VARIABLES ALÉATOIRES DISCRÈTES

I. Généralités

I.1. Définition

Déf 1:

Soit $(\Omega, \mathcal{A})$, un espace probabilisable. On appelle **variable aléatoire discrète** définie sur $(\Omega, \mathcal{A})$ à valeurs dans un ensemble E toute application X de Ω dans E telle que :

- son image $X(\Omega)$ est au plus dénombrable (on pourra donc écrire $X(\Omega) = \{x_i,\ i \in I\}$ avec $I \subset \mathbb{N}$);
- pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$ appartient à $\mathcal{A}$ (c'est-à-dire est un évènement).
  Cet ensemble se note, en abrégé : $(X = x)$.

Si $X(\Omega)$ est un ensemble fini, on dit que X est une **variable aléatoire finie**.
Si X est constante, on dit que c'est une variable aléatoire **certaine**.

Exemples

1. Un joueur lance deux fois de suite un dé et note les deux nombres obtenus sous la forme d'un couple. L'univers de cette expérience est $\Omega = [\![1\,;6]\!] \times [\![1\,;6]\!]$.

   On peut alors définir la variable aléatoire finie X qui à chaque couple associe la somme des deux nombres obtenus. Ici on a $X(\Omega) = [\![2\,;12]\!]$.

2. On effectue une succession de lancers d'un dé cubique jusqu'à obtenir 6. Soit X le nombre de lancers effectués.
   L'univers associé à cette expérience est $[\![1\,;6]\!]^{\mathbb{N}}$, qui n'est pas dénombrable. On ne sait donc pas décrire la tribu et la probabilité associée. Il ne nous est donc pas possible de prouver rigoureusement que X est une variable aléatoire.

   On peut tout de même donner très clairement $X(\Omega)$:

   En étant rigoureux on devrait écrire : $X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$ car il se peut que l'on n'obtienne jamais 6. Mais on peut démontrer que la probabilité de ne jamais obtenir 6 est égale à 0, c'est-à-dire que l'on obtiendra presque sûrement 6. On peut alors choisir de considérer que $X(\Omega) = \mathbb{N}^*$ et X est une variable aléatoire discrète infinie.

Rem: Lorsque Ω est dénombrable, on prend traditionnellement $\mathcal{A} = \mathcal{P}(\Omega)$ comme tribu d'évènements; dans ce cas, toute application X de Ω dans E est une variable aléatoire discrète,car son image est forcément au plus dénombrable, et l'image réciproque de tout élément de E est une partie de Ω donc de la tribu $\mathcal{A}$.

Théorème 1:

Soit X une variable aléatoire discrète définie sur $(\Omega, \mathcal{A})$, à valeurs dans E.
Pour toute partie $A \subset E$, l'ensemble $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$ est un événement de $\mathcal{A}$ que l'on notera $(X \in A)$.

☞ *Démonstration:*

En effet, si $A = \bigcup_{i \in I} \{x_i\}$ avec $I \subset \mathbb{N}$, on aura $X^{-1}(A) = \bigcup_{i \in I} X^{-1}(\{x_i\})$: c'est donc une réunion dénombrable d'éléments de $\mathcal{A}$, donc c'est un élément de $\mathcal{A}$ par définition d'une tribu.

Rem: Dans le cas où $E = \mathbb{R}$, on pourra donc considérer, pour tout x réel, les évènements $(X \geqslant x)$, $(X > x)$ etc...
Par exemple : $(X > x) = \{\omega \in \Omega \mid X(\omega) > x\}$.

Le résultat suivant est extrêmement important et permet en particulier de vérifier la cohérence des résultats obtenus.

Théorème 2:

Soit X une variable aléatoire discrète. Si $X(\Omega) = \{x_i \mid i \in I\}$, alors la famille d'événements $(X = x_i)_{i\in I}$ est un système complet d'événements.
En particulier on a : $\sum_{i\in I} \mathbb{P}(X = x_i) = 1$.

☞ *Démonstration:*

- Si $i \neq j$, $(X = x_i) \cap (X = x_j) = \emptyset$ puisque un élément $\omega \in \Omega$ ne peut avoir deux images différentes.
- Tout $\omega \in \Omega$ appartient à $(X = X(\omega))$ donc $\Omega = \bigcup_{i\in I} (X = x_i)$.

Rem: Comme $(X = x_i)_{i\in I}$ est un système complet d'événements, on peut appliquer la formule des probabilités totales pour n'importe quel événement A :

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}((X = x_i) \cap A) = \sum_{i\in I} \mathbb{P}(X = x_i)\mathbb{P}(A \mid (X = x_i)).$$

I.2. Loi d'une variable aléatoire

Déf 2:

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et X une variable aléatoire discrète à valeurs dans E.
L'application $\mathbb{P}_X : \begin{array}{ccc} \mathcal{P}(X(\Omega)) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & \mathbb{P}(X \in A) \end{array}$ s'appelle la **loi de probabilité** de X.

Remarques

1. Puisque, pour toute partie A de $X(\omega)$ on a $\mathbb{P}(X \in A) = \sum_{x\in A} \mathbb{P}(X = x)$, il suffit de connaître toutes les probabilités $\mathbb{P}(X = x)$ pour $x \in X(\Omega)$ pour connaître la loi de X.
   L'ensemble des $\mathbb{P}(X = x)$ pour $x \in X(\Omega)$ s'appelle la **distribution de probabilité** de X.
2. Pour répondre à la question « déterminer la loi de X », il faut *commencer par donner clairement* $X(\Omega)$. Puis pour chaque élément x_i de cet ensemble $X(\Omega)$ il faut donner $\mathbb{P}(X = x_i)$ (d'après la remarque précédente, cela suffit).

Prop 1:

$\mathbb{P}_X$ est une probabilité sur $\Big(X(\Omega), \mathcal{P}(X(\Omega))\Big)$.

☞ *Démonstration:*

Vérification facile.

Le théorème suivant (*admis*) montre qu'il suffit de connaître la distribution de probabilité pour caractériser une variable aléatoire.

Théorème 3: Germe de probabilité

Soit $(\Omega, \mathcal{A})$ un espace probabilisable, et $\{x_i,\ i \in I\}$ un ensemble au plus dénombrable.
Soit $(p_i)_{i\in I}$ une famille de réels positifs telle que $\sum_{i\in I} p_i = 1$.
Alors il existe une probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ et une variable aléatoire discrète X sur $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\{x_i,\ i \in I\}$, tels que

$$\forall\, i \in I\ ,\ \mathbb{P}(X = x_i) = p_i\,.$$

Rem : L'écriture $\sum_{i\in I} p_i$ sous-entend, lorsque I est infini, que la famille est sommable !

I.3. Fonction d'une variable aléatoire

Prop 2:

Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un ensemble E et f une fonction définie sur $X(\Omega)$ à valeurs dans un ensemble F.
Alors $f \circ X$ est une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ (on la notera simplement $f(X)$).

☞ *Démonstration:*

Si $X(\Omega) = \{x_i,\, i \in I\}$ avec $I \subset \mathbb{N}$, alors $f \circ X(\Omega) = \{f(x_i),\, i \in I\}$ donc $f \circ X(\Omega)$ est au plus dénombrable.
De plus soit $x \in f \circ X(\Omega)$ alors

$$A = (f \circ X)^{-1}(\{x\}) = \{\omega \in \Omega \text{ tq } f(X(\omega)) = x\} = \left\{\omega \in \Omega \text{ tq } X(\omega) \in f^{-1}(\{x\})\right\} = (X \in f^{-1}(\{x\}))\,.$$

Puisque $f^{-1}(\{x\})$ est un sous-ensemble de E, A est bien un évènement ce qui prouve que $f \circ X$ est bien une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Prop 3:

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un ensemble E et f une fonction définie sur $X(\Omega)$ à valeurs dans un ensemble F.
Si X et Y suivent la même loi, alors $f(X)$ et $f(Y)$ suivent aussi la même loi.

☞ *Démonstration:*

Par définition, la loi de $f(X)$ est la donnée des $\mathbb{P}(f(X) \in B)$ pour toute partie B de $f(X)(\Omega)$.
Or $\mathbb{P}(f(X) \in B) = \mathbb{P}(X \in f^{-1}(B)) = \mathbb{P}_X(f^{-1}(B))$ donc il est clair que si $\mathbb{P}_X = \mathbb{P}_Y$, les lois de $f(X)$ et de $f(Y)$ sont les mêmes.

Notation : on note $X \sim Y$ lorsque X et Y ont même loi.

II. Lois usuelles

II.1. Lois discrètes finies

II.1.1. Loi uniforme

Situation : Tirage « au hasard » d'un entier parmi $[\![1\,;n]\!]$ (ou une boule dans une urne...), tous les tirages étant équiprobables.

Déf 3:

On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$ suit la **loi uniforme sur $[\![1\,;n]\!]$**, et l'on note $X \sim \mathcal{U}([\![1\,;n]\!])$ (ou $X \hookrightarrow \mathcal{U}([\![1\,;n]\!])$) si :

a) $X(\Omega)) = [\![1\,;n]\!]$;

b) $\forall\, x \in X(\Omega),\ \mathbb{P}(X = x) = \dfrac{1}{n}$.

II.1.2. Loi de Bernoulli

Situation : Expérience de type succès-échec (pile ou face, tirage d'une boule noire ou blanche...)

Déf 4:

Soit p un réel dans $[0\,;1]$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$ suit la **loi de Bernoulli de paramètre p**, et l'on note $X \sim \mathcal{B}(p)$ (ou $X \hookrightarrow \mathcal{B}(p)$) si :

a) $X(\Omega) = \{0;1\}$;

b) $\mathbb{P}(X = 1) = p \quad$ et $\quad \mathbb{P}(X = 0) = 1 - p$.

Rem: Les cas $p = 0$ et $p = 1$ ne sont guère intéressants, c'est pourquoi on considère souvent $p \in\,]0\,;1[$.

II.1.3. Loi binomiale (ou tirages avec remise)

Situation : On considère une expérience de type succès-échec que l'on effectue n fois dans les même conditions. On considère X le nombre de fois où l'expérience a amené un succès.
X prend donc les valeurs $0, 1, ..., n$.

Soit $k \in [\![0\,;n]\!]$. On cherche à calculer $\mathbb{P}(X=k)$. Parmi les n expériences, il y a $\binom{n}{k}$ façons de placer les k fois où il y a eu un succès. Chacun de ces événements est réalisé avec la probabilité $p^k(1-p)^{n-k}$, où p est la probabilité d'un succès.

On a donc : $\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Déf 5:

Soit n un entier non nul et p un réel dans $[0\,;1]$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$ suit la **loi binomiale de paramètres n et p**, et l'on note $X \sim \mathcal{B}(n,p)$ (ou $X \hookrightarrow \mathcal{B}(n,p)$) si :

a) $X(\Omega) = [\![0\,;n]\!]$;

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

Remarques

1. Les cas $p=0$ et $p=1$ ne sont guère intéressants, c'est pourquoi on considère souvent $p \in]0\,;1[$.
2. On note traditionnellement $q = 1-p$ de sorte que $\mathbb{P}(X=k) = \binom{n}{k} p^k q^{n-k}$.
3. Une loi de Bernoulli de paramètre p est aussi une loi binomiale de paramètres 1 et p, c'est pourquoi l'on note souvent $\mathcal{B}(1,p)$ au lieu de $\mathcal{B}(p)$.

II.2. Lois discrètes infinies

II.2.1. Loi géométrique (ou temps d'attente d'un premier succès)

Situation : On considère une expérience de type succès-échec et on répète l'expérience dans des conditions identiques. On note X le nombre d'épreuves effectuées jusqu'à ce que l'on obtienne un succès pour la première fois.
Notons p la probabilité d'un succès, avec $p \in]0\,;1[$ (sinon, inutile de jouer !), et A_i l'événement « on a obtenu un succès lors de la i^{eme} expérience ».

