Lycée Lakanal

2025 - 2026

PC

Variables aléatoires discrètes

I – Préliminaires

1. Séries géométriques et séries exponentielles

On rappelle les points suivants :

a. ***Séries exponentielles***

Pour tout $x \in \mathbb{R}$, la série $\sum_{k \geq 0} \frac{x^k}{k!}$ converge absolument, et l'on a : $\boxed{e^x = \sum_{k=0}^{+\infty} \frac{x^k}{k!}}$.

b. ***Séries géométriques, géométriques dérivées, géométriques dérivées secondes***

Pour tout $q \in]-1, 1[$, les séries $\sum_{n \geq 0} q^n$, $\sum_{n \geq 1} n\, q^{n-1}$, $\sum_{n \geq 2} n(n-1)\, q^{n-2}$ convergent absolument, et l'on a :

$$\boxed{\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}},\ \boxed{\sum_{n=1}^{+\infty} n q^n = \frac{1}{(1-q)^2}} \text{ et } \boxed{\sum_{n=2}^{+\infty} n(n-1)\, q^{n-2} = \frac{2}{(1-q)^3}}.$$

2. Ensembles dénombrables, familles sommables

a. Rappel sur les ensembles dénombrables

Définition

Soit E un ensemble. Alors :

1_ E est dit ***dénombrable*** s'il est ***équipotent*** à $\mathbb{N}$, *i.e.* s'il existe une bijection de E sur $\mathbb{N}$.

On peut alors énumérer ses éléments sous la forme d'une suite $(e_n)_{n \in \mathbb{N}}$.

2_ E est dit ***fini, de cardinal p*** si et seulement s'il est équipotent à $\{1, ..., p\}$.

3_ E est dit ***au plus dénombrable*** s'il est fini ou dénombrable.

On montre que $\mathbb{Z}$, $\mathbb{Q}$ sont dénombrables. Tout produit cartésien d'un nombre fini d'ensemble dénombrables, toute réunion d'un nombre au plus dénombrable d'ensemble dénombrables, tout sous-ensemble d'un ensemble dénombrable, sont dénombrables.

b. Rappel sur les familles sommables

Proposition

Soit $S = \sum_{n \geq 0} a_n$ une série absolument convergente.

Alors la nature et la somme de cette série ne dépend pas de l'ordre de la sommation : si l'on permute les termes de cette série, on garde l'absolue convergence, et la somme reste inchangée.

Dit correctement :

pour toute bijection φ de $\mathbb{N}$ dans $\mathbb{N}$, la série $\sum_{n \geq 0} a_{\varphi(n)}$ converge, et $\sum_{n=0}^{+\infty} a_{\varphi(n)} = \sum_{n=0}^{+\infty} a_n$.

Corolaire - définition

Soit I un ensemble dénombrable. Soit $\left(u_i\right)_{i\in\mathbb{N}}$ une famille de réels ou de complexes indexée par I.

On dit que la famille $\left(u_i\right)_{i\in\mathbb{N}}$ est **sommable** lorsqu'il existe une bijection $\varphi : \mathbb{N} \to I$ telle que la série $\sum_{n\geq 0} a_{\varphi(n)}$ converge **absolument**. Dans ce cas, pour toute autre bijection $\psi : \mathbb{N} \to I$: $\sum_{n\geq 0} a_{\psi(n)}$ converge, et l'on a $\sum_{n=0}^{+\infty} a_{\varphi(n)} = \sum_{n=0}^{+\infty} a_{\psi(n)}$.

On note alors : $\sum_{i\in I} a_i = \sum_{n=0}^{+\infty} a_{\varphi(n)}$.

Ainsi on ne modifie ni la nature, ni la somme, d'une « série »
absolument convergente en modifiant l'ordre de ses termes. Pour une partie infinie I de $\mathbb{N}$ ou $\mathbb{Z}$, ou de tout autre ensemble dénombrable, on s'autorisera alors l'écriture $\sum_{i\in I} u_i$ lorsque la série en question est absolument convergente (ou plus exactement, lorsque la famille $\left(u_i\right)_{i\in I}$ est sommable…). *Traduction mathématique* :

Théorème 1 ***(convergence commutative)***

Soit $\left(u_i\right)_{i\in I}$ une famille sommable de réels. Alors pour toute permutation σ de I :

$$\boxed{\sum_{i\in I} u_{\sigma(i)} = \sum_{i\in I} u_i}.$$

c. Sommation par paquets

Théorème 2 ***(sommation par paquets)***

Soit I un ensemble dénombrable, et $\left(u_i\right)_{i\in I}$ une famille de réels. *Alors :*

La famille $\left(u_i\right)_{i\in I}$ est sommable, *i.e.* $\sum_{i\in I} u_i$ converge absolument,

si, et seulement si, pour une (*resp.* toute) partition $\left(A_k\right)_{k\in\mathbb{N}}$ de I, on a :

- • La famille $\left(u_i\right)_{i\in A_k}$ est sommable, de somme $\sum_{i\in A_k} u_i$.
- •• La série $\sum_{k\in\mathbb{N}}\left(\sum_{i\in A_k} u_i\right)$ converge absolument.

On a de plus, dans ce cas :

••• $$\boxed{\sum_{i\in I} u_i = \sum_{k=0}^{+\infty}\left(\sum_{i\in A_k} u_i\right) = \sum_{k\in\mathbb{N}}\left(\sum_{i\in A_k} u_i\right)}.$$

Application :

Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $\left(\Omega, \mathcal{T}, \mathbb{P}\right)$, avec $X\left(\Omega\right) = \left\{x_k,\ k\in\mathbb{N}\right\}$.

Soit $s \in \mathbb{N}^*$. On suppose que X admet un moment d'ordre s (*rappeler ce point…*)

Montrer que pour tout entier $r \in \{1, ..., s\}$, X admet un moment d'ordre r.

hint *séparer les* x_k *selon que* $|x_k| \leq 1$ *ou* $|x_k| > 1$.

d. Interversion des ordres de sommation : théorème de Fubini

Proposition 1 & Définition 1 ***(Fubini – Tonelli)***

Soit $(u_{i,j})_{(i,j) \in \mathbb{N}^2}$ une famille double de réels ***positifs ou nuls***.

On suppose que :

- Pour tout entier naturel $i \in \mathbb{N}$, la série $\sum_{j \in \mathbb{N}} u_{i,j}$ converge.

  On note alors A_i *sa somme.*

- •• La série $\sum_{i \in \mathbb{N}} A_i$ converge, et a pour somme S.

Alors :

- Pour tout entier naturel $j \in \mathbb{N}$, la série $\sum_{i \in \mathbb{N}} u_{i,j}$ converge.

  On note alors B_j *sa somme.*

- •• La série $\sum_{j \in \mathbb{N}} B_j$ converge, et a pour somme S.

On dit alors que la ***série double*** $\sum_{(i,j) \in \mathbb{N}^2} u_{i,j}$ ***converge***, ou encore que la famille $(u_{i,j})_{(i,j) \in \mathbb{N}^2}$ est ***sommable***, et admet pour ***somme*** :

$$\boxed{S = \sum_{i=0}^{+\infty} \left(\sum_{j=0}^{+\infty} u_{i,j} \right) = \sum_{j=0}^{+\infty} \left(\sum_{i=0}^{+\infty} u_{i,j} \right)}.$$

Remarque *Autre manière de formuler les choses, le théorème de Fubini – Tonelli assure que, pour une suite double* $(u_{i,j})_{(i,j) \in \mathbb{N}^2}$ *à termes positifs, les séries numériques* $\sum_{j \in \mathbb{N}} \left(\sum_{i=0}^{+\infty} u_{i,j} \right)$ *et* $\sum_{i \in \mathbb{N}} \left(\sum_{j=0}^{+\infty} u_{i,j} \right)$ *sont de même nature, et ont même somme en cas de convergence.*

Exercice - type :

Etudier la convergence de la série double $\sum_{(i,j) \in \mathbb{N}^2} u_{i,j}$, où :

$$\forall (i,j) \in \mathbb{N}^2,\ u_{i,j} = \frac{i^j}{i!\, j!}.$$

Le cas échéant, calculer sa somme.

e. Théorème de comparaison pour des séries doubles

Théorème 1 ***(sous-familles des familles sommables à termes positifs)***
(convergence d'une « sous-série » d'une série double convergente à termes positifs)

Soit I et J des ensembles finis ou dénombrables.

Soit $\left(u_{i,j} \right)_{(i,j) \in I \times J}$ une famille double de réels ***positifs ou nuls*** .

On suppose que la série double $\sum_{(i,j) \in I \times J} u_{i,j}$ converge et admet pour somme S .

Alors, pour toute partie $I' \times J' \subset I \times J$, la série double $\sum_{(i,j) \in I' \times J'} u_{i,j}$ converge,

et a pour somme un réel S' vérifiant : $\boxed{S' \leq S}$.

Théorème 2 ***(théorème de comparaison pour les séries doubles à termes positifs)***

Soit I et J des ensembles finis ou dénombrables.

Soit $\left(u_{i,j} \right)_{(i,j) \in I \times J}$ et $\left(v_{i,j} \right)_{(i,j) \in I \times J}$ deux familles doubles de réels tels que :

- $\forall \left(i, j \right) \in I \times J$, $0 \leq u_{i,j} \leq v_{i,j}$.
- •• La série double $\sum_{(i,j) \in I \times J} v_{i,j}$ converge.

Alors :

- La série double $\sum_{(i,j) \in I \times J} u_{i,j}$ converge.
- •• $0 \leq \sum_{(i,j) \in I \times J} u_{i,j} \leq \sum_{(i,j) \in I \times J} v_{i,j}$.

f. Séries doubles absolument convergentes (ou sommables)

Définition 1

Soit I et J des ensembles finis ou dénombrables.

Soit $\left(u_{i,j} \right)_{(i,j) \in I \times J}$ une famille double de réels ***quelconques*** .

On dit que ***la série double*** $\sum_{(i,j) \in I \times J} u_{i,j}$ ***converge absolument*** , ou encore que la famille $\left(u_{i,j} \right)_{(i,j) \in I \times J}$ est ***sommable***, lorsque la famille $\left(\left| u_{i,j} \right| \right)_{(i,j) \in I \times J}$ est sommable, *i.e.* lorsque la série double $\sum_{(i,j) \in I \times J} \left| u_{i,j} \right|$ converge.

Théorème 1 ***(de Fubini)***

Soit I et J des ensembles finis ou dénombrables.

Soit $\left(u_{i,j} \right)_{(i,j) \in I \times J}$ une famille double de réels ***quelconques*** .

On suppose que la série double $\sum_{(i,j)\in I\times J} u_{i,j}$ converge **absolument**. *Alors :*

- $\bullet$ Pour tout entier $i \in I$, la famille $(u_{i,j})_{j\in J}$ est sommable (*i.e.* pour le programme, la « série » $\sum_{j\in J} u_{i,j}$ converge). *On note alors* A_i *sa somme.*

- $\bullet\bullet$ La famille $(A_i)_{i\in I}$ est sommable, et a pour somme S

  (*i.e.* pour le programme, la « série » $\sum_{i\in I} A_i$ converge, et a pour somme S).

Mézoci…

- $\bullet$ Pour tout entier $j \in J$, la famille $(u_{i,j})_{i\in I}$ est sommable (*i.e.* pour le programme, la « série » $\sum_{i\in I} u_{i,j}$ converge). *On note alors* B_j *sa somme.*

- $\bullet\bullet$ La famille $(B_j)_{j\in J}$ est sommable, et a pour somme S

  (*i.e.* pour le programme, la « série » $\sum_{i\in I} B_j$ converge, et a pour somme S).

On a de plus dans ce cas :

$$\boxed{S = \sum_{i\in I}\left(\sum_{j\in J} u_{i,j}\right) = \sum_{j\in J}\left(\sum_{i\in I} u_{i,j}\right) = \sum_{(i,j)\in I\times J} u_{i,j}}\ .$$

Exercice - type :

Etudier la convergence de la série double $\sum_{(i,j)\in\mathbb{N}^2} u_{i,j}$, où :

$$\forall\,(i,j)\in\mathbb{N}^2,\ u_{i,j} = \frac{(-1)^{i+j}\, i^j}{i\,!\,j\,!}.$$

Le cas échéant, calculer sa somme.

Théorème 2 ***(de sommation par paquets)***

Soit I et J des ensembles finis ou dénombrables.

Soit $(u_{i,j})_{(i,j)\in I\times J}$ une famille double de réels ***quelconques***.

On suppose que la série double $\sum_{(i,j)\in I\times J} u_{i,j}$ converge **absolument**. *Alors :*

Pour toute partition $(A_k)_{k\in\mathbb{N}}$ de $I\times J$, on a :

- $\bullet$ La famille $(u_{i,j})_{(i,j)\in A_k}$ est sommable, de somme $\sum_{(i,j)\in A_k} u_{i,j}$.

- $\bullet\bullet$ La série $\sum_{k\in\mathbb{N}}\left(\sum_{(i,j)\in A_k} u_{i,j}\right)$ converge absolument.

••• $\boxed{\displaystyle\sum_{(i,j)\in I\times J} u_{i,j} = \sum_{k=0}^{+\infty}\left(\sum_{(i,j)\in A_k} u_{i,j}\right) = \sum_{k\in\mathbb{N}}\left(\sum_{(i,j)\in A_k} u_{i,j}\right)}$.

Exercice - type :

Etudier la convergence de la série double $\displaystyle\sum_{(i,j)\in\mathbb{N}^2} u_{i,j}$, où :

$$\forall\,(i,j)\in\mathbb{N}^2,\ u_{i,j} = \frac{1}{(i+j)!}.$$

Le cas échéant, calculer sa somme.

II – Variables aléatoires réelles discrètes

1. Variable aléatoire réelle discrète

Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

On appelle ***variable aléatoire réelle discrète*** sur $(\Omega, \mathcal{A}, \mathbb{P})$ toute application $X:\ \Omega\to\mathbb{R}$ telle que :

i – $X(\Omega)$ est un sous – ensemble fini ou dénombrable de $\mathbb{R}$.

ii – Pour tout sous – ensemble A de $X(\Omega)$, $(X\in A)$ est un évènement.

On dit alors que X suit une loi finie (respectivement une loi infinie discrète) lorsque $X(\Omega)$ est fini (respectivement, lorsque $X(\Omega)$ est infini dénombrable).

2. Loi d'une variable aléatoire réelle discrète

Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

a. *Proposition – définition (loi d'une vard)*

L'application :

$$\mathbb{P}_X : \begin{cases} \mathcal{P}(X(\Omega)) \to \mathbb{R} \\ A\subset X(\Omega) \mapsto \mathbb{P}(X\in A) = \mathbb{P}(X^{-1}\in A) \end{cases}$$

est une mesure de probabilités sur $\left(X(\Omega), \mathcal{P}(X(\Omega))\right)$.

On dit que $\mathbb{P}_X$ est la loi de probabilités de la vard X.

b. *Conséquence pratique (détermination de la loi d'une vard)*

D'après ce qui précède, préciser la loi de la vard X revient à décrire l'application $\mathbb{P}_X$, ce qui revient encore à déterminer son ensemble de départ, à savoir $X(\Omega)$, et à donner, pour tout sous – ensemble A de $X(\Omega)$, la valeur de $\mathbb{P}(X\in A)$. Ecrire rigoureusement ce dernier point serait a priori technique et fastidieux (si l'on suppose par exemple que $X(\Omega)=\mathbb{N}$, il faudrait que la description écrite donne, par exemple aussi, $\mathbb{P}(X \in 2\mathbb{N})$, $\mathbb{P}('X$ est le produit de cinq nombres premiers$')$, etc.).

Heureusement, la remarque qui suit permet de simplifier le travail à fournir :

Constatation (caractérisation de la loi d'une variable aléatoire réelle discrète)

Pour tout $A \in X(\Omega)$, on a $\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x)$.

On en vient doucement à la conséquence pratique annoncée

D'après la constatation précédente, si l'on connaît $X(\Omega)$, et si l'on connaît l'application

$$\mathcal{L}_X : \begin{cases} X(\Omega) & \to \mathbb{R} \\ x \in X(\Omega) & \mapsto \mathbb{P}(X = x) \end{cases},$$

alors on connaît la loi de X.

L'application $\mathcal{L}_X$ ci – dessus est d'ailleurs parfois appelée, pour cette raison mais par abus de langage, loi de X.

Et alors :

Lorsqu'il est demandé de déterminer la loi de la vard X, il s'agit en pratique de :

- déterminer l'ensemble $X(\Omega)$;
- •• donner, pour tout x de $X(\Omega)$, la valeur de la probabilité $\mathbb{P}(X = x)$.

c. ***Un premier exemple***

Application 1 : une vard de loi géométrique

On lance un dé équilibré à six faces numérotées de 1 à 6. Notons X la variable aléatoire représentant le nombre de lancers nécessaires pour obtenir un premier 6.

Déterminer la loi de X.

Réponse

- X est une variable aléatoire discrète, car son univers – image est l'ensemble dénombrable $\mathbb{N}^*$.
- •• Notons, pour tout $k \in \mathbb{N}^*$, A_k l'évènement : « Un 6 est obtenu lors du $k^{\text{ième}}$ lancer ».

Pour tout $n \in \mathbb{N}^*$, l'évènement $(X = n)$ est réalisé si et seulement si un 6 est obtenu, pour la première fois, lors du $n^{\text{ième}}$ lancer, donc si et seulement si le $n^{\text{ième}}$ lancer amène effectivement un 6, alors que les $n-1$ jets précédents n'en ont pas apporté.

On a donc $(X = n) = \left(\bigcap_{k=1}^{n-1} \overline{A_k}\right) \cap A_n$, et, par indépendance des lancers, on en déduit que

$$\mathbb{P}(X = n) = \left(\prod_{k=1}^{n-1} \mathbb{P}(\overline{A_k})\right).\mathbb{P}(A_n) = \boxed{\left(\frac{5}{6}\right)^{n-1} \frac{1}{6}}.$$

d. ***Variables aléatoires discrètes de même loi***

"Définition" (variables aléatoires discrètes de même loi)

Soient deux vards X et Y, définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On dit que X et Y ont la même loi lorsque l'une des propositions équivalentes suivantes est vérifiée :

i – $\forall x \in X(\Omega), \ \mathbb{P}(X = x) = \mathbb{P}(Y = x)$.

ii – Pour tout x réel tel que $\mathbb{P}(X = x) \neq 0, \mathbb{P}(X = x) = \mathbb{P}(Y = x)$.

iii – Pour tout x réel, $\mathbb{P}(X = x) = \mathbb{P}(Y = x)$.

Si l'on se réfère à la définition de la loi d'une variable aléatoire discrète, les vards X et Y ont la même loi si et seulement si $X(\Omega) = Y(\Omega)$ et pour tout $x \in X(\Omega)$, $\mathbb{P}(X = x) = \mathbb{P}(Y = x)$; il s'agit donc ici d'une extension de définition, qui s'avérera souvent nécessaire lors de futurs exemples pratiques

Remarque

Deux variables aléatoires ayant la même loi n'ont aucune raison d'être égales : considérons, par exemple, l'exemple débile d'un lancer d'une pièce équilibrée, et des variables aléatoires X et Y données par

$$\begin{cases} X = 1 \text{ si on obtient " Face ", et } X = 0 \text{ sinon} \\ Y = 1 \text{ si on obtient " Pile ", et } Y = 0 \text{ sinon} \end{cases}.$$

Il est clair que X et Y ont la même loi ($X(\Omega) = Y(\Omega) = \{0;1\}$ et $\mathbb{P}(X = 0) = \mathbb{P}(Y = 0) = \frac{1}{2}$, $\mathbb{P}(X = 1) = \mathbb{P}(Y = 1) = \frac{1}{2}$), et il est non moins clair que X et Y ne sont pas égales (remarquons, par exemple, que $|X - Y|$ est une variable certaine égale à 1 …).

3. Caractérisation des lois de variables aléatoires discrètes

En première année, on a vu que les lois de variables aléatoires discrètes d'univers – image fini sont caractérisées de la manière suivante :

Proposition 1

Soient $n \in \mathbb{N}^*$, $E = \{x_k, 1 \leq k \leq n\}$ un ensemble de n réels, et $(p_k)_{1 \leq k \leq n}$ une suite finie de n réels tels que :

- pour tout $k \in \{1, ..., n\}$, $p_k \geq 0$;
- • $\sum_{k=1}^{n} p_k = 1$.

