Variables aléatoires discrètes

1. Définitions . . . 1
 1.1 Variable aléatoire . . . 1
 1.2 Loi d'une variable aléatoire . . . 2
 1.3 Lois usuelles . . . 3
  1.3.1 Loi uniforme finie (rappel) . . . 3
  1.3.2 Loi de Bernoulli (rappel) . . . 3
  1.3.3 Loi binomiale (rappel) . . . 4
  1.3.4 Loi géométrique . . . 4
  1.3.5 Loi de Poisson . . . 4
2. Couples de variables aléatoires, indépendance. . . . 5
 2.1 Définition, lois . . . 5
 2.2 Indépendance . . . 6
 2.3 Opérations . . . 6
3. Espérance et variance . . . 7
 3.1 Espérance . . . 7
  3.1.1 Définition et espérance des lois usuelles . . . 7
  3.1.2 Formule de transfert et propriétés de l'espérance . . . 8
 3.2 Variance . . . 9
  3.2.1 Deux inégalités . . . 9
  3.2.2 Définition et propriétés . . . 10
 3.3 Covariance . . . 11
 3.4 Inégalités probabilistes . . . 12
4. Fonctions génératrices . . . 13

Dans tout le chapitre, $(\Omega, \mathscr{A}, P)$ désigne un espace probabilisé, et E un ensemble. Sauf mention contraire, toutes les variables aléatoires considérées sont définies sur cet espace probabilisé $(\Omega, \mathscr{A}, P)$.

1 Définitions

1.1 Variable aléatoire

Définition 1 – Variable aléatoire

On appelle *variable aléatoire discrète* une application définie sur Ω telle que :

① l'ensemble $X(\Omega)$ des valeurs prises par X est au plus dénombrable ;

② pour tout $x \in X(\Omega)$, l'image réciproque $X^{-1}(\{x\}) \in \mathscr{A}$, *i.e.* est un événement. On le note généralement $(X = x)$, parfois $\{X = x\}$.

Si $X : \Omega \to E$ est une variable aléatoire discrète, pour tout $A \subset E$, $X^{-1}(A)$ est un événement[1] que l'on note $(X \in A)$.

Lorsque X est à valeurs réelles, on parle de variable aléatoire réelle, et pour $x \in \mathbb{R}$, on peut noter des événements sous la forme $(X \geqslant x)$ au lieu de $\big(X \in [x\,;+\infty[\big)$, $(X < x)$ au lieu de $\big(X \in \left]-\infty\,;x\right[\big)$, et autres analogues.

1. En effet, c'est la réunion des $X^{-1}\big(\{x\}\big)$ pour $x \in A$. Or cette réunion est au plus dénombrable puisque $X(\Omega)$ l'est et chaque élément est un événement d'après la définition. On conclut par stabilité par union dénombrable d'une tribu.

Exemple 2 – *Première variable aléatoire*

On lance deux dés discernables à quatre faces. On définit la variable aléatoire X comme la somme des résultats des deux dés.

L'ensemble des valeurs prises par X est $X(\Omega) = \dots$

Définition 3 – Composition

Soient X une variable aléatoire discrète et f une application définie sur $X(\Omega)$. Alors $f \circ X$ est une variable aléatoire, notée $f(X)$.

1.2 Loi d'une variable aléatoire

Définition 4 – Loi d'une variable aléatoire

Soit X une variable aléatoire discrète.
On appelle *loi de X* l'application $P_X : \begin{array}{ccc} \mathcal{P}(X(\Omega)) & \longrightarrow & [0\,;1] \\ A & \longmapsto & P(X \in A). \end{array}$
C'est une probabilité sur l'espace probabilisable $\big(X(\Omega), \mathcal{P}(X(\Omega))\big)$.

Pour connaître P_X, il suffit de connaître $P(X = x)$ pour tout $x \in X(\Omega)$.

Méthode : « Donner la loi de la variable aléatoire X » consiste donc à donner :

① l'ensemble $X(\Omega)$ des valeurs prises par X ;

② la probabilité $P(X = x)$ de chaque valeur $x \in X(\Omega)$.

Exemple 5 – *Première loi*

On reprend la situation de l'exemple 2 et on suppose les deux dés équilibrés.
Pour $i \in [\![1\,;4]\!]$, on note A_i : « obtenir i avec le premier dé » et B_i : « obtenir i avec le second dé ».