Alors, pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(X=n) = \mathbb{P}(\overline{A_1} \cap \overline{A_2} \cap ... \cap \overline{A_{n-1}} \cap A_n) = (1-p)^{n-1} p\,.$$

Soit R l'événement « on n'obtient que des échecs ». On a $R = \bigcap_{i=1}^{+\infty} \overline{A_i}$. Donc

$$\mathbb{P}(X=+\infty) = \mathbb{P}(R) = \lim_{n\to+\infty} P\left(\bigcap_{i=1}^{n} \overline{A_i}\right) = \lim_{n\to+\infty} (1-p)^n = 0\,.$$

On peut donc considérer que X prend ses valeurs dans $\mathbb{N}^*$.

Déf 6:

Soit p un réel dans $]0,1[$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$ suit la **loi géométrique de paramètre p**, et l'on note $X \sim \mathcal{G}(p)$ (ou $X \hookrightarrow \mathcal{G}(p)$) si :

a) $X(\Omega) = \mathbb{N}^*$;

b) $\forall n \in \mathbb{N}^*,\quad \mathbb{P}(X=n) = (1-p)^{n-1} p$.

Rem: Il est facile de vérifier que $\sum_{n=1}^{+\infty} \mathbb{P}(X=n)$ est bien égal à 1 (série géométrique).

Prop 4:

Si $X \hookrightarrow \mathcal{G}(p)$, alors pour tout $k \in \mathbb{N}^*$:

$$\mathbb{P}(X > k) = (1-p)^k.$$

☞ *Démonstration:*

On peut écrire que, puisque $(X > k) = \bigsqcup_{n>k} (X = n)$:

$$\mathbb{P}(X > k) = \sum_{n=k+1}^{+\infty} \mathbb{P}(X = n)$$

puis utiliser les formules sur les séries géométriques, ou bien faire une démonstration probabiliste en remarquant que l'évènement $(X > k)$ est tout simplement l'évènement : « les k premières épreuves ont été un échec ».

II.2.2. Loi de Poisson

La loi de Poisson est attribuée à Siméon D. Poisson, mathématicien français (1781-1840). Cette loi fut proposée par Poisson dans un ouvrage qu'il publia en 1837 sous le titre "Recherche sur la probabilité de jugements en matière criminelle et en matière civile".

Situation concrète :

Beaucoup de situations sont liées à l'étude de la réalisation d'un événement dans un intervalle de temps donné (arrivée de clients qui se présentent à un guichet d'une banque en une heure, apparitions de pannes d'un réseau informatique en une année, arrivée de malades aux urgences d'un hôpital en une nuit,....).
Pour décrire les réalisations dans le temps d'un événement donné, on peut

- soit chercher le nombre de réalisations de l'événement dans un intervalle de temps donné qui est distribué suivant une loi de Poisson.
- soit chercher le temps entre deux réalisations successives de l'événement qui est distribué suivant une loi exponentielle (hors programme).

On va voir que la loi de Poisson peut être interprétée comme un cas limite d'une loi binomiale.

Processus de Poisson :

Précisons les hypothèses faites relativement à la réalisation de l'événement qui nous intéresse.

1. Les nombres de réalisations de l'événement au cours d'intervalles de temps disjoints sont des variables aléatoires indépendantes, c'est-à-dire que le nombre de réalisations au cours d'un intervalle de temps est indépendant du nombre de réalisations au cours d'intervalles de temps antérieurs.
2. La probabilité pour que l'événement se réalise une fois, au cours d'un petit intervalle de temps Δt, est proportionnelle à l'amplitude de l'intervalle et vaut $\alpha\Delta t$, où α est une valeur positive que l'on suppose constante tout au long de la période d'observation.
3. Il est très rare d'observer plus d'une fois l'événement au cours d'un petit intervalle de temps Δt, c'est-à-dire que la probabilité pour que l'événement se réalise plus d'une fois au cours de l'intervalle de temps Δt est négligeable.

Les hypothèses 1., 2., 3. caractérisent ce qu'on appelle un *processus de Poisson*. α est une constante du processus qui représente le nombre moyen de réalisations par unité de temps et que l'on appelle l'*intensité* du processus.
Sous ces hypothèses, la variable aléatoire $X =$ « nombre de fois où l'événement considéré se réalise au cours d'un intervalle de temps de durée t » est distribuée suivant une loi de Poisson de paramètre $\lambda = \alpha t$.

Loi de Poisson :

Nous connaissons déjà la loi de probabilités de la variable $Y =$ « nombre de réalisations d'un événement donné, de probabilité p, au cours de n essais ». Il s'agit d'une loi binomiale $\mathcal{B}(n,p)$.
Divisons alors l'intervalle de temps de longueur t en n petits intervalles de temps disjoints de longueur $\Delta t = t/n$.
L'hypothèse 3. permet d'affirmer que dans chacun de ces n petits intervalles il n'y a principalement que deux possibilités : l'événement se réalise une fois ou ne se réalise pas (cela sera d'autant plus vrai que n est grand).
L'hypothèse 2. permet d'affirmer que dans chacun de ces n petits intervalles, la probabilité de réalisation de l'événement est constante et égale à $\alpha\Delta t = \alpha t/n = p$.
D'après l'hypothèse 1., ces évènements sont indépendants, donc le nombre de réalisations de l'évènement dans l'intervalle de temps t suit la loi binomiale $\mathcal{B}(n, \alpha t/n)$.
Pour trouver la loi que l'on cherche, c'est-à-dire la loi de Poisson de paramètre αt, il suffit de faire tendre n vers $+\infty$.

Faisons le calcul :
Si Y suit une loi $\mathcal{B}(n,\lambda/n)$, on sait que

$$\begin{aligned}\mathbb{P}(Y=k) &= \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} &(1)\\ &= \left(1-\frac{\lambda}{n}\right)^{n-k}\frac{\lambda^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k} &(2)\\ &= \left(1-\frac{\lambda}{n}\right)^{n}\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^{-k}\left[\frac{n}{n}\times\frac{n-1}{n}\times\cdots\times\frac{n-k+1}{n}\right].\end{aligned}$$

Chaque terme du produit entre crochets tend vers 1 lorsque n tend vers l'infini. Il y a k termes, c'est-à-dire un nombre fini ne dépendant pas de n donc le crochet tend vers 1. De même, $\left(1-\frac{\lambda}{n}\right)^{-k}$ tend vers 1. De plus,

$$\left(1-\frac{\lambda}{n}\right)^{n}=\mathrm{e}^{n\ln\left(1-\frac{\lambda}{n}\right)} \quad\text{et}\quad \ln\left(1-\frac{\lambda}{n}\right)\underset{n\to\infty}{\sim}-\frac{\lambda}{n}$$

donc $\left(1-\frac{\lambda}{n}\right)^{n}$ tend vers $\mathrm{e}^{-\lambda}$. On en conclut que $\mathbb{P}(Y=k)$ tend vers $\mathrm{e}^{-\lambda}\lambda^k/k!$.

Déf 7:

Soit λ un réel strictement positif. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},\mathcal{P})$ suit la **loi de Poisson de paramètre** λ, et l'on note $X\sim\mathcal{P}(\lambda)$ (ou $X\hookrightarrow\mathcal{P}(\lambda)$) si :

a) $X(\Omega)=\mathbb{N}$;

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

III. Couples de variables aléatoires discrètes

III.1. Définitions

Prop 5: (et définition 8)

Soit $(\Omega,\mathcal{A},\mathcal{P})$, un espace probabilisé et E un ensemble.
Soient X et Y deux variables aléatoires discrètes sur Ω à valeurs dans E. Alors l'application :

$$\begin{array}{rcl}Z:\Omega &\longrightarrow& E^2\\ \omega &\longmapsto& (X(\omega),Y(\omega))\end{array}$$

est une variable aléatoire discrète.
On dit que le couple (X,Y) est un **couple de variables aléatoires discrètes**.

☞ *Démonstration:*

Supposons que X et Y sont des variables aléatoires discrètes. Alors $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables ; il en est donc de même du produit $X(\Omega)\times Y(\Omega)$, donc de $Z(\Omega)$ qui en est une partie.
Ensuite, pour tout $(x,y)\in\mathbb{Z}(\Omega)$:

$$Z^{-1}(\{(x,y)\})=\{\omega\in\Omega\text{ tq }(x,y)=Z(\omega)=(X(\omega),Y(\omega))\}=\{\omega\in\Omega\text{ tq }X(\omega)=x\text{ et }Y(\omega)=y\}=X^{-1}(\{x\})\cap Y^{-1}(\{y\})$$

donc $Z^{-1}(\{(x,y)\})\in\mathcal{A}$ puisque par hypothèse $X^{-1}(\{x\})$ et $Y^{-1}(\{y\})$ sont dans $\mathcal{A}$.

Déf 9:

On appelle **loi du couple de variables aléatoires discrètes** (X,Y), ou encore **loi conjointe des variables aléatoires** X **et** Y la donnée l'ensemble des couples $\big((x_i,y_j),p_{i,j}\big)$ où

$$x_i\in X(\Omega)\quad,\quad y_j\in Y(\Omega)\quad\text{et}\quad p_{i,j}=\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)=\mathbb{P}\big((X,Y)=(x_i,y_j)\big).$$

Les lois de X et Y s'appellent les **lois marginales** du couple X et Y.

Prop 6: Lien entre lois conjointes et lois marginales

Soit $(\Omega, \mathcal{A}, \mathcal{P})$, un espace probabilisé et (X, Y) un couple de variables aléatoires discrètes sur Ω. Notons $X(\Omega) = \{x_i,\ i \in I\}$ et $Y(\Omega) = \{y_j,\ j \in J\}$ où I, J sont deux parties de $\mathbb{N}$. Alors

$$\forall i \in I,\ \mathbb{P}(X = x_i) = \sum_{j\in J} \mathbb{P}\big((X = x_i) \cap (Y = y_j)\big) = \sum_{j\in J} p_{i,j}$$

$$\forall j \in J,\ \mathbb{P}(Y = y_j) = \sum_{i\in I} \mathbb{P}\big((Y = y_j) \cap (X = x_i)\big) = \sum_{i\in I} p_{i,j}\,.$$

☞ *Démonstration:*

formule des probabilités totales appliqué au système complet d'évènements $(Y = y_j)_{j\in J}$.

D'après le théorème de Fubini, la proposition 6 implique que la suite double $(p_{i,j})_{(i,j)\in I\times J}$ est sommable et :

$$\sum_{i\in I, j\in J} p_{i,j} = \sum_{i\in I}\left(\sum_{j\in J} p_{i,j}\right) = \sum_{j\in J}\left(\sum_{i\in I} p_{i,j}\right) = 1.$$

Déf 10:

Soit $(\Omega, \mathcal{A}, \mathcal{P})$, un espace probabilisé et (X, Y) un couple de variables aléatoires discrètes sur Ω. Notons $X(\Omega) = \{x_i,\ i \in I\}$ et $Y(\Omega) = \{y_j,\ j \in J\}$ où I, J sont deux parties de $\mathbb{N}$.

- Pour tout indice $j \in J$ tel que $\mathbb{P}(Y = y_j)$ soit non nul, on appelle **loi conditionnelle de X sachant $(Y = y_j)$** l'application de $X(\Omega)$ dans $[0\,;1]$ définie par la relation

$$x_i \longmapsto \mathbb{P}_{(Y=y_j)}(X = x_i) = \mathbb{P}\big((X = x_i) \mid (Y = y_j)\big) = \frac{\mathbb{P}\big((X = x_i) \cap (Y = y_j)\big)}{\mathbb{P}(Y = y_j)}\,.$$

- Pour tout indice $i \in I$ tel que $\mathbb{P}(X = x_i)$ soit non nul, on appelle **loi conditionnelle de Y sachant $(X = x_i)$** l'application de $Y(\Omega)$ dans $[0\,;1]$ définie par la relation

$$y_j \longmapsto \mathbb{P}_{(X=x_i)}(Y = y_j) = \mathbb{P}\big((Y = y_j) \mid (X = x_i)\big) = \frac{\mathbb{P}\big((X = x_i) \cap (Y = y_j)\big)}{\mathbb{P}(X = x_i)}\,.$$

Il est important de bien savoir jongler entre loi du couple, loi marginale et loi conditionnelle. Il s'agit en fait uniquement d'appliquer la définition des probabilités conditionnelles ou alors d'appliquer la formule des probabilités totales. Cela est résumé dans la proposition suivante.

