Chapitre 16

Variables aléatoires discrètes

16.1 Variables aléatoires discrètes.

16.1.1 Préliminaires : rappel sur les images réciproques.

Définition 16.1.1 *Soit $X : E \to F$ une application.*
Pour $B \subset F$ (ou $B \in \mathcal{P}(F)$) l'image réciproque de B par X est par définition le sous ensemble de E noté $X^{-1}(B) = \{e \in E / X(e) \in B\}$

Exemple 16.1.2 — *Pour $X : \mathbb{R} \to \mathbb{R}$ définie par $x \to x^2$. Donner $X^{-1}([1,4])$*
— *$X = \sin$. Déterminer $X^{-1}([0,1])$*

Proposition 16.1.3 *Avec les mêmes notations*

1. $X^{-1}(\emptyset) = \emptyset$
2. $X^{-1}(F) = E$
3. $A = X^{-1}(B) \not\Rightarrow X(A) = B$
   En fait, $A = X^{-1}(B) \Rightarrow X(A) \subset B$
4. $B \subset F$, $X^{-1}(B^c) = (X^{-1}(B))^c$
5. $X^{-1}(\bigcup\limits_{i\in I} B_i) = \bigcup\limits_{i\in I} X^{-1}(B_i)$
6. $X^{-1}(\bigcap\limits_{i\in I} B_i) = \bigcap\limits_{i\in I} X^{-1}(B_i)$

Pour 3) : $X : x \to x^2$. Alors $X^{-1}([-4,4]) = [-2,2]$, mais $X([-2,2]) = [0,4] \subsetneq [-4,4]$
Démontrons 6) (faire les autres)
$e \in X^{-1}(\bigcap\limits_{i\in I} B_i) \iff X(e) \in \bigcap\limits_{i\in I} B_i \iff \forall i \in I,\ X(e) \in B_i \iff \forall i \in I,\ e \in X^{-1}(B_i)$
$\iff e \in \bigcap\limits_{i\in I} X^{-1}(B_i)$

16.1.2 Variables aléatoires.

En fait, définir l'ensemble Ω de tous les résultats possibles de l'expérience peut être compliqué. Dans la pratique, ce qui nous intéresse quand on a une expérience aléatoire, c'est de quantifier certaines valeurs qui dépendent de l'expérience.

Pour $(\Omega, \mathcal{A}, P)$ donné, une variable aléatoire est une fonction $X : \Omega \to F$, $\omega \to X(\omega)$.

Par exemple, lors d'un lancer de pièces, ce qui va nous intéresser ce peut être le nombre de Piles.
Lors du tirage de boules dans une urne, on va s'intéresser au nombre de boules d'une certaine sorte.
En fait, très souvent, ce n'est pas l'expérience aléatoire qui nous importe, mais certaines valeurs liées à cette expérience quantifiable.
Cette approche est très féconde.

Remarque 16.1.4 *Mauvais choix de vocabulaire : une variable aléatoire n'est pas une variable, mais c'est une fonction.*

L'énorme intérêt de cette approche c'est que pour X variable aléatoire, l'espace $X(\Omega)$ est connu et simple (sous-ensemble de $\mathbb{R}$ ou de $\mathbb{R}^d$ ou de $\mathbb{Z}$) alors que Ω est un ensemble souvent difficile à décrire, voire très abstrait.

Plutôt que de travailler sur l'espace $(\Omega, \mathcal{A}, P)$ on va s'intéresser à la probabilité que certaines valeurs de X soient réalisées.

Notation 16.1.5 *Pour* $B \subset F$ *on notera* $(X \in B) = \{\omega \in \Omega / X(\omega) \in B\} = X^{-1}(B)$.

La définition d'une variable aléatoire ci-dessus n'est pas complète. En effet, pour que $P(X \in B)$ ait un sens il est nécessaire que $\{\omega \in \Omega / X(\omega) \in B\} = X^{-1}(B) \in \mathcal{A}$.

Proposition 16.1.6 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $X : \Omega \to F$.
Notons $F = X(\Omega)$. *L'ensemble* $\mathcal{B}$ *des parties* B *de* F *telles que* $X^{-1}(B) \in \mathcal{A}$ *est une tribu sur* F.

En effet
· $X^{-1}(F) = \Omega \in \mathcal{A}$ donc $F \in \mathcal{B}$
· Si $B \in \mathcal{B}$ alors $X^{-1}(B^c) = (X^{-1}(B))^c \in \mathcal{A}$ donc $B^c \in \mathcal{B}$.
· Soit $(B_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{B}$. Alors $X^{-1}(\bigcup_n B_n) = \bigcup_n X^{-1}(B_n) \in \mathcal{A}$ et donc $\bigcup_n B_n \in \mathcal{B}$.
Ceci conduit à la définition suivante :

Définition 16.1.7 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $(F, \mathcal{B})$ *un espace probabilisable.*
Une variable aléatoire X *de* Ω *dans* F *est une application* $X : \Omega \to F$ *telle que*

$$\forall B \in \mathcal{B}, \; X^{-1}(B) \in \mathcal{A}$$

16.1.3 Variables aléatoires discrètes.

On se limitera cette année au cas où l'espace image $X(\Omega)$ est fini ou dénombrable. Dans ce cas, la tribu sur $X(\Omega)$ est $\mathcal{P}(X(\Omega))$.

Définition 16.1.8 *Soit* F *un ensemble et* $(\Omega, \mathcal{A})$ *un espace probabilisable.*
On appelle ***variable aléatoire discrète*** *définie sur* $(\Omega, \mathcal{A})$ *toute application* X *de* Ω *dans* F *telle que :*
— $X(\Omega)$ *est fini ou dénombrable*
— $\forall x \in X(\Omega), \; X^{-1}(\{x\}) \in \mathcal{A}$

*Si $X(\Omega)$ est fini on dit que X est une **variable aléatoire finie***
*Si $X(\Omega)$ est infini et dénombrable on dit que X est une **variable aléatoire discrète infinie***
*Si F est une partie de $\mathbb{R}$ on dit que la **variable aléatoire discrète X est réelle**.*

Notation 16.1.9 • *Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$. Pour $B \subset X(\Omega)$ l'ensemble $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$ se note $(X \in B)$. Donc $(X \subset B) = X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$.*
• *D'où, pour $x \in F$, $(X = \{x\}) = \{\omega \in \Omega / X(\omega) = x\}$. Pour alléger les notations cet ensemble se notera $(X = x) = X^{-1}(\{x\})$*
• *De même si X est une variable aléatoire réelle on note $(X \leqslant x) = \{\omega \in \Omega / X(\omega) \leqslant x\}$*

Proposition 16.1.10 *Les évènements $(X = x)_{x \in X(\Omega)}$ forment un système complet d'évènements de Ω.*

car ...

Proposition 16.1.11 *Soit X une variable aléatoire discrète et f une fonction définie sur $X(\Omega)$ à valeurs dans $\mathbb{R}$. Alors $Y = f \circ X$ est une variable aléatoire discrète réelle, notée $f(X)$.*

En effet, $X(\Omega)$ étant au plus dénombrable, alors $Y(\Omega)$ aussi.
Soit $y \in Y(\Omega)$, alors
$(Y = y) = \{\omega \in \Omega / f(X(\omega)) = y\} = \{\omega \in \Omega / X(\omega) \in f^{-1}(\{y\})\} = \displaystyle\bigcup_{x \in f^{-1}(\{y\})} (X = x)$
Or $f^{-1}(\{y\})$ est une partie de $X(\Omega)$ donc, est au plus dénombrable, alors $(Y = y)$ est une réunion au plus dénombrable d'éléments de la tribu $\mathcal{A}$, c'est donc un évènement.

Remarque 16.1.12 *En particulier si X est une variable aléatoire réelle discrète, la fonction X^2 est aussi une variable aléatoire, de même que $2X + 3$ ou e^X ou $|X|$. .*

16.1.4 Loi d'une variable aléatoire discrète.

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une variable aléatoire discrète définie sur Ω. On sait donc que $X(\Omega)$ est fini ou dénombrable.