Alors il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et une variable aléatoire discrète X définie sur cet espace, tels que $X(\Omega) = E$ et $\forall\ k \in \{1, ..., n\}$, $\mathbb{P}(X = x_k) = p_k$.

Les lois infinies discrètes d'univers – image infini (dénombrable) sont caractérisées de manière analogue :

Proposition 2

Soient $E = \{x_k, k \in \mathbb{N}\}$ un ensemble de réels, et $(p_k)_{k \in \mathbb{N}}$ une suite de réels tels que :

- pour tout $k \in \mathbb{N}$, $p_k \geq 0$;
- • la série $\sum_{k \geq 0} p_k$ converge, et $\sum_{k=0}^{+\infty} p_k = 1$.

Alors il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et une variable aléatoire discrète X sur cet espace tels que $X(\Omega) = E$ et $\forall\ k \geq 0, \mathbb{P}(X = x_k) = p_k$.

Ce résultat peut s'étendre à un ensemble de la forme $E = \{x_n,\ n \in \mathbb{N} \text{ et } n \geq n_0\}$, où $n_0 \in \mathbb{N}$.

On admettra ce résultat, dont la démonstration est parfaitement inintéressante.

On admettra également la proposition suivante :

Proposition 3

> Soient $(\Omega, \mathcal{A})$ un espace probabilisable, et X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{A})$, d'univers-image dénombrable $E = \{x_n,\ n \in \mathbb{N}\}$.
>
> Soit par ailleurs $(p_n)_{n \in \mathbb{N}}$ une suite de réels tels que :
>
> - pour tout $n \in \mathbb{N}$, $p_n \geq 0$;
> - la série $\sum\limits_{n \geq 0} p_n$ converge, et $\sum\limits_{n=0}^{+\infty} p_n = 1$.
>
> Alors il existe une mesure de probabilités $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ telle que : $\forall\ n \geq 0$, $\mathbb{P}(X = x_n) = p_n$.

***Exemple* 1**

L'énoncé : " Soit X une variable aléatoire à valeurs dans $\{1, ..., n\}$, et de loi donnée par

$$\forall\ k \in \{1, ..., n\},\ P(X = k) = \frac{2k}{n(n+1)} \text{ "}$$

a bien un sens, car il existe une telle variable aléatoire. En effet :

- pour tout $k \in \{1, ..., n\}$, le réel $p_k = \dfrac{2k}{n(n+1)}$ est positif ou nul ;
- on a $\sum\limits_{k=1}^{n} p_k = \dfrac{2}{n(n+1)} \sum\limits_{k=1}^{n} k = \dfrac{2}{n(n+1)} \dfrac{n(n+1)}{2} = 1$,

et il s'ensuit qu'il existe une variable aléatoire X, d'univers – image égal à $\{1, ..., n\}$, telle que pour tout $k \in \{1, ..., n\}$, $\mathbb{P}(X = k) = p_k$.

Exemple 2

Soit la suite $(p_k)_{k \in \mathbb{N}^*}$ définie par : $\forall\ k \in \mathbb{N}^*$, $p_k = \dfrac{1}{k(k+1)}$. Montrons qu'il existe une variable aléatoire X, d'univers – image $\mathbb{N}^*$, telle que pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = p_k$.

Pour tout $k \in \mathbb{N}^*$, on a $p_k \geq 0$. De plus, pour tout $N \in \mathbb{N}^*$:

$$\sum_{k=1}^{N} p_k = \sum_{k=1}^{N} \frac{1}{k(k+1)} = \sum_{k=1}^{N} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \sum_{k=1}^{N} \frac{1}{k} - \sum_{k=1}^{N} \frac{1}{k+1},$$

et en changeant d'indice dans la somme $\sum\limits_{k=1}^{N} \dfrac{1}{k+1}$ on obtient : $\sum\limits_{k=1}^{N} p_k = \sum\limits_{k=1}^{N} \dfrac{1}{k} - \sum\limits_{k=2}^{N+1} \dfrac{1}{k}$.

On en déduit que $\sum\limits_{k=1}^{N} p_k = 1 - \dfrac{1}{N+1}$, puis que $\lim\limits_{N \to +\infty} \left(\sum\limits_{k=1}^{N} p_k\right) = 1$.

Alors, par définition de la convergence d'une série, la série de terme général p_k est convergente, et l'on a $\sum_{k=1}^{+\infty} p_k = 1$. Il existe donc une variable discrète X définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ telle que

$X(\Omega) = \mathbb{N}^*$, et pour tout $k \in \mathbb{N}^*$: $\mathbb{P}(X = k) = p_k$.

4. Système complet d'évènements associé à une variable aléatoire discrète

a. *Proposition*

> Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Alors la famille $((X = x))_{x \in X(\Omega)}$ est un système complet d'évènements.

On dit que la famille $((X = x))_{x \in X(\Omega)}$ est le système complet d'évènements associé à la variable aléatoire X.

Démonstration

- Pour tout $(x, y) \in (X(\Omega))^2$ tel que $x \neq y$, les évènements $(X = x)$ et $(X = y)$ sont clairement incompatibles.
- • On a :

$$\begin{aligned} \bigcup_{x \in X(\Omega)} (X = x) &= \bigcup_{x \in X(\Omega)} X^{-1}(\{x\}) && \text{par définition des évènements } (X = x) \\ &= X^{-1}\left(\bigcup_{x \in X(\Omega)} \{x\}\right) && \text{d'après les propriétés des images réciproques} \\ &= X^{-1}(X(\Omega)) = \Omega, \end{aligned}$$

$\boxed{\text{la famille } ((X = x))_{x \in X(\Omega)} \text{ est donc bien un système complet d'évènements}}$.

b. *Conséquences*

Etant donnée une variable aléatoire discrète X :

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

Lorsque $X(\Omega)$ est infini, de la forme $X(\Omega) = \{x_n, n \geq n_0\}$, $\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = 1$ signifie que la série $\sum_{n \geq n_0} \mathbb{P}(X = x)$ converge absolument et a pour somme 1, et ceci ne dépend pas de l'indexation de $X(\Omega)$ choisie.

- • Pour tout sous – ensemble A de $\mathbb{R}$: $P(X \in A) = \sum_{x \in A \cap X(\Omega)} P(X = x)$.

- • • Pour tout évènement B :

$$\begin{aligned} \mathbb{P}(B) = \sum_{x \in X(\Omega)} \mathbb{P}(B \cap (X = x)) &= \sum_{x \in X(\Omega),\, \mathbb{P}(X = x) \neq 0} \mathbb{P}(B \cap (X = x)) \\ &= \sum_{x \in X(\Omega),\, \mathbb{P}(X = x) \neq 0} \mathbb{P}_{(X = x)}(B) \cdot P(X = x), \end{aligned}$$

que l'on écrit plus simplement, en posant $\mathbb{P}_{(X = x)}(B) \cdot P(X = x) = 0$ lorsque $(X = x)$ est un évènement négligeable :

$$\boxed{\mathbb{P}(B) = \sum_{x \in X(\Omega)} \mathbb{P}_{(X = x)}(B) \cdot P(X = x)} \quad :$$

c'est la formule des probabilités totales naturellement associée à B, et au système complet $((X = x))_{x \in X(\Omega)}$.

5. Fonction de répartition d'une variable aléatoire réelle discrète

a. *Définition*

Soit X une vard, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On appelle fonction de répartition de X, et l'on note F_X, l'application $F_X : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \mathbb{P}(X \leq x) \end{cases}$.

Premières propriétés

Soit X une variable aléatoire réelle discrète, et soit F_X la fonction de répartition de X. Alors :

i _ F_X est croissante.

ii _ $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(x) = 1$.

b. *Expression de la fonction de répartition à l'aide de la loi de probabilité*

Soit X une vard de fonction de répartition F_X.

On sait que pour tout sous – ensemble A de $\mathbb{R}$, $\mathbb{P}(X \in A) = \sum\limits_{x \in A \cap X(\Omega)} \mathbb{P}(X = x)$.

En considérant les ensembles $A_x =]-\infty, x]$, pour $x \in \mathbb{R}$, on obtient le résultat suivant :

Proposition

Pour tout $x \in \mathbb{R}$: $F_X(x) = \sum\limits_{\substack{u \in X(\Omega) \\ u \leq x}} \mathbb{P}(X = u)$.

En particulier, lorsque $X(\Omega) \subset \mathbb{N}$: $\forall\, n \in \mathbb{N},\ F_X(n) = \sum\limits_{k=0}^{n} \mathbb{P}(X = n)$.

Exemple (légèrement anticipé)

Considérons une variable X suivant la loi géométrique $\mathcal{G}(p)$, et déterminons sa fonction de répartition F_X.

Soit $x \in \mathbb{R}$.

- Si $x < 1$, alors : $\{n \in X(\Omega) \ / \ n \leq x\} = \{n \in \mathbb{N}^* \ / \ n \leq x\} = \varnothing$, et $F_X(x) = \sum\limits_{n \in \varnothing} P(X = n) = 0$;

- • Si $x \geq 1$, alors : $\{n \in X(\Omega) \ / \ n \leq x\} = \{n \in \mathbb{N}^* \ / \ n \leq x\} = [\![1, \lfloor x \rfloor]\!]$, donc :

$$F_X(x) = \sum_{n=1}^{\lfloor x \rfloor} P(X = n) = \sum_{n=1}^{\lfloor x \rfloor} p\, q^{n-1} = p\, \frac{1 - q^{\lfloor x \rfloor}}{1 - q} = 1 - q^{\lfloor x \rfloor}.$$

$\lfloor x \rfloor$ désigne ici la partie entière de x.

c. *Expression de la loi de probabilité à l'aide de la fonction de répartition*

On se borne ici au cas d'une variable aléatoire à valeurs dans $\mathbb{N}$. On a alors le résultat suivant :

Proposition

Soit X une vard de fonction de répartition F_X, et à valeurs dans $\mathbb{N}$. Alors pour tout $n \in \mathbb{N}$:

$$\mathbb{P}(X = n) = F_X(n) - F_X(n-1).$$

Démonstration

Puisque $X(\Omega) \subset \mathbb{N}$, on a pour tout $n \in \mathbb{N}$: $F_X(n) = \sum\limits_{k=0}^{n} \mathbb{P}(X = k)$ et $F_X(n-1) = \sum\limits_{k=0}^{n-1} \mathbb{P}(X = k)$. Le résultat en découle.

La démonstration reste valable lorsque $n = 0$, car alors $F_X(n-1) = F_X(-1) = 0$.

d. ***La fonction de répartition caractérise la loi d'une vard***

D'après ce qui précède, lorsque $X(\Omega) \subset \mathbb{N}$, la fonction de répartition de X permet de déterminer la loi de probabilité de cette variable aléatoire. On en déduit immédiatement que deux variables discrètes, à valeurs dans $\mathbb{N}$ et ayant même fonction de répartition, ont même loi. On admettra que ce résultat reste valable si ces variables aléatoires discrètes ne sont pas à valeurs dans $\mathbb{N}$:

Proposition

> Deux variables aléatoires X et Y définies sur le même espace probabilisé ont la même fonction de répartition si et seulement si elles ont la même loi.

e. ***Exemple d'utilisation des fonctions de répartition***

Les fonctions de répartition sont un outil important, en particulier, lorsque l'on recherche la loi d'un maximum de variables aléatoires.

Si l'on recherche la loi d'un minimum, on s'intéressera à sa fonction d'antirépartition $x \mapsto \mathbb{P}(X > x)$.

Exemple

On lance n fois un dé usuel. On note X le plus grand des résultats obtenus ; déterminer la loi de X.

Il est clair que la variable aléatoire X est à valeurs dans $\{1, ..., 6\}$.

Pour $k \in \{1, ..., n\}$, notons X_k le résultat du $k-$ième lancer. X_k suit la loi uniforme sur $\{1, ..., 6\}$, et l'on en déduit que pour tout $i \in \{1, ..., 6\}$, $\mathbb{P}(X_k \leq i) = \sum_{j=1}^{i} \mathbb{P}(X_k = j) = \sum_{j=1}^{i} \frac{1}{6} = \frac{i}{6}$.

Maintenant : pour $i \in \{1, ..., 6\}$, le plus grand des résultats obtenus est inférieur ou égal à i si et seulement si *tous* les résultats sont inférieurs ou égaux à i ; nous avons donc : $\mathbb{P}(X \leq i) = \mathbb{P}\left(\bigcap_{k=1}^{n} (X_k \leq i)\right)$. Les lancers sont indépendants ; il en résulte que :

$\mathbb{P}(X \leq i) = \prod_{k=1}^{n} \mathbb{P}(X_k \leq i) = \prod_{k=1}^{n} \frac{i}{6} = \left(\frac{i}{6}\right)^n$. Notons en outre que cette formule reste correcte lorsque $i = 0$.

On en conclut que pour tout $i \in \{1, ..., 6\}$: $\boxed{\mathbb{P}(X = i) = \mathbb{P}(X \leq i) - \mathbb{P}(X \leq i-1) = \left(\frac{i}{6}\right)^n - \left(\frac{i-1}{6}\right)^n}$.

III – Lois discrètes usuelles

On considère ici une variable aléatoire X, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1. Loi d'une variable aléatoire quasi – certaine

- On dit que X est la variable aléatoire certaine égale à a lorsque $X(\Omega) = \{a\}$, et dans ce cas : $\mathbb{P}(X = a) = 1$.
- • On dit que X est une variable quasi – certaine égale à a lorsque $\mathbb{P}(X = a) = 1$. La var. X est alors, presque sûrement, à valeurs dans $\{a\}$.

2. Loi de Bernoulli

a. ***Définition***

> Soit p un réel tel que $p \in [0, 1]$. Posons $q = 1 - p$.

On dit que X suit la loi de Bernoulli de paramètre p, et l'on note $X \hookrightarrow \mathcal{B}(1, p)$,

lorsque $X(\Omega) = \{0, 1\}$, $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = q$.

En particulier, si A est un élément de $\mathcal{A}$, l'application $\mathbf{1}_A : \begin{cases} \Omega \to \mathbb{R} \\ \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases} \end{cases}$ est une variable de Bernoulli, de paramètre $p = \mathbb{P}(X \in A)$. $\mathbf{1}_A$ est appelée variable indicatrice de l'évènement A.

b. ***Epreuves de Bernoulli***

Considérons une épreuve aléatoire ayant pour issues possibles le succès avec probabilité $p \in [0, 1]$, et l'échec avec probabilité $q = 1 - p$. Soit X la var. valant 1 si l'on a un succès, et 0 sinon. Alors $X \hookrightarrow \mathcal{B}(1, p)$.

Ceci justifie le nom donné d'épreuve de Bernoulli de paramètre p à toute épreuve aléatoire ayant deux issues possibles.

3. Lois uniformes discrètes

a. ***Définition***

Soit $F = \{x_1, x_2, ..., x_n\}$ un sous – ensemble fini de $\mathbb{R}$, de cardinal n.

On dit que X suit la loi uniforme sur F lorsque

- $X(\Omega) = F$;
- •• $\forall k \in \{1, ..., n\}$, $\mathbb{P}(X = x_k) = \dfrac{1}{n}$.

Cas particulier

Lorsque X suit la loi uniforme sur l'ensemble $\{1, 2, ..., n\}$, on note : $\boxed{X \hookrightarrow \mathcal{U}(n)}$.

b. ***Lois uniformes et choix aléatoires***

La loi uniforme permet de donner un sens à la locution « au hasard » : lorsque l'on dit, sans autre précision, que l'on choisit un élément « au hasard » dans l'ensemble $F = \{x_1, x_2, ..., x_n\}$, il est sous – entendu que la variable aléatoire représentant le nombre choisi suit la loi uniforme sur F.

4. Lois binomiales

a. ***Définition***

Soient n un entier naturel, et p un réel tel que $p \in [0, 1]$. Posons $q = 1 - p$.

On dit que X suit la loi binomiale de paramètres n et p, et l'on note $X \hookrightarrow \mathcal{B}(n, p)$, lorsque

- $X(\Omega) = \{0, ..., n\}$;
- •• $\forall k \in \{0, ..., n\}$, $\mathbb{P}(X = k) = \dbinom{n}{k} p^k q^{n-k}$.

On vérifie aisément que, lorsque $n = 1$, la loi $\mathcal{B}(n, p)$ est la loi de Bernoulli de paramètre p : il était donc légitime de noter $\mathcal{B}(1, p)$ cette loi.

b. ***Modèle usuel et schéma théorique***

La loi binomiale apparaît naturellement lors de tirages indépendants et avec remise :

Modèle usuel

Considérons une urne contenant une proportion p de boules blanches, où $p \in [0,1]$, et une proportion $q = 1 - p$ de boules noires. Effectuons, successivement et avec remise, n tirages au hasard d'une boule de cette urne, et notons X la variable aléatoire représentant le nombre de tirages amenant une boule blanche.

Alors X suit la loi binomiale de paramètres n et p.

De manière analogue :

Schéma théorique

Considérons une suite de n épreuves de Bernoulli indépendantes, de même paramètre p.

Soit X la variable aléatoire représentant le nombre de succès enregistrés lors de ces n épreuves.

Alors X suit la loi binomiale $\mathcal{B}(n, p)$.

5. Lois géométriques

a. *Définition*

Définition

Soit p un réel tel que $p \in \left]0,1\right[$; comme d'habitude, on notera $q = 1 - p$.

On dit que la variable aléatoire X suit la loi géométrique de paramètre p, et l'on note $\boxed{X \hookrightarrow \mathcal{G}(p)}$, lorsque

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

Vérification

Vérifions que l'on a bien défini la loi d'une variable aléatoire discrète. Pour cela, posons pour tout $k \in \mathbb{N}^*$: $p_k = pq^{k-1}$.

- On a pour tout $k \in \mathbb{N}^*$: $p_k \geq 0$;
- •• La série géométrique $\sum\limits_{k \geq 1} p_k = \sum\limits_{k \geq 1} pq^{k-1}$ converge, car sa raison, q, est strictement inférieure à 1 en valeur absolue, et

  l'on a : $\sum\limits_{k=1}^{+\infty} p_k = p\dfrac{1}{1-q} = 1$.

On sait alors qu'il existe bien des variables aléatoires X, à valeurs dans $\mathbb{N}^*$ et telles que pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = n) = p\,q^{k-1}$.

b. *Schéma théorique*

Considérons une suite infinie d'épreuves de Bernoulli de paramètre p, mutuellement indépendantes. Pour tout $n \in \mathbb{N}^*$, on note X_n la var. égale à 1 si un succès est enregistré lors de la n – ième épreuve, et à 0 sinon.

Notons X la variable aléatoire représentant le numéro de l'épreuve lors de laquelle on obtient le premier succès. Alors X suit la loi géométrique $\mathcal{G}(p)$. *En effet :*

- $X(\Omega) = \mathbb{N}^*$;
- •• Pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned}\mathbb{P}(X = n) &= \mathbb{P}\left[\left(\bigcap_{k=1}^{n-1}(X_k = 0)\right)\cap(X_n = 1)\right]\\ &= \left(\prod_{k=1}^{n-1}\mathbb{P}(X_k = 0)\right)P(X_n = 1) \qquad \text{car les épreuves sont indépendantes,}\\ &= \left(\prod_{k=1}^{n-1} q\right)p = pq^{n-1}, \qquad \text{ce qui achève la démonstration.}\end{aligned}$$

Remarque

Soulevons ici une difficulté courante : l'évènement " le résultat de chacune des épreuves est l'échec" n'est pas un évènement *impossible*, de sorte que la variable aléatoire X décrite ci – dessus n'est pas définie de manière entièrement correcte. Cette difficulté est classiquement levée en convenant d'attribuer à X la valeur 0 lorsqu'aucun succès n'est enregistré. On a alors :

$$\mathbb{P}(X = 0) = 1 - \sum_{n=1}^{+\infty} \mathbb{P}(X = n) = 1 - \sum_{n=1}^{+\infty} p q^{n-1} = 1 - \frac{p}{1-q} = 0 :$$

L'évènement $(X = 0)$ n'est pas *impossible*, mais il est *presque impossible*. Ainsi, l'ensemble des valeurs atteintes par X avec une probabilité non nulle est égal à $\mathbb{N}^*$, et pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(X = n) = p\,q^{n-1}$: X suit bien la loi $\mathcal{G}(p)$.

c. ***Caractérisation des lois géométriques par la propriété d'amnésie***

La proposition suivante établit l'une des propriétés fondamentales des lois géométriques, leur caractérisation par la propriété d'amnésie :

Proposition

Soit X une VAR discrète définie sur un univers probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ telle que :

- • $X(\Omega) \subset \mathbb{N}^*$.
- •• $\mathbb{P}(X = 1) \in \,]0,1[$. On pose $p = \mathbb{P}(X = 1)$.
- ••• $\forall (s,t) \in (\mathbb{N}^*)^2$, $\mathbb{P}_{[X > s]}(X > s+t) = \mathbb{P}(X > t)$.

