrm{P}(\mathrm{X} \geqslant n+1)\big)$$

en faisant astucieusement apparaître un terme télescopique

soit : $\sum_{n=0}^{\mathrm{N}} n\mathrm{P}(\mathrm{X}=n) = \sum_{n=0}^{\mathrm{N}} \big(n\mathrm{P}(\mathrm{X} \geqslant n) - (n+1)\mathrm{P}(\mathrm{X} \geqslant n+1)\big) + \sum_{n=0}^{\mathrm{N}} \mathrm{P}(\mathrm{X} \geqslant n+1)$ où on repère une première somme télescopique. Quitte à changer d'indice dans la seconde, on a :

$\sum_{n=0}^{\mathrm{N}} n\mathrm{P}(\mathrm{X}=n) = 0 - (\mathrm{N}+1)\mathrm{P}(\mathrm{X} \geqslant \mathrm{N}+1) + \sum_{k=1}^{\mathrm{N}+1} \mathrm{P}(\mathrm{X} \geqslant k) = -\mathrm{N}\mathrm{P}(\mathrm{X} \geqslant \mathrm{N}+1) + \sum_{k=1}^{\mathrm{N}} \mathrm{P}(\mathrm{X} \geqslant k)$ en simplifiant le terme $-\mathrm{P}(\mathrm{X} \geqslant \mathrm{N}+1)$ avec le terme d'indice $k = \mathrm{N}+1$ de la somme.

Pour obtenir le résultat, il suffit alors de prouver que : $\lim_{\mathrm{N}\to+\infty} \mathrm{N}\mathrm{P}(\mathrm{X} \geqslant \mathrm{N}+1) = 0$ Or, par σ-additivité

$$0 \leqslant \mathrm{N}\mathrm{P}(\mathrm{X} \geqslant \mathrm{N}+1) = \sum_{n=\mathrm{N}+1}^{+\infty} \mathrm{N}\mathrm{P}(\mathrm{X}=n) \leqslant \sum_{n=\mathrm{N}+1}^{+\infty} n\mathrm{P}(\mathrm{X}=n) \xrightarrow[\mathrm{N}\to+\infty]{} 0$$ (reste de la série convergente de somme E(X))

Une astuce à retenir! :
Même si ce n'est pas un résultat du cours, l'astuce suivante est fréquemment réclamée sans forcément être rappelée.

Si X est à valeurs dans $\mathbb{N}$ alors : $\forall n \in \mathbb{N},\ (\mathrm{X} \leqslant n+1) = (\mathrm{X} \leqslant n) \cup (\mathrm{X} = n+1)$

Les deux événements étant incompatibles on a : $\mathrm{P}(\mathrm{X} = n+1) = \mathrm{P}(\mathrm{X} \leqslant n+1) - \mathrm{P}(\mathrm{X} \leqslant n)$

Définition et propositions : Série génératrice d'une variable aléatoire à valeurs dans $\mathbb{N}$

Si X est une variable aléatoire à valeurs dans $\mathbb{N}$ définie sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$,

on appelle fonction génératrice de X la fonction $\mathrm{G_X}$ définie par : $\mathrm{G_X}(t) = \mathrm{E}(t^{\mathrm{X}}) = \sum_{n=0}^{+\infty} \mathrm{P}(\mathrm{X}=n)\, t^n$

- $\mathrm{G_X}$ est la fonction somme d'une série entière dont le rayon de convergence est $\mathrm{R} \geqslant 1$

  En effet, la série $\mathrm{G_X}(1) = \boxed{\sum_{n=0}^{+\infty} \mathrm{P}(\mathrm{X}=n) = \mathrm{P}(\mathrm{X} \in \mathbb{N}) = \mathrm{P}(\Omega) = 1}$ est convergente.

- La fonction $\mathrm{G_X}$ caractérise totalement la loi de X puisque : $\forall n \in \mathbb{N},\ \mathrm{P}(\mathrm{X}=n) = \boxed{\dfrac{\mathrm{G_X}^{(n)}(0)}{n!}}$

  aussi : Si X et Y sont des variables aléatoire à valeurs dans $\mathbb{N}$ avec $\mathrm{G_X} = \mathrm{G_Y}$ alors X et Y ont la même loi.