Définition 16.1.13 *Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$. On définit la fonction P_X sur $\mathcal{P}(X(\Omega))$ par*

$$P_X : \begin{cases} \mathcal{P}(X(\Omega)) & \to [0, 1] \\ B & \to P_X(B) = P(X \in B) \end{cases}$$

Cette application est appelée loi de probabilité de X. C'est une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$.

En effet :
- $P_X(X(\Omega)) = P(X^{-1}(X(\Omega))) = P(\Omega) = 1$
- Pour $B \in \mathcal{P}(X(\Omega))$, $P_X(B) = P(X^{-1}(B)) \in [0, 1]$.
- Soit $(B_n)_{n \in \mathbb{N}} \in \mathcal{P}(X(\Omega))^{\mathbb{N}}$ avec les B_n deux à deux disjoints.
  $P_X(\displaystyle\bigcup_{n=1}^{+\infty} B_n) = P\left(X^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right)\right) = P\left(\bigcup_{n=1}^{+\infty} \left(X^{-1}(B_n)\right)\right)$
  Les ensembles $X^{-1}(B_n)$ sont deux à deux disjoints, car les B_n sont deux à deux disjoints.
  Donc $P_X(\displaystyle\bigcup_{n=1}^{+\infty} B_n) = \sum_{n=1}^{+\infty} P(X^{-1}(B_n)) = \sum_{n=1}^{+\infty} P_X(B_n)$.

Remarque 16.1.14 *En pratique, P_X est plus facile à caractériser que P car $F = X(\Omega)$ est connu et simple. C'est ça qui fait l'intérêt de cette approche en loi.*

Conséquence 16.1.15 ***Important :***
Pour déterminer la loi de X on détermine les valeurs x_i susceptibles d'être prises par X puis les probabilités $p_i = P(X = x_i)$.

Conséquence 16.1.16 *Soit X une variable aléatoire réelle discrète définie sur Ω et notons $X(\Omega) = \{x_i / i \in I\}$ avec I au plus dénombrable et les x_i deux à deux distincts*
Alors $\forall i \in I, p_i \geqslant 0$ et $\sum_{i \in I} p_i = 1$ où $p_i = P(X = x_i)$

Notation 16.1.17 *La mesure de probabilité P_X ainsi définie, appelée loi de X est aussi notée $\mathcal{L}(X)$.*
Si X a une loi de probabilité $\mathcal{L}$ on écrira $X \sim \mathcal{L}$
Si deux variables aléatoires X et Y suivent la même loi, on notera $X \sim Y$

Exemple 16.1.18 *On lance deux fois un dé parfait. On note X la somme des résultats obtenus. Quelle est la loi de X ?*

Ici $\Omega = \{1, 2, 3, 4, 5, 6\}^2$ et $X(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
L'évènement $(X = 4)$ se produit ssi $\omega \in \{(1, 3), (2, 2), (3, 1)\}$. Chacun des couples est de probabilité $\frac{1}{36}$ donc $P(X = 4) = \frac{3}{36} = \frac{1}{12}$
En raisonnant de la même manière on obtient la loi de X, présentée dans le tableau ci dessous :

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

et par exemple, $(X = 8) = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$

Exemple 16.1.19 *On lance un dé autant de fois que nécessaire. On note X la variable aléatoire égale au nombre de lancers nécessaires pour obtenir pour la première fois le chiffre 6. Donner la loi de X*

Soit $n \in \mathbb{N}^*$. Notons $p_n = P(X = n)$. Notons S_k l'évènement : le kième lancer donne 6, et N_k l'évènement : le kième lancer ne donne pas le chiffre 6 ($N_k = S_k^c$). Pour tout i, $P(S_i) = \frac{1}{6}$ et $P(N_i) = \frac{5}{6}$.
Les lancers sont indépendants, donc les S_i et les N_j sont indépendants.
$P(X = 1) = P(S_1) = \frac{1}{6}$
$P(X = 2) = P(N_1 \cap S_2) = P(N_1)P(S_2) = \frac{5}{6} \times \frac{1}{6} = \frac{5}{6^2}$
$\vdots$
$P(X = n) = P(N_1 \cap N_2 \cap ... \cap N_{n-1} \cap S_n) = \left(\frac{5}{6}\right)^{n-1} \frac{1}{6} = \frac{5^{n-1}}{6^n}$
Remarquons que $\sum_{n=1}^{+\infty} p_n = \sum_{n=1}^{+\infty} \frac{5^{n-1}}{6^n} = \frac{1}{6} \sum_{n=1}^{+\infty} \left(\frac{5}{6}\right)^{n-1} = \frac{1}{6} \frac{1}{1 - \frac{5}{6}} = 1$. Donc la probabilité que $X = +\infty$, c'est à dire que le la probabilité que le jeu ne s'arrête jamais est nulle.
On reverra cette loi de probabilité plus tard.

Remarque 16.1.20 *Deux variables aléatoires distinctes peuvent avoir la même loi. La connaissance de la loi ne permet pas de reconstituer la variable aléatoire.*

Par exemple, dans l'exemple ci-dessus, notons Y le nombre de lancers nécessaires jusqu'à l'apparition du chiffre 3. X et Y ont la même loi. Mais la connaissance de cette loi ne permet pas de retrouver la variable aléatoire : si on sait qu'une var a cette même loi, rien ne permet de dire si $Z = X$ ou $Z = Y$ ou Z est une autre variable aléatoire.

16.2 Lois usuelles.

16.2.1 Rappels : loi uniforme, loi de Bernoulli, loi binomiale

Définition 16.2.1 *Soit X une variable aléatoire telle que $X(\Omega)$ est un ensemble fini de cardinal n.*
Notons $X(\Omega) = \{a_1, a_2, ..., a_n\}$ où les a_k sont distincts 2 à 2.
*X soit une **loi uniforme** sur $(X(\Omega))$ si et seulement si*

$$\forall k \in \{1, .., n\}, \forall i \in \{1, .., n\},\ P(X = a_k) = P(X = a_i)$$

Dans ce cas, $\forall k \in \{1, 2, .., n\},\ P(X = a_k) = \dfrac{1}{n}$
X soit une loi uniforme sur un ensemble de cardinal n se notera $X \sim U_n$ ou $\mathcal{L}(X) = U_n$.

L'exemple type d'une variable aléatoire suivant une loi uniforme est celui où X désigne le résultat d'un lancer de dé.

Définition 16.2.2 *X suit une **loi de Bernoulli de paramètre** p ssi $X(\Omega)$ est de cardinal 2, en général $X(\Omega) = \{0, 1\}$ et $P(X = 1) = p$ et donc $P(X = 0) = 1 - p = q$*
X soit une loi de Bernoulli de paramètre p se notera $X \sim \mathcal{B}(1, p)$ (voir Binomiale)

L'exemple type d'une variable aléatoire suivant une loi de Bernoulli est celui d'une expérience à 2 issues, le succès $(X = 1)$ et l'échec $(X = 0)$. Le paramètre p est la probabilité du succès de l'expérience.

Définition 16.2.3 *X suit une **loi binomiale** $\mathcal{B}(n, p)$ **de paramètres** n **et** p avec $0 < p < 1$ si et seulement si :*

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

X suit une loi binomiale de paramètres n et p se notera $X \sim \mathcal{B}(n, p)$

C'est bien une loi car en utilisant la formule du binôme,
$\displaystyle\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = (p + 1 - p)^n = 1$
Le schéma d'une expérience débouchant sur une loi binomiale : on considère une expérience pouvant déboucher soit sur un succès avec la probabilité p, soit sur un échec. On répète n fois cette expérience, les expériences étant indépendantes. X désignera le nombre de succès. Alors $X \sim B(n, p)$
Par exemple, un lancer de dé n fois, avec X la variable aléatoire décomptant le nombre de "5" sortis.

Alors $X \sim \mathcal{B}(n, 1/6)$

Remarquons qu'une loi binomiale de paramètre 1 et p est en fait une loi de Bernoulli, ce qui justifie la notation $\mathcal{B}(1, p)$ pour la loi de Bernoulli.