Lorsqu'une VAR Y vérifie la propriété ••• , elle est dite ***sans mémoire*** *, ou* ***amnésique*** .

Alors, X *suit la loi géométrique de paramètre* p.

Réciproquement, toute var. de loi géométrique possède la propriété d'amnésie

Preuve :

Sens direct

Soit X une variable aléatoire vérifiant les propriétés •, •• et •••.

Soit $n \in \mathbb{N}$. En appliquant la relation ••• avec $s = 1$ et $t = n$, on obtient $\mathbb{P}_{[X>1]}(X > n+1) = \mathbb{P}(X > n)$, soit :

$\dfrac{\mathbb{P}((X > n+1) \cap (X > 1))}{\mathbb{P}(X > 1)} = \mathbb{P}(X > n)$. Mais $(X > n+1) \cap (X > 1) = (X > n+1)$ (si X est plus grande que $n+1$, elle est plus grande que 1 …). On en déduit que $\dfrac{\mathbb{P}(X > n+1)}{\mathbb{P}(X > 1)} = \mathbb{P}(X > n)$:

$\boxed{\text{la suite } (\mathbb{P}(X > n))_{n \in \mathbb{N}} \text{ est géométrique, de raison } \mathbb{P}(X > 1)}$.

Comme X est à valeurs dans $\mathbb{N}^*$, $\mathbb{P}(X > 1) = 1 - \mathbb{P}(X = 1) = q$, en notant comme d'habitude $q = 1 - p$, et $\boxed{\text{la suite } (\mathbb{P}(X > n))_{n \in \mathbb{N}} \text{ est géométrique de raison } q}$.

On en déduit que pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(X > n) = q^n\, \mathbb{P}(X > 0) = q^n$ (X étant à valeurs dans $\mathbb{N}^*$, $\mathbb{P}(X > 0) = 1$).

On a ensuite $\mathbb{P}(X = n) = \mathbb{P}(X > n-1) - \mathbb{P}(X > n) = q^{n-1} - q^n = p\,q^{n-1}$:

$\boxed{\text{la variable aléatoire } X \text{ suit la loi géométrique de paramètre } p}$.

Sens réciproque

Considérons réciproquement une variable aléatoire Y de loi $\mathcal{G}(p)$, avec $p \in \,]0,1[$. Alors Y est bien à valeurs dans $\mathbb{N}^*$, et l'on a $\mathbb{P}(Y = 1) = p \in \,]0,1[$; montrons maintenant que Y est amnésique. Pour tout $(s,t) \in (\mathbb{N}^*)^2$,

$$\mathbb{P}_{[X>s]}(X>s+t) = \frac{\mathbb{P}\big((X>s+t)\cap(X>s)\big)}{\mathbb{P}(X>s)}$$

$$= \frac{\mathbb{P}(X>s+t)}{\mathbb{P}(X>s)}, \quad \text{car } (X>s+t)\cap(X>s)=(X>s+t),$$

et comme on connaît la fonction d'antirépartition des lois géométriques, on en déduit que

$$\mathbb{P}_{[X>s]}(X>s+t) = \frac{q^{s+t}}{q^{s}} = q^{t} = \mathbb{P}(X>t) :$$

$\boxed{\text{Les variables aléatoires suivant des lois géométriques sont amnésiques}}$.

Application 2

n joueurs lancent en même temps une pièce honnête. Chaque joueur est gagnant s'il obtient une face de la pièce que tous les autres joueurs n'ont pas.

1. Quelle est la probabilité qu'il y ait un gagnant à une partie donnée ?

2. Soit X le nombre de parties nécessaires pour obtenir un gagnant.
Déterminer la loi de X.

1. Notons Y le nombre de « Face » obtenus lors de la partie. Les n lancers sont indépendants, et, pour chacun d'entre eux, la probabilité d'obtenir « Face » vaut $\frac{1}{2}$, donc Y suit la loi $\mathcal{B}\left(n,\frac{1}{2}\right)$.

Il y a un gagnant à la partie si un seul joueur obtient « Face », ou si un seul joueur obtient « Pile », la probabilité qu'il y ait un gagnant à la partie est donc $p=\mathbb{P}\big((Y=1)\cup(Y=n-1)\big)$. On distingue alors deux cas :

- Si $n=2$, les évènements $(Y=1)$ et $(Y=n-1)$ sont égaux, d'où :

$$\boxed{p=\mathbb{P}(Y=1)=\binom{2}{1}\frac{1}{2^{2}}=\frac{1}{2}}.$$

- •• Si $n\neq 2$, $(Y=1)$ et $(Y=n-1)$ sont deux évènements incompatibles, et ainsi

$$p=\mathbb{P}(Y=1)+\mathbb{P}(Y=n-1)=\binom{n}{1}\frac{1}{2}\frac{1}{2^{n-1}}+\binom{n}{1}\frac{1}{2}\frac{1}{2^{n-1}}, \text{ soit } \boxed{p=\frac{n}{2^{n-1}}}.$$

2. Les parties sont indépendantes, et, pour chacune d'entre elles, la probabilité qu'il y ait un gagnant est la même, à savoir p. On sait alors que la variable aléatoire X égale au nombre de parties nécessaires pour qu'il y ait un gagnant suit la loi géométrique de paramètre p :

$\boxed{X \text{ est à valeurs dans } \mathbb{N}^*, \text{ et pour tout } n\in\mathbb{N}^*,\ \mathbb{P}(X=n)=p\,q^{n-1}}$, où $q=1-p$.

7. *Lois de Poisson*

a. *Définition*

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$. Soit λ un réel strictement positif.

On dit que X suit la loi de Poisson de paramètre λ, et l'on note $X \hookrightarrow \mathcal{P}(\lambda)$, lorsque

- $X(\Omega)=\mathbb{N}$;
- •• $\forall\, k\in\mathbb{N},\ \mathbb{P}(X=k)=\mathrm{e}^{-\lambda}\dfrac{\lambda^{k}}{k!}$.

Vérification

Vérifions que l'on a bien défini la loi d'une variable aléatoire discrète.

On pose pour tout $k \in \mathbb{N}$: $p_k = \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!}$.

- On a pour tout $k \in \mathbb{N}$: $p_k \geq 0$;
- • la série exponentielle $\sum\limits_{k \geq 0} p_k = \mathrm{e}^{-\lambda} \sum\limits_{k \geq 0} \dfrac{\lambda^k}{k!}$ converge, et : $\sum\limits_{k=0}^{+\infty} p_k = \mathrm{e}^{-\lambda} \mathrm{e}^{\lambda} = 1$.

Il existe donc bien un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, et une variable aléatoire X définie sur cet espace, tels que $X(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}, \mathbb{P}(X = 0) = \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!}$.

b. ***Lois de Poisson, lois des évènements rares***

On ne donnera pas à proprement parler de modèle théorique pour les lois de Poisson. Notons cependant que ces lois apparaissent, sous certaines conditions, comme lois limites de lois binomiales… nous préciserons ce fait en fin de chapitre. Retenons dès maintenant que les lois de Poisson permettent de modéliser le nombre de succès enregistrés lors d'un grand nombre d'expériences aléatoires indépendantes, ayant toutes une même probabilité de succès, faible ; c'est pourquoi on qualifie parfois ces lois de *lois des évènements rares*.

Exemples

Soit X le nombre d'appels enregistrés, pendant une durée donnée, par le standard téléphonique du syndicat d'initiative de Saint – Firmin – Les Bains :

On peut considérer qu'un grand nombre d'individus est susceptible de contacter ce standard, mais que la probabilité qu'une personne donnée appelle effectivement pendant la période considérée est très faible. Si l'on suppose de plus que chaque individu a la même probabilité d'appeler le standard, et que les appels sont indépendants les uns des autres, alors on modélisera le nombre X par une variable aléatoire suivant une loi de Poisson (de paramètre à préciser).

De la même façon, on pourra modéliser par des variables aléatoires suivant une loi de Poisson :

-- Le nombre de véhicules se présentant à un péage pendant une période fixée,

-- Le nombre d'assiettes cassées, un jour donné, par un serveur à l'adresse moyenne ;

-- le nombre de pièces défectueuses sortant, pendant une période donnée, d'une chaîne de montage…

c. ***Conditionnement Poisson / binomial : un exercice – type***

Exercice – type 1

Le nombre N de yahourts mangés par Erwann en TD après rafle au réfectoire suit une loi de Poisson de paramètre λ. Ces yahourts sont expédiés indépendamment les uns des autres.

La probabilité pour que l'un de ces yahourts soit aux fruits est égale à t.

On note X la variable aléatoire représentant le nombre de yahourts aux fruits ingurgités ; Y est la var. égale au nombre des autres yahourts. On a donc $N = X + Y$.

1. Calculer, pour tout $(n, k) \in \mathbb{N}^2$, la probabilité conditionnelle $\mathbb{P}_{(N=n)}(X = k)$.
2. En déduire que X suit la loi de Poisson de paramètre λt .
3. Déterminer la loi de Y.
4. Montrer que les variables aléatoires X et Y sont indépendantes.

1. Supposons l'évènement $(N = n)$ réalisé. Alors Erwann se tape n yahourts ; chacun d'entre eux est aux fruits avec la même probabilité t, et ce indépendamment des autres. Conditionnellement à $(N = n)$, la variable aléatoire X représentant le nombre de yahourts aux fruits engloutis suit donc la loi $\mathcal{B}(n, t)$:

$$\boxed{\forall\, k \in [\![0, n]\!],\ \mathbb{P}\left(X = k \,/\, N = n \right) = \binom{n}{k} t^{k} \left(1 - t \right)^{n - k}},$$

et pour tout $k > n$, $\mathbb{P}\left(X = k \,/\, N = n \right) = 0$.

2. $\left(N = n \right)_{n \in \mathbb{N}}$ est un système complet d'évènements (c'est le système complet d'évènements associé à la var N).

La formule des probabilités totales assure que la série $\sum\limits_{n \geq 0} \mathbb{P}\left(X = k \,/\, N = n \right) \mathbb{P}\left(N = n \right)$ converge, et que

$$\mathbb{P}\left(X = k \right) = \sum_{n = 0}^{+\infty} \mathbb{P}\left(X = k \,/\, N = n \right) \mathbb{P}\left(N = n \right).$$

D'après la question précédente, la probabilité conditionnelle $\mathbb{P}\left(X = k \,/\, N = n \right)$ est nulle dès que $n < k$, d'où :

$\mathbb{P}\left(X = k \right) = \sum\limits_{n = k}^{+\infty} \mathbb{P}\left(X = k \,/\, N = n \right) \mathbb{P}\left(N = n \right)$. En explicitant les probabilités mises en jeu dans cette somme, on obtient

$\mathbb{P}\left(X = k \right) = \sum\limits_{n = k}^{+\infty} \binom{n}{k} t^{k} \left(1 - t \right)^{n - k} \mathrm{e}^{-\lambda} \dfrac{\lambda^{n}}{n!}$, d'où :

$$\begin{aligned}
\mathbb{P}\left(X = k \right) &= \frac{\mathrm{e}^{-\lambda}}{k!} t^{k} \sum_{n = k}^{+\infty} \frac{n!}{\left(n - k \right)!} \left(1 - t \right)^{n - k} \frac{\lambda^{n}}{n!} = \frac{\mathrm{e}^{-\lambda} \lambda^{k}}{k!} t^{k} \sum_{n = k}^{+\infty} \frac{1}{\left(n - k \right)!} \left(1 - t \right)^{n - k} \lambda^{n - k} \\
&= \frac{\mathrm{e}^{-\lambda}}{k!} \left(\lambda t \right)^{k} \sum_{n = k}^{+\infty} \frac{\left(\left(1 - t \right) \lambda \right)^{n - k}}{\left(n - k \right)!},
\end{aligned}$$

puis en changeant d'indice : $\mathbb{P}\left(X = k \right) = \dfrac{\mathrm{e}^{-\lambda}}{k!} \left(\lambda t \right)^{k} \sum\limits_{j = 0}^{+\infty} \dfrac{\left(\left(1 - t \right) \lambda \right)^{j}}{j!}$

On reconnaît la somme d'une série exponentielle, et l'on en déduit que : $\mathbb{P}\left(X = k \right) = \dfrac{\mathrm{e}^{-\lambda}}{k!} \left(\lambda t \right)^{k} \mathrm{e}^{\left(1 - t \right) \lambda} = \dfrac{\mathrm{e}^{-\lambda t}}{k!} \left(\lambda t \right)^{k}$.

On en conclut que $\boxed{X \text{ suit la loi de Poisson de paramètre } \lambda t}$.

3. La var Y joue le même rôle que X, lorsque l'on change t en $\left(1 - t \right)$. Par symétrie, $\boxed{Y \text{ suit donc la loi } \mathcal{P}\left(\lambda \left(1 - t \right) \right)}$.

4. Pour tout $\left(k, \ell \right) \in \mathbb{N}^{2}$:

$$\begin{aligned}
\mathbb{P}\left(X = k, Y = \ell \right) &= \mathbb{P}\left(X = k, N = k + \ell \right) && \text{car } N = X + Y \\
&= \mathbb{P}_{\left(N = k + \ell \right)}\left(X = k, N = k + \ell \right) \mathbb{P}\left(N = k + \ell \right).
\end{aligned}$$

Par hypothèse, N suit la loi $\mathcal{P}\left(\lambda \right)$, et l'on sait que la loi conditionnelle de X sachant $\left(N = k + \ell \right)$ est la loi $\mathcal{B}\left(n, k + \ell \right)$. On obtient donc $\mathbb{P}\left(X = k, Y = \ell \right) = \dbinom{k + \ell}{k} t^{k} \left(1 - t \right)^{\ell} \dfrac{\mathrm{e}^{-\lambda}}{\left(k + \ell \right)!} \lambda^{k + \ell} = \dfrac{t^{k} \left(1 - t \right)^{\ell} \mathrm{e}^{-\lambda} \lambda^{k + \ell}}{k!\,\ell!}$.

D'autre part, puisque X et Y suivent respectivement les lois $\mathcal{P}\left(\lambda t \right)$ et $\mathcal{P}\left(\lambda \left(1 - t \right) \right)$:

$$\mathbb{P}\left(X = k \right) \mathbb{P}\left(Y = \ell \right) = \left[\mathrm{e}^{-\lambda t} \frac{\left(\lambda t \right)^{k}}{k!} \right] \cdot \left[\mathrm{e}^{-\lambda \left(1 - t \right)} \frac{\left(\lambda \left(1 - t \right) \right)^{\ell}}{\ell!} \right] = \frac{\left(\lambda t \right)^{k} \left(\lambda \left(1 - t \right) \right)^{\ell} \mathrm{e}^{-\lambda}}{k!\,\ell!}.$$

On constate ainsi que pour tout $\left(k, \ell \right) \in \mathbb{N}^{2}$, $\mathbb{P}\left(X = k, Y = \ell \right) = \mathbb{P}\left(X = k \right) \mathbb{P}\left(Y = \ell \right)$:

$\boxed{\text{Les variables aléatoires } X \text{ et } Y \text{ sont indépendantes}}$.

Les lois présentées dans le paragraphe suivant ne sont pas au programme ; il n'est donc pas obligatoire de retenir les formules les concernant ; toutefois, il faut savoir retrouver de manière automatique une loi hypergéométrique ; les lois de Pascal ou les lois binomiales négatives sont des sous – produits directs des lois géométriques, que l'on rencontrera (probablement) fréquemment dans les épreuves de concours, et c'est pourquoi un minimum de savoir – faire est nécessaire les concernant.

8. **(*HP, mais en exercice*) Lois hypergéométriques ; lois de Pascal et lois binomiales négatives**

a. ***Lois hypergéométriques***

> ***Exercice – type 2***
>
> Une urne contient initialement N boules, parmi lesquelles une proportion $p \in \left]0,1\right[$ de boules noires, et une proportion $q = 1 - p$ de boules blanches.
>
> **1.** Soit $n \in \mathbb{N}^*$, $n < N$. On tire, successivement et sans remise, n boules de l'urne. Soit X_n une variable aléatoire modélisant le nombre de boules noires obtenues.
>
> **a.** Déterminer l'univers – image $X_n(\Omega)$ de X_n.
>
> *On distinguera entre autres les cas* $Np < n$ et $Np \geq n$.
>
> **b.** Justifier l'égalité : $\forall\ k \in \{0, ..., n\}$, $\mathbb{P}(X_n = k) = \dfrac{\dbinom{Np}{k}\dbinom{Nq}{n-k}}{\dbinom{N}{n}}$.
>
> ***On dit que*** X_n ***suit la loi hypergéométrique de paramètres N , n et p , notée*** $\mathcal{H}(N, n, p)$**.**
>
> **2.** On note Y le nombre de tirages nécessaires pour obtenir toutes les boules noires.
> Déterminer la loi de Y.

1.a. Le nombre maximal de boules noires que l'on peut obtenir est égal à $\begin{cases} n & \text{si } Np \geq n \\ Np & \text{si } Np < n \end{cases}$. De même, le nombre minimal de boules noires que l'on peut tirer est égal à $\begin{cases} 0 & \text{si } Nq \geq n \\ n - Nq & \text{si } Np < n \end{cases}$. Entre ces deux valeurs, tout est possible… Ainsi, l'univers – image de X est $X(\Omega) = \{\max(0, n - Nq), ..., \min(n, Np)\}$.

1.b. • Il y a n façons de choisir les boules obtenues lors des n tirages (sans se préoccuper de l'ordre d'obtention).

• • Soit $k \in \ 0, n\ $. L'évènement $(X_n = k)$ est réalisé si et seulement si, au cours des n tirages, on obtient k boules noires (et donc, fatalement, $n - k$ boules blanche). Il y a $\dbinom{Np}{k}$ façons de choisir, parmi les Np boules noires initialement dans l'urne, les k boules qui seront tirées ; de même, il y a $\dbinom{Nq}{n-k}$ façons de choisir les $n - k$ boules blanches tirées. Finalement, la probabilité d'obtenir k boules noires est :

$$\mathbb{P}(X_n = k) = \frac{\dbinom{Np}{k}\dbinom{Nq}{n-k}}{\dbinom{N}{n}}.$$

Notons que cette formule reste correcte lorsque $k \notin X(\Omega)$: en effet, si $k > p$, alors $\dbinom{Np}{k}$ est nul, et, lorsque $k < n - Nq$, on a $n - k > Nq$, d'où $\dbinom{Nq}{n-k} = 0$.

2. Il est clair que Y est à valeurs dans $\{Np, ..., N\}$.

Pour tout $k \in \{Np, ..., N\}$, l'évènement $(Y = k)$ est réalisé si et seulement si la dernière des Np boules noires est obtenue lors du k – ème tirage, donc si et seulement si : $\begin{cases} Np - 1 \text{ boules noires sont obtenues lors des } k - 1 \text{ premiers tirages} \\ \text{une boule noire est obtenue lors du tirage N°} k \end{cases}$.