Exemple : Dans le cas d'une variable aléatoire finie, la fonction $\mathrm{G_X}$ est polynomiale et donc $\mathrm{R} = \boxed{+\infty}$

Théorème : Calcul de l'espérance et de la variance avec la série génératrice (ADMIS)

Si X est une variable aléatoire à valeurs dans $\mathbb{N}$ définie sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$,

- X est d'espérance finie $\Leftrightarrow$ G_X est dérivable en 1 et on a alors : $E(X) = G'_X(1)$ $\boxed{= \sum_{n=0}^{+\infty} nP(X=n)}$
- X est de variance finie $\Leftrightarrow$ G_X est deux fois dérivable en 1 Dans ce cas (**raisonnement à savoir retrouver!**) :

$G''_X(1) = E\big(X(X-1)\big)$ $\boxed{= \sum_{n=0}^{+\infty} n(n-1)P(X=n)}$ donc $E(X^2) = \boxed{G''_X(1) + G'_X(1)}$ puis $V(X) = \boxed{= G''_X(1) + G'_X(1) - G'_X(1)^2}$

Proposition : Série génératrice d'une somme de variable aléatoire indépendante

- Si X et Y sont des variables aléatoires à valeurs dans $\mathbb{N}$ définie sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$,

<u>si</u> X et Y sont indépendantes <u>alors</u> $\forall t \in]-R, R[,\ G_{X+Y}(t) = G_X(t) \times G_Y(t)$ avec $R = \min(R_X, R_Y)$

où G_X (resp. G_Y et G_{X+Y}) est la fonction génératrice de X (resp. Y et X + Y) de rayon de convergence R_X (resp. R_Y)

- Si $X_1, \dots, X_n$ sont des variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ dans $(\Omega, \mathscr{A}, P)$,

alors $G_{X_1+X_2+\dots+X_n} = G_{X_1} \times G_{X_2} \times \dots \times G_{X_n}$ définie à minima sur $]-1, 1[$

On note R_X et R_Y les rayons de convergences. Pour t avec $|t| < \min(R_X, R_Y)$

$G_X(t)G_Y(t) = \left(\sum_{n=0}^{+\infty} P(X=n)t^n\right)\left(\sum_{n=0}^{+\infty} P(Y=n)t^n\right)$ est un produit de Cauchy de 2 séries ACV

Aussi : $G_X(t)G_Y(t) = \sum_{n=0}^{+\infty}\left(\sum_{k=0}^{n} P(X=k)P(Y=n-k)\right)t^n$

Or : $P(X+Y=n) = P\left(\bigcup_{k=0}^{n}(X=k)\cap(Y=n-k)\right) =_{\text{evts incomp}} = \sum_{k=0}^{n} P\big((X=k)\cap(Y=n-k)\big) = \sum_{k=0}^{n} P(X=k)P(Y=n-k)$ par indépendance. Donc finalement : $G_X(t)G_Y(t) = G_{X+Y}(t)$

IV-5) Espérance, variance et série génératrice des lois usuelles

$X \sim$	Loi	$E(X) =$	$V(X) =$	$G_X(t) =$	Schémas types et propriétés
$\mathscr{U}(\llbracket 1,n \rrbracket)$ Loi uniforme	$X(\Omega) = \llbracket 1,n \rrbracket$ $P(X=k) = \dfrac{1}{n}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$	$\dfrac{t}{n} \times \dfrac{1-t^n}{1-t}$	Lancer d'un dé Tirage dans une urne
$\mathscr{B}(p)$ Loi de Bernoulli	$X(\Omega) = \{0,1\}$ $P(X=1) = p = 1-P(X=0)$	p	$p(1-p)$	$pt+1-p$	Lancer pile/face Obtenir un 6 avec un dé Tirer un as dans un jeu
$\mathscr{B}(n,p)$ Loi binomiale	$X(\Omega) = \llbracket 0,n \rrbracket$ $P(X=k) = \dbinom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	$(pt+1-p)^n$	nb pile sur n lancers pile/face nb 6 sur n lancers de dé
$\mathscr{G}(p)$ Loi géométrique	$X(\Omega) = \mathbb{N}^*$ $P(X=k) = p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$\dfrac{pt}{1-(1-p)t}$	premier pile sur lancers pile/face premier 6 sur lancers de dé
$\mathscr{P}(\lambda)$ Loi de Poisson	$X(\Omega) = \mathbb{N}$ $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$	λ	λ	$e^{\lambda(t-1)}$	**Loi des événements rares** Compteur Geiger