16.2.2 Loi géométrique

Définition 16.2.4 *Soit* $p \in]0, 1[$
On dit qu'une variable aléatoire X *suit la* ***loi géométrique de paramètre*** p *si et seulement si*

$$X(\Omega) = N^* \text{ et } \forall k \geqslant 1,\ P(X = k) = (1-p)^{k-1}p = q^{k-1}p \text{ où } q = 1-p$$

X suit une loi géométrique de paramètre p se notera $X \sim \mathcal{G}(p)$

C'est bien une loi car :
$\sum\limits_{k=1}^{+\infty} P(X = k) = \sum\limits_{k=1}^{+\infty}(1-p)^{k-1}p$ (Remarquons $0 < 1-p < 1 \Longrightarrow \sum (1-p)^n$ converge).
Donc $\sum\limits_{k=1}^{+\infty} P(X = k) = p\sum\limits_{k=1}^{+\infty}(1-p)^{k-1} = p\sum\limits_{k=0}^{+\infty}(1-p)^k = p\dfrac{1}{1-(1-p)} = 1$
Le schéma d'une expérience débouchant sur une loi géométrique : on considère une expérience pouvant déboucher soit sur un succès (noté S) avec la probabilité p, soit sur un échec (avec la probabilité $q = 1-p$). On répète cette expérience, les expériences étant indépendantes (i.e. on répète des expériences de Bernoulli jusqu'à l'apparition du premier succès). X désignera le nombre d'épreuves effectuées jusqu'au premier succès. On dit que X est le temps d'attente du premier succès.
X suit une loi géométrique : vérifions le
Soit A_k l'évènement "S est réalisé au cours de la kième épreuve". On a dit que les épreuves sont indépendantes, donc les A_k le sont aussi.
Pour $n > 1$, $P(X = n) = P(A_1^c \cap .... \cap A_{n-1}^c \cap A_n) = (1-p)^{n-1}p$ (pour $n \geqslant 2$)
$P(X = 1) = P(A_1) = p = (1-p)^0 p$. Cqfd
C'est ce qu'on avait fait dans l'exemple 16.1.4

Remarque 16.2.5 $p = P(X = 1)$

Théorème 16.2.6 *Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}^*$.*
X suit une loi géométrique si et seulement si

$$\forall (k, n) \in \mathbb{N}^{*2}, P(X > n + k | X > n) = P(X > k) \text{ et } P(X = 1) \in]0, 1[$$

Dans ce cas, le paramètre de la loi géométrique est $p = P(X = 1)$

En effet, • $P(X = 1) = p \in]0, 1[$
• Supposons que X suit une loi géométrique de paramètre p. Évaluons $P(X > j)$ avec $j \in \mathbb{N}^*$.
$(X > j) = \bigcup\limits_{i=j+1}^{+\infty} (X = i)$ et les évènements $(X = i)$ sont disjoints 2 à 2.
$P(X > j) = \sum\limits_{i=j+1}^{+\infty} (1-p)^{i-1}p = p\sum\limits_{i=j}^{+\infty}(1-p)^i$. On reconnaît le reste d'une série géométrique.

Donc $P(X > j) = p\dfrac{(1-p)^j}{1-(1-p)} = (1-p)^j$.

Alors $P(X > n+k|X > n) = \dfrac{P((X > n+k)\cap(X > n))}{P(X > n)} = \dfrac{P(X > n+k)}{P(X > n)} = \dfrac{(1-p)^{n+k}}{(1-p)^n} = (1-p)^k = P(X > k)$.

• Réciproquement. Vues les hypothèses, en reprenant le calcul du dessus,
$\forall(k,n)\in\mathbb{N}^{*2}, P(X > n+k|X > n) = \dfrac{P((X > n+k)\cap(X > n))}{P(X > n)} = \dfrac{P(X > n+k)}{P(X > n)} = P(X > k)$.
Donc $P(X > n+k) = P(X > n)P(X > k)$
Posons $p = P(X = 1)$, alors $p \in]0,1[$.
$P(X > 1) = 1 - P(X = 1) = 1 - p$ car P prend ses valeurs dans $\mathbb{N}^*$
donc $P(X > 1+k) = P(X > 1)P(X > k) = (1-p)P(X > k)$.
Donc la suite $(P(X > k))_k$ est une suite géométrique de raison $(1-p)$ donc
$\exists\alpha$ tel que $P(X > k) = \alpha(1-p)^k$
Or $P(X > 1) = 1 - p$ donc $\alpha = 1$ et donc $\forall k \geqslant 1,\ P(X > k) = (1-p)^k$.
On termine en remarquant que pour $k > 1$,
$P(X = k) = P(k-1 < X \leqslant k) = P(X > k-1) - P(X > k) = (1-p)^{k-1} - (1-p)^k = p.(1-p)^{k-1}$
Cette formule est encore vraie pour $k = 1$ et donc, X suit effectivement une loi géométrique de paramètre $p = P(X = 1)$.

Conséquence 16.2.7 *Si X suit une loi géométrique, alors $\forall(k,n)\in\mathbb{N}^{*2},\ P(X > n+k) = P(X > n)P(X > k)$*

16.2.3 Loi de Poisson

Définition 16.2.8 *Soit $\lambda > 0$.*
On dit qu'une variable aléatoire X suit une loi de Poisson de paramètre λ si et seulement si :

$$X(\Omega) = \mathbb{N} \text{ et } P(X = n) = \frac{e^{-\lambda}\lambda^n}{n!}$$

X suit une loi de Poisson de paramètre λ se notera $X \sim \mathcal{P}(\lambda)$

C'est bien une loi car la série $\sum \dfrac{e^{-\lambda}\lambda^n}{n!}$ converge et $\displaystyle\sum_{n=0}^{+\infty}\frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!} = e^{-\lambda}e^{\lambda} = 1$, ce qui était attendu.

Théorème 16.2.9 ***Approximation de la loi binomiale par la loi de Poisson***
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telles que
$\forall n \in \mathbb{N},\ X_n \sim \mathcal{B}(n,p_n)$ avec $\lim\limits_n np_n = \lambda > 0$.

Alors $\forall k \in \mathbb{N},\ \lim\limits_{n\to+\infty} P(X_n = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$
Dans ce cas on dit qu'on peut approximer la loi binomiale $\mathcal{B}(n,p_n)$ par la loi de Poisson $\mathcal{P}(\lambda)$ où $\lambda = \lim\limits_{n\to+\infty} np_n$.

En effet : $\lim\limits_n np_n = \lambda \Longrightarrow p_n = O(\dfrac{\lambda}{n})$ et donc $\lim p_n = 0$.
$P(X_n = k) = \binom{n}{k}p_n^k(1-p_n)^{n-k}$. Or $(1-p_n)^{n-k} = e^{(n-k)\ln(1-p_n)}$.
Mais $\lim p_n = 0 \Longrightarrow (n-k)\ln(1-p_n) \sim (n-k)(-p_n) \sim -\lambda$, donc $\lim\limits_n(1-p_n)^{n-k} = e^{-\lambda}$

$P(X_n = k) \sim \dfrac{n!}{k!(n-k)!} p_n^k e^{-\lambda} \sim \dfrac{e^{-\lambda}}{k!} n(n-1)...(n-k+1)p_n^k$. Or $n(n-1)...(n-k+1)p_n^k \sim n^k p_n^k$ car k est un entier fixé, donc finalement $\lim\limits_{n\to+\infty} P(X_n = k) = \dfrac{e^{-\lambda}}{k!}\lambda^k$.

En pratique, on approxime souvent une loi binomiale $\mathcal{B}(n,p)$ par une loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda = np$ lorsque n est assez grand ($n \geqslant 50$) et np pas très grand ($np \leqslant 15$).

Exemple 16.2.10 *Dans une population, la probabilité qu'une personne soit atteinte par une maladie M est de 1%. Évaluer la probabilité pour que sur un échantillon de 500 personnes il y en ait exactement 7 qui soient malades ?*

16.3 Couples de variables aléatoires, variables aléatoires indépendantes

16.3.1 Introduction

Définition 16.3.1 *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et X et Y deux variables aléatoires réelles discrètes définies sur Ω*
On appelle couple (X,Y) l'application de Ω dans $\mathbb{R}^2$ définie par

$$\forall \omega \in \Omega,\ (X,Y)(\omega) = (X(\omega), Y(\omega))$$

Le couple (X,Y) est une variable aléatoire discrète sur Ω

En effet, l'ensemble image est inclus dans $X(\Omega) \times Y(\Omega)$ qui est au plus dénombrable, comme produit cartésien de deux ensembles au plus dénombrables ; de plus $\forall (x,y) \in X(\Omega) \times Y(\Omega)$, $((X,Y) = (x,y)) = \{\omega \in \Omega / X(\omega) = x \text{ et } Y(\omega) = y\} = (X = x) \cap (Y = y)$ intersection de 2 évènements de Ω, donc est dans $\mathcal{A}$.