Ainsi, en notant A_k l'évènement "le $k^{\text{ème}}$ tirage amène une boule noire" :

$$\mathbb{P}(Y=k)=\mathbb{P}\left(\left(X_{k-1}=Np-1\right)\cap A_k\right)=\mathbb{P}\left(X_{k-1}=Np-1\right).\mathbb{P}_{\left(\left(X_{k-1}=Np-1\right)\right)}\left(A_k\right).$$

On sait déjà d'après **Q1.b.** que $\mathbb{P}\left(X_{k-1}=Np-1\right)=\dfrac{\binom{Np}{Np-1}\binom{Nq}{k-Np}}{\binom{N}{k-1}}$. D'autre part, si l'évènement $\left(X_{k-1}=Np-1\right)$ est réalisé, alors il reste dans l'urne avant le $k^{\text{ème}}$ tirage $N-(k-1)$ boules, parmi lesquelles une seule boule noire ; par conséquent, $\mathbb{P}_{\left(\left(X_{k-1}=Np-1\right)\right)}\left(A_k\right)=\dfrac{1}{N-k+1}$.

On en déduit que $\mathbb{P}(Y=k)=\dfrac{\binom{Np}{Np-1}\binom{Nq}{k-Np}}{\binom{N}{k-1}}.\dfrac{1}{N-k+1}$, expression que l'on peut arranger un peu : on obtient

$$\begin{aligned}\mathbb{P}(Y=k)&=\frac{1}{Np-1}\frac{(Nq)!}{(k-Np)!\left(Nq-(k-Np)\right)!}\frac{(k-1)!(N-k+1)!}{N!}\frac{1}{N-k+1}\\&=\frac{1}{Np-1}\frac{(Nq)!}{(k-Np)!(N-k)!}\frac{(k-1)!(N-k)!}{N!}=\boxed{\frac{1}{Np-1}\frac{(Nq)!(k-1)!}{(k-Np)!N!}}\end{aligned}$$

b. Lois de Pascal

Exercice – type 3

Soient $r\in\mathbb{N}^*$ et $p\in\left]0,1\right[$. On considère une suite d'épreuves de Bernoulli indépendantes telles que pour chacune d'entre elles, la probabilité de succès soit égale à p.

On note X_r le nombre d'épreuves qu'il faut réaliser pour obtenir, pour la première fois, r succès, non forcément consécutifs (X_r est donc le numéro de l'épreuve où l'on obtient le $r^{\text{ème}}$ succès). On convient que $X_r=0$ si l'on n'obtient jamais r succès.

1.a. Soit n un entier supérieur ou égal à r.

Quelle est la probabilité d'obtenir moins de r succès lors des n premières épreuves ?

b. A l'aide de **1.a.**, montrer que $\left\{\omega\in\Omega / X_r(\omega)=0\right\}$ est un événement négligeable.

2. Déterminer la loi de X_r.

3. Ecrire qu'il s'agit bien d'une loi de probabilité.

En déduire que, pour tout entier $s\geq 1$, la série $\sum_{m\geq s}\dfrac{m!}{(m-s)!}q^{m-s}$ converge, et

que $\displaystyle\sum_{m=s}^{+\infty}\frac{m!}{(m-s)!}q^{m-s}=\frac{s!}{(1-q)^{s+1}}$.

4. (***anticipé***) Déterminer $\mathbb{E}\left(X_r\right)$ et $\mathbb{V}\left(X_r\right)$.

5. Déterminer la fonction génératrice G_{X_r} de X_r.

Les fonctions génératrices de variables aléatoires réelles discrètes seront étudiées plus tard ; pour l'instant, il suffit de savoir que G_{X_r} est la fonction définie sur $[-1,1]$ par :

$$\forall x\in[-1,1],\ G_{X_r}(x)=\sum_{n=r}^{+\infty}\mathbb{P}\left(X_r=n\right)x^n.$$

1.a. Notons Y_n le nombre de succès enregistrés lors des n premières épreuves. Lesdites épreuves sont indépendantes, et lors de chacune d'entre elles, la probabilité d'obtenir un succès est égale à p, la variable aléatoire Y_n suit donc la loi binomiale $\mathcal{B}(n, p)$; par conséquent, la probabilité d'obtenir (strictement) moins de r succès lors des n premières épreuves est :

$$\boxed{\mathbb{P}\left(Y_n \leq r-1\right) = \sum_{k=0}^{r-1} \binom{n}{k} p^k q^{n-k}},$$

où l'on a posé comme d'habitude $q = 1 - p$.

1.b. Si l'on n'obtient jamais r succès, alors on obtient toujours moins de r succès lors des n premières épreuves. Par conséquent, l'évènement $\left(X_r = 0\right)$ est inclus dans l'évènement $\left(Y_n \leq r-1\right)$, pour tout $n \geq r$, et l'on a donc :

$\forall\, n \geq r,\ 0 \leq \mathbb{P}\left(X_r = 0\right) \leq \mathbb{P}\left(Y_n \leq r-1\right)$. Or pour tout $n \geq r$, pour tout $k \in [\![0, r-1]\!]$,

$$\binom{n}{k} p^k q^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!} p^k q^{n-k} \leq \frac{n(n-1)\ldots(n-k+1)\ldots(n-r+1)}{1} q^{n-r+1} \leq n^r q^{n-r+1},$$

donc : $0 \leq \mathbb{P}\left(X_r = 0\right) \leq \mathbb{P}\left(Y_n \leq r-1\right) = \sum_{k=0}^{r-1} \binom{n}{k} p^k q^{n-k} \leq r\, n^r q^{n-r+1}$. On a $\lim_{n \to +\infty} r\, n^r q^{n-r+1} = 0$ par croissances comparées, donc par encadrement, $\mathbb{P}\left(X_r = 0\right) = 0$: $\boxed{\text{L'évènement } \left(X_r = 0\right) = \left\{\omega \in \Omega \,/\, X_r(\omega) = 0\right\} \text{ est négligeable}}$.

2. • Si l'on obtient r succès, il faut au moins r épreuves pour le faire, et, si l'on n'en obtient pas r, X_r est nulle : $\boxed{\text{la var. } X_r \text{ est donc à valeurs dans } \{0\} \cup [\![r, +\infty [\![}$.

•• D'après **1.b.**, $\boxed{\mathbb{P}\left(X_r = 0\right) = 0}$. Pour tout $n \geq r$, l'évènement $\left(X_r = n\right)$ est réalisé si et seulement si le $r^{\text{ième}}$ succès est enregistré lors de la $n^{\text{ième}}$ épreuve, donc ssi $\begin{cases} \text{exactement } r-1 \text{ succès sont enregistrés lors des } n-1 \text{ premières épreuves} \\ \text{un succès est enregistré lors de la } n^{\text{ième}} \text{ épreuve} \end{cases}$. En notant A_n l'évènement " un succès est enregistré lors de la $n^{\text{ième}}$ épreuve ", on a donc : $\left(X_r = n\right) = \left(Y_{n-1} = r-1\right) \cap A_n$.

Les évènements $\left(Y_{n-1} = r-1\right)$ et A_n étant indépendants, on en tire : $\mathbb{P}\left(X_r = n\right) = \mathbb{P}\left(Y_{n-1} = r-1\right) \cdot \mathbb{P}\left(A_n\right)$.

On a vu en **1.b.** que Y_{n-1} suit la loi $\mathcal{B}(n-1, p)$, et bien sûr $\mathbb{P}\left(A_n\right) = p$; on en conclut que

$$\boxed{\mathbb{P}\left(X_r = n\right) = \binom{n-1}{r-1} p^{r-1} q^{(n-1)-(r-1)} \cdot p = \binom{n-1}{r-1} p^r q^{n-r}}.$$

3. Dire que la loi de X_r est bien une loi de probabilité revient à dire que pour tout $n \in X_r(\Omega)$, $\mathbb{P}\left(X_r = n\right)$ est positif ou nul, et que la série $\sum_{n \in X_r(\Omega)} \mathbb{P}\left(X_r = n\right)$ est convergente, de somme égale à 1. Comme $\mathbb{P}\left(X_r = 0\right) = 0$, ce dernier point équivaut à la convergence de la série $\sum_{n \geq r} \mathbb{P}\left(X_r = n\right)$, et à l'égalité $\sum_{n=r}^{+\infty} \mathbb{P}\left(X_r = n\right) = 1$, soit : $\sum_{n=r}^{+\infty} \binom{n-1}{r-1} p^r q^{n-r} = 1$.

En arrangeant la somme précédente, on obtient $\sum_{n=r}^{+\infty} \frac{(n-1)!}{(n-r)!(r-1)!} (1-q)^r q^{n-r} = 1$, puis en posant $m = n-1$:

$$\sum_{m=r-1}^{+\infty} \frac{m!}{\big(m-(r-1)\big)!(r-1)!} (1-q)^r q^{m-(r-1)} = 1,$$

soit : $\sum_{m=r-1}^{+\infty} \frac{m!}{\big(m-(r-1)\big)!} q^{m-(r-1)} = \frac{(r-1)!}{(1-q)^r}$, et encore, en notant $s = r-1$: $\sum_{m=s}^{+\infty} \frac{m!}{(m-s)!} q^{m-s} = \frac{s!}{(1-q)^{s+1}}$.

On a retrouvé la convergence des séries géométriques dérivées $s-$ ièmes (de raison $q \in \,]0,1[$ ici), ainsi que la valeur de leur somme.

4. • On a $\mathbb{E}\left(X_r\right) = \sum_{n=r}^{+\infty} n\, \mathbb{P}\left(X_r = n\right)$, sous réserve de convergence absolue de la série correspondante.

Or pour tout $n \geq r$, $n\,\mathbb{P}\left(X_r = n\right) = n\binom{n-1}{r-1} p^r q^{n-r} = \dfrac{p^r}{(r-1)!}\dfrac{n}{(n-r)!} q^{n-r}$.

D'après **3.**, la série $\dfrac{p^r}{(r-1)!}\sum\limits_{n \geq r} \dfrac{n}{(n-r)!} q^{n-r}$ converge ; sa convergence est absolue, puisque tous ses termes sont positifs.

On en déduit que X_r possède une espérance, et, toujours d'après **3.** :

$$\boxed{\mathbb{E}\left(X_r\right) = \frac{p^r}{(r-1)!}\sum_{n=r}^{+\infty} \frac{n}{(n-r)!} q^{n-r} = \frac{p^r}{(r-1)!}\frac{r!}{(1-q)^{r+1}} = \frac{r}{p}}.$$

• • D'après le théorème du transfert, et sous réserve de convergence absolue :

$\mathbb{E}\left(\left(X_r + 1\right)X_r\right) = \sum\limits_{n=r}^{+\infty} (n+1)\,n\,\mathbb{P}\left(X_r = n\right)$. Or pour tout $n \geq r$,

$$\begin{aligned}(n+1)\,n\,\mathbb{P}\left(X_r = n\right) &= n(n+1)\binom{n-1}{r-1} p^r q^{n-r} = \frac{p^r}{(r-1)!}\frac{(n+1)!}{(n-r)!} q^{n-r} \\ &= \frac{p^r}{(r-1)!}\frac{(n+1)!}{\left((n+1)-(r+1)\right)!} q^{(n+1)-(r+1)}.\end{aligned}$$

Or la série $\sum\limits_{n \geq r} \dfrac{(n+1)!}{\left((n+1)-(r+1)\right)!} q^{(n+1)-(r+1)} = \sum\limits_{m \geq r+1} \dfrac{m!}{\left(m-(r+1)\right)!} q^{m-(r+1)}$ converge d'après **3.**, et

$\sum\limits_{n=r}^{+\infty} \dfrac{(n+1)!}{\left((n+1)-(r+1)\right)!} q^{(n+1)-(r+1)} = \dfrac{(r+1)!}{(1-q)^{r+2}}$. La convergence étant évidemment absolue, $\mathbb{E}\left(\left(X_r+1\right)X_r\right)$

existe, et l'on a : $\mathbb{E}\left(\left(X_r+1\right)X_r\right) = \dfrac{p^r}{(r-1)!}\dfrac{(r+1)!}{(1-q)^{r+2}} = \dfrac{r(r+1)}{p^2}$. On en déduit, par linéarité de l'espérance, que

$\mathbb{E}\left(X_r^2\right)$ existe, et que $\mathbb{E}\left(X_r^2\right) = \mathbb{E}\left(\left(X_r+1\right)X_r\right) - \mathbb{E}\left(X_r\right)$. La var. X_r possède donc une variance, et l'on a

$$\mathbb{V}\left(X_r\right) = \mathbb{E}\left(X_r^2\right) - \mathbb{E}\left(X_r\right)^2 = \mathbb{E}\left(\left(X_r+1\right)X_r\right) - \mathbb{E}\left(X_r\right) - \mathbb{E}\left(X_r\right)^2 = \frac{r(r+1)}{p^2} - \frac{r}{p} - \frac{r^2}{p^2},$$

d'où $\boxed{\mathbb{V}\left(X_r\right) = \dfrac{r}{p^2} - \dfrac{rp}{p^2} = \dfrac{rq}{p^2}}$.

5. • Pour tout $x \in [-1,1]$: $\forall\, n \geq r$, $\left|\mathbb{P}\left(X_r = n\right) x^n\right| \leq \mathbb{P}\left(X_r = n\right)$, et l'on sait que la série $\sum\limits_{n \geq r} \mathbb{P}\left(X_r = n\right)$ converge.

D'après le théorème de convergence par majoration pour les séries à terme général positif, $\sum\limits_{n \geq r} \left|\mathbb{P}\left(X_r = n\right) x^n\right|$ converge ; la série

$\sum\limits_{n \geq r} \left|\mathbb{P}\left(X_r = n\right) x^n\right|$ est donc absolument convergente, et de ce fait elle converge : la fonction génératrice G_{X_r} de X_r est bien

définie sur $[-1,1]$.

• • Pour tout $x \in [-1,1]$:

$$\begin{aligned}G_{X_r}(x) &= \sum_{n=r}^{+\infty} \mathbb{P}\left(X_r = n\right) x^n = \sum_{n=r}^{+\infty} \binom{n-1}{r-1} p^r q^{n-r} x^n = \frac{(px)^r}{(r-1)!}\sum_{n=r}^{+\infty} \frac{(n-1)!}{(n-r)!}(qx)^{n-r} \\ &= \frac{(px)^r}{(r-1)!}\sum_{m=r-1}^{+\infty} \frac{m!}{\left(m-(r-1)\right)!}(qx)^{m-(r-1)}.\end{aligned}$$

On reconnaît à nouveau une somme de série géométrique dérivée $(r-1)^{\text{ième}}$, et l'on en tire :

$$\boxed{G_{X_r}(x) = \frac{(px)^r}{(r-1)!}\frac{(r-1)!}{(1-qx)^r} = \left(\frac{px}{1-qx}\right)^r}.$$

c. *Lois binomiales négatives (exercice corrigé en fin de poly)*

Exercice – type 4

On rappelle, que, pour tout entier $r \in \mathbb{N}^*$ et pour tout réel q tel que $|q| < 1$, la série

$\sum_{m \geq r} \dfrac{m!}{(m-r)!} q^{m-r}$ converge, et que $\sum_{m=r}^{+\infty} \dfrac{m!}{(m-s)!} q^{m-r} = \dfrac{r!}{(1-q)^{r+1}}$.

Pour tout $k \in \mathbb{N}$, et pour tout $r \in \mathbb{Z}$, définissons le coefficient binomial généralisé $\begin{pmatrix} r \\ k \end{pmatrix}$ par :

$$\begin{pmatrix} r \\ k \end{pmatrix} = \frac{r(r-1)\dots(r-k+1)}{k!}.$$

On remarquera que, lorsque $r \in \mathbb{N}$, on retrouve le bon vieux coefficient du binôme, donné par

$$\begin{pmatrix} r \\ k \end{pmatrix} = \begin{cases} \dfrac{r!}{k!(r-k)!} & \text{si } 0 \leq k \leq r \\ 0 & \text{sinon} \end{cases}.$$

1. Montrer que pour tout $r \in \mathbb{Z}$, pour tous réels x et y tels que $y \neq 0$ et $\left|\dfrac{x}{y}\right| < 1$, la série

$\sum_{k \geq 0} \begin{pmatrix} r \\ k \end{pmatrix} x^k y^{r-k}$ converge, et $\sum_{k=0}^{+\infty} \begin{pmatrix} r \\ k \end{pmatrix} x^k y^{r-k} = (x+y)^r$.

Cette formule est appelée formule du binôme généralisée.

2. On considère un entier $r \geq 1$, ainsi qu'un réel a strictement négatif.

Montrer qu'il existe une variable aléatoire X à valeurs dans $\mathbb{N}$ et telle que pour tout $k \in \mathbb{N}$,

$\mathbb{P}(X = k) = \begin{pmatrix} -r \\ k \end{pmatrix} a^k (1-a)^{-r-k}$. *On dit que X suit la loi binomiale négative $\mathcal{B}(-r, a)$.*

3. Soit r un entier strictement positif. On considère une suite d'épreuves de Bernoulli indépendantes, ayant toutes même probabilité de succès $p \in \left]0, 1\right[$. On note X la variable aléatoire représentant le nombre d'échecs subis avant l'obtention du $r^{\text{ième}}$ succès. Montrer que X suit la loi binomiale négative $\mathcal{B}\left(-r, -\dfrac{q}{p}\right)$.

IV – FONCTIONS D'UNE OU PLUSIEURS VARD(S)

On admettra que

- • Toute fonction à valeurs réelles d'une variable aléatoire réelle discrète, ou de plusieurs variables aléatoires réelles discrètes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, est encore une variable aléatoire réelle discrète définie sur $(\Omega, \mathcal{A}, \mathbb{P})$.
- • • Par exemple, tout produit, tout quotient défini de vards définies sur le même espace probabilisé, est encore une vard.
- • • • De la même façon, toute combinaison linéaire de variables aléatoires réelles discrètes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, est encore une variable aléatoire réelle discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$. On remarquera alors que, muni des lois naturelles, l'ensemble des vards sur $(\Omega, \mathcal{A}, \mathbb{P})$ est un $\mathbb{R}$ – espace vectoriel.

V – INDÉPENDANCE

A_ INDÉPENDANCE

1. Indépendance de deux variables aléatoires discrètes

Définition

Soient X et Y deux vards définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On dit que X et Y sont indépendantes (pour la probabilité $\mathbb{P}$) lorsque

$$\boxed{\forall (A, B) \in (\mathcal{P}(\mathbb{R}))^2, \ \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A).\mathbb{P}(Y \in B)}.$$

Proposition

Soient X et Y deux vards définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Les vards X et Y sont indépendantes si et seulement si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,

les événements $(X = x)$ et $(Y = y)$ sont $\mathbb{P}$ – indépendants, *i.e.* si et seulement si :

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

2. Indépendance mutuelle d'une famille de variables aléatoires discrètes

Définition 1 (indépendance mutuelle d'une famille finie de vards)

Soient $X_1, ..., X_n$ des vards définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

$X_1, ..., X_n$ sont dites ***(mutuellement) indépendantes*** si et seulement si :

$$\boxed{\forall (x_1, ..., x_n) \in \prod_{i=1}^{n} X_i(\Omega), \ \mathbb{P}\left(\bigcap_{i=1}^{n} [X_i = x_i]\right) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)}.$$

Définition 2 (indépendance mutuelle d'une famille dénombrable de vards)

X_n, $n \in \mathbb{N}$, sont dites ***(mutuellement) indépendantes*** si et seulement si pour toute partie finie $I \subset \mathbb{N}$, la famille $(X_i)_{i \in I}$ est indépendante au sens de la définition précédente.

Remarque

On dit également que n variables aléatoires $X_1, ..., X_n$ sont indépendantes deux à deux lorsque pour tout $i \neq j$, X_i et X_j sont indépendantes.

On fera attention de ne pas confondre les notions d'indépendance mutuelle et d'indépendance deux à deux (la première est fondamentale, la deuxième ne sert pas à grand – chose).

3. Propriétés

Proposition 1

Soient X et Y deux vards **indépendantes** définies sur un même probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$,

f et g deux applications numériques définies respectivement sur $X(\Omega)$ et $Y(\Omega)$.

Alors, $f(X)$ et $g(Y)$ sont deux vards indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition 2 (lemme des sous – familles)

Soit $\left(X_n \right)_{n \in \mathbb{N}}$ une suite de vards (mutuellement) indépendantes, et $I \subset \mathbb{N}$.

Alors la sous – famille $\left(X_n \right)_{n \in I}$ est encore une famille de vards indépendantes.