L'essentiel sur les séries génératrices :

- On peut toujours définir la série génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$ (ou une partie de $\mathbb{N}$)
  c'est une série entière de rayon de somme $G_X(t) = E(t^X) = \sum_{n=0}^{+\infty} P(X=n)t^n$
- Le rayon de convergence R vérifie $R \geqslant 1$ car la somme $G_X(1) = \sum_{n=0}^{+\infty} P(X=n) = P(X \in \mathbb{N}) = 1$ converge.
  La fonction génératrice G_X est donc, à minima, définie sur $]-1,1[$ (et même $]-1,1]$)
  Comme somme d'une série entière, elle est donc C^∞ au moins sur $]-1,1[$.
- La fonction génératrice caractérise totalement la loi de X car : $\forall n \in \mathbb{N},\ P(X=n) = \dfrac{G_X^{(n)}(0)}{n!}$
- Si X et Y sont indépendantes, alors $G_{X+Y}(t) = G_X(t) \times G_Y(t)$ **Attention pas de réciproque!**
  Par récurrence : $G_{X_1+\cdots+X_n}(t) = G_{X_1}(t) \times \cdots \times G_{X_n}(t)$ lorsque $X_1, \ldots, X_n$ sont mutuellement indépendantes
- Sur $]-R,R[$, par dérivation termes à termes : $G_X'(t) = \sum_{n=0}^{+\infty} nP(X=n)t^{n-1}$ et $G_X''(t) = \sum_{n=0}^{+\infty} n(n-1)P(X=n)t^{n-2}$
  Si on peut substituer $t=1$: $G_X'(1) = \sum_{n=0}^{+\infty} nP(X=n) = E(X)$ et $G_X''(1) = \sum_{n=0}^{+\infty} n(n-1)P(X=n) = E\big(X(X-1)\big)$
- L'existence d'une espérance et d'une variance pour X est relié à la fonction génératrice via les théorèmes :
  $\boxed{\text{X admet une espérance si et seulement si } G_X \text{ est dérivable en 1 et alors } E(X) = G_X'(1)}$

  $\boxed{\text{X admet une variance si et seulement si } G_X \text{ est deux fois dérivable en 1}}$ qu'on obtient car
  $V(X) = E(X^2) - E(X)^2 = E\big(X(X-1)+X\big) - E(X)^2 =_{\text{par linéarité}} E\big(X(X-1)\big) + E(X) - E(X)^2 = G_X''(1) + G_X'(1) - G_X'(1)^2$

Calcul des séries génératrices pour les lois usuelles et applications aux calculs d'espérance et variance

- Si X est une variable aléatoire finie càd $X(\Omega)$ est un ensemble fini de $\mathbb{N}$ alors $G(t)$ est un polynôme en t

  La fonction G est donc définie sur $\mathbb{R}$ (car $R = +\infty$ puisque $G(t)$ converge pour tout t réel) et on retrouve l'existence certaine d'une espérance et d'une variance

- Si $X \sim \mathscr{U}([\![1,n]\!])$ alors $G_X(t) = \dfrac{t+t^2+\cdots+t^n}{n} = \dfrac{t}{n} \times \dfrac{1-t^n}{1-t}$ et $E(X) = \dfrac{n+1}{2}$ $V(X) = \dfrac{n^2-1}{12}$