Définition 16.3.2 *On appelle loi conjointe du couple (X,Y) la loi du couple $Z = (X,Y)$ entièrement déterminée par la donnée de $(p_{i,j})_{i\in I, j\in J}$*
où $p_{i,j} = P(Z = (x_i, y_j)) = P((X = x_i) \cap (Y = y_j))$ et $X(\Omega) = \{x_i / i \in I\}$ et $Y(\Omega) = \{y_j / j \in J\}$

Conséquence 16.3.3 $\displaystyle\sum_{(i,j)\in I\times J} p_{i,j} = 1$

Dans le cas où X et Y sont finies, alors la loi du couple est souvent donnée sous forme d'un tableau à double entrée.

Exemple 16.3.4 *On lance un dé autant de fois que nécessaire. On note X le numéro d'apparition du premier 6, et Y le numéro d'apparition du premier 1. Déterminer la loi du couple (X,Y).*

Remarquons d'abord que X et Y ne peuvent pas prendre la même valeur.
Donc pour $n \geqslant 1$, $P((X,Y) = (n,n)) = 0$.

Justifier que pour $k < n$ alors $P((X = n) \cap (Y = k)) = \left(\dfrac{4}{6}\right)^{k-1} \left(\dfrac{1}{6}\right) \left(\dfrac{5}{6}\right)^{n-k-1} \left(\dfrac{1}{6}\right) = \dfrac{5^{n-2}}{6^n} \left(\dfrac{4}{5}\right)^{k-1}$

Idem pour $k > n$ alors $P((X = n) \cap (Y = k)) = \left(\dfrac{4}{6}\right)^{n-1} \left(\dfrac{1}{6}\right) \left(\dfrac{5}{6}\right)^{k-n-1} \left(\dfrac{1}{6}\right) = \dfrac{5^{k-2}}{6^k} \left(\dfrac{4}{5}\right)^{n-1}$

16.3.2 Lois marginales

Définition 16.3.5 *Les var X et Y sont appelées variables marginales du couple (X,Y)*
La loi marginale de X est la loi de la variable aléatoire X du couple (X,Y).

Proposition 16.3.6 *Supposons connue la loi du couple (X,Y), on retrouve la loi de X en déterminant*

$$P(X=x_i)=\sum_{j\in J}p_{i,j}$$

En effet l'ensemble des $(Y=y_j)$ quand j décrit J est un système complet d'évènements. Alors la formule des probabilités totales conduit à $\forall i\in I,\ P(X=x_i)=\sum_{j\in J}P((X=x_i)\cap(Y=y_j))=$ $\sum_{j\in J}p_{i,j}$

Remarque 16.3.7 *Si la loi du couple (X,Y) est présentée dans un tableau dans lequel l'élément $p_{i,j}$ est situé ligne i colonne j, alors $P(X=x_i)$ est la somme des éléments de la ligne i, de même que $P(Y=y_j)$ est la somme des éléments de la colonne j.*

Remarque 16.3.8 *Mais la connaissance des **lois** de X et de Y ne suffit pas pour connaître la loi du couple (X,Y).*

En effet Soit X la variable aléatoire correspondant au nombre de Pile lors de 4 lancers indépendants d'une pièce, et Y la variable aléatoire correspondant au nombre de Face obtenus lors de ces lancers, c'est à dire $Y=4-X$. Y a la même loi que X (à vérifier). Déterminons la loi du couple (X,Y) : pour $(i,j)\in\{0,1,2,3,4\}^2$, si $j\neq 4-i$ alors $P((X,Y)=(i,j))=0$, et sinon, $P(X,Y)=(i,4-i))=P(X=i)=\binom{4}{i}(\frac{1}{2})^4$.
Notons $X'=X$ alors bien évidemment Y et X' ont la même loi.
Par contre la loi de (X,X') est telle que pour $(i,j)\in\{0,1,2,3,4\}^2$,
si $i\neq j$ alors $P((X,X')=(i,j))=0$. Les couples (X,Y) et (X,X') n'ont pas la même loi.

Exemple 16.3.9 *Reprenons l'exemple 16.3.4*

On connaît par ailleurs la loi de X et de Y : $X\sim\mathcal{G}(1/6)$ et $Y\sim\mathcal{G}(1/6)$
Retrouver la loi de X en cherchant la loi marginale X du couple (X,Y)

16.3.3 Loi conditionnelle

Définition 16.3.10 *Soit (X,Y) un couple de variables aléatoires discrètes réelles définies sur $(\Omega,\mathcal{A},P)$ et soit $x\in\mathbb{R}$ tel que $P(X=x)\neq 0$*
On note $\mathcal{B}=\mathcal{P}(Y(\Omega))$. Alors l'application $\mathcal{B}\to[0,1]$ définie par $B\to P_{(X=x)}(Y\in B)$ est une probabilité sur $(Y(\Omega),\mathcal{B})$ appelée loi de Y conditionnée par $(X=x)$.

Remarque 16.3.11 *Ainsi* $P_{(X=x)}(Y=y_j)=\dfrac{P((X=x)\cap(Y=y_j))}{P(X=x)}$.

Remarque 16.3.12 $P_{(X=x)}(Y=y_j)$ *se note aussi* $P(Y=y_j|X=x)$

Exemple 16.3.13 *Une urne contient 4 boules numérotées de 1 à 4. On tire successivement et sans remise 2 boules. On note X_1 le numéro de la première boule tirée et X_2 le numéro de la seconde. On note Y le maximum des numéros des boules 1 et 2.*
Donner la loi du couple (X_1,X_2), la loi marginale de X_2, la loi du couple (X_1,Y), la loi de Y, la loi de (X_2,Y) , et enfin la loi de X_1 sachant que $(Y=3)$.

16.3.4 Couple de variables indépendantes

Définition 16.3.14 *Soit (X,Y) un couple de variables aléatoires discrètes sur Ω. On dira que X et Y sont indépendantes*
$\Longleftrightarrow$ $\forall (x,y) \in X(\Omega)\times Y(\Omega)$, *les évènements $(X=x)$ et $(Y=y)$ sont indépendants, i.e.*
$\Longleftrightarrow$ $\forall (x,y) \in X(\Omega)\times Y(\Omega),\ P((X=x)\cap(Y=y)) = P(X=x)P(Y=y)$

Exemple 16.3.15 *En reprenant l'exemple 16.3.13 on constate que le couple (X_1,X_2) ne prend pas toutes les valeurs de $X_1(\Omega)\times X_2(\Omega)$, et donc X_1 et X_2 ne sont pas indépendantes (ce qu'on pouvait deviner). Autre manière de voir, $P((X_1=n)\cap(X_2=n)) = 0 \neq P(X_1=n)P(X_2=n)$.*

Exemple 16.3.16 *On lance un dé autant de fois que nécessaire. On note X le numéro de la première apparition du 6. On note Y le nombre de lancers nécessaires après la première apparition du 6 pour voir apparaître un 1. Les variables X et Y sont-elles indépendantes ?*

car ...