Proposition 3 (lemme des coalitions)

Soit $\left(X_n \right)_{n \in \mathbb{N}}$ une suite de vards (mutuellement) indépendantes.

Alors toute vard fonction de certaines de ces variables est indépendante de toute fonction d'autres de ces variables aléatoires.

4. Schémas de Bernoulli

On appelle *schéma de Bernoulli* toute suite (finie ou infinie) d'épreuves aléatoires mutuellement indépendantes, ayant toutes deux issues possibles : succès et échec, et telles que, lors de chacune d'entre elles, la probabilité de succès est la même. Autrement dit, un schéma de Bernoulli est une suite d'épreuves de Bernoulli, indépendantes et ayant toutes le même paramètre. On appellera encore schéma de Bernoulli toute suite finie $\left(X_k \right)_{1 \leq k \leq n}$, ou infinie $\left(X_k \right)_{k \in \mathbb{N}^*}$, de variables de Bernoulli indépendantes et de même paramètre p ; on parlera alors de schéma de Bernoulli de paramètre p, (et de taille n dans le cas d'une suite finie $\left(X_k \right)_{1 \leq k \leq n}$). On a la proposition suivante :

Proposition

Soit $\left(X_k \right)_{1 \leq k \leq n}$ un schéma de Bernoulli de paramètre p et de taille $n \in \mathbb{N}^*$.

Posons $S_n = \sum_{k=1}^{n} X_k$. Alors S_n suit la loi $\mathcal{B}\left(n, p \right)$.

" Preuve "

S_n représente le nombre de succès enregistrés lors d'une succession d'épreuves de Bernoulli indépendantes, ayant toutes p pour probabilité de succès, on sait alors que S_n suit la loi $\mathcal{B}\left(n, p \right)$.

B_ THÉORÈMES DE STABILITÉ

1. Loi de la somme de deux vards à valeurs entières : la formule de convolution discrète

Soient X et Y deux variables aléatoires définies sur le même espace probabilisé, et à valeurs dans $\mathbb{N}$. On cherche la loi de la var $Z = X + Y$. Notons déjà que Z est, de manière évidente, à valeurs dans $\mathbb{N}$.

Pour tout $n \in \mathbb{N}$, on peut écrire, en utilisant le système complet d'évènements $\left(X = k \right)_{k \in \mathbb{N}}$ associé à la var. X, que la série $\sum_{k \geq 0} \mathbb{P}\left(Z = n, X = k \right)$ converge, et que $\mathbb{P}\left(Z = n \right) = \sum_{k=0}^{+\infty} \mathbb{P}\left(Z = n, X = k \right)$. On en déduit que

$$\mathbb{P}\left(Z = n \right) = \sum_{k=0}^{+\infty} \mathbb{P}\left(X + Y = n, X = k \right) = \sum_{k=0}^{+\infty} \mathbb{P}\left(X = k, Y = n - k \right).$$

On note de plus que, lorsque $k > n$, $n - k$ est strictement négatif, donc $\left(Y = n - k \right)$ est l'évènement impossible ; il en résulte qu'alors $\mathbb{P}\left(X = k, Y = n - k \right) = 0$. Par conséquent, $\boxed{\mathbb{P}\left(Z = n \right) = \sum_{k=0}^{n} \mathbb{P}\left(X = k, Y = n - k \right)}$.

Si l'on suppose de plus X et Y indépendantes, on obtient le résultat suivant :

$$\boxed{\mathbb{P}(Z=n)=\sum_{k=0}^{n}\mathbb{P}(X=k)\,\mathbb{P}(Y=n-k)}\,.$$

Cette formule doit être redémontrée systématiquement ; elle est appelée ***formule de convolution discrète***, et elle est à la base des démonstrations des théorèmes de stabilités énoncés ci – dessous.

2. Stabilité des lois binomiales

Théorème 1 (théorème de stabilité des lois binomiales)

> Soient m et n deux entiers naturels, et p un réel appartenant à $[0,1]$.
>
> Soient X et Y deux variables aléatoires indépendantes, suivant respectivement les lois binomiales $\mathcal{B}(m,p)$ et $\mathcal{B}(n,p)$.
>
> Alors la variable aléatoire $Z=X+Y$ suit la loi binomiale $\mathcal{B}(m+n,p)$.

Théorème 2 (généralisation)

> Soit p un réel appartenant à $[0,1]$. Soit un entier $r\in\mathbb{N}^*$, et r entiers naturels $n_1,\dots,n_r$; soient $X_1,\dots,X_r$ r variables aléatoires mutuellement indépendantes, telles que pour tout $k\in\{1,\dots,r\}$, $X_k \hookrightarrow \mathcal{B}(n_k,p)$.
>
> Alors la variable aléatoire $S_r=\sum_{k=1}^{r}X_k$ suit la loi $\mathcal{B}\left(\sum_{k=1}^{r}n_k,p\right)$.

Preuve

C'est une conséquence directe du théorème précédent : pour tout $j\in\{1,\dots,r\}$, on pose $S_j=\sum_{k=1}^{j}X_k$, et l'on définit la propriété $\mathcal{H}(j)$ par $\mathcal{H}(j)\Leftrightarrow S_j\hookrightarrow\mathcal{B}\left(\sum_{k=1}^{j}n_k,p\right)$. Montrons par récurrence que pour tout $j\in\{1,\dots,r\}$, $\mathcal{H}(j)$ est vérifiée.

Initialisation

$S_1=X_1$ suit évidemment la loi $\mathcal{B}(n_1,p)$.

Hérédité

Soit $j\in\{1,\dots,r-1\}$; supposons $\mathcal{H}(j)$. On a $S_{j+1}=\sum_{k=1}^{j+1}X_k=\sum_{k=1}^{j}X_k+X_{j+1}=S_j+X_{j+1}$, et :

- • par hypothèse, X_{j+1} suit la loi binomiale $\mathcal{B}(n_{j+1},p)$;
- •• par hypothèse de récurrence, $S_j\hookrightarrow\mathcal{B}\left(\sum_{k=1}^{j}n_k,p\right)$;
- ••• d'après le lemme des coalitions, S_j et X_{j+1} sont indépendantes.

D'après le théorème de stabilité des lois binomiales, S_{j+1} suit donc la loi binomiale de paramètres $\sum_{k=1}^{j}n_k+n_{j+1}=\sum_{k=1}^{j+1}n_k$ et p, d'où $\mathcal{H}(j+1)$.

Conclusion

On a $\mathcal{H}(1)$ et pour tout $j\in\{1,\dots,r-1\}$, $\mathcal{H}(j)\Rightarrow\mathcal{H}(j+1)$, d'où $\mathcal{H}(j)$ pour tout $j\in\{1,\dots,r\}$.

En particulier, on a $\mathcal{H}(r)$: $\boxed{S_r=\sum_{k=1}^{r}X_k \text{ suit la loi } \mathcal{B}\left(\sum_{k=1}^{r}n_k,p\right)}$.

3. Théorème de stabilité des lois de Poisson

Théorème 1 (somme de deux variables aléatoires poissonniennes indépendantes)

> Soient λ et μ deux réels strictement positifs. Soient X et Y deux variables aléatoires indépendantes, suivant respectivement les lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$.
>
> Alors la variable aléatoire $Z = X + Y$ suit la loi de Poisson $\mathcal{P}(\lambda + \mu)$.

Théorème 2 (généralisation)

> Soit un entier $n \in \mathbb{N}^*$, et n réels strictement positifs $\lambda_1, \dots, \lambda_n$; soient $X_1, \dots, X_n$ n variables aléatoires mutuellement indépendantes, telles que pour tout $k \in \{1, \dots, n\}$, $X_k \hookrightarrow \mathcal{P}(\lambda_k)$.
>
> Alors la variable aléatoire $S_n = \sum_{k=1}^{n} X_k$ suit la loi $\mathcal{P}\left(\sum_{k=1}^{n} \lambda_k \right)$.

4. Stabilité des lois de Pascal (exercice)

Exercice – type 5

Étant donnée une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, on dit que X suit la loi de Pascal de paramètres $r \in \mathbb{N}^*$ et $p \in]0, 1[$, et l'on note $X \hookrightarrow \mathcal{P}a(r, p)$, lorsque

- $X(\Omega) = \llbracket r, +\infty \llbracket$;
- • pour tout $n \geq r$, $\mathbb{P}(X = n) = \binom{n}{r} p^r q^{n-r}$ (avec $q = 1 - p$).

1. Soit s un entier positif ou nul. Démontrer que pour tout $n \in \llbracket s, +\infty \llbracket$: $\sum_{k=s}^{n} \binom{k}{s} = \binom{n+1}{s+1}$.

2. Soit p un réel tel que $p \in]0, 1[$, et soit $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes, et identiquement distribuées selon la loi $\mathcal{G}(p)$.

Montrer que, pour tout $r \in \mathbb{N}^*$, la variable aléatoire $S_r = \sum_{i=1}^{r} X_i$ suit la loi $\mathcal{P}a(r, p)$.

3. Soient X et Y deux variables aléatoires indépendantes, de lois respectives $\mathcal{P}a(r, p)$ et $\mathcal{P}a(s, p)$.

Montrer que la variable aléatoire $Z = X + Y$ suit la loi $\mathcal{P}a(r + s, p)$.

VI – MOMENTS DE VARIABLES ALÉATOIRES DISCRÈTES

A_ MOMENTS D'UNE VARIABLE ALÉATOIRE RÉELLE DISCRÈTE

1. Espérance d'une vard

Soit X une variable aléatoire réelle discrète, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

a. ***Cas d'un univers – image fini***

On suppose que $X(\Omega)$ est un ensemble fini.

On définit alors l'espérance de X, que l'on note $\mathbb{E}(X)$, par : $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\,\mathbb{P}(X = x)$.

Lorsque X prend un nombre fini de valeurs, cette variable aléatoire admet toujours une espérance.

b. ***Cas d'un univers – image dénombrable***

On suppose que $X(\Omega)$ est un ensemble infini de la forme : $X(\Omega) = \{x_n, n \in \mathbb{N}, n \geq n_0\}$.

On dit alors que X admet une espérance, ou que X est d'espérance finie, lorsque la série $\sum_{n \geq n_0} x_n\,\mathbb{P}(X = x_n)$ est absolument convergente, et dans ce cas on définit l'espérance de X, notée $\mathbb{E}(X)$, par :

$$\mathbb{E}(X) = \sum_{n = n_0}^{+\infty} x_n\,\mathbb{P}(X = x_n).$$

2. Premières propriétés

a. ***Linéarité***

Proposition

> Soient X et Y deux variables aléatoires discrètes, définies sur le même espace probabilisé, et soit λ un réel. On suppose que X et Y possèdent toutes deux une espérance.
>
> Alors la variable aléatoire $\lambda X + Y$ admet une espérance, et l'on a : $\mathbb{E}(\lambda X + Y) = \lambda\,\mathbb{E}(X) + \mathbb{E}(Y)$.

Cette propriété s'étend naturellement à une combinaison linéaire quelconque de variables aléatoires discrètes, admettant toutes une espérance.

La démonstration de ce résultat, faisant intervenir des séries doubles, n'est pas si évidente…

Exemple (espérance d'une loi hypergéométrique)

Soient $N \in \mathbb{N}^*$, $n \in \{0, ..., N\}$, et $p \in]0,1[$ tel que Np soit entier. On pose $q = 1 - p$. On considère une urne contenant Np boules blanches, et Nq boules noires. On effectue un tirage simultané de n boules de cette urne, au hasard, et l'on note X le nombre de boules blanches obtenues. Déterminons l'espérance de X.

Pour tout $i \in \{1, ..., Np\}$, notons X_i la variable aléatoire égale à 1 si la boule blanche N°i a été obtenue, et à 0 sinon.

- On choisit n boules parmi N : chaque boule est donc tirée avec une probabilité de $\frac{n}{N}$. Il s'ensuit que les X_i sont des variables de Bernoulli de paramètre $\frac{n}{N}$, et que $\forall i \in \{1, ..., Np\}$, $\mathbb{E}(X_i) = \frac{n}{N}$.

•• La somme $\sum_{i=1}^{Np} X_i$ représente le nombre total de boules blanches obtenues : $\sum_{i=1}^{Np} X_i = X$.

On a alors, par linéarité de l'espérance : $\mathbb{E}(X) = \sum_{i=1}^{Np} \mathbb{E}(X_i) = \sum_{i=1}^{Np} \frac{n}{N} = \frac{n}{N} Np = np$.

••• On sait que X suit la loi $\mathcal{H}(N, n, p)$. On a donc montré, sans calcul, que si X suit cette loi, alors $\mathbb{E}(X) = np$.

b. ***Croissance***

> Soient X et Y deux variables aléatoires discrètes, définies sur le même espace probabilisé et admettant une espérance.
>
> On suppose que $X \geq Y$ presque sûrement (p.s.) : $\mathbb{P}(X \geq Y) = 1$.
>
> Alors : $\mathbb{E}(X) \geq \mathbb{E}(Y)$.

Remarque

En choisissant Y constante égale à a, on en déduit que si $X \geq a$ p.s. , alors $\mathbb{E}(X) \geq a$. De même, si $X \leq a$ p.s. , alors $\mathbb{E}(X) \leq a$.

3. Théorème du transfert

Le résultat présenté ici, connu sous le nom de *théorème **de** transfert*, est essentiel. Il permet en effet de déterminer l'espérance de variables aléatoires, sans avoir nécessairement à en préciser la loi.
Le programme l'appelle théorème ***du*** transfert, nous en ferons de même désormais.

Théorème (du transfert)

Soit X une variable aléatoire réelle discrète, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et soit f une application définie sur $X(\Omega)$, à valeurs dans $\mathbb{R}$.

- On suppose que $X(\Omega)$ est un ensemble fini.

  Alors la variable aléatoire $f(X)$ possède une espérance, et l'on a :

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

•• On suppose que $X(\Omega)$ est un ensemble infini de la forme : $X(\Omega) = \{x_n, n \in \mathbb{N}, n \geq n_0\}$.

La variable aléatoire $f(X)$ possède une espérance si et seulement si la série $\sum_{n \geq n_0} f(x_n)\,\mathbb{P}(X = x_n)$ est absolument convergente, et dans ce cas on a :

$$\mathbb{E}(f(X)) = \sum_{n = n_0}^{+\infty} f(x_n)\,\mathbb{P}(X = x_n).$$

On admettra ce résultat (démonstration nettement hors – programme).

Exemple 1

Soit $\lambda > 0$; considérons une variable X de loi $\mathcal{P}(\lambda)$, et posons $Y = \mathrm{e}^X$.

La série exponentielle $\sum_{n \geq 0} \mathrm{e}^n\,\mathbb{P}(X = n) = \mathrm{e}^{-\lambda} \sum_{n \geq 0} \frac{(\mathrm{e}\lambda)^n}{n!}$ converge absolument, et a pour somme $\mathrm{e}^{-\lambda} \sum_{n=0}^{+\infty} \frac{(\mathrm{e}\lambda)^n}{n!} = \mathrm{e}^{-\lambda}\,\mathrm{e}^{\mathrm{e}\lambda}$. D'après le théorème du transfert, la variable aléatoire Y admet donc une espérance, et $\mathbb{E}(Y) = \mathrm{e}^{-\lambda}\,\mathrm{e}^{\mathrm{e}\lambda} = \mathrm{e}^{\lambda(\mathrm{e}-1)}$.

Exemple 2

On lance indéfiniment un dé usuel, et on note X le numéro du lancer amenant le premier 6.
Si X est pair, on gagne X euros ; sinon, on perd X euros.
Soit Y la variable aléatoire représentant le gain algébrique (positif ou négatif) ainsi obtenu. Déterminons l'espérance de Y, après avoir établi son existence.

La variable aléatoire X suit la loi $\mathcal{G}\left(\frac{1}{6}\right)$: elle représente en effet le temps d'attente du premier succès lors d'une succession infinie d'épreuves de Bernoulli indépendantes, toutes de paramètre $\frac{1}{6}$.

On a $Y = X$ lorsque X est paire, et $Y = -X$ si X est impaire, donc $Y = (-1)^X X$.

Or la série $\sum_{n\geq 1}(-1)^n n\,\mathbb{P}(X=n) = -\frac{1}{6}\sum_{n\geq 1} n\left(-\frac{5}{6}\right)^{n-1}$ converge absolument (c'est une série dérivée de série géométrique de raison $-\frac{5}{6}$, strictement inférieure à 1 en valeur absolue), et $-\frac{1}{6}\sum_{n=1}^{+\infty} n\left(-\frac{5}{6}\right)^{n-1} = -\frac{1}{6}\frac{1}{\left(1-(-5/6)\right)^2} = -\frac{6}{121}$.

Le théorème du transfert assure alors l'existence de $\mathbb{E}(Y)$, et fournit l'égalité : $\mathbb{E}(Y) = -\frac{6}{121}$.

4. Moments d'une vard

a. *Définition*

Définition

Soit $r \in \mathbb{N}^*$. On dit que X admet un moment d'ordre r lorsque la variable aléatoire X^r possède une espérance ; dans ce cas, le moment d'ordre r de X, noté $m_r(X)$, est défini par :

$$m_r(X) = \mathbb{E}(X^r).$$

En particulier, le moment d'ordre 1 de X est son espérance, sous réserve d'existence.

Cas d'un univers – image fini

Supposons $X(\Omega)$ de cardinal fini n, et posons $X(\Omega) = \{x_k\}_{1\leq k\leq n}$.

Pour tout $r \in \mathbb{N}^*$, la VAR X^r possède une espérance, car son univers image est lui aussi fini. Il en résulte que X admet un moment de tout ordre, et d'après le théorème du transfert :

$$\forall\, r \in \mathbb{N}^*,\ m_r(X) = \mathbb{E}(X^r) = \sum_{k=1}^{n} (x_k)^r\,\mathbb{P}(X = x_k).$$

Cas d'un univers – image dénombrable

Supposons maintenant $X(\Omega)$ dénombrable, et notons $X(\Omega) = \{x_n\}_{n\geq n_0}$. Pour tout $r \in \mathbb{N}^*$, d'après le théorème du transfert, X admet un moment d'ordre r si et seulement si la série $\sum_{n\geq n_0} (x_n)^r\,\mathbb{P}(X = x_n)$ converge absolument, et dans ce cas : $m_r(X) = \mathbb{E}(X^r) = \sum_{n=n_0}^{+\infty} (x_n)^r\,\mathbb{P}(X = x_n)$.

b. *Proposition*

On suppose que X admet un moment d'ordre $r \in \mathbb{N}^*$. Alors pour tout $s \in \{1, ..., r\}$, X admet un moment d'ordre s.

5. Variance et écart – type

a. *Définition*

Soit X une variable aléatoire réelle discrète.

On dit que X possède une variance lorsque la vard. $(X - \mathbb{E}(X))^2$ est définie, et est d'espérance finie. Dans ce cas, on définit la variance de X, notée $\mathbb{V}(X)$, par :

$$\mathbb{V}(X) = \mathbb{E}\left(\left(X - \mathbb{E}(X)\right)^2\right).$$

Remarque

Dire que la variable aléatoire X possède une variance revient donc à dire que X possède une espérance, et que $\left(X-\mathbb{E}(X)\right)^2$ possède une espérance.

b. *Premières propriétés*

Soit X une variable aléatoire réelle discrète, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition 1 : formule de Huygens – Koenig.

La variable aléatoire X possède une variance si et seulement si elle admet un moment d'ordre 2, et dans ce cas :

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2.$$

Proposition 2 : quasi définie – positivité de la variance

On suppose que X possède une variance. Alors :

- $\mathbb{V}(X) \geq 0$;
- •• $\mathbb{V}(X) = 0$ si et seulement si X est une variable quasi – certaine.

Proposition 3 : variance d'une fonction affine de VARD

On suppose que X possède une variance. Alors pour tout $(a, b) \in \mathbb{R}^2$, la VARD $aX + b$ possède une variance, et $\boxed{\mathbb{V}(aX+b) = a^2\,\mathbb{V}(X)}$.

c. Ecart – type

Définition 1 : écart – type d'une vard

Soit X une variable aléatoire discrète possédant une variance.

On définit l'écart – type de X, noté $\sigma(X)$, par $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Cette définition a bien un sens, car la variance de X est positive.