On rappelle $X(\Omega) = [\![1,n]\!]$ et $\forall k \in \mathbb{N}, P(X=k) = \dfrac{1}{n}$ Comme c'est une variable aléatoire finie,

G_X est un polynôme défini sur $\mathbb{R}$ avec : $G_X(t) = E(t^X) = \displaystyle\sum_{k=1}^{n} P(X=k)\,t^k = \sum_{k=1}^{n} \frac{1}{n} t^k = \frac{1}{n}\sum_{k=1}^{n} t^k = \frac{1}{n} \times t \times \frac{1-t^n}{1-t}$

L'écriture sans le symbole somme laisse croire que G_X n'est pas définie en $t=1$ ce qui n'est pas le cas : $G_X(1) = 1$

L'écriture avec une somme permet mieux d'identifier un polynôme pour G_X : $G_X(t) = \frac{1}{n}t + \frac{1}{n}t^2 + \cdots + \frac{1}{n}t^n$

Pour une variable aléatoire finie, la fonction génératrice n'est pas forcément le meilleur outil pour calculer espérance/variance car ici on peut mener rapidement le calcul sur les sommes finies issues des définitions :

$E(X) = \displaystyle\sum_{k=1}^{n} kP(X=k) = \sum_{k=1}^{n} \frac{k}{n} = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2}$ $\qquad E(X^2) = \displaystyle\sum_{k=1}^{n} k^2P(X=k) = \sum_{k=1}^{n} \frac{k^2}{n} = \frac{1}{n} \times \frac{n(n+1)(2n+1)}{6}$

d'où : $V(X) = E(X^2) - E(X)^2 = \dfrac{(n+1)(2n+1)}{6} - \dfrac{(n+1)^2}{4} = \dfrac{(n+1)}{12}(2(2n+1) - 3(n+1)) = \dfrac{n^2-1}{12}$

- Si $X \sim \mathscr{B}(p)$ alors $G_X(t) = (1-p) + pt$ et $E(X) = p$ $V(X) = p(1-p)$

On rappelle $X(\Omega) = \{0,1\}$ et $P(X=1) = p = 1 - P(X=0)$ C'est une variable aléatoire finie,
G_X est un polynôme défini sur $\mathbb{R}$ avec : $G_X(t) = t^0P(X=0) + t^1P(X=1) \Leftrightarrow G_X(t) = 1-p+pt$

On obtient alors facilement espérance/variance
- soit par la définition : $E(X) = 1 \times P(X=1) + 0 = p$
  et, puisque $X^2 = X$ vu que $X(\Omega) = \{0,1\}$, $E(X^2) = E(X)$ donc $V(X) = E(X^2) - E(X)^2 = p - p^2$
- soit via la fonction génératrice : G est deux fois dérivable en 1 avec $G_X'(t) = p$, $G_X''(t) = 0$
  donc $E(X) = G_X'(1) = p$ et $V(X) = G_X''(1) + G_X'(1) - G_X'(1)^2 = p - p^2$

- Si $X \sim \mathscr{B}(n,p)$ alors $G_X(t) = (pt+1-p)^n$ et $E(X) = np$ $V(X) = np(1-p)$

On rappelle que $X(\Omega) = [\![0,n]\!]$ et $\forall k \in [\![0,n]\!], P(X=k) = \dbinom{n}{k} p^k(1-p)^{n-k}$

C'est une variable aléatoire finie donc E(X) et V(X) existent et $G_X(t)$ est une expression polynômiale

$G_X(t) = \displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k(1-p)^{n-k} t^k = \sum_{k=0}^{n} \binom{n}{k} (pt)^k (1-p)^{n-k} = \big(pt+1-p\big)^n$ avec la formule du binôme de Newton

<u>Méthode n°1</u> : en utilisant la loi

On obtient alors, très rapidement, l'espérance et la variance (beaucoup plus vite qu'en revenant à la définition!) :

$$G_X'(t) = p \times n(pt+1-p)^{n-1} \quad \text{et} \quad G_X''(t) = p^2 \times n(n-1)(pt+1-p)^{n-1}$$

de sorte que $E(X) = G_X'(1) = p \times n$ et $V(X) = \cdots = G_X''(1) + G_X'(1) - G_X'(1)^2 = p^2 \times n(n-1) + np - n^2p^2 = np(1-p)$