Proposition 16.3.17 *Si X et Y sont des variables aléatoires sur Ω indépendantes, alors pour toute partie $A \subset X(\Omega)$ et toute partie $B\subset Y(\Omega)$, $P(X\in A, Y\in B) = P(X\in A)P(Y\in B)$*

Car $(X\in A, Y\in B) = \displaystyle\bigcup_{(x,y)\in A\times B}((X=x)\cap(Y=y))$; les évènements $(X=x)\cap(Y=y))$ sont disjoints 2 à 2, donc $P(X\in A, Y\in B) = \displaystyle\sum_{(x,y)\in A\times B} P((X=x)\cap(Y=y))$
$= \displaystyle\sum_{(x,y)\in A\times B} P(X=x)P(Y=y) = \sum_{x\in A}\sum_{y\in B} P(X=x)P(Y=y)$. Donc, par le corollaire du théorème de Fubini, $P(X\in A, Y\in B) = \displaystyle\sum_{x\in A} P(X=x)\sum_{y\in B}P(Y=y) = P(X\in A)P(Y\in B)$

16.3.5 Suites de variables aléatoires discrètes indépendantes

Définition 16.3.18 *Soient $X_1,\ldots,X_n$ n variables aléatoires discrètes définies sur Ω. On dit que ces variables aléatoires sont mutuellement indépendantes*
ssi $\forall x_1\in X_1(\Omega),\ldots,\forall x_n\in X_n(\Omega)$, les évènements $(X_1=x_1),\ldots,(X_n=x_n)$ sont indépendants
ssi $\forall x_1\in X_1(\Omega),\ldots,\forall x_n\in X_n(\Omega),\ P((X_1=x_1)\cap\ldots\cap(X_n=x_n)) = \displaystyle\prod_{k=1}^{n} P(X_k=x_k)$

Donc

Proposition 16.3.19 *Soit $X_1,\ldots,X_n$ n variables aléatoires discrètes définies sur Ω mutuellement indépendantes, alors*
pour tout $A_1\subset X_1(\Omega),\ldots,A_n\subset X_n(\Omega)$, $P\left(\displaystyle\bigcap_{k=1}^{n}(X_k\in A_k)\right) = \prod_{k=1}^{n} P(X_k\in A_k)$

Par récurrence à partir de la propriété 16.3.17

Définition 16.3.20 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur Ω. On dira que ces variables aléatoires sont mutuellement indépendantes si et seulement si, pour tout entier n, les n variables aléatoires $X_1,\ldots,X_n$ sont mutuellement indépendantes, ce qui est équivalent à dire que toute famille finie extraite de la suite est constituée de variables mutuellement indépendantes.*

Théorème 16.3.21 *Si $X_1, ..., X_n$ sont des variables aléatoires sur Ω mutuellement indépendantes, alors pour tout $m \in [\![1, n-1]\!]$ et pour toute fonction f et g, les variables aléatoires $f(X_1, ..., X_m)$ et $g(X_{m+1}, ..., X_n)$ sont indépendantes.*

Théorème admis.
Une idée de la démonstation dans le cas de 2 variables : X et Y deux variables indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.
En effet, f définie sur $X(\Omega)$ à valeurs dans $\mathbb{R}$ et g définie sur $Y(\Omega)$ à valeurs dans $\mathbb{R}$. Notons $X' = f(X)$ et $Y' = g(Y)$. Soit $x' \in X'(\Omega) = f(X(\Omega))$ et $y' \in Y'(\Omega) = g(Y(\Omega))$.
$(X' = x') = \{\omega \in \Omega / f(X(\omega)) = x'\} = (X \in f^{-1}(\{x'\})$ et de même $(Y' = y') = (Y \in g^{-1}(\{y'\})$.
Alors $P((X' = x') \cap (Y' = y')) = P((X \in f^{-1}(\{x'\}) \cap (Y \in g^{-1}(\{y'\})$. L'indépendance de X et Y conduit alors à l'indépendance de X' et Y'.

Une suite de variables aléatoires mutuellement indépendantes permet de modéliser la répétition indépendante d'une même épreuve, par exemple un jeu infini de Pile ou Face est modélisé par une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes suivant toutes une loi de Bernoulli $\mathcal{B}(1, 1/2)$.
La question qui se pose alors : est-on certain qu'il existe toujours un espace probabilisé associé à cette suite infinie dénombrable d'expériences ? On en a construit un dans le cas du jeu de Pile ou Face. Mais dans d'autres situations ? La réponse est positive, en vertu du théorème suivant :

Théorème 16.3.22 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes. Notons $(\Omega_n, \mathcal{A}_n, P_n)$ l'espace probabilisé associé à X_n et notons Ω l'ensemble des suites $(\omega_n)_{n\in\mathbb{N}}$ telles que pour tout n, $\omega_n \in \Omega_n$. Alors il existe une tribu $\mathcal{A}$ de Ω et une probabilité P sur Ω telle que*
$\forall n > 0, \forall x_1, ..., x_n \in \mathbb{R}^n,\ P\left(\bigcap_{k=1}^{n}(X_k = x_k)\right) = \prod_{k=1}^{n} P_k(X_k = x_k)$

Ce théorème est admis (hors programme)
C'est ce théorème qui valide la démarche employée dans ce cours sur tout ce qui concerne les jeux infinis (lancers de dés etc)

16.4 Espérance, variance

16.4.1 Espérance d'une variable aléatoire

Vous avez vu en première année la définition de l'espérance d'une variable aléatoire finie. On va maintenant généraliser aux variables aléatoires discrètes

Définition 16.4.1 .

1. *Soit X une variable aléatoire réelle discrète à valeurs positives. L'espérance de X est la somme, finie ou $+\infty$ égale à $E(X) = \sum_{x\in X(\Omega)} xP(X = x)$*
2. *Si X est à valeurs réelles, on dira que X est d'espérance finie si et seulement si la famille $(xP(X = x))_{x\in X(\Omega)}$ est sommable.*
   Dans ce cas, sa somme est appelée espérance de la variable aléatoire réelle discrète X, notée $E(X) = \sum_{x\in X(\Omega)} xP(X = x)$.
3. *On dira que la variable aléatoire réelle discrète est centrée si et seulement si elle est d'espérance nulle.*

Remarque 16.4.2 *Si $X(\Omega) = \{x_n, n \in \mathbb{N}\}$ alors X est d'espérance finie si et seulement si la série $\sum_n x_n P(X = x_n)$ est absolument convergente.*

Remarque 16.4.3 *Une variable aléatoire discrète bornée est d'espérance finie*

car si M est un majorant de $|X|$ alors $\forall x \in X(\Omega)$, $|xP(X = x)| \leqslant MP(X = x)$ et on sait que la famille $(P(X = x))_{x \in X(\Omega)}$ est sommable de somme 1.

Exemple 16.4.4 *Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$ où $\forall n \in \mathbb{N}$, $P(X = n) = \dfrac{1}{(n+1)(n+2)}$. Après avoir vérifié qu'il s'agit bien d'une variable aléatoire réelle discrète étudier son espérance.*

Exemple 16.4.5 *Calculer l'espérance de la variable aléatoire X définie par la somme des faces lors du lancer de 3 dés.*

16.4.2 Propriétés de l'espérance

Proposition 16.4.6 ***Linéarité***
X et Y deux variables aléatoires réelles discrètes d'espérances finies.
Alors $\forall \alpha, \beta \in \mathbb{R}$, $\alpha X + \beta Y$ est d'espérance finie et $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$
Cas particulier : $\alpha X + \beta$ est d'espérance finie et $E(\alpha X + \beta) = \alpha E(X) + \beta$

En effet : Cas de la somme de deux variables aléatoires
• (1) Montrons que la famille $((x+y)P(X = x, Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable et calculons sa somme.
$|x + y|\, P(X = x, Y = y) \leqslant |x|\, P(X = x, Y = y) + |y|\, P(X = x, Y = y)$.
Or $\displaystyle\bigcup_{y \in Y(\Omega)} ((X = x) \cap (Y = y)) = (X = x)$. Donc $\displaystyle\sum_{y \in Y(\Omega)} |x|\, P(X = x, Y = y) = |x|\, P(X = x)$.
Pour finir, X étant d'espérance finie, la famille $(|x|\, P(X = x))_{x \in X(\Omega)}$ est sommable.
On fait de même pour $|y|\, P(X = x, Y = y)$,
on conclut donc que la famille $((x + y)P(X = x, Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable.
De plus en reprenant maintenant le même raisonnement mais sans les valeurs absolues, on déduit que

$$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (x + y)P(X = x, Y = y) = E(X) + E(Y)$$