1. $P(X = 2) = \dots$

2. $P(X = 3) = \dots$

3. Comme $X(\Omega)$ est petit, on peut représenter le tout sous forme d'un tableau :

k	2	3	4	5	6	7	8
$P(X = k)$							

4. On a alors $P(X \geqslant 6) = \dots$

Notation : Si deux variables aléatoires X et Y ont même loi, *i.e.* $P_X = P_Y$, on note $X \sim Y$.

Exemple 6 – *On lance une pièce équilibrée, on note X_1 la variable aléatoire égale à 0 si on obtient pile et 1 si on obtient face.*
On lance un dé non pipé, on note X_2 la variable aléatoire égale à 0 si le résultat est pair et à 1 sinon.
Donner les lois de X_1 et X_2.

Proposition 7 – Variables de même loi et composition

Soient X et Y deux variables aléatoires discrètes et f une application définie sur $X(\Omega)$.
Si $X \sim Y$ alors $f(X) \sim f(Y)$.

1.3 Lois usuelles

1.3.1 Loi uniforme finie (rappel)

Définition 8 – Loi uniforme

On suppose E fini non vide et on note $n = \mathrm{Card}(E) \in \mathbb{N}^*$.
On dit que X suit la *loi uniforme* sur E, et on note $X \sim \mathscr{U}(E)$, si

① $X(\Omega) = E$;

② $\forall x \in E$, $P(X = x) = \dfrac{1}{n}$.

Exemple 9 – *Donner quelques exemples typiques de v.a.r. suivant une loi uniforme.*

1.3.2 Loi de Bernoulli (rappel)

Définition 10 – Loi de Bernoulli

Soit $p \in [0\,;1]$. On dit que X suit la *loi de Bernoulli de paramètre* p, et on note $X \sim \mathscr{B}(p)$, si

① $X(\Omega) = \{0, 1\}$;

② $P(X = 1) = p$ (et donc $P(X = 0) = 1 - p$).

Interprétation : Lors d'une épreuve de Bernoulli, *i.e.* une expérience aléatoire ayant deux issues « succès » (de probabilité p) et « échec » (de probabilité $1 - p$), $X = \begin{cases} 1 & \text{si succès,} \\ 0 & \text{si échec.} \end{cases}$

Exemple 11 – *Donner quelques exemples typiques de v.a.r. suivant une loi de Bernoulli.*

1.3.3 Loi binomiale (rappel)

Définition 12 – Loi binomiale

Soient $n \in \mathbb{N}^*$ et $p \in [0\,;1]$. On dit que X suit la *loi binomiale de paramètres n et p*, et on note $X \sim \mathscr{B}(n,p)$, si

① $X(\Omega) = \{0, 1, \ldots, n\} = [\![0\,;n]\!]$;

② $\forall k \in [\![0\,;n]\!]$, $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Interprétation : Lors d'une succession de n épreuves de Bernoulli **identiques** et indépendantes, chacune de probabilité p de succès, X est le nombre de succès.

Exemple 13 – *Donner deux exemples typiques de v.a.r. suivant une loi binomiale.*

1.3.4 Loi géométrique

Définition 14 – Loi géométrique

Soit $p \in]0\,;1[$. On dit que X suit la *loi géométrique de paramètre p*, et on note $X \sim \mathscr{G}(p)$, si

① $X(\Omega) = \mathbb{N}^* = [\![1\,;+\infty[\![$;

② $\forall k \in \mathbb{N}^*$, $P(X=k) = p(1-p)^{k-1}$.

Interprétation : Lors d'une succession illimitée d'épreuves de Bernoulli **identiques** et indépendantes, chacune de probabilité p de succès, X est le rang du premier succès.

Remarque. On vérifie que $\displaystyle\sum_{k \in X(\Omega)} P(X=k) = \ldots$

Proposition 15

Si $X \sim \mathscr{G}(p)$ alors, pour tout $k \in \mathbb{N}^*$, $P(X>k) = (1-p)^k$.