Prop 7:

Avec les mêmes notations que dans la proposition précédente :

- Loi conditionnelle à partir de la loi du couple :

$$\mathbb{P}_{(X=x_i)}(Y=y_j)=\frac{\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)}{\mathbb{P}(X=x_i)}$$
$$\mathbb{P}_{(Y=y_j)}(X=x_i)=\frac{\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)}{\mathbb{P}(Y=y_j)}$$

- Loi du couple à partir de la loi conditionnelle :

$$\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)=\mathbb{P}(X=x_i)\mathbb{P}_{(X=x_i)}(Y=y_j)$$
$$\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)=\mathbb{P}(Y=y_j)\mathbb{P}_{(Y=y_j)}(X=x_i)$$

- Lois marginales à partir de la loi du couple :

$$\mathbb{P}(X=x_i)=\sum_{j\in J}\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)$$
$$\mathbb{P}(Y=y_j)=\sum_{i\in I}\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)$$

- Lois marginales à partir des lois conditionnelles :

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

Exemples

1. On réalise une suite de lancers indépendants d'une pièce ayant la probabilité $p\in\,]0;1[$ de tomber côté Pile. On note X la longueur de la première série de lancers identiques et Y la longueur de la seconde série. Par exemple, les successions de lancers PPFFFP... et FFPPPF... correspondent à $X=2$ et $Y=3$.

Déterminer la loi conjointe du couple (X,Y) et en déduire les lois de X et de Y.

Solution:

Comme d'habitude, on peut considérer que X et Y prennent leurs valeurs dans $\mathbb{N}^*$, les évènements $(X=+\infty)$ et $(Y=+\infty)$ étant quasi impossibles.
On a alors, pour $(k,\ell)\in\mathbb{N}^{*2}$:

$$\mathbb{P}\big((X=k)\cap(Y=\ell)\big)=p^k(1-p)^\ell p+(1-p)^k p^\ell(1-p)$$

(à expliquer...) d'où l'on tire après calcul (sommes de séries géométriques) :

$$\mathbb{P}(X=k)=\sum_{\ell=1}^{+\infty}\mathbb{P}\big((X=k)\cap(Y=\ell)\big)=(1-p)p^k+p(1-p)^k$$

et

$$\mathbb{P}(Y=\ell)=\sum_{k=1}^{+\infty}\mathbb{P}\big((X=k)\cap(Y=\ell)\big)=p^2(1-p)^{\ell-1}+(1-p)^2p^{\ell-1}.$$

2. Le nombre X de voitures (ou de motos !) se présentant à une barrière de péage en 1h suit une loi de Poisson de paramètre $\lambda>0$.
Cette barrière comporte N postes de péage. Chaque véhicule arrivant se présente à un péage choisi au hasard. On note Y le nombre de véhicules se présentant au poste n°1.
Quelle est la loi de Y ?

Solution:

X et Y sont des variables aléatoires à valeurs dans $\mathbb{N}$.
Par hypothèse, la loi de Y conditionnée par $(X=n)$ est la loi binomiale $\mathcal{B}(n,\frac{1}{N})$:

$$\forall n\in\mathbb{N},\ \forall k\in[\![0;n]\!],\ \mathbb{P}((Y=k)\mid(X=n))=\binom{n}{k}\frac{1}{N^k}\left(1-\frac{1}{N}\right)^{n-k}.$$

On en déduit la loi du couple (X,Y) :

$$\forall (n,k)\in\mathbb{N}^2,\ \mathbb{P}((X=n)\cap(Y=k)) = \begin{cases} \mathbb{P}((Y=k)\mid(X=n))\mathbb{P}(X=n) = \binom{n}{k}\frac{1}{N^k}\left(1-\frac{1}{N}\right)^{n-k}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} & \text{si } k\leqslant n \\ 0 & \text{sinon} \end{cases}.$$

Rem : Compte tenu de la convention $\binom{n}{k}=0$ si $k>n$, la 1ère expression suffit.

La loi de Y s'obtient alors en sommant (les séries écrites sont forcément convergentes) :

$$\begin{aligned}\forall k\in\mathbb{N},\ \mathbb{P}(Y=k) &= \sum_{n=0}^{+\infty}\mathbb{P}((X=n)\cap(Y=k)) = \sum_{n=k}^{+\infty}\binom{n}{k}\frac{1}{N^k}\left(1-\frac{1}{N}\right)^{n-k}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}\\ &= \mathrm{e}^{-\lambda}\left(\frac{\lambda}{N}\right)^k\frac{1}{k!}\sum_{n=k}^{+\infty}\frac{1}{(n-k)!}\left(\lambda\left(1-\frac{1}{N}\right)\right)^{n-k} = \mathrm{e}^{-\lambda}\left(\frac{\lambda}{N}\right)^k\frac{1}{k!}\sum_{i=0}^{+\infty}\frac{1}{i!}\left(\lambda\left(1-\frac{1}{N}\right)\right)^{i}\\ &= \frac{1}{k!}\mathrm{e}^{-\lambda}\left(\frac{\lambda}{N}\right)^k\mathrm{e}^{\lambda\left(1-\frac{1}{N}\right)} = \frac{1}{k!}\mathrm{e}^{-\frac{\lambda}{N}}\left(\frac{\lambda}{N}\right)^k\end{aligned}$$

donc Y suit une loi de Poisson de paramètre $\frac{\lambda}{N}$. (on pouvait s'y attendre ! Pourquoi ?)

Extension :

Si $X_1,\dots,X_n$ sont n variables aléatoires discrètes définies sur un espace probabilisé $(\Omega,\mathcal{A},\mathcal{P})$ à valeurs dans un ensemble E, on peut considérer de la même façon l'application $(X_1,\dots,X_n)$ de Ω dans E^n qui à tout $\omega\in\Omega$ associe le n-uplet $\big(X_1(\omega),\dots,X_n(\omega)\big)$.

Ce n-uplet s'appelle un **n-uplet de variables aléatoires discrètes** ou, parfois, un **vecteur aléatoire**.

III.2. Indépendance de deux variables aléatoires

Déf 11:

Soit (X,Y) un couple aléatoire discret défini sur un espace probabilisé $(\Omega,\mathcal{A},\mathcal{P})$. Notons $X(\Omega)=\{x_i,\ i\in I\}$ et $Y(\Omega)=\{y_j,\ j\in J\}$ où I,J sont deux parties de $\mathbb{N}$.
On dit que les variables X et Y sont **indépendantes** si

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

(en termes savants : la loi conjointe est le produit des lois marginales).
On note alors $X\perp\!\!\!\perp Y$.

Exemples :

1. Une urne contient a boules noires et b boules blanches. On tire deux fois une boule et on note X_i la variable qui vaut 1 si la boule tirée au i-ème tirage est blanche et 0 sinon.
Étudier l'indépendance de X_1 et de X_2 si le tirage se fait avec remise puis dans le cas où le tirage se fait sans remise.

Solution:

- tirage avec remise : $\mathbb{P}(X_1=1)=\frac{b}{a+b}$ et $\mathbb{P}(X_2=1)=\frac{b}{a+b}$ et $\mathbb{P}(X_1=1,X_2=1)=(\frac{b}{a+b})^2$ et de même pour les autres donc X_1 et X_2 sont indépendants.
- tirage sans remise : $\mathbb{P}(X_1=1)=\frac{b}{a+b}$ et $\mathbb{P}(X_2=1)=\frac{b(b-1)+ab}{(a+b)(a+b-1)}=\frac{b}{a+b}$ et $\mathbb{P}(X_1=1,X_2=1)=\frac{b(b-1)}{(a+b)(a+b-1)}$ donc X_1 et X_2 ne sont pas indépendants.

2. Étudier l'indépendance des lois X et Y connaissant la loi du couple (X,Y) dans les cas suivants :
   - $p_{ij}=\frac{1}{e^2}\frac{1}{i!j!}$ pour tous entiers $i,j\in\mathbb{N}$.
   - $p_{ij}=\frac{1}{2e^2}\frac{i+j}{i!j!}$ pour tous entiers $i,j\in\mathbb{N}$.

Solution:

Dans le premier cas il y a indépendance, $p_{ij}=p_ip_j$.
Dans le second cas, il n'y a pas indépendance $p_i=\frac{i+1}{2i!e}$ et $p_j=\frac{j+1}{2j!e}$.

Théorème 4:

Deux variables aléatoires discrètes X et Y sont indépendantes si et seulement si :

$$\forall A \subset X(\Omega),\ \forall B \subset Y(\Omega),\ \mathbb{P}\big((X \in A) \cap (Y \in B)\big) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B)\,.$$

☞ *Démonstration:*

Supposons X et Y indépendantes (la réciproque est immédiate). Alors si A et B sont des parties comme dans l'énoncé elles sont au plus dénombrables. On peut donc écrire

$$A = \{x_i,\ i \in I\} \quad , \quad B = \{y_j,\ j \in J\} \quad \text{avec } I, J \subset \mathbb{N}.$$

où les x_i sont distincts, ainsi que les y_j. On aura alors

$$\begin{aligned}
\mathbb{P}((X \in A) \cap (Y \in B)) &= P\left(\bigsqcup_{i\in I}\bigsqcup_{j\in J}(X = x_i) \cap (Y = y_j)\right) \\
&= \sum_{i\in I}\sum_{j\in J}\mathbb{P}((X = x_i) \cap (Y = y_j)) \qquad \text{par } \sigma - \text{additivité} \\
&= \sum_{i\in I}\sum_{j\in J}\mathbb{P}(X = x_i)\mathbb{P}(Y = y_j) \qquad \text{par indépendance} \\
&= \left(\sum_{i\in I}\mathbb{P}(X = x_i)\right)\left(\sum_{j\in J}\mathbb{P}(Y = y_j)\right) \qquad \text{cas particulier du théorème de Fubini} \\
&= \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B)\,.
\end{aligned}$$

Cela démontre $X \perp\!\!\!\perp Y$.

Théorème 5: Fonction de deux variables aléatoires indépendantes

Soient X et Y deux variables aléatoires discrètes un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$, indépendantes. Soient f et g deux applications définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement. Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

(En abrégé : $X \perp\!\!\!\perp Y \Longrightarrow f(X) \perp\!\!\!\perp f(Y)$.)

☞ *Démonstration:*

On utilise le th. précédent. Soient $A \subset f(X(\Omega))$ et $B \subset g(Y(\Omega))$. On a

$$\begin{aligned}
\mathbb{P}((f(X) \in A) \cap (g(Y) \in B)) &= \mathbb{P}((X \in f^{-1}(A)) \cap (Y \in g^{-1}(B))) \\
&= \mathbb{P}(X \in f^{-1}(A))\mathbb{P}(Y \in g^{-1}(B)) \qquad \text{car } X, Y \text{ indépendantes} \\
&= \mathbb{P}(f(X) \in A)\mathbb{P}(g(Y) \in B)).
\end{aligned}$$

On a donc bien $f(X) \perp\!\!\!\perp g(Y)$.

III.3. Indépendance de n variables aléatoires

Déf 12:

Soient $X_1, ..., X_n$ n variables aléatoires discrètes ($n \geqslant 2$) définies un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$, à valeurs dans un ensemble E.