• (2) Maintenant qu'on sait que la famille $((x + y)P(X = x, Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable, on va la sommer en choisissant une autre partition :
Notons $A_z = \{(x, y) \in X(\Omega) \times Y(\Omega) / x + y = z\}$. Ces ensembles $(A_z)_{z \in Z(\Omega)}$ forment un système complet d'évènements de $X(\Omega) \times Y(\Omega)$
Notons $Z = X + Y$. $(Z = z) = \displaystyle\bigcup_{(x,y) \in A_z} ((X = x) \cap (Y = y))$ réunion disjointe, donc
$P(Z = z) = \displaystyle\sum_{(x,y) \in A_z} P((X = x) \cap (Y = y))$ et donc pour $z \in Z(\Omega)$,
$zP(Z = z) = \displaystyle\sum_{(x,y) \in A_z} (x + y)P((X = x) \cap (Y = y))$
Ainsi $E(X) + E(Y) = \displaystyle\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (x + y)P(X = x, Y = y)$

$$= \sum_{z \in \mathbb{Z}(\Omega)} \sum_{(x,y) \in A_z} (x+y)P((X=x) \cap (Y=y)) = \sum_{z \in \mathbb{Z}(\Omega)} zP(Z=z) = E(Z) = E(X+Y)$$

• Pour ce qui concerne αX,
si α est nul alors la variable aléatoire réelle discrète est nulle, donc $E(\alpha X) = 0 = \alpha E(X)$.
Si $\alpha \neq 0$, $P(\alpha X = \alpha x) = P(X = x)$. X est d'espérance finie, donc la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable et donc la famille $(\alpha x P(X = x))_{x \in X(\Omega)}$ est sommable,
avec $\alpha x P(X = x)) = \alpha x P(\alpha X = \alpha x))$, donc αX est d'espérence finie et $E(\alpha X) = \alpha E(X)$.
• Cas particulier : il suffit de calculer $E(Y)$ dans le cas où Y est la variable aléatoire réelle discrète ne prenant que la valeur 1. Alors $P(Y = y) = 0$ si $y \neq 1$ et $P(Y = 1) = 1$, d'où $E(Y) = 1$.

Proposition 16.4.7 ***Positivité et croissance***
Soit X et Y deux variables aléatoires réelles discrètes d'espérances finies.
a. Si $X(\Omega) \subset \mathbb{R}_+$ alors $E(X) \geqslant 0$
b. $X \leqslant Y \Longrightarrow E(X) \leqslant E(Y)$

a. $E(X) = \sum_{x \in X(\Omega)} xP(X = x)$ avec $x \geqslant 0$, donc $E(X) \geqslant 0$.
b. $X \leqslant Y \Longrightarrow Y - X \geqslant 0 \Longrightarrow E(Y - X) \geqslant 0 \Longrightarrow E(Y) - E(X) \geqslant 0$

Proposition 16.4.8 ***Comparaison***
Soient X et Y deux variables aléatoires réelles discrètes .
Si $|X| \leqslant Y$ avec Y d'espérance finie, alors X est d'espérance finie.

Car pour $x \in X(\Omega)$,

$$|x| P(X = x) = |x| \sum_{y \in Y(\Omega)} P((X = x) \cap (Y = y)) \leqslant \sum_{y \in Y(\Omega)} yP((X = x) \cap (Y = y))$$

Or la famille $((X = x) \cap (Y = y))_{x \in X(\Omega)}$ constitue une partition au plus dénombrable de $(Y = y)$.
Donc $\sum_{x \in X(\Omega)} |y| P((X = x) \cap (Y = y)) = |y| P(Y = y)$ et la famille $(|y| P(Y = y))_{y \in Y(\Omega)}$ est sommable car Y est d'espérance finie. Donc par le théorème de sommation par paquets, la famille $(|y| P((X = x) \cap (Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable et donc la famille
$(\sum_{y \in Y(\Omega)} |y| P((X = x) \cap (Y = y)))_{x \in X(\Omega)}$ est sommable, d'où la famille $(|x| P(X = x))_{x \in X(\Omega)}$ est sommable, i.e ; X est d'espérance finie.

Théorème 16.4.9 ***Théorème de transfert***
Soit X une variable aléatoire réelle discrète et soit f une fonction définie sur $X(\Omega)$ à valeurs dans $\mathbb{R}$.
$f(X)$ est d'espérance finie si et seulement si la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable.
Dans ce cas, $E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x)$

Démonstration non exigible.
La démonstration repose sur le même genre d'arguments que celle sur la linérité de l'espérance.

Proposition 16.4.10 ***Inégalité de Markov***
Soit X une variable aléatoire réelle discrète positive et d'espérance finie.
Alors $\forall t > 0,\ P(X \geqslant t) \leqslant \dfrac{E(X)}{t}$

En effet : $E(X) = \sum\limits_{x\in X(\Omega)} xP(X=x) \geqslant \sum\limits_{x\in X(\Omega) \text{ et } x\geqslant t} xP(X=x)$

$E(X) \geqslant t \sum\limits_{x\in X(\Omega) \text{ et } x\geqslant t} P(X=x) = tP(X\geqslant t)$

16.4.3 Espérance du produit de variables aléatoires indépendantes

Théorème 16.4.11 ***Produit de 2 variables indépendantes***
Soient X et Y deux variables aléatoires réelles discrètes indépendantes et d'espérances finies. Alors XY est d'espérance finie et $E(XY) = E(X)E(Y)$

En effet, notons $X(\Omega) = \{x_i/i \in I\}$ et $Y(\Omega) = \{y_j/j \in J\}$ où I et J sont au plus dénombrables. La variable aléatoire (X, Y) est discrète à valeurs dans $\mathbb{R}^2$. Considérons $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = xy$.
XY est d'espérance finie si et seulement si la famille
$(x_iy_jP((X,Y) = (x_i, y_j))_{(i,j)\in I\times J}$ est sommable .
Or $x_iy_jP((X,Y) = (x_i, y_j)) = x_iy_jP((X = x_i) \cap (Y = y_j)) = x_iy_jP(X = x_i)P(Y = y_j)$ car X et Y sont indépendantes.
X et Y sont d'espérances finies donc les familles $(x_iP(X = x_i))_{i\in I}$ et $(y_jP(Y = y_j))_{j\in J}$ sont sommables. Donc par le corollaire du théorème de Fubini, la famille $(x_iy_jP((X,Y) = (x_i, y_j))_{(i,j)\in I\times J}$ est sommable, i.e. XY est d'espérance finie.
De plus $E(XY) = \sum\limits_{(i,j)\in I\times J} x_iy_jP((X,Y) = (x_i, y_j)) = \sum\limits_{i\in I} x_iP(X = x_i) \times \sum\limits_{j\in J} y_jP(Y = y_j) = E(X)E(Y)$

16.4.4 Espérance des variables aléatoires usuelles

Théorème 16.4.12 ***Lois finies***
a. Loi constante : $X = a$ alors $E(X) = a$
b. Loi uniforme : $X \sim U_n$ alors $E(X) = \dfrac{n+1}{2}$
c. Loi de bernoulli : $X \sim \mathcal{B}(1,p)$ alors $E(X) = p$
d. Loi binomiale : $X \sim \mathcal{B}(n,p)$ alors $E(X) = np$

Toutes ces lois sont finies, donc elles sont toutes d'espérance finie.
a. $E(X) = a.1 = a$
b. $E(X) = \sum\limits_{k=1}^{n} kP(X=k) = \sum\limits_{k=1}^{n} k\dfrac{1}{n} = \dfrac{1}{n}\dfrac{n(n+1)}{2} = \dfrac{n+1}{2}$
c. $E(X) = 0(1-p) + 1p = p$
d. $E(X) = \sum\limits_{k=0}^{n} k.\binom{n}{k}(1-p)^{n-k}p^k$. Or pour $1 \leqslant k \leqslant n,\ k.\binom{n}{k} = n\binom{n-1}{k-1}$
Donc $E(X) = \sum\limits_{k=1}^{n} k.\binom{n}{k}(1-p)^{n-k}p^k = n\sum\limits_{k=1}^{n}\binom{n-1}{k-1}(1-p)^{n-k}p^k$
$= np\sum\limits_{k=1}^{n}\binom{n-1}{k-1}(1-p)^{(n-1)-(k-1)}p^{k-1}$
$E(X) = np\sum\limits_{j=0}^{n-1}\binom{n-1}{j}(1-p)^{(n-1)-j}p^j = np((1-p)+p)^{n-1} = np$