Démonstration. Décomposer $(X > k)$, utiliser la σ-additivité et reconnaître une somme géométrique. ❑

Exemple 16 – *Donner deux exemples typiques de v.a.r. suivant une loi géométrique.*

1.3.5 Loi de Poisson

Définition 17 – Loi de Poisson

Soit $\lambda > 0$. On dit que X suit la *loi de Poisson de paramètre λ*, et on note $X \sim \mathscr{P}(\lambda)$, si

① $X(\Omega) = \mathbb{N} = [\![0\,;+\infty[\![$;

② $\forall k \in \mathbb{N}$, $P(X=k) = \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!}$.

Interprétation : C'est la loi « des événements rares » : il s'agit souvent de décrire le nombre d'occurrences d'un phénomène dans un laps de temps ou d'espace donné. Par exemple, le nombre de connexions à un serveur en une minute, le nombre de coquilles par page dans mon cours, le nombre de soldats de l'armée prussienne tués accidentellement chaque année par des chevaux (Bortkiewicz 1898), etc.

Remarque. On vérifie que $\sum_{k \in X(\Omega)} P(X = k) = \dots$

2 Couples de variables aléatoires, indépendance

2.1 Définition, lois

Définition 18 – Couple de variables aléatoires, loi conjointe, lois marginales

- Un *couple de variables aléatoires* est une variable aléatoire à valeurs dans un ensemble produit. On note $P\big((X,Y) = (x,y)\big)$ par $P(X = x, Y = y)$ ou $P\big((X = x) \cap (Y = y)\big)$.
- On appelle *loi conjointe* de X et Y (ou *loi du couple* (X,Y)) la loi de la variable aléatoire (X,Y), c'est-à-dire la donnée de :

  ① les ensembles $X(\Omega)$ et $Y(\Omega)$;

  ② la probabilité $P\big((X = x) \cap (Y = y)\big)$ pour chaque couple $(x,y) \in X(\Omega) \times Y(\Omega)$.
- On appelle *lois marginales* du couple (X,Y), les lois de X et de Y.

Si on connaît la loi du couple (X,Y), on peut toujours déterminer les lois marginales :

Méthode : Pour déterminer la loi de X à partir de celle du couple (X,Y), il suffit d'utiliser la formule des probabilités totales avec le s.c.e. $\big\{(Y = y) \mid y \in Y(\Omega)\big\}$.
De façon symétrique, on obtient la loi de Y en considérant le s.c.e. $\big\{(X = x) \mid x \in X(\Omega)\big\}$.

⚠ La réciproque est fausse : sauf en cas d'indépendance (cf paragraphe suivant), on ne peut pas déterminer la loi d'un couple à partir de ses lois marginales.

Exemple 19 – *Soit le couple (X,Y) de loi donnée par :* $\forall (i,j) \in \mathbb{N}^2$, $P\big((X = i) \cap (Y = j)\big) = \dfrac{1}{\mathrm{e}\, 2^{i+1} j!}$.
Déterminer la loi de X.

Définition 20 – Loi conditionnelle

Soit A un événement. On appelle *loi conditionnelle de Y sachant A* la loi de Y sur l'espace probabilisé $(\Omega, \mathscr{A}, P_A)$.
Autrement dit, elle est définie par la donnée de :

1. l'ensemble $Y(\Omega)$ des valeurs prises par Y ;
2. les probabilités $P_A(Y = y)$ pour tout $y \in Y(\Omega)$.

2.2 Indépendance

Définition 21 – Variables aléatoires indépendantes

- Deux variables aléatoires discrètes X et Y sont dites *indépendantes*, et on note $X \perp\!\!\!\perp Y$, si l'une des deux conditions suivantes est vérifiée :
  1. $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, $P\big((X = x) \cap (Y = y)\big) = P(X = x)P(Y = y)$;
  2. Pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants.
- Plus généralement, n variables aléatoires discrètes $(X_1, \dots, X_n)$ sont dites (mutuellement) *indépendantes* si l'une des deux conditions suivantes est vérifiée :
  1. $\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$, $P\left(\bigcap\limits_{i=1}^{n} (X_i = x_i)\right) = \prod\limits_{i=1}^{n} P(X_i = x_i)$.
  2. Pour toutes parties $A_1 \subset X_1(\Omega), \dots, A_n \subset X_n(\Omega)$, on a l'égalité $P\big((X_1 \in A_1) \cap \cdots \cap (X_n \in A_n)\big) = P(X_1 \in A_1) \times \cdots \times P(X_n \in A_n)$.