On dit que les variables $X_1, ..., X_n$ sont **mutuellement indépendantes** (ou tout simplement **indépendantes**) lorsque pour tout $(x_1, ..., x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$:

$$\mathbb{P}\big((X_1 = x_1) \cap \cdots \cap (X_n = x_n)\big) = \mathbb{P}(X_1 = x_1) \times \cdots \times \mathbb{P}(X_n = x_n)$$

Comme dans le cas de deux variables aléatoires (voir théorème 4), on démontre le résultat suivant :

Prop 8:

Soient $X_1, ..., X_n$ n variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$.
Alors pour tous sous-ensembles $A_1 \subset X_1(\Omega), \ldots, A_n \subset X_n(\Omega)$ on a :

$$P\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) = \prod_{i=1}^{n}\mathbb{P}(X_i \in A_i).$$

***Corollaire 8.1*:**

L'indépendance de n variables aléatoires $X_1, \dots, X_n$ entraîne l'indépendance de k quelconque d'entre elles (et donc en particulier l'indépendance deux à deux).

☞ *Démonstration:*

Supposons $X_1, \dots, X_n$ indépendantes. On veut démontrer l'indépendance de $X_1, \dots, X_k$.
On vient de voir que pour tous sous-ensembles $A_1 \subset X_1(\Omega), \dots, A_n \subset X_n(\Omega)$ on a :

$$\mathbb{P}\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) = \prod_{i=1}^{n}\mathbb{P}(X_i \in A_i).$$

Il suffit alors dans cette relation de prendre pour $A_1, \dots, A_k$ un singleton puis $A_i = X_i(\Omega)$ pour $i \geqslant k+1$ pour retrouver la définition de l'indépendance pour $(X_1, \dots, X_k)$.

Le résultat du théorème 5 s'étend à plus de deux variables aléatoires :

Théorème 6: Lemme des coalitions

Soient $X_1, \cdots, X_n$ n variables aléatoires discrètes (mutuellement) indépendantes et soit $k \in [\![2\,; n-1]\!]$. Alors toute variable aléatoire fonction des variables $X_1, \cdots, X_k$ est indépendante de toute variable aléatoire fonction des variables $X_{k+1}, \dots, X_n$.

☞ *Démonstration:*

On considère les deux vecteurs aléatoires :

$$Y = (X_1, \cdots, X_k) \quad \text{et} \quad Z = (X_{k+1}, \cdots, X_n).$$

Montrons que Y et Z sont indépendantes :
Soient $y = (x_1, \dots, x_k) \in Y(\Omega)$ et $z = (x_{k+1}, \dots, x_n) \in Z(\Omega)$. Alors :

$$(Y = y) = \bigcap_{i=1}^{k}(X_i = x_i) \quad \text{et} \quad (Z = z) = \bigcap_{i=k+1}^{n}(X_i = x_i)$$

donc

$$\mathbb{P}\big((Y = y) \cap (Z = z)\big) = \mathbb{P}\left(\bigcap_{i=1}^{n}(X_i = x_i)\right) = \prod_{i=1}^{n}\mathbb{P}(X_i = x_i)$$

par indépendance de X_i.
D'autre part, encore par indépendance des X_i :

$$\mathbb{P}(Y = y) = \mathbb{P}\left(\bigcap_{i=1}^{k}(X_i = x_i)\right) = \prod_{i=1}^{k}\mathbb{P}(X_i = x_i) \quad \text{et, de même} \quad \mathbb{P}(Z = z) = \prod_{i=k+1}^{n}\mathbb{P}(X_i = x_i)$$

donc on a bien $\mathbb{P}(Y = y, Z = z) = \mathbb{P}(Y = y)\mathbb{P}(Z = z)$, Y et Z sont indépendantes.
D'après le théorème 5, $Y \perp\!\!\!\perp Z \Longrightarrow f(Y) \perp\!\!\!\perp g(Z)$, cqfd.

Déf 13: Extension

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires discrètes définies un espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$, à valeurs dans un ensemble E.
Ces variables aléatoires sont dites **indépendantes** si, pour toute partie finie $I \subset \mathbb{N}^*$, la famille $(X_i)_{i\in I}$ est une famille finie de variables aléatoires indépendantes.
Si, de plus, les X_n suivent la même loi, on parlera de **suite de variables aléatoires indépendantes identiquement distribuées**, en abrégé : « i.i.d ».

III.4. Loi d'une fonction de variables aléatoires

Prop 9:

Soit X et Y deux variables aléatoires discrètes et g une fonction quelconque définie sur $X(\Omega) \times Y(\Omega)$. Alors $Z = g(X, Y)$ est une variable aléatoire discrète.

☞ *Démonstration:*

On applique la proposition 2 et la proposition 5 .

Un cas particulier important de loi d'un couple de variables aléatoires est le cas de la somme, lorsque ces variables sont discrètes à valeurs complexes.

Soient donc $(\Omega, \mathcal{A}, \mathcal{P})$, un espace probabilisé et (X, Y) un couple de variables aléatoires discrètes sur Ω, à valeurs dans $\mathbb{R}$. Notons $X(\Omega) = \{x_i,\ i \in I\}$ et $Y(\Omega) = \{y_j,\ j \in J\}$ où I, J sont deux parties de $\mathbb{N}$.

Soit $Z = X + Y$. On sait déjà que Z est une variable aléatoire discrète. Pour tout $z \in Z(\Omega)$, l'évènement $(Z = z)$ est la réunion disjointe et au plus dénombrable des évènements $((X = x_i) \cap (Y = y_j))$ pour tous les couples $(i, j) \in I \times J$ tels que $x_i + y_j = z$ On aura donc

$$\mathbb{P}(Z = z) = \sum_{(i,j) \text{ tq } x_i + y_j = z} \mathbb{P}((X = x_i) \cap (Y = y_j)).$$

Dans le cas où les variables X et Y sont indépendantes on aura donc aussi

$$\mathbb{P}(Z = z) = \sum_{(i,j) \text{ tq } x_i + y_j = z} \mathbb{P}(X = x_i)\,\mathbb{P}(Y = y_j).$$

Exemple

On réalise une suite d'expériences de type succès-échec, indépendantes, chacune avec la même probabilité de succès $p \in]0;1[$. On appelle X_1 la variable aléatoire : « rang du premier succès obtenu » et X_2 la variable aléatoire : « rang du second succès ».

Montrer que les variables aléatoires X_1 et $X_2 - X_1$ sont indépendantes.

Solution:

La loi de X_1 est la loi géométrique $\mathcal{G}(p)$.
On va déterminer la loi du couple (X_1, X_2) :

$$\forall i \in \mathbb{N}^*,\ \forall j \geqslant i+1,\ \mathbb{P}((X_1 = i) \cap (X_2 = j)) = q^{i-1}p \times q^{j-i-1}p = q^{j-2}p^2.$$

La variable aléatoire $Y = X_2 - X_1$ est à valeurs dans $\mathbb{N}^*$ et on a

$$\forall k \in \mathbb{N}^*,\ \mathbb{P}(Y = k) = \sum_{\substack{j-i=k \\ i \geqslant 1, j \geqslant 2}} \mathbb{P}((X_1 = i) \cap (X_2 = j)) = \sum_{i=1}^{+\infty} \mathbb{P}((X_i = i) \cap (X_2 = i+k))$$

$$= \sum_{i=1}^{+\infty} q^{i+k-2}p^2 = p^2 q^{k-1} \sum_{i=1}^{+\infty} q^{i-1} = \frac{p^2 q^{k-1}}{1-q} = pq^{k-1}.$$

Ainsi $X_2 - X_1$ suit aussi une loi géométrique de paramètre p (on pouvait s'en douter...).
Il est alors facile de vérifier que Y et X_1 sont indépendantes :

$$\mathbb{P}((X_1 = x) \cap (Y = y)) = \mathbb{P}((X_1 = x) \cap (X_2 - X_1 = y)) = \mathbb{P}((X_1 = x) \cap (X_2 = x + y)) = q^{x+y-2}p^2$$

et

$$\mathbb{P}(X_1 = x)\mathbb{P}(Y = y) = q^{x-1}p \times q^{y-1}p$$

donc on a bien $\mathbb{P}((X_1 = x) \cap (Y = y)) = \mathbb{P}(X_1 = x)\mathbb{P}(Y = y)$.

Dans ce qui suit, on étudie les sommes de variables aléatoires indépendantes suivant les lois usuelles.

Théorème 7:

Si X et Y sont deux variables aléatoires **indépendantes** suivant des lois binomiales de paramètres respectifs (n, p) et (m, p) alors $X + Y$ suit une loi binomiale de paramètre $(n + m, p)$.

Démonstration:

Puisque $X(\Omega) = [\![0;n]\!]$ et $Y(\Omega) = [\![0;m]\!]$, on a $(X+Y)(\Omega) = [\![0;n+m]\!]$. Et pour tout $k \in [\![0;n+m]\!]$:

$$\mathbb{P}(X+Y=k) = \sum_{i+j=k} \mathbb{P}((X=i) \cap (Y=j)) = \sum_{i+j=k} \mathbb{P}(X=i)\mathbb{P}(Y=j) = \sum_{i+j=k} \binom{n}{i} p^i q^{n-i} \times \binom{m}{j} p^j q^{m-j}$$

$$= p^k q^{n+m-k} \sum_{i+j=k} \binom{n}{i}\binom{m}{j} = p^k q^{n+m-k} \binom{n+m}{k}$$

en utilisant la formule de Vandermonde (rappels du chapitre précédent).
Cela démontre que $X + Y$ suit la loi $\mathcal{B}(n+m, p)$.

Rem: Ce résultat est conforme à l'intuition : tirer n boules avec remise dans une urne puis en tirer de nouveau m revient à en tirer $n + m$...

Corollaire 7.1:

Si $X_1, \cdots, X_k$ sont k variables aléatoires **mutuellement indépendantes** suivant des lois binomiales de paramètres respectifs $(n_1, p), \cdots, (n_k, p)$ alors la variable aléatoire $X_1 + \cdots + X_k$ suit la loi binomiale de paramètre $\left(\sum_{i=1}^{k} n_i, p\right)$.

☞ *Démonstration:*

banale récurrence sur k en remarquant que, si $X_1, \ldots, X_k$ sont mutuellement indépendantes, alors X_k et $X_1 + \cdots X_{k-1}$ sont indépendantes en vertu du lemme des coalitions (théorème 6).

Corollaire 7.2:

Si $X_1, \cdots, X_n$ sont n variables aléatoires **mutuellement indépendantes** suivant la loi de Bernoulli de paramètre p, la variable aléatoire $X = X_1 + X_2 + \cdots + X_n$ suit la loi binomiale $\mathcal{B}(n, p)$.

☞ *Démonstration:*

On applique le corollaire précédent puisque la loi de Bernoulli de paramètre p n'est rien d'autre que la loi $\mathcal{B}(1, p)$.

Rem: On peut aussi dire que, puisque chaque X_i vaut 0 ou 1 selon échec/succès à chaque épreuve, la somme $X_1 + \cdots + X_n$ représente tout simplement le nombre de succès dans n épreuves de Bernoulli indépendantes.

Théorème 8:

Si X et Y sont deux variables aléatoires **indépendantes** suivant des lois de Poisson de paramètres respectifs λ et μ alors $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$.

☞ *Démonstration:*

Soit $Z = X + Y$. On a tout d'abord $Z(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(Z = k) &= \sum_{(i,j)\in\mathbb{N}^2/i+j=k} \mathbb{P}((X = i) \cap (Y = j)) = \sum_{i=0}^{k} \mathbb{P}((X = i) \cap (Y = k - i) \\
&= \sum_{i=0}^{k} \mathbb{P}(X = i)\mathbb{P}(Y = k - i) \qquad \text{car } X \text{ et } Y \text{ sont indépendantes} \\
&= \sum_{i=0}^{k} \frac{\mathrm{e}^{-\lambda}\lambda^i}{i!}\frac{\mathrm{e}^{-\mu}\mu^{k-i}}{(k-i)!} = \mathrm{e}^{-(\lambda+\mu)} \sum_{i=0}^{k} \frac{\lambda^i\mu^{k-i}}{i!(k-i)!} \\
&= \frac{\mathrm{e}^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!}\lambda^i\mu^{k-i} = \frac{\mathrm{e}^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \binom{k}{i}\lambda^i\mu^{k-i} = \frac{\mathrm{e}^{-(\lambda+\mu)}(\lambda+\mu)^k}{k!}.
\end{aligned}$$

Corollaire 8.1:

Si $X_1, \cdots, X_n$ sont n variables aléatoires **mutuellement indépendantes** suivant des lois de Poisson de paramètres respectifs $\lambda_1, \ldots, \lambda_n$, alors la variable aléatoire $X_1 + \cdots + X_n$ suit la loi de Poisson de paramètre $\sum_{i=1}^{n} \lambda_i$.