Définition 2 : variable centrée – réduite

Soit X une variable aléatoire discrète.

- On dit que X est une variable *centrée* lorsque X possède une espérance et $\mathbb{E}(X) = 0$;
- On dit que X est une variable *réduite* lorsque X possède une variance et $\mathbb{V}(X) = 1$.

Proposition

Soit X une variable discrète non quasi – certaine et possédant une variance. Alors la VARD $X^* = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée et réduite. On dit parfois que X^* est la variable centrée réduite issue de X.

B_ ESPÉRANCE ET VARIANCE DES LOIS DISCRÈTES USUELLES

Théorème

Soit X une variable aléatoire réelle discrète, définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, suivant l'une des lois usuelles au programme de PC. Alors X possède une espérance et une variance, et :

i – ***variable aléatoire quasi – certaine***

Soit $a \in \mathbb{R}$. Si X est presque sûrement égale à a, alors $\mathbb{E}(X) = a$ et $\mathbb{V}(X) = 0$.

ii – ***variable de Bernoulli***

Soit $p \in [0, 1]$. Si $X \hookrightarrow \mathcal{B}(1, p)$, alors $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = p\,q$.

En particulier, pour tout $A \in \mathcal{T}$, l'espérance et la variance de la variable indicatrice de A sont données par $\mathbb{E}(\mathbf{1}_A) = \mathbb{P}(A)$ et $\mathbb{V}(\mathbf{1}_A) = \mathbb{P}(A)\,\mathbb{P}(\overline{A})$.

iii – ***variable aléatoire de loi uniforme sur*** $\{1, \ldots, n\}$.

Soit $n \in \mathbb{N}^*$. Si $X \hookrightarrow \mathcal{U}(n)$, alors $\mathbb{E}(X) = \dfrac{n+1}{2}$ et $\mathbb{V}(X) = \dfrac{n^2 - 1}{12}$.

iv – ***variable aléatoire de loi binomiale***

Soit $n \in \mathbb{N}^*$, et soit $p \in]0, 1[$. On pose $q = 1 - p$. Si $X \hookrightarrow \mathcal{B}(n, p)$, alors $\mathbb{E}(X) = n\,p$ et $\mathbb{V}(X) = n\,p\,q$.

v – ***Loi de Poisson***

Soit $\lambda > 0$. Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors $\mathbb{E}(X) = \lambda$ et $\mathbb{V}(X) = \lambda$.

vi – ***Loi géométrique***

Soit $p \in]0, 1[$. Si $X \hookrightarrow \mathcal{G}(p)$, alors $\mathbb{E}(X) = \dfrac{1}{p}$ et $\mathbb{V}(X) = \dfrac{q}{p^2}$.

VII – VECTEURS ALÉATOIRES RÉELS DISCRETS

A_ VECTEURS ALÉATOIRES RÉELS DISCRETS

1. Vecteur aléatoire discret

a. *Définition*

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On appelle vecteur aléatoire réel discret tout n – uplet $(X_1, \ldots, X_n)$ de variables aléatoires réels discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, où n est un entier strictement positif.

Dans le cas où $n = 2$, on dit que (X_1, X_2) est un ***couple*** aléatoire réel discret.

b. *Univers – image*

Définition

Soit $(X_1, \ldots, X_n)$ un vecteur aléatoire discret, défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle *univers – image* de $(X_1, \ldots, X_n)$ l'ensemble $(X_1, \ldots, X_n)(\Omega)$ de tous les vecteurs $(X_1(\omega), \ldots, X_n(\omega))$, lorsque ω décrit Ω :

$$(X_1, \ldots, X_n)(\Omega) = \{(X_1(\omega), \ldots, X_n(\omega)), \omega \in \Omega\} \subset \mathbb{R}^n .$$

Exemple

Une urne contient n jetons numérotés de 1 à n. On tire un jeton de cette urne, on l'y remet, mais on supprime tous les jetons portant un numéro strictement supérieur à celui qui a été obtenu. Après quoi, on effectue un deuxième tirage d'un jeton de cette urne. On note X le numéro du premier jeton obtenu, et Y le numéro du deuxième.

Alors :

- $X(\Omega)$ est, de manière évidente, égal à $\{1, ..., n\}$;
- •• Pour tout $k \in \{1, ..., n\}$, si l'on suppose $(X = k)$ réalisé, le deuxième tirage s'effectue dans une urne contenant k jetons numérotés de 1 à k, le deuxième numéro obtenu peut donc prendre n'importe quelle valeur entre ces deux extrêmes ; mais, comme k peut lui – même prendre n'importe quelle valeur entre 1 et n, on en conclut, finalement, que $Y(\Omega) = \{1, ..., n\}$.
- ••• Il résulte de tout ceci que $(X(\Omega)) \times (Y(\Omega)) = \{1, ..., n\}^2$, alors que $(X, Y)(\Omega) = \left\{(k, \ell) \in \{1, ..., n\}^2 \ / \ k \leq \ell\right\}$. On constate, sur cet exemple, que $(X, Y)(\Omega) \subset (X(\Omega)) \times (Y(\Omega))$, mais que $(X, Y)(\Omega) \neq (X(\Omega)) \times (Y(\Omega))$.

En réfléchissant un peu, on se persuadera facilement que, de manière générale, pour tout vecteur aléatoire réel discret $(X_1, ..., X_n)$: $(X_1, ..., X_n)(\Omega)$ est une partie de $X_1(\Omega) \times ... \times X_n(\Omega)$, mais n'a pas de raison particulière de lui être égal.

c. *Loi d'un vecteur aléatoire discret*

Notation

Soient $n \in \mathbb{N}^*$, et $X = (X_1, ..., X_n)$ un n – uplet de variables aléatoires réelles discrètes.

- Pour tous sous – ensembles $A_1, ..., A_n$ de $\mathbb{R}$, on note $(X_1 \in A_1, ..., X_n \in A_n)$ l'événement :

$$(X_1 \in A_1, ..., X_n \in A_n) = (X_1 \in A_1) \cap ... \cap (X_n \in A_n) = \bigcap_{k=1}^{n} (X_k \in A_k).$$

- •• Pour tout $(x_1, ..., x_n) \in \mathbb{R}^n$, l'événement $\bigcap_{k=1}^{n} (X_k = x_k)$ est noté :

$$\bigcap_{k=1}^{n} (X_k = x_k) = (X_1 = x_1, ..., X_n = x_n).$$

Loi conjointe d'un couple aléatoire discret

Soit (X_1, X_2) un couple aléatoire réel discret.

On appelle *loi conjointe* du couple (X_1, X_2) l'application $\begin{pmatrix} (X_1, X_2)(\Omega) & \to & \mathbb{R} \\ (x_1, x_2) & \mapsto & \mathbb{P}(X_1 = x_1, X_2 = x_2) \end{pmatrix}$.

Cette loi conjointe pourra être notée $\mathcal{L}_{(X_1, X_2)}$.

Déterminer la loi du couple (X_1, X_2) revient donc à :

- Déterminer l'ensemble $(X_1, X_2)(\Omega)$;
- •• Préciser, pour tout (x_1, x_2) de $(X_1, X_2)(\Omega)$, la valeur de $\mathbb{P}(X_1 = x_1, X_2 = x_2)$.

Généralisation : loi conjointe d'un vecteur aléatoire discret

Plus généralement, pour tout entier $n \geq 2$, et pour tout n – uplet $\left(X_1, ..., X_n \right)$ de variables aléatoires réelles discrètes définies sur $\left(\Omega, \mathcal{A}, \mathbb{P} \right)$, on appelle *loi* de $\left(X_1, ..., X_n \right)$ l'application :

$$\begin{cases} \left(X_1, ..., X_n \right)\left(\Omega \right) & \to & \mathbb{R} \\ \left(x_1, ..., x_n \right) & \mapsto & \mathbb{P}\left(X_1 = x_1, ..., X_n = x_n \right) \end{cases}.$$

Donner la loi de $\left(X_1, ..., X_n \right)$ revient à préciser $\left(X_1, ..., X_n \right)\left(\Omega \right)$, et à déterminer, pour tout $\left(x_1, ..., x_n \right)$ de $\left(X_1, ..., X_n \right)\left(\Omega \right)$, la valeur $\mathbb{P}\left(X_1 = x_1, ..., X_n = x_n \right)$.

On dispose, à l'image de ce que l'on a déjà rencontré pour une variable aléatoire discrète, d'un théorème de caractérisation des lois conjointes :

d. ***Caractérisation d'une loi conjointe***

Proposition 1

Soit $\left(X, Y \right)$ un couple aléatoire discret défini sur un espace probabilisé $\left(\Omega, \mathcal{A}, \mathbb{P} \right)$.

Alors la famille $\left(\left(X = i, Y = j \right) \right)_{(i, j) \in (X, Y)(\Omega)}$ est un système complet d'événements.

Corollaire

$$\sum_{(i, j) \in (X, Y)(\Omega)} \mathbb{P}\left(X = i, Y = j \right) = 1 .$$

On admettra que la somme de cette série double convergente, à termes positifs, ne dépend pas de l'ordre de sommation (Fubini).

La réciproque de ce résultat est vraie, et c'est le théorème de caractérisation annoncé :

Proposition 2

Soit $\left(p_{i, j} \right)_{(i, j) \in \mathbb{N}^2}$ une famille de réels telle que :

- pour tout $\left(i, j \right) \in \mathbb{N}^2$, $p_{i, j}$ est positif ou nul ;
- •• la série double $\sum_{(i, j) \in \mathbb{N}^2} p_{i, j}$ converge, et $\sum_{(i, j) \in \mathbb{N}^2} p_{i, j} = 1$.

Alors il existe un espace probabilisé $\left(\Omega, \mathcal{A}, \mathbb{P} \right)$ et un couple de variables discrètes $\left(X, Y \right)$ définies sur cet espace, tels que : $\forall \left(i, j \right) \in \mathbb{N}^2, \mathbb{P}\left(X_1 = i, X_2 = j \right) = p_{i, j}$.

Le résultat se généralise au cas d'une famille $\left(p_{i, j} \right)_{(i, j) \in \Theta}$, où Θ est une partie de $\mathbb{N}^2$ ou de $\mathbb{Z}^2$.

Exemple 1

Soit n un entier naturel non nul, et soit $\Theta = \left\{ \left(i, j \right) \in \mathbb{N}^2 \ / \ 0 \leq j \leq i \leq n \right\}$. Montrons qu'il existe un couple aléatoire discret $\left(X, Y \right)$ défini sur un espace probabilisé $\left(\Omega, \mathcal{A}, \mathbb{P} \right)$ tel que $\left(X, Y \right)\left(\Omega \right) = \Theta$ et $\forall \left(i, j \right) \in \Theta, \mathbb{P}\left(X_1 = i, X_2 = j \right) = \dfrac{1}{\left(n+1 \right)\left(i+1 \right)}$.

Pour tout $\left(i, j \right) \in \Theta$, on pose $p_{i, j} = \dfrac{1}{\left(n+1 \right)\left(i+1 \right)}$.

- Pour tout $(i, j) \in \Theta$, $p_{i,j}$ est positif ou nul ;
- •• La somme (finie) $\sum_{(i,j) \in \Theta} p_{i,j}$ est évidemment bien définie, et l'on a :

$$\sum_{(i,j) \in \Theta} p_{i,j} = \frac{1}{n+1} \sum_{i=0}^{n} \left(\frac{1}{i+1} \sum_{j=0}^{i} 1 \right) = \frac{1}{n+1} \sum_{i=0}^{n} 1 = 1 .$$

Le théorème de caractérisation d'une loi conjointe vu ci – dessus prouve alors qu'il existe un espace probabilisé et un couple de variables discrètes (X, Y) sur cet espace tels que

$$(X, Y)(\Omega) = \Theta, \text{ et } \forall (i, j) \in \Theta, \mathbb{P}(X_1 = i, X_2 = j) = p_{i,j} .$$

Exemple 2

Soit λ un réel. On pose, pour tout $(i, j) \in \mathbb{N}^2$, $p_{i,j} = \lambda \dfrac{i^j}{i!\, j!}$.

Pour quelle(s) valeur(s) du réel λ la famille $(p_{i,j})_{(i,j) \in \mathbb{N}^2}$ définit – elle la loi de probabilité d'un couple aléatoire discret ?

- $p_{i,j}$ est positif ou nul pour tout $(i, j) \in \mathbb{N}^2$ si et seulement si $\lambda \in \mathbb{R}^+$.
- •• Ici, $\sum_{(i,j) \in \mathbb{N}^2} p_{i,j}$ désigne la série double $\sum_{i \geq 0} \sum_{j \geq 0} p_{i,j} = \lambda \sum_{i \geq 0} \left(\frac{1}{i!} \sum_{j \geq 0} \frac{i^j}{j!} \right)$. Or :
  - ◦ Pour tout $i \in \mathbb{N}$ fixé, la série $\sum_{j \geq 0} \frac{i^j}{j!}$ (série exponentielle) converge, et $\sum_{j=0}^{+\infty} \frac{i^j}{j!} = e^i$.
  - ◦ ◦ La série exponentielle $\lambda \sum_{i \geq 0} \frac{e^i}{i!}$ converge, et $\lambda \sum_{i=0}^{+\infty} \frac{e^i}{i!} = \lambda\, e^e$.

  Ainsi, la série double $\sum_{(i,j) \in \mathbb{N}^2} p_{i,j}$ est bien convergente ; sa somme est égale à 1 si et seulement si $\lambda = e^{-e}$. Notons que dans ce cas, la condition $\lambda \in \mathbb{R}^+$ est vérifiée. Par conséquent :

Il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et un couple de variables discrètes (X, Y) sur cet espace tels que $(X, Y)(\Omega) = \mathbb{N}^2$ et $\forall (i, j) \in \mathbb{N}^2$, $\mathbb{P}(X_1 = i, X_2 = j) = p_{i,j}$ si et seulement si $\lambda = e^{-e}$.

On a utilisé en fait le *théorème de Fubini* évoqué ci – dessus.

e. *Fonction de répartition*

Soit (X_1, X_2) un couple aléatoire réel discret. On appelle *fonction de répartition* du couple (X_1, X_2) l'application

$$\begin{cases} (X_1, X_2)(\Omega) & \to & \mathbb{R} \\ (x_1, x_2) & \mapsto & \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \end{cases} .$$

La généralisation de cette définition au cas d'un n – uplet $X = (X_1, \ldots, X_n)$ de variables aléatoires réelles discrètes est immédiate, comme ci – dessus pour la loi.

2. Loi conjointe et lois marginales

Définitions

Soit (X, Y) un couple de variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Les var (discrètes) X et Y sont appelées les *marginales* du couple (X, Y).

Leurs lois sont appelées les *lois marginales* de la loi conjointe du couple (X, Y).

On suppose dans ce qui suit que $X(\Omega) = \{x_i, i \in I\}$, et $Y(\Omega) = \{y_j, j \in J\}$, où I et J désignent des parties finies ou infinies de $\mathbb{N}$ ou de $\mathbb{Z}$. On notera alors dans ce cours :

$$\forall i \in I,\ p_{i,\bullet} = \mathbb{P}(X = x_i) \text{ et } \forall j \in J,\ p_{\bullet,j} = \mathbb{P}(Y = y_j).$$

Proposition

Dans la même situation que ci – dessus, et avec les mêmes notations :

$$\forall i \in I,\ p_{i,\bullet} = \sum_{j \in J} p_{i,j} \text{ et } \forall j \in J,\ p_{\bullet,j} = \sum_{i \in I} p_{i,j}.$$

Démonstration

Soit $i \in I$. $(Y = y_j)_{j \in J}$ est un système complet d'évènements, on sait alors que la somme $\sum_{j \in J} \mathbb{P}(X = x_i, Y = y_j)$ est convergente, et que $\mathbb{P}(X = x_i) = \sum_{j \in J} \mathbb{P}(X = x_i, Y = y_j)$. Autrement dit : $\forall i \in I,\ p_{i,\bullet} = \sum_{j \in J} p_{i,j}$.

Dans le cas (pour des var discrètes finies) d'une présentation en tableau de la loi conjointe du couple (X, Y), les lois de X et Y s'obtiennent donc en sommant sur chaque ligne et sur chaque colonne les valeurs du tableau. On écrit alors les résultats dans une colonne et une ligne supplémentaires, les *marges*, d'où le nom de variables et lois *marginales*…

Nous pouvons alors généraliser au cas de vecteurs aléatoires qui ne sont pas forcément des couples.

Lois marginales d'un vecteur aléatoire réel discret

Soit $n \in \mathbb{N}^*$, et soit $(X_1, \dots, X_n)$ un vecteur aléatoire réel discret. Pour $i \in \{1, ..., n\}$, la loi de probabilité de la variable aléatoire X_i est appelée *loi marginale* de la variable aléatoire X_i.

Les lois marginales s'obtiennent encore par sommation ; par exemple :

$$\forall x_1 \in X_1(\Omega),\ \mathbb{P}(X_1 = x_1) = \sum_{\substack{x_2 \in X_2(\Omega) \\ \vdots \\ x_n \in X_n(\Omega)}} \mathbb{P}\left(\bigcap_{i=1}^{n} (X_i = x_i)\right).$$

3. Lois conditionnelles

a. *Définition*

On suppose dans ce paragraphe que (X, Y) est un couple de VAR discrètes défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, avec $X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$, où I et J désignent des parties finies ou infinies de $\mathbb{N}$ ou de $\mathbb{Z}$.

On notera pour tout $j \in J$, $\mathbb{P}(Y = y_j) = p_{\bullet,j}$, et pour tout $(i, j) \in I \times J$: $\mathbb{P}(X = x_i, Y = y_j) = p_{i,j}$.

Conditionnement d'une VAR discrète par un événement non négligeable

Soit A un événement non négligeable. On appelle *loi de* X *conditionnellement à* A (ou *sachant* A) l'application $\begin{pmatrix} X(\Omega) \to \mathbb{R} \\ x_i \mapsto \mathbb{P}_A(X = x_i) \end{pmatrix}$.

On a bien sûr $\sum_{i \in I} \mathbb{P}_A(X = x_i) = 1$.

Cas particulier

Soit $j \in J$ tel que $\mathbb{P}(Y = y_j) \neq 0$. La loi de X conditionnée par l'événement $(Y = y_j)$ est l'application $\left\{ \begin{array}{lcl} X(\Omega) & \to & \mathbb{R} \\ x_i & \mapsto & \dfrac{p_{i,j}}{p_{\bullet,j}} \end{array} \right.$.

Démonstration

C'est une évidence au vu de la définition précédente, puisque l'on a, pour tout $i \in I$:

$$\mathbb{P}_{(Y = y_j)}(X = x_i) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)} = \frac{p_{i,j}}{p_{\bullet,j}}.$$

Conditionnement d'une variable aléatoire discrète par une autre VAR discrète

On appelle *loi de X conditionnée par Y (ou sachant Y)* l'ensemble des couples

$\left((x_i, y_j), \mathbb{P}_{(Y = y_j)}(X = x_i) \right)_{(i,j) \in I \times J}$ *i.e.* l'ensemble des couples $\left((x_i, y_j), \dfrac{p_{i,j}}{p_{\bullet,j}} \right)_{(i,j) \in I \times J}$.

Pour $I = \{1, ..., n\}$ et $J = \{1, ..., p\}$ avec n et p petits, cette loi est souvent représentée par un tableau à double entrée, dont la somme des éléments situés sur chaque ligne (ou chaque colonne) est égale à 1.

b. *Lois conditionnelles et indépendance*

On démontre facilement le résultat suivant :

Proposition

Soit (X, Y) un couple de variables aléatoires discrètes. Les assertions suivantes sont équivalentes :

i – Les variables aléatoires X et Y sont indépendantes

ii – Pour tout $y \in \mathbb{R}$ tel que $\mathbb{P}(Y = y) \neq 0$, la loi de X et la loi de X sachant $(Y = y)$ sont égales.

iii – Pour tout $x \in \mathbb{R}$ tel que $\mathbb{P}(X = x) \neq 0$, la loi de Y et la loi de Y sachant $(X = x)$ sont égales.

c. Exemples

Application 3

Soient a, b et y trois réels différents de 0 et 1. Soient X et Y deux variables aléatoires réelles définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et dont la loi conjointe est donnée par :

X \ Y	0	1	y
0	a	$1/8$	$1/4$
1	b	$1/10$	$1/5$

1. Déterminer les valeurs de a et b pour que X et Y soient indépendantes.
2. Quelles sont alors les lois conditionnelles de X pour les différentes valeurs de Y ?