⚠ Si n variables aléatoires discrètes sont indépendantes, elles sont deux à deux indépendantes, mais la réciproque est fausse.

Exemple 22 – *Somme de loi de Bernoulli (vu en PCSI)*

Si $X_1, \dots, X_n$ sont n variables aléatoires indépendantes, toutes de loi de Bernoulli de même paramètre p, alors $X_1 + \cdots + X_n \sim \mathscr{B}(n, p)$. Cela justifie l'interprétation de la loi binomiale comme le nombre de succès lors de la répétition de n épreuves de Bernoulli indépendantes ayant chacune la probabilité p de succès.

Définition 23 – Suites de variables indépendantes, suites i.i.d.

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires discrètes sur Ω.
- On dit que c'est une *suite de variables aléatoires indépendantes* si toute famille finie extraite de cette suite est indépendante, ou de façon équivalente, si pour tout $k \in \mathbb{N}$, la famille $(X_1, \dots, X_k)$ est indépendante.
- On dit que c'est une *suite i.i.d.* (**i**ndépendantes **i**dentiquement **d**istribuées) si c'est une suite de v.a. indépendantes qui suivent toutes la même loi.

Exemple 24 – *Jeu de pile ou face infini*

La modélisation d'un jeu de pile ou face infini requiert l'utilisation d'une suite i.i.d. de variables aléatoires de même loi de Bernoulli.

2.3 Opérations

Proposition 25 – Fonctions de variables indépendantes

- Soient X et Y deux variables aléatoires discrètes, f une fonction définie sur $X(\Omega)$ et g une fonction définie sur $Y(\Omega)$. Si $X \perp\!\!\!\perp Y$ alors $f(X) \perp\!\!\!\perp f(Y)$.
- Plus généralement, si $X_1, \dots, X_n$ sont des variables aléatoires discrètes indépendantes et $f_1, \dots, f_n$ des fonctions définies respectivement sur $X_1(\Omega), \dots, X_n(\Omega)$, alors les variables aléatoires $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes.

Proposition 26 – Lemme des coalitions

- Si les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, alors pour tout $r \in \{1, \dots, n\}$, les variables aléatoires $f(X_1, \dots, X_r)$ et $g(X_{r+1}, \dots, X_n)$ le sont aussi.
- Ce résultat se généralise au cas de plus de deux coalitions.

Exemple 27 – *Soient X_1, X_2, X_3, X_4 quatre v.a.r. indépendantes qui suivent toutes une loi $\mathscr{B}(p)$ avec $p \in]0\,;1[$. Pour $i \in \{1,2,3\}$, on définit $Y_i = X_iX_{i+1}$.*

1. Quelle est la loi des Y_i ?

2. Y_1 et Y_3 sont-elles indépendantes ?

3. Y_1 et Y_2 sont-elles indépendantes ?

3 Espérance et variance

Dans cette partie, toutes les variables aléatoires considérées sont à valeurs réelles ou complexes.

3.1 Espérance

3.1.1 Définition et espérance des lois usuelles

Définition 28 – Espérance d'une variable aléatoire - cas positif

Soit X une variable aléatoire discrète positive, *i.e.* à valeurs dans $[0\,;+\infty]$.
On appelle *espérance de X* le nombre

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x),$$

avec la convention $xP(X = x) = 0$ lorsque $x = +\infty$ et $P(X = +\infty) = 0$.

Remarque. Dans ce cas positif, on peut éventuellement avoir $E(X) = +\infty$. Selon le programme, « *les étudiants peuvent découper, calculer et majorer leurs sommes directement, la finitude valant preuve de sommabilité* ».

Définition 29 – Espérance d'une variable aléatoire - cas général

Soit X une variable aléatoire discrète à valeurs réelles ou complexes.
On dit que X est *d'espérance finie* si la famille $\big(xP(X = x)\big)_{x \in X(\Omega)}$ est sommable, *i.e.* la série $\sum_{x \in X(\Omega)} |x|P(X = x)$ converge.
Dans ce cas on appelle *espérance de X* la somme de cette famille, *i.e.* le nombre

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

Remarque. Dans ce cas général, il faut montrer que l'espérance est finie avant d'entreprendre son calcul.