IV. Moments d'une variable aléatoire discrète

Les variables aléatoires considérées dans ce paragraphe sont à valeurs réelles ou complexes.

IV.1. Espérance

Déf 14:

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

- si $X(\Omega)$ est fini, $X(\Omega) = \{x_1, \dots, x_n\}$, alors **l'espérance** de X est

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i).$$

- si $X(\Omega)$ est infini , $X(\Omega) = \{x_i,\ i \in I\}$ avec I dénombrable, on dira que X **admet une espérance** ou **est d'espérance finie** si la famille $\left(x\mathbb{P}(X = x)\right)_{x \in X(\Omega)}$ est sommable. Dans ce cas **l'espérance** de X est :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) = \sum_{i \in I} x_i \mathbb{P}(X = x_i)$$

(par convention $x\mathbb{P}(X = x) = 0$ lorsque $x = +\infty$ et $\mathbb{P}(X = x) = 0$).

Remarques

- Une variable aléatoire discrète finie a toujours une espérance ce qui n'est pas le cas pour une variable aléatoire discrète infinie.
- Lorsque $X(\Omega) = (x_n)_{n \in \mathbb{N}}$, dire que X est d'espérance finie équivaut à dire que la série de terme général $x_n\mathbb{P}(X = x_n)$ est *absolument* convergente et alors

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} x_n \mathbb{P}(X = x_n).$$

- L'absolue convergence est nécessaire car l'ordre dans lequel les termes sont additionnés ne doit pas intervenir (en effet, lors d'une expérience aléatoire, les résultats peuvent apparaître dans n'importe quel ordre), cette condition est liée à l'absolue convergence (voir chapitre précédent).

Exemple

Soit X une variable aléatoire prenant les valeurs $n \in \mathbb{N}^*$ avec les probabilités :

$$\mathbb{P}(X = n) = \frac{1}{n(n+1)}\cdot$$

Cette définition a bien un sens puisque $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1$.

Cette variable aléatoire n'a pas d'espérance finie.

Pour toute la suite, on se limitera au cas où l'ensemble $X(\Omega)$ est infini (le cas fini étant plus facile et ayant été étudié en 1ère année). On pourra donc écrire $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$.

Le théorème qui suit facilite dans certains cas le calcul de l'espérance.

Théorème 9:

Soit X une variable aléatoire *à valeurs dans* $\mathbb{N}$ (ou dans $\mathbb{N} \cup \{+\infty\}$).
X admet une espérance finie si et seulement si la série de terme général $\mathbb{P}(X \geqslant k)$ converge et, dans ce cas :

$$\mathbb{E}(X) = \sum_{k=1}^{+\infty} \mathbb{P}(X \geqslant k) = \sum_{k=0}^{+\infty} \mathbb{P}(X > k).$$

☞ *Démonstration:*

La démonstration repose sur le calcul ci-dessous (c'est une transformation d'Abel..) :

$$\sum_{k=1}^{N} k\mathbb{P}(X = k) = \sum_{k=1}^{N} k\left[\mathbb{P}(X \geqslant k) - \mathbb{P}(X \geqslant k+1)\right] = \sum_{k=1}^{N} k\mathbb{P}(X \geqslant k) - \sum_{\cancel{k=0}}^{N} k\mathbb{P}(X \geqslant k+1)$$

$$= \sum_{k=1}^{N} k\mathbb{P}(X \geqslant k) - \sum_{k=1}^{N+1} (k-1)\mathbb{P}(X \geqslant k) = \sum_{k=1}^{N+1} \mathbb{P}(X \geqslant k) - (N+1)\mathbb{P}(X \geqslant N+1).$$

- Supposons que la série de terme général $\mathbb{P}(X \geqslant k)$ converge. D'après le calcul ci-dessus, on a, pour tout entier $N \in \mathbb{N}^*$:

$$\sum_{k=1}^{N} k\mathbb{P}(X = k) \leqslant \sum_{k=1}^{N+1} \mathbb{P}(X \geqslant k) \leqslant \sum_{k=1}^{+\infty} \mathbb{P}(X \geqslant k)$$

donc la série à termes positifs $\sum k\mathbb{P}(X=k)$ converge, et $\mathbb{E}(X)$ existe.

- Supposons que $\mathbb{E}(X)$ existe. Alors pour tout entier N :

$$(N+1)\mathbb{P}(X\geqslant N+1)=(N+1)\sum_{k=N+1}^{+\infty}\mathbb{P}(X=k)\leqslant\sum_{k=N+1}^{+\infty}kP(X=k)$$

donc $\lim\limits_{N\to+\infty}(N+1)\mathbb{P}(X\geqslant N+1)=0$.
D'après le calcul préliminaire, on en déduit

$$\lim_{N\to+\infty}\left(\sum_{k=1}^{N+1}\mathbb{P}(X\geqslant k)\right)=\lim_{N\to+\infty}\left(\sum_{k=1}^{N}k\mathbb{P}(X=k)\right)=\mathbb{E}(X),$$

ce qui prouve que la série $\sum\mathbb{P}(X\geqslant k)$ converge et donne l'égalité annoncée.

- Enfin, la relation $\sum\limits_{k=1}^{+\infty}\mathbb{P}(X\geqslant k)=\sum\limits_{k=0}^{+\infty}\mathbb{P}(X>k)$ est immédiate à l'aide d'un changement d'indice, et puisque, X ne prenant que des valeurs entières, on a $(X>k)=(X\geqslant k+1)$.

Le théorème qui suit est un des plus importants :

Théorème 10: de transfert

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans un ensemble E. Notons $X(\Omega)=\{x_i,\ i\in I\}$ (I au plus dénombrable).

Soit f une application de $X(\Omega)$ dans $\mathbb{C}$. La variable aléatoire $f(X)$ admet une espérance si et seulement si la famille $\big(f(x)\mathbb{P}(X=x)\big)_{x\in X(\Omega)}$ est sommable, et dans ce cas :

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

L'intérêt de ce théorème est de pouvoir calculer l'espérance de $f(X)$ sans avoir besoin de calculer sa loi, seulement en connaissant la loi de X.

☞ *Démonstration:*

Notons $f(X)(\Omega)=\{f(x_i),\ i\in I\}=\{y_j,\ j\in J\}$.
Supposons que $f(X)$ admet une espérance finie. Alors

$$\mathbb{E}(f(X))=\sum_{j\in J}y_j\mathbb{P}(f(X)=y_j),$$

la famille $\big(y_j\mathbb{P}(f(X)=y_j)\big)_{j\in J}$ étant sommable.
Soit $I_j=\{i\in I\text{ tq }f(x_i)=y_j\}$. La famille $(I_j)_{j\in J}$ est une partition de I et on a $(f(X)=y_j)=\bigcup\limits_{i\in I_j}(X=x_i)$, cette réunion étant disjointe. Donc $\mathbb{P}(f(X)=y_j)=\sum\limits_{i\in I_j}\mathbb{P}(X=x_i)$.
En remplaçant dans la formule précédente :

$$\mathbb{E}(f(X))=\sum_{j\in J}y_j\left(\sum_{i\in I_j}\mathbb{P}(X=x_i)\right)=\sum_{j\in J}\left(\sum_{i\in I_j}f(x_i)\mathbb{P}(X=x_i)\right).$$

Par le théorème sommation par paquets (possible car la famille de départ est sommable), puisque $\bigsqcup\limits_{j\in J}I_j=I$ on obtient

$$\mathbb{E}(f(X))=\sum_{i\in I}f(x_i)\mathbb{P}(X=x_i).$$

L'implication réciproque se traite de la même façon (dans le théorème d'associativité généralisé, il y a une équivalence).

Cas particulier important :
Ce théorème s'applique aussi aux couples (et plus généralement aux n-uplets) de variables aléatoires.
Soit $(\Omega,\mathcal{A},\mathcal{P})$, un espace probabilisé et (X,Y) un couple de variables aléatoires discrètes sur Ω. Notons $X(\Omega)=\{x_i,\ i\in I\}$ et $Y(\Omega)=\{y_j,\ j\in J\}$ où I,J sont deux parties de $\mathbb{N}$.
Soit f une application de $X(\Omega)\times Y(\Omega)$ dans $\mathbb{C}$, et Z la variable aléatoire $Z=f(X,Y)$. Alors Z est d'espérance finie si et seulement si la famille des $\Big(f(x_i,y_j)\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big)\Big)_{(i,j)\in I\times J}$ est sommable, et alors :

$$\mathbb{E}(Z)=\sum_{(i,j)\in I\times J}f(x_i,y_j)\mathbb{P}\big((X=x_i)\cap(Y=y_j)\big).$$

En particulier, sous réserve qu'elle existe, on aura, avec les mêmes notations :

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

Théorème 11: Linéarité de l'espérance

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, à valeurs dans $\mathbb{C}$, et soit $\lambda \in \mathbb{C}$.
Si les espérances de X et de Y existent, alors l'espérance de $\lambda X + Y$ existe et

$$\mathbb{E}(\lambda X + Y) = \lambda \mathbb{E}(X) + \mathbb{E}(Y).$$

☞ *Démonstration:*

Notons $X(\Omega) = \{x_i,\ i \in I\}$ et $Y(\Omega) = \{y_j,\ j \in J\}$ où I, J sont deux parties de $\mathbb{N}$.
Les évènements $(Y = y_j)$ pour $j \in J$ formant un système complet, on a :

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

puis en utilisant le théorème de Fubini (la famille définissant $\mathbb{E}(X)$ étant sommable) :

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

De même :

$$\mathbb{E}(Y) = \sum_{(i,j) \in I \times J} y_j \mathbb{P}\big((X = x)_i \cap (Y = y_j)\big),$$

donc, par addition de deux familles sommables :

$$\lambda \mathbb{E}(X) + \mathbb{E}(Y) = \sum_{(i,j) \in I \times J} (\lambda x_i + y_j) \mathbb{P}\big((X = x_i) \cap (Y = y_j)\big).$$

En considérant alors l'application f définie sur $X(\Omega) \times Y(\Omega)$ par $f : (x, y) \mapsto \lambda x + y$, le théorème de transfert donne directement que $\lambda X + Y$ possède une espérance et que

$$\mathbb{E}(\lambda X + Y) = \sum_{(i,j) \in I \times J} (\lambda x_i + y_j) \mathbb{P}\big((X = x_i) \cap (Y = y_j)\big) = \lambda \mathbb{E}(X) + \mathbb{E}(Y).$$

Déf 15:

Soit X une variable aléatoire discrète dont l'espérance existe. Puisque $\mathbb{E}(1) = 1$ il résulte du théorème précédent que la variable aléatoire $X - \mathbb{E}(X)$ est d'espérance nulle. On l'appelle **variable aléatoire centrée associée à X.**

Prop 10: Positivité de l'espérance

Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Si $X \geqslant 0$ (c'est-à-dire si $X(\omega) \geqslant 0$ pour tout $\omega \in \Omega$) et si son espérance existe, alors $\mathbb{E}(X) \geqslant 0$.

De plus, si $X \geqslant 0$ et si $\mathbb{E}(X) = 0$, alors $\mathbb{P}(X = 0) = 1$ (on dit que X est **quasi certainement** nulle ou **presque sûrement** nulle).

