COUPLES DE VARIABLES ALÉATOIRES

Dans tout ce chapitre, toutes les variables aléatoires sont discrètes et sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$; $X(\Omega)$ et $Y(\Omega)$ désigne l'ensemble des valeurs prises respectivement par X et Y.
En pratique l' ensemble des valeurs prises par ces variables aléatoires sera fini ou bien un sous-ensemble de $\mathbb{N}$.

1 Loi d'un couple de variables aléatoires

Définition 1.1. *Étant donné un couple de variables aléatoires (X,Y), l'application de $X(\Omega) \times Y(\Omega)$ dans $\mathbb{R}$ définie par $(x,y) \mapsto P([X=x] \cap [Y=y])$ est appelée loi du couple (X,Y), ou encore loin conjointe.*

Remarques 1.2.

- *On notera souvent $P(X=x, Y=y)$ au lieu de $P([X=x] \cap [Y=y])$.*
- *Lorsque $X(\Omega)$ et $Y(\Omega)$ sont finis, on pourra représenter la loi du couple par un tableau.*

Exemple 1.3. *Le tableau suivant définit bien la loi d'un couple (X,Y) :*

$Y \diagdown X$	-1	0	1
1	1/8	2/8	0
2	3/8	1/8	1/8

Exemple 1.4. *On lance deux dés tétraédriques (quatre faces) équilibrés. On note X la variable aléatoire représentant le numéro du premier et Y la variable aléatoire représentant le numéro du second. Déterminer la loi du couple (X,Y).*

Remarque 1.5. *La somme des probabilités dans le tableau d'une loi conjointe est égale à 1.*

Réciproquement, étant donné un tel tableau, on définit la loi d'un couple :

Propriété 1.6 (admise). *Soit I, J deux sous-ensembles de $\mathbb{Z}$. Alors un ensembles de valeurs $\{p_{ij},\ (i,j) \in I \times J\}$ définit la loi d'un couple de variables aléatoires si et seulement si :*

$$\forall (i,j) \in I \times J,\ p_{ij} \geq 0 \quad \text{et} \quad \sum_{(i,j) \in I \times J} p_{ij} = 1 \ .$$

Remarque 1.7. *Sur un tableau fini tel que celui de l'exemple 1.3, il n'y a pas de problème pour définir $\sum_{(i,j) \in I \times J} p_{ij}$; et il est clair que l'on peut effectuer cette somme dans l'ordre voulu :*

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

Lorsque l'ensemble des valeurs prises est infini, cela n'est pas si simple. On admettra que les formules ci-dessus restent valables sous réserve de convergence absolue des séries en question.

2 Lois marginales, lois conditionnelles

Les lois marginales sont les lois de chaque élément du couple :

Définition 2.1 (Lois marginales). *On appelle lois marginales du couple (X,Y) les lois des v.a.r X et Y.*

Propriété 2.2 (Obtention des marginales à partir de la loi du couple). *Étant donnée la loi d'un couple (X,Y),*

- *la loi de X est donnée par : pour tout $x \in X(\Omega)$, $P(X=x) = \displaystyle\sum_{y \in Y(\Omega)} P(X=x, Y=y)$;*
- *la loi de Y est donnée par : pour tout $y \in Y(\Omega)$, $P(Y=y) = \displaystyle\sum_{x \in X(\Omega)} P(X=x, Y=y)$.*

Démonstration :

Exemple 2.3. *Déterminer les lois marginales dans le cas de l'exemple 1.3.*

Dans les exercices se basant sur une situation pratique, on sera très souvent dans la situation de l'exemple suivant.

Exemple 2.4. *On lance un dé tétraédrique équilibré. On note X la variable aléatoire représentant le nombre obtenu. Si l'on a obtenu le nombre i, on pioche une boule dans une urne contenant i boules numérotées de 1 à i. On note Y la variable aléatoire représentant le numéro de la boule piochée. Déterminer la loi de Y.*

Définition 2.5 (Loi conditionnelle). *Si $P(X=x) \neq 0$, on appelle loi conditionnelle de Y sachant $[X=x]$ la loi de probabilité définie sur $Y(\Omega)$ par les valeurs de $P_{[X=x]}(Y=y)$. Lorsque l'on ne veut pas préciser x, on parle de loi conditionnelle de Y sachant X.*