Théorème 16.4.13 ***Lois discrètes infinies***
a. Loi géométrique : $X \sim \mathcal{G}(p)$ *alors* $E(X) = \dfrac{1}{p}$
b. Loi de Poisson : $X \sim \mathcal{P}(\lambda)$ *alors* $E(X) = \lambda$

a. Pour $k \in \mathbb{N}^*$, $kP(X = k) = kpq^{k-1}$ avec $0 < q < 1$. Alors, D'Alembert montre que la série $\sum_k kP(X = k)$ converge (absolument), donc la variable X est d'espérance finie. Alors $E(X) = \sum_{k=1}^{+\infty} pkq^{k-1}$. Or la série entière $\sum_{k=1}^{+\infty} kx^{k-1}$ est la série dérivée de $\sum_{k\in\mathbb{N}} x^k$.
Donc $\sum_{k=1}^{+\infty} kx^{k-1} = \dfrac{1}{(1-x)^2}$.
D'où $E(X) = p\dfrac{1}{(1-(1-p))^2} = \dfrac{1}{p}$.
b. Pour $k \in \mathbb{N}^*$, $kP(X = k) = k.e^{-\lambda}\dfrac{\lambda^k}{k!} = \lambda e^{-\lambda}\dfrac{\lambda^{k-1}}{(k-1)!}$. Cette série est clairement convergente donc X est d'espérance finie et $E(X) = \lambda e^{-\lambda}e^{\lambda} = \lambda$

16.5 Variance, covariance, écart type

16.5.1 Moments

Définition 16.5.1 *Soit X une variable aléatoire réelle discrète et $m \in \mathbb{N}^*$. On dira que X admet un moment d'ordre m si et seulement si X^m est d'espérance finie, c'est à dire si la famille $(x^m P(X = x))_{x\in X(\Omega)}$ est sommable.*

Remarque 16.5.2 *X a un moment d'ordre 1 signifie que X est d'espérance finie.*

Proposition 16.5.3 *Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2, alors X est d'espérance finie.*

En effet, $2|X| \leqslant X^2 + 1$. Or X^2 est d'espérance finie, donc $X^2 + 1$ aussi, et donc $|X|$ d'espérance finie. OK

Théorème 16.5.4 ***Inégalité de Cauchy-Schwarz***
X et Y deux variables aléatoires réelles discrètes admettant chacune un moment d'ordre 2.
Alors XY est d'espérance finie et $(E(XY))^2 \leqslant E(X^2)E(Y^2)$

En effet, $|XY| \leqslant \dfrac{1}{2}(X^2 + Y^2)$, donc XY est d'espérance finie.
Soit alors f la fonction définie sur $\mathbb{R}$ par $f(t) = E((tX + Y)^2)$. Cette fonction est définie car X^2, Y^2 et XY sont d'espérances finies.
$f(t) = t^2E(X^2) + 2tE(XY) + E(Y^2)$.
Si $E(X^2) = 0$ alors f est affine de signe constant, donc est constante, i.e. $E(XY) = 0$ donc l'inégalité est vraie.
Si $E(X^2) \neq 0$, alors la fonction f qui est polynomiale de degré 2 est de signe constant, donc son discriminant est négatif, ce qui conduit à l'inégalité souhaitée.

Conséquence 16.5.5 *l'ensemble des variables aléatoires réelles discrètes admettant un moment d'ordre 2 est un sev de l'ensemble des variables aléatoires réelles discrètes d'espérances finies.*

Clair car d'une part cet ensemble est inclus dans l'ensemble des variables aléatoires réelles discrètes d'espérances finies (voir propriété 16.5.3)
Par ailleurs, $(X+Y)^2 = X^2 + Y^2 + 2XY$ et donc, $X+Y$ admet un moment d'ordre 2. Et il en est de même de λX.

16.5.2 Variance et écart type

Définition 16.5.6 *Soit X une variable aléatoire discrète ayant un moment d'ordre 2.*
On définit la variance de X, notée $V(X)$ par : $V(X) = E((X - E(X))^2)$.
On appelle écart type de X le nombre souvent noté $\sigma(X) = \sqrt{V(X)}$

Cette définition a bien un sens car d'une part, X ayant un moment d'ordre 2, alors $(X - E(X))^2 = X^2 - 2XE(X) + E(X)^2$ a un moment d'ordre 1, donc est d'espérance finie.
De plus, $(X - E(X))^2 \geqslant 0$ donc son espérance est positive, i.e. $V(X) \geqslant 0$, donc $\sigma(X)$ existe.

Remarque 16.5.7 *X a un moment d'ordre 2 $\Longrightarrow V(X) \geqslant 0$*

Théorème 16.5.8 *Soit X variable aléatoire réelle discrète ayant un moment d'ordre 2. Alors*
$V(X) = E(X^2) - E(X)^2$ et
$\forall \alpha, \beta \in \mathbb{R},\ V(\alpha X + \beta) = \alpha^2 V(X)$

Car ...

Remarque 16.5.9 *Si X est une variable aléatoire réelle discrète ayant un moment d'ordre 2, alors $E(X^2) \geqslant (E(X))^2$.*

Définition-propriété 16.5.10 .
Soit X variable aléatoire réelle discrète ayant un moment d'ordre 2.
- *Si $\sigma(X) > 0$ alors la variable aléatoire réelle discrète $\dfrac{X}{\sigma(X)}$ a un écart type égal à 1. on dit alors qu'elle est réduite*
et la variable aléatoire réelle discrète $\dfrac{X - E(X)}{\sigma(X)}$ est d'espérance nulle et d'écart type égal à 1. On dit qu'elle est centrée réduite.
- *Si $\sigma(X) = 0$ alors X est presque surement égale à $E(X)$, i.e. $P(X = E(X)) = 1$.*

Car ...

Théorème 16.5.11 *Inégalité de Bienaymé-Tchebychev*
Soit X variable aléatoire réelle discrète ayant un moment d'ordre 2.
Alors, $\forall \varepsilon > 0,\ P(|X - E(X)| \geqslant \varepsilon) \leqslant \dfrac{V(X)}{\varepsilon^2}$

Car on applique l'inégalité de Markov (16.4.10) à la variable aléatoire réelle discrète $Z = X - E(X)$:
$P(|Z| \geqslant \varepsilon) = P(Z^2 \geqslant \varepsilon^2) \leqslant \dfrac{E(Z^2)}{\varepsilon^2} = \dfrac{V(X)}{\varepsilon^2}$

16.5.3 Variance des lois usuelles

Théorème 16.5.12 .
a. Loi constante : $X = a$ alors $V(X) = 0$

b. Loi uniforme : $X \sim U_n$ alors $V(X) = \dfrac{n^2-1}{12}$
c. Loi de bernoulli : $X \sim \mathcal{B}(1,p)$ alors $V(X) = pq$
d. Loi binomiale : $X \sim \mathcal{B}(n,p)$ alors $V(X) = npq$
e. Loi géométrique : $X \sim \mathcal{G}(p)$ alors $V(X) = \dfrac{q}{p^2}$
f. Loi de Poisson : $X \sim \mathcal{P}(\lambda)$ alors $V(X) = \lambda$

Les 4 premières lois sont finies et ont été étudiées en 1ère année. Refaire la démonstration.
e. $X(\Omega) = \mathbb{N}^*$. Notons $u_n = n^2P(X=n) = n^2q^{n-1}p \neq 0$. $\dfrac{u_{n+1}}{u_n} = \left(\dfrac{n+1}{n}\right)^2 q$ a pour limite $q < 1$. On en déduit que la série $\sum n^2P(X=n)$ converge, c'est à dire que X a un moment d'ordre 2 donc une variance.
Vérifier qu'alors $V(X) = E(X^2) - E(X)^2$ vaut q/p^2 (exercice sur les sommes de séries)
f. Justifier que X a un moment d'ordre 2 et montrer que $V(X) = \lambda$.

16.5.4 Covariance

Définition 16.5.13 *Soit X et Y deux variables aléatoires réelles discrètes ayant un moment d'ordre 2.*
On appelle covariance de X et Y le réel noté $Cov(X,Y) = E((X-E(X))(Y-E(Y)))$

Cette définition a bien un sens car $(X-E(X))(Y-E(Y)) = XY - XE(Y) - YE(X) + E(X)E(Y)$. Or XY est d'espérance finie (CS 16.5.4), X d'espérance finie donc $XE(Y)$ aussi de même que $YE(X)$ ainsi que la variable aléatoire réelle discrète constante $E(X)E(Y)$. Donc $Cov(X,Y)$ existe.