☞ *Démonstration:*

C'est quasi immédiat :

- En notant $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$:
$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} x_n \mathbb{P}(X = x_n).$$
donc si $X \geqslant 0$ alors les x_n sont $\geqslant 0$ d'où $\mathbb{E}(X) \geqslant 0$.
- Et si de plus $\mathbb{E}(X) = 0$ alors pour tout n on a $x_n \mathbb{P}(X = x_n) = 0$ donc si $x_n \neq 0$ on a $\mathbb{P}(X = x_n) = 0$; et puisque $\sum_{n=0}^{+\infty} \mathbb{P}(X = x_n) = 1$, il existe un entier n_0 tel que $x_{n_0} = 0$ et alors $\mathbb{P}(X = x_{n_0}) = 1$.

Prop 11: Croissance de l'espérance

Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, dont l'espérance existe.

Si $X \leqslant Y$ sur Ω, alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

☞ *Démonstration:*

On utilise la positivité de l'espérance, appliqué à $Y-X$, et la linéarité de l'espérance.

Prop 12:

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, X à valeurs dans $\mathbb{C}$ et Y à valeurs dans $\mathbb{R}$.
On suppose que l'espérance de Y existe, et que $|X|\leqslant Y$ sur Ω.
Alors l'espérance de X existe.

☞ *Démonstration:*

On reprend toutes les notations du th. de linéarité.
Pour tout $i\in\mathbb{N}$ on a

$$|x_i|\,\mathbb{P}(X=x_i)=|x_i|\sum_{j\in J}\mathbb{P}((X=x_i)\cap(Y=y_j))=\sum_{j\in J}|x_i|\,\mathbb{P}((X=x_i)\cap(Y=y_j)).$$

Or $|x_i|\,\mathbb{P}((X=x_i)\cap(Y=y_j))\leqslant y_j\mathbb{P}((X=x_i)\cap(Y=y_j))$! En effet, soit la probabilité de l'évènement $(X=x_i)\cap(Y=y_j)$ est nulle, auquel cas c'est vérifié, soit il existe un ω tel que $X(\omega)=x_i$ et $Y(\omega)=y_j$ et puisque $|X(\omega)|\leqslant Y(\omega)$ on a encore l'inégalité.
Ainsi $|x_i|\,\mathbb{P}(X=x_i)\leqslant\sum_{j\in J}y_j\mathbb{P}((X=x_i)\cap(Y=y_j))$.

De plus, $\mathbb{E}(Y)=\sum_{j\in J}y_j\mathbb{P}(Y=y_j)=\sum_{j\in J}y_j\left(\sum_{i\in I}\mathbb{P}((X=x_i)\cap(Y=y_j))\right)$, (formule des probabilités totales) , et la suite double $\left(y_j\mathbb{P}((X=x_i)\cap(Y=y_j))\right)_{(i,j)\in\mathbb{N}^2}$ est donc sommable puisque $\mathbb{E}(Y)$ existe.

Par Fubini, $\sum_{i\in I}|x_i|\,\mathbb{P}(X=x_i)\leqslant\sum_{i\in I}\sum_{j\in J}y_j\mathbb{P}((X=x_i)\cap(Y=y_j))=\mathbb{E}(Y)<+\infty$, ce qui permet alors de conclure que l'espérance de X existe.

***Corollaire 11.1*:**

Si X est une variable aléatoire discrète bornée, son espérance existe.

Théorème 12: Produit de deux variables aléatoires indépendantes

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, dont l'espérance existe.
Si X et Y sont indépendantes, la variable aléatoire $Z=XY$ possède une espérance finie et

$$\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y).$$

☞ *Démonstration:*

Avec les notations désormais habituelles, puisque X et Y admettent une espérance, les séries $\sum_i x_i\mathbb{P}(X=x_i)$ et $\sum_j y_j\mathbb{P}(Y=y_j)$ sont absolument convergentes. Les théorèmes sur les familles sommables permettent alors d'écrire :

$$\begin{aligned}\mathbb{E}(X)\mathbb{E}(Y)&=\left(\sum_i x_i\mathbb{P}(X=x_i)\right)\left(\sum_j y_j\mathbb{P}(Y=y_j)\right)\\&=\sum_{(i,j)\in\mathbb{N}^2}x_iy_j\mathbb{P}(X=x_i)\mathbb{P}(Y=y_j)\qquad\text{famille double sommable}\\&=\sum_{(i,j)\in\mathbb{N}^2}x_iy_j\mathbb{P}((X=x_i)\cap(Y=y_j))\qquad\text{indépendance}\\&=\mathbb{E}(XY)\qquad\text{théorème de transfert.}\end{aligned}$$

Rem: Ce résultat s'étend bien sûr au cas de n variables aléatoires discrètes indépendantes.

Théorème 13: Inégalité de Markov

Soit X une variable aléatoire discrète réelle positive, dont l'espérance existe.
Alors, pour tout réel $a>0$:

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

☞ *Démonstration:*

Soit $X(\Omega)=\{x_n,\ n\in\mathbb{N}\}$ avec ici les x_n positifs ou nuls. On a $\mathbb{E}(X)=\sum_{n=0}^{+\infty} x_n\mathbb{P}(X=x_n)$.
Notons I_1 l'ensemble des entiers n tels que x_n est supérieur ou égal à a, et I_2 l'ensemble des indices n tels que $x_n<a$.
Donc, la série étant absolument convergente, par sommation par paquets :

$$\mathbb{E}(X)=\sum_{n\in I_1} x_n\mathbb{P}(X=x_n)+\underbrace{\sum_{n\in I_2} x_n\mathbb{P}(X=x_n)}_{\geqslant 0}\geqslant \sum_{n\in I_1} x_n\mathbb{P}(X=x_n)\geqslant a\sum_{n\in I_1}\mathbb{P}(X=x_n)$$

Mais $\sum_{n\in I_1}\mathbb{P}(X=x_n)=\mathbb{P}(X\geqslant a)$ puisque $(X\geqslant a)$ est la réunion disjointe des évènements $(X=x_n)$ avec $x_n\geqslant a$, d'où l'inégalité cherchée.

IV.2. Variance d'une variable aléatoire réelle

Les variables aléatoires de ce paragraphe sont toutes à valeurs réelles.

Déf 16:

Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et X une variable aléatoire discrète à valeurs dans $\mathbb{R}$, et soit $r\geqslant 1$ un entier.

- si $X(\Omega)$ est fini, $X(\Omega)=\{x_1,\dots,x_n\}$, alors le **moment d'ordre r de X** est

$$m_r(X)=\mathbb{E}(X^r)=\sum_{i=1}^{n} x_i^r\mathbb{P}(X=x_i).$$

- si $X(\Omega)$ est infini , $X(\Omega)=\{x_i,\ i\in I\}$ avec I dénombrable, on dira que X admet un **moment d'ordre r** si la variable aléatoire X^r admet une espérance, c'est-à-dire si la famille $\left(x^r\mathbb{P}(X=x)\right)_{x\in X(\Omega)}$ est sommable. Dans ce cas le moment d'ordre r de X est

$$m_r(X)=\mathbb{E}(X^r)=\sum_{x\in X(\Omega)} x^r\mathbb{P}(X=x)=\sum_{i\in I} x_i^r\mathbb{P}(X=x_i).$$

Remarques

- Une variable aléatoire discrète finie admet des moments de tous ordres, ce qui n'est pas forcément le cas pour une variable aléatoire discrète infinie.
- Lorsque $X(\Omega)=(x_n)_{n\in\mathbb{N}}$, dire que X est d'espérance finie équivaut à dire que la série de terme général $x_n^r\mathbb{P}(X=x_n)$ est *absolument* convergente et alors

$$\mathbb{E}(X)=\sum_{n=0}^{+\infty} x_n^r\mathbb{P}(X=x_n).$$

Pour toute la suite, on se limitera au cas où l'ensemble $X(\Omega)$ est infini. On pourra donc écrire $X(\Omega)=\{x_n,\ n\in\mathbb{N}\}$.

Théorème 14:

Si X admet un moment d'ordre $r\geqslant 2$, elle admet des moments d'ordre s pour tout $s\in[\![1\,;r]\!]$.

☞ *Démonstration:*

Pour tout x réel on a $|x|^{r-1}\leqslant 1+|x|^r$ (il suffit de distinguer les cas $|x|\leqslant 1$ et $|x|\geqslant 1$). Pour tout n notons $p_n=\mathbb{P}(X=x_n)$. Si $m_r(X)$ existe la série $\sum_{n\in\mathbb{N}} p_n|x_n|^r$ est convergente donc aussi $\sum_{n\in\mathbb{N}} p_n(1+|x_n|^r)$ puisque $\sum_{n=0}^{+\infty} p_n=1$. Par comparaison de séries à termes positifs, la série de terme général $p_n x_n^{r-1}$ sera donc aussi absolument convergente, c'est-à-dire que $m_{r-1}(X)$ existe.

Théorème 15: (et définition 17)

Soit X une variable aléatoire réelle dont le moment d'ordre 2 (c'est-à-dire $\mathbb{E}(X^2)$) existe. Alors

1. La variable aléatoire $X - \mathbb{E}(X)$ admet un moment d'ordre 2, appelé **variance de** X. Ainsi
$$\mathbb{V}(X) = E\left(\left(X - \mathbb{E}(X)\right)^2\right).$$
2. On a aussi
$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \qquad \text{(formule de Koenig-Huygens)}$$

On appelle alors **écart type** de X le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

☞ *Démonstration:*

Puisque X admet un moment d'ordre 2 elle admet aussi une espérance d'après le théorème précédent. Puisque :
$$\left(X - \mathbb{E}(X)\right)^2 = X^2 - 2\mathbb{E}(X).X + \mathbb{E}(X)^2,$$
la variable aléatoire $\left(X - \mathbb{E}(X)\right)^2$ admet aussi une espérance et par linéarité de l'espérance :
$$\mathbb{E}\left(\left(X - \mathbb{E}(X)\right)^2\right) = \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Prop 13:

Soit X une variable aléatoire réelle dont la variance $\mathbb{V}(X)$ existe.
Si a et b sont deux réels, la variable $aX + b$ admet aussi une variance et
$$\mathbb{V}(aX + b) = a^2\mathbb{V}(X).$$

☞ *Démonstration:*

Déjà, puisque $\mathbb{V}(X)$ existe il en est de même de $\mathbb{E}(X)$, par le théorème 14
On a : $(aX + b)^2 = a^2X^2 + 2abX + b^2$ donc par linéarité $\mathbb{E}\left((aX + b)^2\right)$ existe puis
$$\begin{aligned}\mathbb{V}(aX + b) &= \mathbb{E}\left((aX + b)^2\right) - \left(\mathbb{E}(aX + b)\right)^2 = \mathbb{E}\left((aX + b)^2\right) - \left(a\mathbb{E}(X) + b\right)^2 \\ &= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - \left(a^2\mathbb{E}(X)^2 + b^2 + 2ab\mathbb{E}(X)\right) = a^2\mathbb{V}(X).\end{aligned}$$

Déf 18:

Une variable aléatoire X qui admet un moment d'ordre 2 est dite **réduite** si $\sigma(X) = 1$.

Prop 14:

Si X est une variable aléatoire admettant un moment d'ordre 2 et si $\sigma(X) \neq 0$, alors la variable aléatoire $\dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée réduite.

☞ *Démonstration:*

Trivial.

Théorème 16: Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Alors, pour tout réel $\varepsilon > 0$
$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

☞ *Démonstration:*

On applique l'inégalité de Markov à la variable aléatoire réelle positive $(X - \mathbb{E}(X))^2$:
$$\mathbb{P}((X - \mathbb{E}(X))^2 \geqslant \varepsilon^2) \leqslant \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{\varepsilon^2}$$

Rem: Cette inégalité montre que la variance permet de mesurer la dispersion de la variable X autour de sa moyenne.