<u>Méthode n°2</u> : en utilisant l'interprétation $X = X_1 + \cdots + X_n$ où $(X_k)_{1 \leqslant k \leqslant n}$ i.i.d de loi $\mathscr{B}(p)$

$G_X(t) =_{\text{par ind}} \displaystyle\prod_{k=1}^{n} \underbrace{G_{X_k}(t)}_{1-p+pt} =_{\text{même loi}} (1-p+pt)^n$

mais aussi $E(X) =_{\text{par linéairité}} \displaystyle\sum_{k=1}^{n} \underbrace{E(X_k)}_{=p} =_{\text{même loi}} np$ et $V(S) =_{\text{par ind}} \displaystyle\sum_{k=1}^{n} \underbrace{V(X_k)}_{p(1-p)} = np(1-p)$

- $\boxed{\text{Si} \quad X \sim \mathscr{G}(p) \quad \text{alors} \quad G_X(t) = \dfrac{pt}{1-(1-p)t} \quad \text{et} \quad E(X) = \dfrac{1}{p} \quad V(X) = \dfrac{1-p}{p^2}}$

On a $X(\Omega) = \mathbb{N}^*$ et $\forall k \in \mathbb{N}^*, P(X=k) = p(1-p)^{k-1}$ (premier succès dans une suite iid de Bernouilli de paramètre p)

On a $G_X(t) = \displaystyle\sum_{n=1}^{+\infty} p(1-p)^{n-1} t^n$ qui sera définie au moins sur $]-R,R[$ où R est le rayon de convergence obtenu par la règle de d'Alembert : $\dfrac{p(1-p)^n|t|^{n+1}}{p(1-p)^{n-1}|t|^n} = (1-p)|t| \xrightarrow[n\to+\infty]{} (1-p)|t|$

La série CVA si $|t| < \dfrac{1}{1-p}$ donc $R \geqslant \dfrac{1}{1-p}$. La série DVG si $|t| > \dfrac{1}{1-p}$ donc $R \leqslant \dfrac{1}{1-p}$. Ainsi : $R = \dfrac{1}{1-p}$

Pour $t \in]-R;R[$ alors $\big|(1-p)t\big| < 1$ et :

$$G_X(t) = \sum_{n=1}^{+\infty} p(1-p)^{n-1} t^n = \frac{p}{1-p}\sum_{n=1}^{+\infty}\big((1-p)t\big)^n =_{\text{série géo CV}} \frac{p}{1-p} \times \underbrace{(1-p)t}_{\text{1er terme}} \times \frac{1-0}{1-(1-p)t} = \frac{pt}{1-(1-p)t}$$

En temps que somme d'une série entière, on sait que G_X est C^∞ sur $]-R,R[$

Or, le paramètre p vérifie $0 < p < 1 \Leftrightarrow 0 < 1-p < 1 \Leftrightarrow R = \dfrac{1}{1-p} > 1$ aussi $1 \in]-R,R[$ et G_X est deux fois C^∞ en 1

on retrouve que X admet une espérance et une variance et on a : $E(X) = G_X'(1)$

puis $E\big(X(X-1)\big) = G_X''(1)$ d'où $V(X) = E(X^2) - E(X) = E\big(X(X-1)\big) + E(X) - E(X)^2 = G_X''(1) + G_X'(1) - G_X(1)''$

On retrouve le calcul de l'espérance (fait par la définition dans l'exemple 3) et on prouve celui de la variance (bcp plus vite qu'avec la définition) :

$G_X'(t) = \dfrac{p(1-(1-p)t) - 1 \times -(1-p) \times pt}{(1-(1-p)t)^2} = \dfrac{p}{(1-(1-p)t)^2}$ d'où $E(X) = G_X'(1) = \dfrac{p}{p^2} = \dfrac{1}{p}$