Remarque. Une v.a. qui prend un nombre fini de valeurs est toujours d'espérance finie.

Vocabulaire 1

Si X est une v.a. telle que $E(X)$ existe et vaut 0, on dit que X est une *variable centrée*.

Exemple 30 – *Soit Y une v.a. à valeurs dans $\mathbb{N}$ telle que pour tout $n \in \mathbb{N}^*$, $P(Y = n) = \dfrac{6}{(n\pi)^2}$. La variable Y est-elle d'espérance finie ?*

Proposition 31 – Espérance des lois usuelles

- Si X est une v.a. constante, *i.e.* $X(\Omega) = \{a\}$ avec $a \in \mathbb{K}$, on a $E(X) = a$.
- Si $X \sim \mathscr{B}(p)$, alors $E(X) = p$.
- Si $X \sim \mathscr{B}(n, p)$, alors $E(X) = np$.
- Si $X \sim \mathscr{G}(p)$, alors $E(X) = \dfrac{1}{p}$.
- Si $X \sim \mathscr{P}(\lambda)$, alors $E(X) = \lambda$.

Démonstration. 1+2. Découlent de la définition. 3. Utiliser la formule dite du capitaine $k\binom{n}{k} = n\binom{n-1}{k-1}$, un glissement d'indice et le binôme de Newton. 4. Considérer la dérivée de la série entière $\sum x^k$ pour $x = 1 - p$. 5. Simplifier avec la factorielle puis glissement d'indice. ❑

Proposition 32 – Une formule dans le cas à valeurs dans $\mathbb{N} \cup \{+\infty\}$

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Alors

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

Démonstration. Théorème de Fubini après avoir écrit $P(X \geqslant n) = \sum\limits_{k=n}^{+\infty} P(X = k)$. ❑

3.1.2 Formule de transfert et propriétés de l'espérance

Théorème 33 – Formule de transfert

Soient X une variable aléatoire discrète et f un application définie sur $X(\Omega)$.
La variable aléatoire $f(X)$ est d'espérance finie si et seulement si la famille $\big(f(x)P(X = x)\big)_{x \in X(\Omega)}$ est sommable. Dans ce cas, on a

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

Remarque. Si X^2 est d'espérance finie, d'après la formule de transfert, $E(X^2) = \sum\limits_{x \in X(\Omega)} x^2 P(X = x)$.

Exemple 34 – *On lance un dé équilibré à n faces et on note X le numéro obtenu. Calcule $E(2^X)$*

Proposition 35 – Propriétés de l'espérance

Soient X et Y deux variables aléatoire discrètes.

1. *Linéarité* : si X et Y sont d'espérance finie et $(\lambda, \mu) \in \mathbb{R}^2$, alors $\lambda X + \mu Y$ est aussi d'espérance finie et on a
$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$
2. *Comparaison* : si $|X| \leqslant Y$ avec Y à valeurs positives et d'espérance finie, alors X est aussi d'espérance finie.
3. *Positivité* : si X est à valeurs positives alors $E(X) \geqslant 0$.
4. *Croissance* : si X et Y sont à valeurs réelles, d'espérance finie et $X \leqslant Y$, alors $E(X) \leqslant E(Y)$.
5. *Caractère quasi-défini* : si X est positive et d'espérance nulle, alors l'événement $(X = 0)$ est presque sûr.

Démonstration. 1. Théorème de transfert avec la v.a. $Z = (X, Y) \colon \Omega \to \mathbb{R}^2$ et $f \colon (x, y) \mapsto \lambda x + \mu y$. 2. Définition et comparaison de séries à termes positifs. 3. Découle de la définition avec $x \geqslant 0$ et $P(X = x) \geqslant 0$. 4. On a $Y - X \geqslant 0$, utiliser 3 et la linéarité. 5. Une somme de termes positifs est nulle ssi chaque terme est nul. On calcule alors $P(X = 0)$ par passage au complémentaire et σ-additivité. ❑

Proposition 36 – Espérance d'un produit de v.a. indépendantes

• Soient X et Y deux variables aléatoire discrètes, **indépendantes** et d'espérance finie.
Alors XY est aussi d'espérance finie et $E(XY) = E(X)E(Y)$.
• Plus généralement, pour n variables aléatoires discrètes **indépendantes** et d'espérance finie, on a $E(X_1 X_2 \cdots X_n) = E(X_1)E(X_2) \cdots E(X_n)$.