Proposition 16.5.14 *Soient X et Y deux variables aléatoires réelles discrètes ayant un moment d'ordre 2.*
Alors $Cov(X,Y) = E(XY) - E(X)E(Y)$

Car ...

Conséquence 16.5.15 *Si X et Y sont indépendantes et ont un moment d'ordre 2,*
alors $Cov(X,Y) = 0$
Attention, la réciproque est fausse.

clair pour le sens direct : c'est le théorème 16.4.11

Exemple 16.5.16 *Soit X et Y dont la loi du couple est définie par le tableau suivant :*

X / Y	-1	0	1
-1	0	1/2	0
0	1/4	0	1/4

(X,Y) est clairement une variable aléatoire réelle discrète (dire pourquoi). On en déduit immédiatement la loi de X et celle de Y (lois marginales)

X	-1	0	1
	1/4	1/2	1/4

et

Y	-1	0
	1/2	1/2

. Alors $E(X) = 0$ et $E(Y) = -1/2$. De plus, la lecture du tableau montre que $P(XY=0) = 1$ et donc $E(XY) = 0$.
Ainsi $E(XY) = E(X)E(Y)$, donc $Cov(X,Y) = 0$.
Mais X et Y ne sont pas indépendantes car $P(X=-1, Y=-1) = 0 \neq P(X=-1)P(Y=-1)$.
La réciproque de la proposition est donc fausse.

Proposition 16.5.17 ***Propriétés de la covariance***
Soient X et Y deux variables aléatoires réelles discrètes ayant un moment d'ordre 2.

1. $Cov(X, X) = V(X)$
2. $Cov(Y, X) = Cov(X, Y)$
3. $Cov(aX, bY) = abCov(X, Y)$ *pour a et b réels*
4. $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$
5. $Cov(X, 1) = 0$

À faire.

Théorème 16.5.18 ***Variance d'une somme***
Soient X et Y deux variables aléatoires réelles discrètes ayant un moment d'ordre 2.
Alors $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$

À faire

Théorème 16.5.19 ***Variance d'une somme de variables indépendantes***
$X_1, ..., X_n$ des variables aléatoires réelles discrètes mutuellement indépendantes.
Alors $V(X_1 + ... + X_n) = \sum_{k=1}^{n} V(X_k)$

Se démontre par récurrence. Tout repose sur le fait que si $X_1, ..., X_n$ sont indépendantes, alors $X_1 + ... + X_{n-1}$ et X_n sont indépendantes (voir théorème 16.3.21)

16.5.5 Loi faible des grands nombres

Théorème 16.5.20 ***Loi faible des grands nombres***
Soit $(X_n)_n$ une suite de variables aléatoires réelles discrètes indépendantes admettant un moment d'ordre 2, de mêmes lois. Notons alors $m = E(X_1)$ et $S_n = \sum_{k=1}^{n} X_k$.
Alors $\forall \varepsilon > 0,\ \lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = 0$

En effet, par linéarité, $E(\frac{S_n}{n}) = \frac{1}{n}E(\sum_{k=1}^{n} X_k) = m$ et en notant $\sigma = \sigma(X_1)$,
$V(\frac{S_n}{n}) = \frac{1}{n^2}V(S_n) = \frac{1}{n^2}\sum_{k=1}^{n} V(X_k) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$ car les variables aléatoires son indépendantes.
Alors le théorème de Bienaymé-Tchebychev (Théorème 16.5.11) donne
$P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{V(S_n/n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$ qui tend vers 0 quand n tend vers $+\infty$

Interprétation : On renouvelle de manière indépendante une même expérience pour laquelle on s'intéresse à une mesure donnée. La loi permet de justifier que fréquemment, une valeur de l'espérance de cette mesure est approximée par la moyenne arithmétique des résultats de cette expérience.

16.6 Fonctions génératrices

16.6.1 Définition

Définition 16.6.1 *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.*
La fonctions génératrice de X, notée G_X est définie par

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

• G_X est donc une série entière
• Pour t réel fixé, t^X est une variable aléatoire dont l'espérance éventuelle est obtenue par le théorème de transfert.

16.6.2 Convergence et propriétés

Théorème 16.6.2 *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Alors*
- *Le rayon R de convergence de la série entière G_X vérifie $R \geqslant 1$*
- *La série entière G_X converge normalement sur $[-1, 1]$*
- *La fonction G_X est continue sur $[-1, 1]$*
- *La fonction G_X est convexe sur $[0, 1]$*

En effet : $\forall t \in [-1, 1]$, $0 \leqslant |P(X = n)t^n| \leqslant P(X = n)$. Or les évènements $(X = n)$ pour $n \in \mathbb{N}$ forment un système complet d'évènements, $\sum_{n\in\mathbb{N}} P(X = n) = 1$, donc la série G_X converge normalement sur $[-1, 1]$, et donc, $R \geqslant 1$, G_X est continue sur $[-1, 1]$, et G_X est infiniment dérivable sur $]-1, 1[$, donc en particulier sur $[0, 1[$ et
$\forall t \in [0, 1[$, $G''_X(t) = \sum_{n=2}^{+\infty} n(n-1)P(X = n)t^{n-2} \geqslant 0$, donc G_X est convexe sur $[0, 1[$, et étant continue en 1, elle est convexe sur $[0, 1]$

Conséquence 16.6.3 *sous ces mêmes hypothèses,* $\forall k \in \mathbb{N},\ P(X = k) = \dfrac{G_X^{(k)}(0)}{k!}$

Car ...

Conséquence 16.6.4 *Donc la connaissance de G_X suffit pour connaître la loi de X.*

Conséquence 16.6.5 $G_X(1) = 1$ *et* $\forall t \in [-1, 1],\ |G_X(t)| \leqslant 1$

16.6.3 Cas des lois usuelles

Théorème 16.6.6 *. a. Loi de bernoulli :* $X \sim \mathcal{B}(1, p) \iff G_X : t \to (1 - p) + pt$ *sur* $\mathbb{R}$
b. Loi binomiale : $X \sim \mathcal{B}(n, p) \iff G_X : t \to (q + pt)^n$ *sur* $\mathbb{R}$
c. Loi géométrique : $X \sim \mathcal{G}(p) \iff G_X : t \to \dfrac{pt}{1 - qt}$ *de rayon de convergence* $R = \dfrac{1}{q}$ *(avec* $q = 1 - p$*)*
f. Loi de Poisson : $X \sim \mathcal{P}(\lambda) \iff G_X : t \to e^{(t-1)\lambda}$ *de rayon* $R = +\infty$

Calculs à savoir faire.

16.6.4 Fonctions génératrices et moments

Théorème 16.6.7 *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.*

— *X est d'espérance finie si et seulement si G_X est dérivable en 1, et dans ce cas,*

$$E(X) = G'_X(1)$$

.

— *X admet un moment d'ordre 2 si et seulement si G_X est deux fois dérivable en 1 et alors*

$$G''_X(1) = E(X(X-1)) \text{ et } V(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$$

car ...

16.6.5 Fonction génératrice et somme

Théorème 16.6.8 — *Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$ et* ***indépendantes****.*
Alors $\forall t \in]-1,1[,\ G_{X+Y}(t) = G_X(t)G_Y(t)$

— *Généralisation : $X_1, ..., X_n$ des variables aléatoires à valeurs dans $\mathbb{N}$ et* ***mutuellement indépendantes****.*
Alors $G_{X_1+...+X_n} = \prod_{k=1}^{n} G_{X_k}$ sur $]-1,1[$.

Car pour $t \in]-1,1[$, $G_{X+Y}(t) = E(t^{X+Y}) = E(t^X t^Y)$. Or X et Y sont indépendantes, donc par le théorème 16.3.21, t^X et t^Y sont indépendantes, donc $E(t^X t^Y) = E(t^X)E(t^Y) = G_X(t)G_Y(t)$ (théorème 16.4.11), cqfd dans le cas de 2 variables. Dans le cas de n variables, récurrence sur n.