Exemple

Combien de lancers d'une pièce équilibrée suffit-il d'effectuer pour pouvoir affirmer avec un risque d'erreur inférieur à 5% que la fréquence d'apparition de pile différera de $\frac{1}{2}$ d'au plus 10^{-2} ?

Solution:

Soit n le nombre de lancers et F_n la fréquence de piles obtenus. On cherche n tel que

$$\mathbb{P}\left(\left|F_n - \frac{1}{2}\right| > 0,01\right) \leqslant 0,05\,.$$

Si on note X_n le nombre de pile obtenus, $F_n = \frac{1}{n}X_n$, et $X_n \hookrightarrow \mathcal{B}(n, \frac{1}{2})$. Donc $\mathbb{E}(X_n) = np = \frac{n}{2}$ et $\mathbb{V}(X_n) = npq = \frac{n}{4}$, de sorte que $\mathbb{E}(F_n) = \frac{1}{2}$ et $\mathbb{V}(F_n) = \frac{1}{4n}$.
L'inégalité de BT s'écrit donc ici

$$\mathbb{P}\left(\left|F_n - \frac{1}{2}\right| \geqslant \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}\,.$$

Pour $\varepsilon = 10^{-2}$, il suffit donc d'avoir $\frac{10^4}{4n} \leqslant 5.10^{-2}$ soit $n \geqslant 50000$.

***Corollaire 16.1*:**

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Notons $m = \mathbb{E}(X)$.
Si $\mathbb{V}(X) = 0$, alors $\mathbb{P}(X = m) = 1$ (X est **quasi certaine**).

Démonstration:

L'événement $(X \neq m)$ est la réunion des évènements $A_n = \left(|X - m| > \frac{1}{n}\right)$. Or les ensembles A_n forment une suite croissante d'événements; par conséquent, $\mathbb{P}(X \neq m) = \lim\limits_{n\to+\infty} \mathbb{P}(A_n)$. Or puisque $\mathbb{V}(X) = 0$, $P\left(|X - m| > \frac{1}{n}\right) = 0$ d'après l'inégalité de BT d'où le résultat.
Remarque : il est plus rapide de remarquer que, la variable $(X - \mathbb{E}(X))^2$ étant positive d'espérance nulle, elle est presque sûrement nulle en vertu de la proposition 10, mais la démonstration précédente est instructive....

IV.3. Moments des lois usuelles

Théorème 17: Loi uniforme

Soit $X \hookrightarrow \mathcal{U}([\![1\,;n]\!])$. Alors :

$$\mathbb{E}(X) = \frac{n+1}{2} \quad \text{et} \quad \mathbb{V}(X) = \frac{n^2-1}{12}\,.$$

Démonstration:

C'est immédiat en utilisant :

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}\,.$$

Théorème 18: Loi binomiale

Soit $X \hookrightarrow \mathcal{B}(n,p)$. Alors

$$\mathbb{E}(X) = np \quad \text{et} \quad \mathbb{V}(X) = npq\,.$$

Démonstration:

$\mathbb{E}(X) = \sum\limits_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}$. Or $k\binom{n}{k} = n\binom{n-1}{k-1}$ donc

$$\mathbb{E}(X) = np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{(n-1)-(k-1)} = n\mathbb{P}(p+(1-p))^{n-1} = np\,.$$

On calcule de même, à l'aide de la formule de transfert :

$$\mathbb{E}(X(X-1)) = \sum_{\substack{k=0\\k=2}}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} = n(n-1)p^2\sum_{k=2}^{n}\binom{n-2}{k-2}p^{k-2}(1-p)^{(n-2)-(k-2)} = n(n-1)p^2$$

d'où
$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X^2) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)\,.$$

Théorème 19: Loi géométrique

Soit $X \hookrightarrow \mathcal{G}(p)$. Alors
$$\mathbb{E}(X) = \frac{1}{p} \quad \text{et} \quad \mathbb{V}(X) = \frac{q}{p^2}\,.$$

☞ *Démonstration:*

$\mathbb{E}(X) = \sum_{n=1}^{+\infty} npq^{n-1}$ (sous réserve de convergence, ce qui est fait ci-après).

Or pour tout $x \in]-1;1[$, $\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$ d'où en dérivant cette série entière (c'est licite, cf. cours correspondant), $\sum_{n=1}^{+\infty} nx^{n-1} = \frac{1}{(1-x)^2}$. On en tire
$$\mathbb{E}(X) = p\frac{1}{(1-q)^2} = \frac{1}{p}\,.$$
Puis $\mathbb{E}(X(X-1)) = \sum_{n=1}^{+\infty} n(n-1)pq^{n-1}$. En dérivant la série entière précédente on a , pour tout $x \in]-1;1[$, $\sum_{n=2}^{+\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$.
On en déduit $\mathbb{E}(X(X-1)) = pq\frac{2}{(1-q)^3} = \frac{2q}{p^2}$ puis $\mathbb{V}(X) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}$.

Théorème 20: Loi de Poisson

Soit $X \hookrightarrow \mathcal{P}(\lambda)$. Alors
$$\mathbb{E}(X) = \lambda \quad \text{et} \quad \mathbb{V}(X) = \lambda\,.$$

☞ *Démonstration:*

$\mathbb{E}(X) = \sum_{n=0}^{+\infty} n\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda}\lambda\sum_{n=1}^{+\infty}\frac{\lambda^{n-1}}{(n-1)!} = \lambda\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = \lambda$. etc...

IV.4. Covariance

Théorème 21:

Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, ayant chacune un moment d'ordre 2.
Alors l'espérance de XY existe et
$$\mathbb{E}(XY)^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2) \qquad \text{(inégalité de Cauchy-Schwarz)}\,.$$
De plus, il y a égalité si et seulement si il existe des réels a et b tels que la variable aléatoire $aX + bY$ est presque surement nulle (c'est-à-dire $\mathbb{P}(aX + bY = 0) = 1$).

☞ *Démonstration:*

• On connaît l'inégalité :
$$\forall\,(x,y) \in \mathbb{R}^2,\ |xy| \leqslant \frac{1}{2}(x^2+y^2)\,.$$
Donc :
$$|XY| \leqslant \frac{1}{2}(X^2+Y^2)\,.$$
Puisque l'espérance de X^2+Y^2 existe, il en est de même de $\mathbb{E}(XY)$ d'après la proposition 12.

• Pour tout $\lambda \in \mathbb{R}$, la variable aléatoire $Z = (\lambda X + Y)^2$ est à valeurs positives. Par combinaison linéaire, son espérance existe et
$$\mathbb{E}(Z) = \lambda^2\mathbb{E}(X^2) + 2\lambda\mathbb{E}(XY) + \mathbb{E}(Y^2).$$
Par positivité de l'espérance on a donc :
$$\forall\,\lambda \in \mathbb{R},\ \lambda^2\mathbb{E}(X^2) + 2\lambda\mathbb{E}(XY) + \mathbb{E}(Y^2) \geqslant 0.$$
- si $\mathbb{E}(X^2) = 0$ alors pour que le signe ne change pas, on doit avoir $\mathbb{E}(XY) = 0$ et l'inégalité est vérifiée.
- sinon, on a un trinôme qui est positif pour tout λ, donc son discriminant est négatif ce qui donne le résultat.

• Il y a égalité si et seulement si $\mathbb{E}(X^2) = 0$ (1er cas), c'est-à-dire si X est presque surement nulle, ou si le discriminant est nul, c'est-à-dire si il existe λ tel que $\mathbb{E}((\lambda X + Y)^2) = 0$, c'est-à-dire $\lambda X + Y$ presque surement nulle.
En rassemblant les 2 cas, on obtient le résultat annoncé.

Corollaire 21.1:

Si X et Y sont deux variables aléatoires réelles admettant un moment d'ordre 2, alors la variable aléatoire $X+Y$ admet aussi un moment d'ordre 2.

☞ *Démonstration:*

découle de $(X+Y)^2 = X^2+Y^2+2XY$ et du théorème précédent.

Avec ces hypothèses, on aura :

$$\begin{aligned}\mathbb{V}(X+Y) &= \mathbb{E}((X+Y)^2) - (\mathbb{E}(X+Y))^2 \\ &= \mathbb{E}(X^2)+\mathbb{E}(Y^2)+2\mathbb{E}(XY)-\mathbb{E}(X)^2-\mathbb{E}(Y)^2-2\mathbb{E}(X)\mathbb{E}(Y) \qquad \text{(l'espérance est linéaire)} \\ &= \mathbb{V}(X)+\mathbb{V}(Y)+2\left(\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)\right).\end{aligned}$$

Cela justifie la définition suivante :

Déf 19:

Si X et Y sont deux variables aléatoires réelles admettant chacune un moment d'ordre 2, on appelle **covariance de X et Y** le réel

$$\text{cov}(X,Y) = \mathbb{E}\big[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\big].$$

Un calcul simple montre que l'on a aussi : $\text{cov}(X,Y) = \mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$.
On a donc la relation :

$$\mathbb{V}(X+Y) = \mathbb{V}(X)+\mathbb{V}(Y)+2\text{cov}(X,Y).$$

Les propriétés qui suivent se vérifient aisément, compte tenu de tout ce qui précède :

Prop 15:

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

1. L'ensemble $\mathcal{V}(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω est un sous-espace vectoriel de $\mathcal{A}(\Omega,\mathbb{R})$;
2. L'ensemble $\mathcal{V}^1(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω admettant une espérance est un sous-espace vectoriel de $\mathcal{V}(\Omega,\mathbb{R})$;
3. L'ensemble $\mathcal{V}^2(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω admettant un moment d'ordre 2 est un sous-espace vectoriel de $\mathcal{V}^1(\Omega,\mathbb{R})$;
4. L'application $(X,Y) \mapsto \text{cov}(X,Y)$ est une forme bilinéaire symétrique positive (mais non définie) sur l'espace vectoriel $\mathcal{V}^2(\Omega,\mathbb{R})$, et, pour tout $X \in \mathcal{V}^2(\Omega,\mathbb{R})$, $\text{cov}(X,X) = \mathbb{V}(X)$;

Prop 16: Inégalité de Cauchy-Schwarz

Si X et Y appartiennent à $\mathcal{V}^2(\Omega,\mathbb{R})$, on a l'inégalité :

$$|\text{cov}(X,Y)| \leqslant \sigma(X)\sigma(Y).$$

☞ *Démonstration:*

Soit λ un réel. On a $\mathbb{V}(X+\lambda Y) = \text{cov}(X+\lambda Y, X+\lambda Y) = \mathbb{V}(X)+\lambda^2\mathbb{V}(Y)+2\lambda\text{cov}(X,Y)$ par bilinéarité. On distingue deux cas :

- si $\mathbb{V}(Y) = 0$, pour que l'expression soit positive pour tout λ, il faut et il suffit que le coefficient de λ soit nul, donc $\text{cov}(X,Y) = 0$, et on a le résultat.
- sinon, on a un trinôme qui est positif pour tout λ, son discriminant est négatif ou nul ce qui donne

$$[\text{cov}(X,Y)]^2 - \mathbb{V}(X)\mathbb{V}(Y) \leqslant 0$$

et l'on obtient l' inégalité.

Prop 17:

Si X et Y sont deux variables aléatoires réelles admettant chacune un moment d'ordre 2 et **indépendantes**, alors leur covariance est nulle.
Il en résulte que l'on a alors $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

☞ *Démonstration:*

Cela résulte directement du théorème 12 ($\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$) et de la définition de la covariance.

Rem : La réciproque de cette propriété est fausse !
Par exemple, soit X une variable aléatoire dont la loi est donnée par $\mathbb{P}(X=-1) = \mathbb{P}(X=1) = 1/4$ et $\mathbb{P}(X=0) = 1/2$, et soit $Y = X^2$. Alors $\mathbb{E}(X) = 0$ et $\mathbb{E}(XY) = \mathbb{E}(X^3) = 0$ donc $\text{cov}(X,Y) = 0$.
Mais $\mathbb{P}((X=0)\cap(Y=1)) \neq \mathbb{P}(X=0)\mathbb{P}(Y=1)$ donc X et Y ne sont pas indépendantes.