Démonstration. Formule de transfert et théorème de Fubini. ❑

⚠ Deux variables aléatoires X et Y peuvent vérifier $E(XY) = E(X)E(Y)$ sans que X et Y soient indépendantes. C'est par exemple le cas si $X \sim \mathscr{U}(-1, 0, 1)$ et $Y = X^2$.

3.2 Variance

Désormais et jusqu'à la fin du chapitre, on considère des variables aléatoires discrète à valeurs **réelles** (on note parfois v.a.r.).

3.2.1 Deux inégalités

Lemme 37 – Espérance du carré

Si la variable aléatoire X^2 est d'espérance finie, alors X est aussi d'espérance finie.

Démonstration. Inégalité classique $|X| \leqslant \frac{1}{2}(1+|X|^2)$, linéarité de l'espérance et comparaison (prop. 35). ❑

Proposition 38 – Inégalité de Cauchy-Schwarz

Soient X et Y deux v.a.r. telles que X^2 et Y^2 sont d'espérance finie, alors XY est aussi d'espérance finie et
$$\big[E(XY)\big]^2 \leqslant E(X^2)E(Y^2).$$
De plus, il y a égalité si et seulement si X et Y sont presque sûrement proportionnelles.

Démonstration. • L'existence de $E(XY)$ se montre de manière analogue au lemme précédent à partir de l'inégalité $|XY| \leqslant \frac{1}{2}(X^2+Y^2)$. • Pour l'inégalité, on considère le polynôme en λ : $P(\lambda) = E((\lambda X+Y)^2)$ qui est toujours positif donc de discriminant négatif ou nul. ❑

3.2.2 Définition et propriétés

Définition 39 – Variance

Si X^2 est d'espérance finie, on définit sa *variance* $V(X)$ et son *écart-type* $\sigma(X)$ par
$$V(X) = E\Big(\big(X - E(X)\big)^2\Big) \qquad \text{et} \qquad \sigma(X) = \sqrt{V(X)}.$$

La variance est un indicateur de dispersion, elle mesure à quel point X est concentrée ou non autour de son espérance.

Vocabulaire 2

On dit que X est *réduite* si elle admet une variance et $V(X) = 1$.

Proposition 40 – Formule de König-Huygens

Si X admet une variance, alors $V(X) = E(X^2) - E(X)^2$.

Démonstration. De la définition de la variance, développer puis utiliser la linéarité de l'espérance. ❑

Proposition 41 – Non linéarité de la variance

Si X admet une variance alors, pour tout $(a, b) \in \mathbb{R}^2$, la variable aléatoire $aX + b$ admet aussi une variance et on a $V(aX + b) = a^2 V(X)$.

Démonstration. À nouveau, partir de la définition de la variance et utiliser la linéarité de l'espérance. ❑

Exemple 42 – *Soit X une v.a.r. telle que $\sigma(X) > 0$. Montrer que la v.a.r. $Y = \dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite.*

Proposition 43 – Variance des lois usuelles

- Si X est une v.a. constante, *i.e.* $X(\Omega) = \{a\}$ avec $a \in \mathbb{K}$, on a $V(X) = 0$.
- Si $X \sim \mathscr{B}(p)$, alors $V(X) = p(1-p)$.
- Si $X \sim \mathscr{B}(n,p)$, alors $V(X) = np(1-p)$.
- Si $X \sim \mathscr{G}(p)$, alors $V(X) = \dfrac{1-p}{p^2}$.
- Si $X \sim \mathscr{P}(\lambda)$, alors $V(X) = \lambda$.

Démonstration. Mêmes techniques respectives que pour l'espérance. Pour les trois dernières, il est plus commode de calculer $E(X(X-1))$ et remarquer que $V(X) = E(X(X-1)) + E(X) - E(X)^2$. ❑

3.3 Covariance

Définition 44 – Covariance

Soient X et Y deux v.a.r. admettant une variance. On appelle *covariance* de X et Y le nombre $\operatorname{Cov}(X,Y) = E([X-E(X)][Y-E(Y)])$.