$G_X''(t) = \dfrac{2p(1-p)}{(1-(1-p)t)^3}$ d'où $E(X^2) = G_X''(1) + G_X'(1) = \dfrac{2(1-p)+p}{p^2} = \dfrac{2-p}{p^2}$ puis : $V(X) = \dfrac{2-p}{p^2} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$

- $\boxed{\text{Si} \quad X \sim \mathscr{P}(\lambda) \quad \text{alors} \quad G_X(t) = e^{\lambda(t-1)} \text{ et } \quad E(X) = \lambda \quad V(X) = \lambda}$

On rappelle que $X(\Omega) = \mathbb{N}$ et $\forall k \in \mathbb{N}, P(X=k) = e^{-\lambda} \times \dfrac{\lambda^k}{k!}$

$G_X(t) = \displaystyle\sum_{n=0}^{+\infty} e^{-\lambda}\frac{\lambda^n}{n!}t^n$ est définie sur $]-R,R[$ où R est le rayon de convergence obtenu par d'Alembert : $\dfrac{P(X=n+1)|t|^{n+1}}{P(X=n)|t|^n} = \dfrac{\lambda}{n+1}|t| \xrightarrow[n\to+\infty]{} 0 < 1$ donc la série CVA pour tout t réel et on a $R = +\infty$

Ainsi, G_X est définie (et même C^∞) sur $\mathbb{R}$ avec : $G_X(t) = e^{-\lambda}\displaystyle\sum_{n=0}^{+\infty}\frac{(\lambda t)^n}{n!} = e^{-\lambda}e^{\lambda t}$

Elle est donc 2 fois dérivable en 1 ce qui assure l'existence de l'espérance et de la variance avec

$E(X) = G_X'(1)$, $E\big(X(X-1)\big) = G_X''(1)$ puis $V(X) = E(X^2) - E(X) = E\big(X(X-1)\big) + E(X) - E(X)^2 = G_X''(1) + G_X'(1) - G_X(1)''$

Ici $G_X'(t) = \lambda e^{-\lambda}e^{\lambda t}$ et $G_X''(t) = \lambda^2 e^{-\lambda}e^{\lambda t}$ aussi $E(X) = G_X'(1) = \lambda$,

$E\big(X(X-1)\big) = G_X''(1) = \lambda^2$ d'où $E(X^2) = G_X''(1) + G_X'(1) = \lambda^2 + \lambda$ et enfin $V(X) = E(X^2) - (E(X))^2 = \lambda$

Nous retrouvons les valeurs déterminées en utilisant directement la définition à l'exemple 3

POINT MÉTHODE : <u>Pour savoir si une variable aléatoire X est d'espérance finie, on peut :</u>

- revenir à la définition et étudier la CVA de $\displaystyle\sum_{n\geqslant 0} x_n P(X = x_n)$ mais il faut alors connaître la loi de X
- si la variable aléatoire X est à valeurs dans $\mathbb{N}$, utiliser un calcul de série génératrice G_X qu'on peut parfois obtenir sans avoir à connaître la loi de X si X construite à partir de variables aléatoires indépendantes de lois connues.

  Si on sait $E(X) < +\infty$, on peut calculer l'espérance avec $E(X) = \displaystyle\sum_{n=1}^{+\infty} P(X \geqslant n)$ si on connaît mieux $P(X \geqslant n)$ que $P(X = n)$ par exemple pour des minima ou des maxima
- utiliser le théorème de transfert lorsque $X = f(Y)$ et qu'on connaît la loi de Y

V) Les théorèmes du programme sur les variables aléatoires

V-1) Inégalités de Markov et de Bienaymé-Tchebychev

Théorème : Inégalité de Markov

Si X est une variable aléatoire discrète d'espérance finie avec $X \geqslant 0$ alors $\forall a > 0, \quad P(X \geqslant a) \leqslant \dfrac{E(X)}{a}$