On suppose désormais que $a = 1/5$.

3. Déterminer y pour que le coefficient de corrélation linéaire de X et Y soit nul.
4. Les v.a.r. X et Y sont – elles alors indépendantes ?

Application 4

Soit (X, Y) un couple de VAR discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs entières.

Pour $(k, \ell) \in \mathbb{N}^2$, on note $p_{k,\ell} = \mathbb{P}(X = k, Y = \ell)$. On suppose que l'on a :

$$p_{k,\ell} = \left| \begin{array}{ll} \dfrac{\lambda^k \, e^{-\lambda} \, \alpha^\ell (1-\alpha)^{k-\ell}}{\ell!\,(k-\ell)!} & \text{si } (k, \ell) \in \mathbb{N}^2 \text{ et } 0 \leq \ell \leq k \\ 0 & \text{sinon} \end{array} \right. ,$$

λ et α étant des réels donnés tels que $\lambda > 0$ et $0 < \alpha < 1$.

1. Vérifier que l'on définit bien ainsi la loi de probabilité d'un couple (X, Y) de VARD.
2. Déterminer la loi de probabilité de la variable aléatoire X.
3. Déterminer la loi de probabilité de la variable aléatoire Y.
4. Les variables X et Y sont-elles indépendantes ?
5. Déterminer la loi de probabilité de la variable aléatoire $Z = X - Y$.
6. Déterminer la probabilité conditionnelle $\mathbb{P}_{[Z = n]}(Y = \ell)$.
7. Qu'en déduire pour les variables aléatoires Y et Z ?

B_ COVARIANCE D'UN COUPLE

1. Théorème du transfert pour l'espérance d'une fonction d'un couple de vards

Soient X et Y deux variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soit φ une application définie sur une partie de $\mathbb{R}^2$ contenant $(X, Y)(\Omega)$, à valeurs dans $\mathbb{R}$, et soit $Z = \varphi(X, Y)$.

On admettra qu'alors Z est une variable aléatoire discrète.

L'espérance d'une telle variable aléatoire est donnée, sous réserve d'existence, par le résultat suivant, que l'on admet également :

Théorème (dit du transfert)

Soient X et Y deux variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Notons $X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$. Soit φ une application définie sur une partie de $\mathbb{R}^2$ contenant $(X, Y)(\Omega)$, à valeurs dans $\mathbb{R}$. Soit Z la variable aléatoire définie par $Z = \varphi(X, Y)$.

Alors :

- La var Z admet une espérance si et seulement si la série double $\sum_{(i,j) \in I \times J} \left| \varphi(x_i, y_j) \right| \mathbb{P}(X = x_i, Y = y_j)$ converge.
- •• Dans ce cas, on a :

$$\begin{aligned} \mathbb{E}(Z) &= \sum_{(i,j) \in I \times J} \varphi(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j) \\ &= \sum_{i \in I} \sum_{j \in J} \varphi(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j) \\ &= \sum_{j \in J} \sum_{i \in I} \varphi(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j). \end{aligned}$$

On signale que, comme toujours, $\mathbb{P}\left(X = x_i, Y = y_j\right)$ est nul lorsque $\left(x_i, y_j\right) \notin (X, Y)(\Omega)$.

Si $X(\Omega)$ et $Y(\Omega)$ sont finis, alors Z admet automatiquement une espérance.

2. Espérance d'un produit

a. *Espérance d'un produit de deux vards*

Proposition

> Soient deux variables aléatoires discrètes X et Y, définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent toutes deux une variance.
> Alors la variable aléatoire $Z = XY$ est d'espérance finie, et l'on a
> $$\begin{aligned}\mathbb{E}(XY) &= \sum_{(i,j) \in (X,Y)(\Omega)} i\, j\, \mathbb{P}(X = i, Y = j) \\ &= \sum_{i \in X(\Omega)} \left(\sum_{j \in Y(\Omega)} i\, j\, \mathbb{P}(X = i, Y = j) \right).\end{aligned}$$

Idée de la preuve

Il suffit d'appliquer le théorème de transfert à la variable aléatoire $Z = XY$: celui – ci assure que l'on a bien

$\mathbb{E}(XY) = \sum_{i \in X(\Omega)} \left(\sum_{j \in Y(\Omega)} i\, j\, \mathbb{P}(X = i, Y = j) \right)$, sous réserve de convergence absolue de cette série double.

Or pour tout $(i, j) \in \mathbb{R}^2$, $|i\, j| \leq \frac{1}{2}\left(i^2 + j^2\right)$, car : $\frac{1}{2}\left(i^2 + j^2\right) - i\, j = \frac{1}{2}(i - j)^2 \geq 0$, et

$\frac{1}{2}\left(i^2 + j^2\right) + i\, j = \frac{1}{2}(i + j)^2 \geq 0$.

De plus, les séries doubles

$\sum_{i \in X(\Omega)} \left(\sum_{j \in Y(\Omega)} i^2\, \mathbb{P}(X = i, Y = j) \right) = \sum_{i \in X(\Omega)} i^2\, \mathbb{P}(X = i, Y = j) = \mathbb{V}(X)$ et

$\sum_{i \in X(\Omega)} \left(\sum_{j \in Y(\Omega)} j^2\, \mathbb{P}(X = i, Y = j) \right) = \sum_{j \in Y(\Omega)} j^2\, \mathbb{P}(Y = j) = \mathbb{V}(Y)$ convergent.

Le théorème de convergence par majoration pour les séries à terme général positif permet d'en déduire que

$\sum_{i \in X(\Omega)} \left(\sum_{j \in Y(\Omega)} i\, j\, \mathbb{P}(X = i, Y = j) \right)$ converge, d'où le résultat.

b. *Espérance d'un produit de n* **vards indépendantes**

Proposition

> Soient X et Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, indépendantes, et admettant chacune une variance. Alors :
> $$\mathbb{E}(XY) = \mathbb{E}(X)\, \mathbb{E}(Y).$$

Idée de la preuve

Avec les notations et hypothèses adoptées dans ce chapitre, on peut écrire, grâce au théorème du transfert, que XY possède une espérance, et que

$$\mathbb{E}(XY) = \sum_{i \in I} \sum_{j \in J} x_i \, y_j \, \mathbb{P}(X = x_i, Y = y_j) = \sum_{i \in I} \sum_{j \in J} x_i \, y_j \, \mathbb{P}(X = x_i) \, \mathbb{P}(Y = y_j),$$

en raison de l'indépendance des variables aléatoires X et Y. On a alors :

$$\mathbb{E}(XY) = \sum_{i \in I} \sum_{j \in J} x_i \, y_j \, \mathbb{P}(X = x_i, Y = y_j) = \sum_{i \in I} x_i \, \mathbb{P}(X = x_i) \sum_{j \in J} y_j \, \mathbb{P}(Y = y_j) = \mathbb{E}(X) \, \mathbb{E}(Y),$$

ce que l'on voulait démontrer.

Notons que ceci n'est pas vraiment une preuve : la justification des opérations effectuées sur les sommes ci-dessus ferait appel à la notion, hors-programme en PC, de familles sommables.

Il suffit en fait que X et Y admettent une espérance et soient indépendantes pour que le produit XY admette une espérance.

De manière plus générale :

Proposition bis

> Soit n un entier naturel supérieur ou égal à 2, et soient $X_1, \ldots, X_n$ n variables aléatoires définies sur un même espace probabilisé, indépendantes, et admettant chacune une espérance. Alors : $\mathbb{E}\left(\prod_{k=1}^{n} X_k \right) = \prod_{k=1}^{n} \mathbb{E}(X_k)$.

3. Covariance, coefficient de corrélation

a. *Définition*

> Soient X et Y deux variables aléatoires discrètes définies sur un même espace probabilisé, et admettant chacune un moment d'ordre deux. On définit la *covariance* de X et Y par la formule de *Huygens – König :*
>
> $$\operatorname{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y).$$

b. *Propriétés*

Proposition

> Soient X, Y, et Z des VAR discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, et admettant chacune un moment d'ordre deux. Soit également $\lambda \in \mathbb{R}$. Alors :
>
> ***i*** – $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$.
>
> ***ii*** – $\operatorname{cov}(X + Y, Z) = \operatorname{cov}(X, Z) + \operatorname{cov}(Y, Z)$.
>
> ***iii*** – $\operatorname{cov}(\lambda X, Y) = \lambda \operatorname{cov}(X, Y)$.
>
> ***iv*** – $\operatorname{cov}(X, X) = \mathbb{V}(X)$.
>
> ***v*** – $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\operatorname{cov}(X, Y)$.
>
> ***vi*** – ***Identités de polarisation***
>
> $$\begin{aligned} \operatorname{cov}(X, Y) &= \frac{1}{2}\big(\mathbb{V}(X + Y) - \mathbb{V}(X) - \mathbb{V}(Y)\big) \\ &= \frac{1}{2}\big(\mathbb{V}(X) + \mathbb{V}(Y) - \mathbb{V}(X - Y)\big) \\ &= \frac{1}{4}\big(\mathbb{V}(X + Y) - \mathbb{V}(X - Y)\big). \end{aligned}$$

Notons qu'il résulte aussi de la proposition précédente que :

$$\operatorname{cov}(X,Y) = \mathbb{E}\Big(\big(X-\mathbb{E}(X)\big)\big(Y-\mathbb{E}(Y)\big)\Big),$$

ce qui est généralement pris comme définition de la covariance.

Le dernier point de la proposition précédente se généralise comme suit :

Proposition ***(variance d'une somme : identité de polarisation étendue)***

Soient n un entier naturel supérieur ou égal à 2, et $X_1, \dots, X_n$ n variables aléatoires définies sur un même espace probabilisé, admettant chacune une variance.

Alors, $\sum_{k=1}^{n} X_k$ admet une variance, et :

$$\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{V}(X_k) + \sum_{(i,j)\in [\![1,n]\!]^2 \,/\, i\neq j} \operatorname{cov}(X_i, X_j)$$

$$= \sum_{k=1}^{n} \mathbb{V}(X_k) + 2\sum_{1\leq i<j\leq n} \operatorname{cov}(X_i, X_j).$$

c. ***Coefficient de corrélation***

Définition

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé, et admettant chacune une variance non nulle. On appelle *coefficient de corrélation linéaire* des VAR X et Y le nombre $\rho(X,Y)$ défini par : $\rho(X,Y) = \dfrac{\operatorname{cov}(X,Y)}{\sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)}}$.

Les résultats suivants sont connus sous le nom d'inégalité de Cauchy – Schwarz :

Proposition

Soient X et Y deux variables aléatoires discrètes admettant chacune un moment d'ordre 2. Alors :

Cauchy – Schwarz version PC

- $\left|\operatorname{cov}(X,Y)\right| \leq \sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)}$.

  Autrement dit, lorsque les variances de X et Y sont non nulles : $\left|\rho(X,Y)\right| \leq 1$.

- •• De plus, $\left|\rho(X,Y)\right| = 1$ si et seulement si les variables X et Y sont quasi – sûrement affinement liées, autrement dit ssi il existe $\lambda_0 \in \mathbb{R}$ et b variable aléatoire certaine, tels que $Y = \lambda_0 X + b$ presque sûrement.

- ••• $\big(\mathbb{E}(XY)\big)^2 \leq \mathbb{E}(X^2)\,\mathbb{E}(Y^2)$.

d. ***Indépendance et non – corrélation***

Définition

Soient X et Y deux vards définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et admettant chacune une variance. On dit que X et Y sont non – corrélées lorsque $\operatorname{cov}(X,Y) = 0$.

Lorsque les variances de X et Y sont non nulles, il revient au même de dire que $\rho(X,Y) = 0$.

Proposition

> Soient X et Y deux variables aléatoires définies sur un même espace probabilisé, indépendantes, et admettant chacune une variance. Alors X et Y sont non – corrélées : $\operatorname{cov}(X, Y) = 0$.

Attention, la réciproque est fausse, et cela est fréquemment illustré dans les exercices.

Conséquence de ce résultat, l'expression de la variance d'une somme se simplifie notablement en situation de non – corrélation, et en particulier en cas d'indépendance :

Corollaire

> Soit $n \in \mathbb{N}^*$, et soient $X_1, \dots, X_n$ n variables aléatoires définies sur un même espace probabilisé, admettant chacune un moment d'ordre 2, et non – corrélées.
>
> Alors $\sum_{k=1}^{n} X_k$ admet une variance, et l'on a $\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{V}(X_k)$.

VIII – SÉRIE GÉNÉRATRICE D'UNE VAR À VALEURS DANS $\mathbb{N}$

1. Définition

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$.

a. Série génératrice d'une variable aléatoire à valeurs entières

La série génératrice de X, notée G_X, est la fonction d'une variable réelle t définie par : $\boxed{G_X : t \mapsto \mathbb{E}\left(t^X\right)}$.

b. Expression comme somme d'une série entière

Soit t un réel. D'après le théorème du transfert, G_X est définie en t si et seulement si la série $\sum_{n \geq 0} \mathbb{P}(X = n)\, t^n$ converge absolument. Lorsque tel est le cas :

$$\boxed{G_X(t) = \mathbb{E}\left(t^X\right) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)\, t^n}.$$

Autrement dit, G_X est la fonction $G_X : t \mapsto \mathbb{E}\left(t^X\right) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)\, t^n$, *définie sur l'ensemble :*

$$\left\{ t \in \mathbb{R},\ \sum_{n \geq 0} \mathbb{P}(X = n)\, t^n \text{ converge absolument} \right\}.$$

2. Séries génératrices des lois géométriques et de Poisson

Proposition

> - Soient $p \in \left]0, 1\right[$, $q = 1 - p$, et soit X une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$. Alors la série génératrice de X a pour rayon de convergence $\boxed{\dfrac{1}{q}}$, et pour tout $t \in \left]-\dfrac{1}{q}, \dfrac{1}{q}\right[$:
>
> $\boxed{G_X(t) = \dfrac{p\,t}{1 - q\,t}}$.

•• Soit $\lambda > 0$, et soit X une variable aléatoire suivant la loi de Poisson $\mathcal{P}(\lambda)$. Alors la série génératrice de X a pour rayon de convergence $\boxed{+\infty}$, et pour tout $t \in \mathbb{R}$: $\boxed{G_X(t) = e^{\lambda(t-1)}}$.

3. Propriétés fondamentales des séries génératrices

On considère ici une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$.

a. Définition et valeur en 1

La série génératrice de X est définie en 1, et l'on a $\boxed{G_X(t) = 1}$.

Preuve : $\big((X = n)\big)_{n \in \mathbb{N}}$ étant un système complet d'évènements, $\sum\limits_{n \geq 0} \mathbb{P}(X = n) 1^n = \sum\limits_{n \geq 0} \mathbb{P}(X = n)$ converge (absolument), et l'on a $\sum\limits_{n=0}^{+\infty} \mathbb{P}(X = n) = 1$.

b. Minoration du rayon de convergence

Le rayon de convergence de G_X est supérieur ou égal à 1.

c. Les séries génératrices caractérisent la loi

Deux var à valeurs dans $\mathbb{N}$ ont la même loi si et seulement si leurs séries génératrices sont égales.

4 Espérance, variance et dérivées de la fonction génératrice

Proposition

Soit X une variables aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$. Alors :

• X admet une espérance s si G_X est dérivable en 1, et lorsque tel est le cas :

$\mathbb{E}(X) = G_X'(1)$.

•• X admet une variance s si $G_X''(1)$ existe, et lorsque tel est le cas :

$G_X''(1) = \mathbb{E}\big(X(X-1)\big)$.

5. Fonction génératrice d'une somme de variables aléatoires (entières) indépendantes

Etant donnée une variable aléatoire X à valeurs entières, on note ici $R(X)$ le rayon de convergence de sa série génératrice.

Théorème

Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$, indépendantes. Alors $R(X+Y) \geq \min\big(R(X), R(Y)\big)$, et :

$$\forall\, t \in \big]-\min\big(R(X), R(Y)\big), \min\big(R(X), R(Y)\big)\big[\ : \ \boxed{G_{X+Y}(t) = G_X(t)\, G_Y(t)}.$$

Autrement dit, G_{X+Y} est le produit de Cauchy de G_X et de *G_Y*.

IX – INTRODUCTION A L'APPROXIMATION

A_ DEUX INÉGALITÉS PROBABILISTES

1. Inégalité de Markov

Théorème (inégalité de Markov)

Soit X une variable aléatoire à valeurs positives, admettant une espérance $\mathbb{E}(X)$.

Alors pour tout $\lambda > 0$: $\mathbb{P}(X \geq \lambda) \leq \dfrac{\mathbb{E}(X)}{\lambda}$.

Démonstration **(à connaître !)**

On considère la variable aléatoire Y définie par : $\begin{cases} Y = 1 & \text{si } X \geq \lambda \\ 0 & \text{sinon} \end{cases}$ *Autrement dit, Y est la var. indicatrice de l'évènement $(X \geq \lambda)$: $Y = \mathbf{1}_{[\lambda, +\infty[}(X)$.*

Alors, Y est une variable de Bernoulli, de paramètre $p = \mathbb{P}(Y = 1) = \mathbb{P}(X \geq \lambda)$; par suite, Y possède une espérance, et $\mathbb{E}(Y) = \mathbb{P}(X \geq \lambda)$.

D'autre part, on note que l'on a $\begin{cases} \lambda Y = \lambda & \text{si } X \geq \lambda \\ 0 & \text{sinon} \end{cases}$, et, comme X est à valeurs positives, il en résulte que λY est inférieure à X.

On en déduit, par propriété de croissance de l'espérance, que $\mathbb{E}(\lambda Y) \leq \mathbb{E}(X)$, puis, par linéarité, que $\lambda \mathbb{E}(Y) \leq \mathbb{E}(X)$.

On a donc $\lambda \mathbb{P}(X \geq \lambda) \leq \mathbb{E}(X)$. □

Application 5

On considère une var X suivant une loi géométrique de paramètre $p = \dfrac{1}{10}$.

1. A l'aide de l'inégalité de Markov, donner une majoration de $\mathbb{P}(X > \lambda)$, pour $\lambda > 0$.
2. Comment choisir λ de manière à ce que le résultat précédent assure que $\mathbb{P}(X > \lambda) < 10^{-4}$?
3. Pour de vrai, pour quelles valeurs de λ a – t – on $\mathbb{P}(X > \lambda) < 10^{-4}$?

Application 6

Une inégalité de Kolmogorov

Soit X une var. définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On suppose X bornée, *i.e.* que : $\exists\, \beta \in \mathbb{R}_+^* \;/\; \forall\, \omega \in \Omega,\ |X(\omega)| \leq \beta$.

Soit $\alpha \in \mathbb{R}_+^*$. Montrer que : $\mathbb{P}(|X| > \alpha) \geq \dfrac{\mathbb{E}(X^2) - \alpha^2}{\beta^2}$.

hint *On copiera honteusement la démonstration faite en cours de l'inégalité de Markov, en faisant intervenir la var $Y = \mathbf{1}_{[\alpha^2, \beta^2]}(X^2)$.*

2. Inégalité de Bienaymé - Tchebychev

Théorème (inégalité de Bienaymé – Tchebychev)

Soit X une variable aléatoire admettant une variance $\mathbb{V}(X)$.

Alors pour tout $\varepsilon > 0$: $\mathbb{P}\left(\left|X - \mathbb{E}(X)\right| \geq \varepsilon\right) \leq \dfrac{\mathbb{V}(X)}{\varepsilon^2}$.