Remarques.

1. Cette définition a bien du sens d'après l'inégalité de Cauchy-Schwarz (prop. 38) appliquée aux variables aléatoires $X - E(X)$ et $Y - E(Y)$ qui donne $\left|\operatorname{Cov}(X,Y)\right| \leqslant \sqrt{V(X)V(Y)}$.
2. $\operatorname{Cov}(X,Y) = \operatorname{Cov}(Y,X)$, *i.e.* la covariance est symétrique.
3. Comme l'espérance est linéaire, la covariance est bilinéaire.
4. $\operatorname{Cov}(X,X) = V(X)$.

Proposition 45 – Formule à la König-Huygens et nullité en cas d'indépendance

Soient X et Y deux v.a.r. admettant une variance. On a

$$\operatorname{Cov}(X,Y) = E(XY) - E(X)E(Y).$$

En particulier, si X et Y sont indépendantes, alors $\operatorname{Cov}(X,Y) = 0$.

Démonstration. • Développement et linéarité de l'espérance à partir de la définition de la covariance.
• En cas d'indépendance, conséquence directe de la prop. 36. ❑

⚠ La réciproque du dernier point est fausse : deux variables aléatoires peuvent être de covariance nulle sans être indépendantes. Même contre-exemple qu'après la prop. 36.

Proposition 46 – Variance d'une somme finie

- Soit X et Y deux v.a.r. admettant une variance. Alors $X+Y$ admet une variance et

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

En particulier, si X et Y sont indépendantes, on a $V(X+Y) = V(X) + V(Y)$.
- Plus généralement, soient $X_1, \dots, X_n$ des v.a.r. admettant une variance.
  1. On a $V(X_1 + \cdots + X_n) = \displaystyle\sum_{k=1}^{n} V(X_k) + 2\sum_{1 \leqslant i < j \leqslant n} \operatorname{Cov}(X_i, X_j)$.
  2. En particulier, si les v.a. sont deux à deux indépendantes, $V(X_1 + \cdots + X_n) = \displaystyle\sum_{k=1}^{n} V(X_k)$.

Démonstration. Simple calcul et dans le cas d'indépendance, conséquence de la prop. 45 ❑

Exemple 47 – *En utilisant l'exemple 22, retrouver la variance d'une loi binomiale.*

3.4 Inégalités probabilistes

Proposition 48 – Inégalité de Markov

Soit X une v.a.r. **positive** et d'espérance finie. Alors

$$\forall a > 0, \quad P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

Démonstration. Partir de la définition de $E(X)$ et séparer les termes de la somme suivant si $x < a$ ou $x \geqslant a$. ❑

Proposition 49 – Inégalité de Bienaymé-Tchebychev

Soit X une v.a.r. admettant une variance. Alors

$$\forall \varepsilon > 0, \quad P(|X - E(X)| \geqslant \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2}.$$

Démonstration. Inégalité de Markov avec la v.a.r. $\left|X - E(X)\right|^2$ qui est bien positive, et $a = \varepsilon^2$. ❑

Cette inégalité permet d'évaluer la probabilité que X prenne une valeur éloignée de sa moyenne.

Proposition 50 – Loi faible des grands nombres

Soit $(X_n)_{n\geqslant 1}$ une suite i.i.d. de variables aléatoires admettant une variance. On note $S_n = \sum_{k=1}^{n} X_k$, $m = E(X_1)$ et $\sigma^2 = V(X_1)$. Alors

$$\forall \varepsilon > 0, \quad P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.$$

Démonstration. Par linéarité de l'espérance, $E(S_n/n) = m$. Par indépendance des X_i et la prop. 46, $V(S_n/n) = \frac{\sigma^2}{n}$. On applique alors l'inégalité de Bienaymé-Tchebychev avec $X = S_n/n$. ❑

En particulier, cela signifie que les fréquences observées convergent vers l'espérance, ce qui justifie l'interprétation de cette dernière comme une moyenne.

Par exemple, si on lance un très grand nombre de fois un dé équilibré, les fréquences d'apparition de chaque face tendent chacune vers 1/6.