On remarque : $X \geqslant a \times \mathbb{1}_{X \geqslant a}$ si $X(\omega) \geqslant a$ (resp $X(\omega) < a$), alors l'indicatrice vaut 1 (resp 0) et l'inégalité est vraie (resp car $X \geqslant 0$).
$Y = a \times \mathbb{1}_{X \geqslant A}$ est une variable aléatoire discrète (puisque $Y(\Omega) = \{0, a\}$)
Par croissance de l'espérance : $E(X) \geqslant E(Y)$ or : $E(Y) = 0 \times P(X < a) + a \times P(X \geqslant a) = aP(X \geqslant a)$ d'où le résultat

Objectif/mise en garde : Majorer des probabilités $P(X \geqslant a)$ **mais ne s'applique que pour** $X \geqslant 0$

Théorème : Inégalité de Bienaymé-Tchebychev

Si X est une variable aléatoire discrète de variance finie alors $\forall \varepsilon > 0,\ P(|X - E(X)| \geqslant \varepsilon) \leqslant \dfrac{V(X)}{\varepsilon^2}$

$P(|X - E(X)| \geqslant \varepsilon) = P\left(|X - E(X)|^2 \geqslant \varepsilon^2\right)$ or $(X - E(X))^2$ est d'espérance finie puisque X est de variance finie aussi, en appliquant, l'inégalité de Markov, on a : $P(|X - E(X)| \geqslant \varepsilon) \leqslant \dfrac{E\left((X - E(X))^2\right)}{\varepsilon^2} = \dfrac{V(X)}{\varepsilon^2}$

Objectif/mise en garde : Majorer des écarts par rapport à la moyenne **s'applique si** X **admet une variance finie**

Cette inégalité permet de comprendre que **la variance est un indicateur de dispersion** :
la probabilité d'un écart de X par rapport à sa moyenne E(X) supérieur à ε est d'autant plus faible que V(X) est petit

V-2) Loi faible des grands nombres

Théorème : Loi faible des grands nombres

Soit $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires discrètes i.i.d. admettant une variance définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$. En notant m l'espérance et σ l'écart-type de X_1 et aussi $S_n = X_1 + X_2 + \cdots + X_n$:

$$\forall \varepsilon > 0, \quad P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2} \quad \text{d'où} \quad P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n \to +\infty]{} 0$$

Attention ! Les étudiants doivent savoir retrouver l'inégalité $P\left(\left|\dfrac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \dfrac{\sigma^2}{n\varepsilon^2}$

On applique l'inégalité de Bienaymé-Tchebychev à $\dfrac{S_n}{n}$

Par linéarité de l'espérance et les X_k étant de même loi : $E\left(\dfrac{S_n}{n}\right) = \dfrac{1}{n} \times nm = m$

On a donc : $\forall \varepsilon > 0, \quad P\left(\left|\dfrac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \dfrac{V\left(\frac{S_n}{n}\right)}{\varepsilon^2}$

Par indépendance et puisque les X_k sont de même loi : $V\left(\dfrac{S_n}{n}\right) = \dfrac{1}{n^2}V(S_n) = \dfrac{1}{n^2} \times nV(X_1) = \dfrac{\sigma^2}{n}$ d'où la majoration

EXEMPLE N° 7 Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoire indépendante et de même loi $\mathscr{B}(p)$ où $0<p<1$

Pour $n \in \mathbb{N}^*$, on pose : $\quad S_n = \dfrac{X_1+X_2+\cdots+X_n}{n}, \quad Y_n = \dfrac{X_n+X_{n+1}}{2} \quad$ et $\quad T_n = \dfrac{Y_1+\cdots+Y_n}{n}$

1. Montrer que, pour tout $\epsilon>0$: $\lim\limits_{n\to+\infty} P(|S_n-p|\geqslant\epsilon)=0$

2. Préciser la loi puis l'espérance de Y_n.

3. Si $n<m$, Y_n et Y_m sont-elles indépendantes ?
On pourra distinguer le cas $m>n+1$ de $m=n+1$ et, dans ce cas, calculer $E(Y_nY_{n+1})$

4. Montrer que, pour tout $\epsilon>0$: $\lim\limits_{n\to+\infty} P(|T_n-p|\geqslant\epsilon)=0$