Généralisation :
Si $X_1, \dots, X_n$ sont n variables aléatoires réelles possédant toutes un moment d'ordre 2, en utilisant la bilinéarité de la covariance on obtient :

$$\begin{aligned}\mathbb{V}(X_1 + \cdots + X_n) &= \text{cov}(X_1 + \cdots + X_n, X_1 + \cdots + X_n)\\ &= \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n) + 2\sum_{1\leqslant i<j\leqslant n} \text{cov}(X_i, X_j)\,.\end{aligned}$$

Il en résulte :

Théorème 22:

Si $X_1, \dots, X_n$ sont n variables aléatoires réelles possédant toutes un moment d'ordre 2 et **indépendantes deux à deux** :

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{V}(X_i)\,.$$

V. Loi faible des grands nombres

Étant donné une pièce équilibrée, il est admis intuitivement que lorsque cette pièce est lancée un certain nombre de fois, la fréquence des faces est proche de $1/2$.
La loi faible des grands nombres est la formulation mathématique de cette intuition. Les lancers successifs sont modélisés par une suite de variables aléatoires X_i où X_i prend la valeur 0 si au i-ème lancer on obtient pile et 1 si l'on obtient face. La fréquence de face sur n lancers est

$$\overline{X_n} = \frac{X_1 + \cdots + X_n}{n}\,.$$

La loi faible des grands nombres exprime que la variable $\overline{X_n}$ tend vers $1/2$ lorsque $n \to \infty$ (en un certain sens).

Théorème 23: Loi faible des grands nombres

Soit $(X_i)_{i\in N}$ une suite de variables aléatoires réelles définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes et de même loi (en abrégé : i.i.d), et de variance finie.

On pose $m = \mathbb{E}(X_1)$, et $\overline{X_n} = \dfrac{\sum_{i=1}^{n} X_i}{n}$.

Alors :

$$\forall\, \varepsilon > 0,\ \lim_{n\to\infty} \mathbb{P}\big(|\overline{X_n} - m| \geqslant \varepsilon\big) = 0.$$

☞ *Démonstration:*

Notons déjà que, puisque les X_i possèdent un moment d'ordre 2, elles possèdent aussi un moment d'ordre 1, donc $m = \mathbb{E}(X_1)$ est fini.
Les variables ayant même loi, on a $\mathbb{E}(X_i) = \mathbb{E}(X_1)$ et $\mathbb{V}(X_i) = \mathbb{V}(X_1)$ pour tout i.
Par linéarité de l'espérance, $\mathbb{E}(\overline{X_n}) = m$ et puisque les variables sont indépendantes, $\mathbb{V}(\overline{X_n}) = \dfrac{\mathbb{V}(X_1)}{n}$.

On applique alors l'inégalité de BT :

$$\mathbb{P}\left(\left|\overline{X_n} - \mathbb{E}(X_n)\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}(\overline{X_n})}{\varepsilon^2} = \frac{\mathbb{V}(\overline{X_1})}{n\varepsilon^2}$$

et le résultat en découle immédiatement.

VI. Variables aléatoires à valeurs dans $\mathbb{N}$

Déf 20:

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$.
La **fonction génératrice** de X est la série entière :

$$G_X(t) = \sum_{n=0}^{+\infty} \mathbb{P}(X=n)t^n .$$

Prop 18:

La fonction génératrice d'une variable aléatoire à valeurs entières a un rayon de convergence R_X supérieur ou égal à 1.

☞ *Démonstration:*

$G_X(1) = \sum_{n=0}^{+\infty} \mathbb{P}(X=n) = 1$ puisque $X(\Omega) \subset \mathbb{N}$, donc la série entière converge pour $t=1$. Puis cf. cours sur les séries entières.

Prop 19:

G_X est de classe $\mathscr{C}^\infty$ sur $]-R_X\,;R_X[$ et pour tout $n \in \mathbb{N}$: $\mathbb{P}(X=n) = \dfrac{G_X^{(n)}(0)}{n!}$.

☞ *Démonstration:*

cf cours sur les séries entières...

Prop 20:

Pour tout $t \in]-R_X\,;R_X[$, $G_X(t) = \mathbb{E}(t^X)$.

☞ *Démonstration:*

C'est le théorème de transfert.

Rem: Cette proposition montre que la loi d'une variable aléatoire à valeurs entières est entièrement caractérisée par sa fonction génératrice.

Prop 21:

1. Si X suit la loi binomiale de paramètres (n,p), sa fonction génératrice est la fonction polynomiale définie par
$$\forall\, t \in \mathbb{R},\ G_X(t) = (pt+q)^n .$$
2. Si X suit la loi géométrique de paramètre p, le rayon de convergence de sa fonction génératrice est égal à $\frac{1}{q}$ et
$$\forall\, t \in \left]-\tfrac{1}{q}\,;\tfrac{1}{q}\right[,\ G_X(t) = \frac{pt}{1-qt} .$$
3. Si X suit la loi de Poisson de paramètre λ, sa série génératrice est de rayon de convergence ∞ et
$$\forall\, t \in \mathbb{R},\ G_X(t) = \mathrm{e}^{\lambda(t-1)} .$$

☞ *Démonstration:*

1. Si $X \hookrightarrow \mathcal{B}(n,p)$ sa série génératrice est en fait la somme finie :
$$G_X(t) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} t^k = (pt+q)^n \quad \text{(formule du binôme).}$$
2. Si $X \hookrightarrow \mathcal{G}(p)$ on a lorsque la série converge :
$$G_X(t) = \sum_{n=1}^{+\infty} pq^{n-1}t^n = pt\sum_{n=1}^{+\infty}(qt)^{n-1}$$
et l'on reconnaît une série géométrique de raison qt.
3. Si $X \hookrightarrow \mathcal{P}(\lambda)$ on a lorsque la série converge :
$$G_X(t) = \mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!}t^n = \mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{(\lambda t)^n}{n!}$$
et on reconnaît le développement en série de $\exp(\lambda t)$.

Rem: Je cite un extrait du programme officiel :

Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.

Théorème 24: Fonction génératrice de la somme de deux v.a. indépendantes

Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$, G_X et G_Y leurs fonctions génératrices respectives, et R_X, R_Y leurs rayons de convergence.
Si X et Y sont indépendantes on a :
$$\text{pour tout } t \text{ tel que } |t| < \min(R_X, R_Y),\ G_{X+Y}(t) = G_X(t)G_Y(t).$$

☞ *Démonstration:*

X et Y étant indépendantes, il en est de même des variables aléatoires t^X et t^Y pour tout t tel que $|t| < \min(R_X, R_Y)$, en vertu du théorème 5.
Puis en vertu du théorème 12 on aura $\mathbb{E}(t^X)\mathbb{E}(t^Y) = \mathbb{E}(t^X t^Y) = \mathbb{E}(t^{X+Y})$ et il ne reste plus qu'à conclure en utilisant la proposition précédente.

Rem: Ce résultat se généralise sans difficulté au cas d'une somme finie de variables aléatoires entières indépendantes (utiliser le lemme des coalitions pour faire la récurrence).

Exemples

Retrouver la loi de la somme de deux variables indépendantes dans les cas de la loi binomiale et de la loi Poisson..

Théorème 25: Fonction génératrice et moments

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice. On suppose que $R_X > 1$, de sorte que G_X sera de classe $\mathscr{C}^\infty$ au voisinage de 1. Alors
$$\forall k \in \mathbb{N}^*,\ G_X^{(k)}(1) = E\big(X(X-1)\dots(X-k+1)\big).$$
En particulier : $\mathbb{E}(X) = G_X'(1)$, $\mathbb{E}(X(X-1)) = G_X''(1)$ donc $\mathbb{E}(X^2) = G_X'(1) + G_X''(1)$.

☞ *Démonstration:*

D'après le th. de dérivation d'une série entière,
$$\forall t \in]-R_X\,;R_X[,\ \forall k \in \mathbb{N},\ G_X^{(k)}(t) = \sum_{n=k}^{+\infty} n(n-1)\dots(n-k+1)\mathbb{P}(X=n)t^{n-k}$$
donc
$$G_X^{(k)}(1) = \sum_{n=k}^{+\infty} n(n-1)\dots(n-k+1)\mathbb{P}(X=n) = \mathbb{E}\big(X(X-1)\dots(X-k+1)\big)$$
en utilisant le théorème de transfert.

Ce théorème admet une réciproque :

Théorème 26:

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

L'espérance de X existe si et seulement si G_X est dérivable à gauche en 1 et la variance de X existe si et seulement si G_X est deux fois dérivable à gauche en 1.

Dans ce cas on aura encore : $\mathbb{E}(X) = G'_X(1)$ et $\mathbb{E}(X^2) = G'_X(1) + G''_X(1)$.

☞ *Démonstration:*

- Si X admet une espérance, la série $\sum_{n\in\mathbb{N}} n\mathbb{P}(X=n)$ est ACV donc la série entière $\sum_{n\in\mathbb{N}} n\mathbb{P}(X=n)t^n$ est normalement convergente, donc uniformément, sur $[-1;1]$. Le théorème de dérivation d'une série de fonctions donne alors le résultat.
- Si X admet une variance alors X et X^2 admettent une espérance donc les séries $\sum_{n\in\mathbb{N}} n\mathbb{P}(X=n)$ et $\sum_{n\in\mathbb{N}} n^2\mathbb{P}(X=n)$ sont ACV.
  La série de fonctions $\sum_{n\in\mathbb{N}^*} n(n-1)\mathbb{P}(X=n)t^n$ est alors normalement donc uniformément convergente sur $[-1;1]$, ce qui permet d'en déduire que G_X est deux fois dérivable à gauche en 1 et que;
  $$G''_X(1) = \sum_{n=1}^{+\infty} n(n-1)\mathbb{P}(X=n) = \sum_{n=1}^{+\infty} n^2\mathbb{P}(X=n) - \sum_{n=1}^{+\infty} n\mathbb{P}(X=n) = \mathbb{E}(X^2) - \mathbb{E}(X)$$
  puis, la relation cherchée.

  La démonstration de la réciproque est admise, mais en voici une partie, pour les curieux :
- Réciproquement, supposons G_X dérivable à gauche en 1, et soit $t \in [0;1[$ Le théorème des accroissements finis montre que :
  $$\exists\, c_t \in]t;1[\text{ tq } \frac{G_X(t)-G_X(1)}{t-1} = G'_X(c_t) = \sum_{n=1}^{+\infty} n\mathbb{P}(X=n){c_t}^{n-1}.$$
  Tous les termes de la série étant positifs on a donc :
  $$\forall N \in \mathbb{N},\ \sum_{n=1}^{N} n\mathbb{P}(X=n){c_t}^{n-1} \leqslant \frac{G_X(t)-G_X(1)}{t-1}$$
  donc en faisant tendre t vers 1 :
  $$\forall N \in \mathbb{N},\ \sum_{n=1}^{N} n\mathbb{P}(X=n) \leqslant G'_X(1).$$
  Cela implique que la série à termes positifs $\sum_{n\in\mathbb{N}} n\mathbb{P}(X=n)$ converge, donc $\mathbb{E}(X)$ existe et $\mathbb{E}(X) \leqslant G'_X(1)$.
  Mais en reprenant la première relation on a alors :
  $$\frac{G_X(t)-G_X(1)}{t-1} = \sum_{n=1}^{+\infty} n\mathbb{P}(X=n){c_t}^{n-1} \leqslant \sum_{n=1}^{+\infty} n\mathbb{P}(X=n)$$
  et en faisant tendre à nouveau t vers 1, on obtient l'autre inégalité.

Exemples

Retrouver les moments d'une variable aléatoire suivant une loi binomiale, géométrique et de Poisson..