4 Fonctions génératrices

Pour cette dernière partie, on ne considère que des variables aléatoires à valeurs dans $\mathbb{N}$.

Définition 51 – Fonction génératrice

Soit X un v.a.r. à valeurs dans $\mathbb{N}$.
On appelle *fonction génératrice* de X la série entière

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

Si X ne prend qu'un nombre fini de valeurs, $G_X(t)$ est polynomiale en t.

Proposition 52 – Convergence et régularité de la fonction génératrice

Soit X un v.a.r. à valeurs dans $\mathbb{N}$. La série entière définissant G_X vérifie :

1. Le rayon de convergence R vérifie $R \geqslant 1$.
2. Sa somme G_X est définie en 1 (et on a $G_X(1) = 1$).
3. Elle convergence normalement sur $[-1\,;1]$. Par conséquent, sa somme G_X est continue sur $[-1\,;1]$.
4. Sa somme G_X est de classe $\mathcal{C}^\infty$ sur son disque ouvert de convergence.
5. En particulier, pour tout $n \in \mathbb{N}$, $P(X=n) = \dfrac{1}{n!}G_X^{(n)}(0)$ donc la loi de X est entièrement déterminée par sa fonction génératrice.

Démonstration. 1+2. Observer que $G_X(1) = 1$. 3. En notant $f_n : t \mapsto P(X=n)t^n$, on a $\|f_n\|_{\infty,[-1;1]} = P(X=n)$ qui est le terme général d'une série convergente. 4+5. Théorème de dérivation terme à terme des séries entières. ❑

Proposition 53 – Fonction génératrice des lois usuelles

- Si $X \sim \mathscr{B}(p)$, alors $G_X(t) = 1-p+pt$.
- Si $X \sim \mathscr{B}(n,p)$, alors $G_X(t) = (1-p+pt)^n$.
- Si $X \sim \mathscr{G}(p)$, alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$ avec $R = \dfrac{1}{1-p} > 1$.
- Si $X \sim \mathscr{P}(\lambda)$, alors $G_X(t) = \mathrm{e}^{\lambda(t-1)}$ avec $R = +\infty$.

Démonstration. Ce n'est pas à proprement parlé un résultat de cours, il faut **savoir les recalculer rapidement** lorsqu'on en a besoin, cela découle à chaque fois de la définition. ❑

Proposition 54 – Espérance et variance via la fonction génératrice

Soit X un v.a.r. à valeurs dans $\mathbb{N}$.

1. On a X d'espérance finie si et seulement si G_X est dérivable en 1. Dans ce cas, $E(X) = G_X'(1)$.
2. Si G_X est deux fois dérivable en 1, alors X admet une espérance et on a

$$G_X''(1) = E\big(X(X-1)\big), \quad \text{d'où} \;\; V(X) = G_X''(1) + G_X'(1) - G_X'(1)^2.$$

Démonstration. 1. $\boxed{\Rightarrow}$ Si $E(X)$ est finie alors la série $\sum nP(X=n)$ converge donc, avec la notation de la prop. 52, $\sum f_n'$ converge normalement sur $[-1\,;1]$. Le théorème de dérivation terme à terme des séries de fonctions conclut. $\boxed{\Leftarrow}$ et 2. Non exigibles. ❑

Proposition 55 – Fonction génératrice d'une somme de v.a. indépendantes

- Soient X et Y deux v.a.r. à valeurs dans $\mathbb{N}$ **indépendantes**.

On a, sous réserve de convergence, $G_{X+Y} = G_X \times G_Y$.

- Plus généralement, si $X_1, \ldots, X_n$ sont des v.a.r. à valeurs dans $\mathbb{N}$ **indépendantes**, alors $G_{X_1+\cdots+X_n} = G_{X_1} \times \cdots \times G_{X_n}$.

Démonstration. Lemme des coalitions donne t^X et t^Y indépendantes et on utilise la prop. 36. ❑

Exemple 56 – *Retrouver la fonction génératrice d'une binomiale à partir de celle d'une loi de Bernoulli.*

Exemple 57 – *Soient $X \sim \mathscr{P}(\lambda)$ et $Y \sim \mathscr{P}(\mu)$ indépendantes. À l'aide des fonctions génératrices, donner la loi de $X+Y$.*