Démonstration

La variable aléatoire X possède un moment d'ordre 2, elle admet donc également une espérance ; on peut alors définir la var. $Y = \left(X - \mathbb{E}(X)\right)^2$. Y est clairement à valeurs positives, et, par définition d'une variance, $\mathbb{E}(Y)$ existe et vaut $\mathbb{V}(X)$. Il est donc licite d'appliquer l'inégalité de Markov à Y. Il vient : $\forall\ \varepsilon > 0,\ \mathbb{P}\left(Y \geq \varepsilon^2\right) \leq \dfrac{\mathbb{E}(Y)}{\varepsilon^2}$,

soit $\forall\ \varepsilon > 0,\ \mathbb{P}\left(\left(X - \mathbb{E}(X)\right)^2 \geq \varepsilon^2\right) \leq \dfrac{\mathbb{V}(X)}{\varepsilon^2}$, et l'on a bien : $\forall\ \varepsilon > 0,\ \mathbb{P}\left(\left|X - \mathbb{E}(X)\right| \geq \varepsilon\right) \leq \dfrac{\mathbb{V}(X)}{\varepsilon^2}$. □

Application 7

Une urne contient des boules blanches et des boules noires ; la proportion de boules blanches présentes dans l'urne est $p \in \left]0, 1\right[$. On effectue n tirages successifs et avec remise d'une boule. Soit X_n la var. égale au nombre de boules blanches obtenues en n tirages.

1. Donner la loi de X_n .
2. Montrer que $\mathbb{V}\left(X_n\right) \leq \dfrac{n}{4}$.
3. A l'aide de l'inégalité de Bienaymé – Tchebychev, donner une majoration de
$$\mathbb{P}\left(\left|\frac{X_n}{n} - p\right| > \varepsilon\right).$$
4. Comment doit – on choisir n pour pouvoir affirmer avec un risque d'erreur inférieur à $5\ \%$ que $\dfrac{X_n}{n}$ est une valeur approchée de p à 10^{-2} près ?

3. *A titre culturel* : la notion de convergence en probabilité

Définition hors – programme

Soit $\left(X_n\right)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un même espace probabilisé $\left(\Omega, \mathcal{A}, \mathbb{P}\right)$, et soit Y une variable aléatoire définie sur ce même espace. On dit que $\left(X_n\right)_{n \in \mathbb{N}^*}$ converge en probabilité vers Y, lorsque :

$$\forall\ \varepsilon > 0,\ \lim_{n \to +\infty} \mathbb{P}\left(\left|X_k - Y\right| \geq \varepsilon\right) = 0.$$

4. Loi faible des grands nombres

a. *Espérance et variance d'une moyenne*

Soit $\left(X_n\right)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un même espace probabilisé, indépendantes, admettant une espérance commune μ et une variance commune σ^2.

Alors pour tout $n \in \mathbb{N}^*$, $\frac{1}{n}\sum_{k=1}^{n} X_k$ admet une espérance et une variance, et l'on a :

$$\boxed{\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \mu \text{, et } \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{\sigma^2}{n}}.$$

Démonstration

Les variables aléatoires X_k admettant toutes μ pour espérance, $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ possède donc elle aussi une espérance, et par linéarité $\mathbb{E}\left(\overline{X}_n\right) = \mu$; de plus, les var. X_k possèdent toutes une variance, $\overline{X}_n$ en a donc une elle aussi. On a $\mathbb{V}\left(\overline{X}_n\right) = \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\mathbb{V}\left(\sum_{k=1}^{n} X_k\right)$, puis par indépendance (donc non – corrélation) :

$\mathbb{V}\left(\overline{X}_n\right) = \frac{1}{n^2}\sum_{k=1}^{n}\mathbb{V}\left(X_k\right) = \frac{1}{n^2}\left(n\sigma^2\right) = \frac{\sigma}{n}$.

b. ***Théorème (loi faible des grands nombres)***

> Soit $\left(X_n\right)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur le même espace probabilisé, indépendantes, admettant une espérance commune μ et une variance commune σ^2.
>
> Alors pour tout $\varepsilon > 0$: $\lim\limits_{n\to+\infty}\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \geq \varepsilon\right) = 0$.

Autrement dit, et sous ces hypothèses, la suite $\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)_{n\in\mathbb{N}^*}$ *converge en probabilité vers la var. certaine égale à* μ.

Démonstration

Il suffit d'appliquer l'inégalité de Bienaymé – Tchebycheff à la var. $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$, en commençant par reprendre ce qui précède : toutes les variables aléatoires X_k admettant μ pour espérance, $\overline{X}_n$ possède elle aussi une espérance, et par linéarité $\mathbb{E}\left(\overline{X}_n\right) = \mu$; de plus, les var. X_k possèdent toutes une variance, $\overline{X}_n$ en a donc une elle aussi.

On a $\mathbb{V}\left(\overline{X}_n\right) = \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\mathbb{V}\left(\sum_{k=1}^{n} X_k\right)$, puis par indépendance (donc non – corrélation) :

$\mathbb{V}\left(\overline{X}_n\right) = \frac{1}{n^2}\sum_{k=1}^{n}\mathbb{V}\left(X_k\right) = \frac{1}{n^2}\left(n\sigma^2\right) = \frac{\sigma}{n}$. Les hypothèses d'application de l'inégalité de Bienaymé – Tchebychev sont bien réunies ; on obtient pour tout $\varepsilon > 0$, pour tout $n \in \mathbb{N}$, $\mathbb{P}\left(\left|\overline{X}_n - \mathbb{E}\left(\overline{X}_n\right)\right| \geq \varepsilon\right) \leq \frac{\mathbb{P}\left(\overline{X}_n\right)}{\varepsilon^2}$, d'où

$0 \leq \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$. On en déduit, par encadrement, que $\lim\limits_{n\to+\infty}\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \geq \varepsilon\right) = 0$.

Remarque

Ce résultat reste vrai avec l'hypothèse plus faible : $\left(X_n\right)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires définies sur le même espace probabilisé, non corrélées, admettant une espérance commune μ et une variance commune σ^2.

Application 8

1. Enoncer et démontrer le théorème de stabilité des lois de Poisson.
2. Montrer que la somme de n variables aléatoires indépendantes, suivant toutes la loi de Poisson de paramètre 1, suit la loi de Poisson de paramètre n.
3. On pose $Q_n(x) = \mathrm{e}^{-x}\sum_{k=0}^{2n}\frac{x^k}{k!}$.

   Montrer en vous aidant de la question précédente que $\lim_{n\to+\infty} Q_n(n) = 1$.

B_ UNE APPROXIMATION DE LOI

1. Convergence en loi *notion hors – programme*

a. *Définition*

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et soit Y une variable aléatoire définie sur ce même espace probabilisé.

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers la variable aléatoire Y, et l'on pourra noter $(X_n)_{n\in\mathbb{N}^*} \xrightarrow[\mathcal{L}]{} Y$, lorsque, en tout point x où la fonction de répartition de Y est continue : $\lim_{n\to+\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(Y \leq x)$.

b. *Cas discret*

Lorsque Y est une variable aléatoire discrète, la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers Y si et seulement si :

$$\forall\ y \in Y(\Omega),\ \lim_{n\to+\infty} \mathbb{P}(X_n \leq y) = \mathbb{P}(Y \leq y)\ .$$

Et pourtant, les éléments de $Y(\Omega)$ sont justement les points où la fonction de répartition de Y peut ne pas être continue… c'est comme ça.

2. Approximation de lois binomiales par des lois de Poisson

a. *Théorème*

Soit $(p_n)_{n\in\mathbb{N}}$ une suite d'éléments de $[0,1]$, et soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. On suppose que :

- Pour tout $n \in \mathbb{N}$, X_n suit la loi $\mathcal{B}(n, p_n)$;
- •• La suite $(n\,p_n)_{n\in\mathbb{N}}$ converge vers un réel λ strictement positif.

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

La suite $(X_n)_{n\in\mathbb{N}}$ converge donc en loi vers une variable aléatoire discrète X de loi $\mathcal{P}(\lambda)$.

Démonstration

Pour tout $k \in \mathbb{N}$:

$$\mathbb{P}(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{1}{k!}\frac{n!}{(n-k)!} p_n^k (1-p_n)^{n-k} = \frac{1}{k!}\left(\prod_{i=0}^{k-1}(n-i)\right) p_n^k (1-p_n)^{n-k},$$

donc :

$$\begin{aligned}
\mathbb{P}(X_n = k) &\underset{n\to+\infty}{=} \frac{1}{k!}\left(n + o(n)\right)^k \left(\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^{n-k} \\
&\underset{n\to+\infty}{=} \frac{1}{k!}\left(1 + o(1)\right)\lambda^k \exp\left((n-k)\ln\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)\right) \\
&\underset{n\to+\infty}{=} \frac{1}{k!}\left(1 + o(1)\right)\lambda^k \exp\left((n-k)\left(-\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)\right) \\
&\underset{n\to+\infty}{=} \frac{1}{k!}\left(1 + o(1)\right)\lambda^k \exp\left(-\lambda + o(1)\right),
\end{aligned}$$

et l'on en conclut que $\mathbb{P}(X_n = k) \underset{n\to+\infty}{\sim} \mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$.

b. ***Conséquence pratique***

Pour $n \in \mathbb{N}^*$ " suffisamment grand" et $p \in [0,1]$ " suffisamment petit ", tels que le produit $\lambda = np$ ne soit " pas trop grand ", on peut approcher la loi d'une variable aléatoire X de loi $\mathcal{B}(n, p)$ par la loi $\mathcal{P}(\lambda)$.

En pratique, on juge souvent cette approximation possible lorsque $n \geq 30$, $p \leq 0{,}1$ et $np \leq 10$; on s'autorisera alors à écrire : $\forall\ k \in \mathbb{N},\ \mathbb{P}(X = k) \approx \mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$.

Annexe : corrigé de l'exercice d'application 4, p. 23

1. Si $r \in \mathbb{N}$, $\displaystyle\sum_{k=0}^{+\infty}\binom{r}{k}x^k y^{r-k}$ est la somme finie $\displaystyle\sum_{k=0}^{n}\binom{r}{k}x^k y^{r-k}$, évidemment convergente, et la formule du binôme habituelle donne $\displaystyle\sum_{k=0}^{n}\binom{r}{k}x^k y^{r-k} = (x+y)^r$, on a donc le résultat. Supposons maintenant que $r \in \mathbb{Z}_-^*$, posons alors $r = -q$.

On a $\displaystyle\binom{r}{k} = \frac{\prod_{j=0}^{k-1}(r-j)}{k!} = \frac{\prod_{j=0}^{k-1}(-q-j)}{k!} = (-1)^k\frac{\prod_{j=0}^{k-1}(q+j)}{k!}$, donc $\displaystyle\binom{r}{k} = (-1)^k\frac{(q+k-1)!}{k!(q-1)!} = (-1)^k\binom{q+k-1}{k}$.

Réécrivons la série $\displaystyle\sum_{k\geq 0}\binom{r}{k}x^k y^{r-k}$ autrement (tout ce qui suit est licite : on ne suppose pas que la série converge, on n'écrit pas de trucs interdits entre séries, on se contente de transformer le terme général).

On a $\displaystyle\sum_{k\geq 0}\binom{r}{k}x^k y^{r-k} = \frac{y^{-q}}{(q-1)!}\sum_{k\geq 0}\frac{(q-1+k)!}{k!}\left(-\frac{x}{y}\right)^k$, puis en posant $m = q-1+k$:

$$\sum_{k\geq 0}\binom{r}{k}x^k y^{r-k} = \frac{y^{-q}}{(q-1)!}\sum_{m\geq q-1}\frac{m!}{(m-q+1)!}\left(-\frac{x}{y}\right)^{m-q+1}.$$

Comme $\left|\dfrac{x}{y}\right| < 1$, d'après le résultat admis en début d'exercice, $\displaystyle\sum_{m\geq q-1}\frac{m!}{(m-q+1)!}\left(-\frac{x}{y}\right)^{m-q+1}$ converge, et $\displaystyle\sum_{m=q-1}^{+\infty}\frac{m!}{(m-q+1)!}\left(-\frac{x}{y}\right)^{m-q+1} = \frac{(q-1)!}{\left(1+\frac{x}{y}\right)^q}$. On en déduit que $\displaystyle\sum_{k\geq 0}\binom{r}{k}x^k y^{r-k}$ converge, et que

$\sum_{k=0}^{+\infty}\binom{r}{k}x^k y^{r-k} = \frac{y^{-q}}{(q-1)!}\frac{(q-1)!}{\left(1+\frac{x}{y}\right)^q}$, d'où l'égalité $\boxed{\sum_{k=0}^{+\infty}\binom{r}{k}x^k y^{r-k} = (x+y)^{-q} = (x+y)^r}$.

2. Notons pour tout $k \geq 0$, $p_k = \binom{-r}{k}a^k(1-a)^{-r-k}$. Notons déjà que les p_k sont bien définis, car $1-a \neq 0$, puisque a est strictement négatif. De plus, on a vu en **1.** que $\binom{-r}{k} = (-1)^k\binom{r+k-1}{k}$; on a donc $p_k = \binom{r+k-1}{k}(-a)^k(1-a)^{-r-k}$, et il est alors clair que p_k est positif, puisque tous les facteurs le sont. De plus, comme $a<0$, $\left|\frac{a}{1-a}\right| < 1$. La formule du binôme assure donc que $\sum p_k$ converge, et que $\sum_{k=0}^{+\infty} p_k = (a+1-a)^{-r} = 1$. On sait alors qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et une variable aléatoire X sur cet espace, tel que $X(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}$, $\mathbb{P}(X=k) = \binom{-r}{k}a^k(1-a)^{-r-k}$.

3. La variable aléatoire X est à valeurs dans $\mathbb{N}$. Pour tout $n \in \mathbb{N}$, notons Y_n la variable aléatoire représentant le nombre d'échecs subis lors des n premières épreuves, et Z_{n+1} la variable aléatoire indicatrice du succès lors de l'épreuve $n+1$; Y_n suit la loi binomiale classique $\mathcal{B}(n,p)$, et Z_{n+1} la loi de Bernoulli de paramètre p.

Soit $k \in \mathbb{N}$. l'évènement $[X=k]$ est réalisé si et seulement si il y a k échecs avant le r ième succès. Notons que ceci est réalisé si et seulement si $\begin{cases}\text{Il y a succès à une certaine épreuve } n\\ \text{Il y a } r-1 \text{ succès et } k \text{ échecs avant cette épreuve}\end{cases}$. Manifestement, on n'a pas le choix quant à la valeur de n : $[X=k]$ est réalisé si et seulement si : $\begin{cases}\text{Il y a succès à l'épreuve } r+k\\ \text{Il y a } r-1 \text{ succès et } k \text{ échecs avant cette épreuve}\end{cases}$.

Autrement dit, $[X=k] = [Z_{r+k}=1] \cap [Y_{r+k-1} = r-1]$, et comme il y a indépendance des épreuves :

$\mathbb{P}(X=k) = \mathbb{P}(Z_{r+k}=1)\mathbb{P}(Y_{r+k-1}=r-1)$. On a donc $\mathbb{P}(X=k) = p\binom{r+k-1}{r-1}p^{r-1}q^k$, puis par symétrie des coefficients binomiaux, $\mathbb{P}(X=k) = \binom{r+k-1}{r}p^r q^k$. On a montré que pour $r \geq 1$, $\binom{-r}{k} = (-1)^k\binom{r+k-1}{k}$. On peut donc réécrire $\mathbb{P}(X=k)$ sous la forme $\mathbb{P}(X=k) = \binom{-r}{k}p^r(-q)^k$, d'où

$$\mathbb{P}(X=k) = \binom{-r}{k}\left(-\frac{q}{p}\right)^k p^{r+k} = \binom{-r}{k}\left(-\frac{q}{p}\right)^k\left(\frac{1}{p}\right)^{-r-k} = \binom{-r}{k}\left(-\frac{q}{p}\right)^k\left(1+\frac{q}{p}\right)^{-r-k} :$$

X suit la loi binomiale négative $\mathcal{B}\left(-r, -\frac{q}{p}\right)$.

Deux questions supplémentaires

4. Calculer, pour p entier naturel, les coefficients $\binom{-1}{p}, \binom{1/2}{p}, \binom{-1/2}{p}$.

5. En utilisant un produit de Cauchy, démontrer la ***formule de Vandermonde généralisée***, à savoir :

$$\forall (\alpha,\beta) \in \mathbb{R}^2,\ \forall n \in \mathbb{N},\ \sum_{i=0}^{n}\binom{\alpha}{i}\binom{\beta}{n-i} = \binom{\alpha+\beta}{n}.$$

1. Pour tout $p \geq 1$:

- On a $\binom{-1}{p} = \frac{-1(-2)\ldots(-p)}{p!} = (-1)^p$.

•• Passons maintenant à $\binom{-1/2}{p}$:

$$\binom{-\frac{1}{2}}{p} = \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\ldots\left(-\frac{2p-1}{2}\right)}{p!} = \left(-\frac{1}{2}\right)^p \frac{1.3.5\ldots(2p-1)}{p!}$$

Simplifier cette expression est archi classique ; on commence par intercaler au numérateur les termes d'indices pairs, en écrivant que $\binom{-\frac{1}{2}}{p} = \left(-\frac{1}{2}\right)^p \frac{1.2.3.4.5\ldots(2p-1).(2p)}{p!\left(2.4.6.\ldots.(2p)\right)}$.

On en déduit que $\binom{-\frac{1}{2}}{p} = \left(-\frac{1}{2}\right)^p \frac{(2p)!}{p!\left(2.4.6.\ldots.(2p)\right)}$, puis : $\binom{-\frac{1}{2}}{p} = \left(-\frac{1}{2}\right)^p \frac{(2p)!}{p!\left(2^p(1.2.3.\ldots.p)\right)}$.

L'essentiel est fait ; il ne reste plus qu'à apporter la touche finale, et l'on obtient :

$$\boxed{\binom{-\frac{1}{2}}{p} = \left(-\frac{1}{2}\right)^p \frac{(2p)!}{p!\left(2^p(p!)\right)} = \left(-\frac{1}{4}\right)^p \frac{(2p)!}{(p!)^2} = \left(-\frac{1}{4}\right)^p \binom{2p}{p}}.$$

••• Il n'est pas nécessaire de recommencer pour $\binom{1/2}{p}$, car

$\binom{1/2}{p} = \frac{\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\ldots\left(-\frac{2p-3}{2}\right)}{p!} = \frac{1}{2}\frac{1}{\left(-\frac{2p-1}{2}\right)}\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\ldots\left(-\frac{2p-1}{2}\right)}{p!} = \frac{-1}{2p-1}\binom{-\frac{1}{2}}{p}$. On déduit alors de ce qui précède que $\boxed{\binom{1/2}{p} = \frac{-1}{2p-1}\left(-\frac{1}{4}\right)^p\binom{2p}{p}}$.

On remarquera que ces trois formules restent valables lorsque $p = 0$.

2. L'égalité $(1+t)^{\alpha+\beta} = (1+t)^\alpha(1+t)^\beta$ donne pour tout $n \in \mathbb{N}$:

$$\sum_{k=0}^{n}\binom{\alpha+\beta}{k}t^k + o(t^n) \underset{0}{=} \left(\sum_{i=0}^{n}\binom{\alpha}{i}t^i + o(t^n)\right)\left(\sum_{j=0}^{n}\binom{\beta}{j}t^j + o(t^n)\right),$$

soit : $\sum_{k=0}^{n}\binom{\alpha+\beta}{k}t^k \underset{0}{=} \left(\sum_{i=0}^{n}\binom{\alpha}{i}t^i\right)\left(\sum_{j=0}^{n}\binom{\beta}{j}t^j\right) + o(t^n)$. La formule du produit polynomial donne alors, en ne gardant que les termes significatifs (ie. de degré inférieur ou égal à n) du terme de droite de cette égalité :

$$\sum_{k=0}^{n}\binom{\alpha+\beta}{k}t^k \underset{0}{=} \sum_{k=0}^{n}\left(\sum_{i=0}^{k}\binom{\alpha}{i}\binom{\beta}{k-i}\right)t^k + o(t^n).$$

L'unicité du développement limité autorise à identifier les coefficients de ces deux écritures, et en particulier les coefficients des termes de degré n. On obtient ainsi la formule de Vandermonde généralisée :

$$\boxed{\forall\, n \in \mathbb{N},\ \binom{\alpha+\beta}{n} \underset{0}{=} \sum_{i=0}^{n}\binom{\alpha}{i}\binom{\beta}{n-i}}.$$

Remarque : au lieu de faire le produit de deux développements limités, on aurait pu faire celui de deux développements en séries entières (produit de Cauchy donc).