Exemple 2.6. *Expliciter la loi conditionnelle de Y sachant X dans l'exemple précédent.*

Propriété 2.7 (Obtention d'une marginale à partir de l'autre marginale et de la loi conditionnelle)**.** *Si l'on connaît la loi de X et la loi conditionnelle de Y sachant X, on retrouve la loi de Y par : pour $y \in Y(\Omega)$,*

$$P(Y = y) = \sum_{x \in X(\Omega)} P(X = x) P_{[X=x]}(Y = y) .$$

Démonstration :

Remarque 2.8. *Dans le cas précédent, on peut bien sûr obtenir la loi du couple car*

$$P(X = x, Y = y) = P(X = x) P_{[X=x]}(Y = y) .$$

3 Indépendance

La notion d'indépendance de variables aléatoire prolonge celle d'indépendance d'événements.

Définition 3.1. *Les variables aléatoire X et Y sont dites indépendantes lorsque :*

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P(X = x, Y = y) = P(X = x)\ P(Y = y) .$$

Exemple 3.2. *Les v.a.r de l'exemple 1.3 sont-elles indépendantes ?*

Remarques 3.3.

- *L'indépendance de deux v.a.r sera souvent déduite directement de l'énoncé (tirage avec remises, expériences indépendantes l'une de l'autre ...).*
- *Si deux variables aléatoires X et Y sont indépendantes, alors pour tous nombres a, b, c, d on aura : $P([a < X \leq b] \cap [c < Y \leq d]) = P(a < X \leq b)\ P(c < Y \leq d)$.*

Propriété 3.4. *Si deux variables aléatoires X et Y sont indépendantes, alors pour tout $x \in X(\Omega)$ tel que $P(X = x) \neq 0$ on a :*

$$\forall y \in Y(\Omega) , \quad P_{[X=x]}(Y = y) = P(Y = y) .$$

Démonstration :

Remarque 3.5. *Lorsque X et Y sont indépendantes, on a donc, pour $x_1 \neq x_2$ et y quelconque,*

$$P_{[X=x_1]}(Y = y) = P_{[X=x_2]}(Y = y) .$$

4 Fonction d'un couple de variables aléatoires

4.1 Cas général

Propriété 4.1. *Soit g une application de $X(\Omega) \times Y(\Omega)$ dans $\mathbb{R}$. Alors l'application Z de Ω dans $\mathbb{R}$ définie par $Z(\omega) = g(X(\omega), Y(\omega))$ est une v.a.r. On a, pour $z \in Z(\Omega)$,*

$$P(Z = z) = \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ g(x,y) = z}} P(X = x, Y = y) \ .$$

Démonstration :

Exemple 4.2. *On reprend l'exemple 1.3. On définit la variable aléatoire Z par $Z = XY$. Déterminer la loi de Z puis son espérance.*

On évitera souvent de déterminer la loi de cette nouvelle variable aléatoire. Si l'on veut juste déterminer son espérance, on utilisera la formule de transfert :

Propriété 4.3 (Formule de transfert, admise)**.** *Soit g une application de $X(\Omega) \times Y(\Omega)$ dans $\mathbb{R}$. Alors, si la v.a.r $g(X, Y)$ admet une espérance,*

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

Exemple 4.4. *Écrire explicitement cette formule lorsque $g(X, Y) = XY$.*
Retrouver ensuite l'espérance de la v.a.r Z de l'exemple précédent.

4.2 Sommes

Exemple 4.5. *On lance deux dés tétraédriques (on reprend la situation et les notations de l'exemple 1.4). On note S la v.a.r représentant la somme des deux dés. Déterminer la loi de S.*

On peut formaliser ce que l'on a fait dans l'exemple précédent par la formule suivante, conséquence directe de la propriété 4.1 :

Propriété 4.6 (Loi d'une somme). *Étant données deux v.a.r X et Y à valeurs dans $\mathbb{N}$ (ou un sous-ensemble de $\mathbb{N}$), on note S leur somme. Sa loi est donnée par : pour tout $n \in \mathbb{N}$,*

$$P(S = n) = \sum_{k=0}^{n} P(X = k, Y = n - k) \ ;$$

certains éléments de cette somme pouvant être nuls.

Pour calculer l'espérance de la somme, on n'utilisera bien sûr pas la loi, mais la linéarité de l'espérance :

Propriété 4.7 (Linéarité de l'espérance, admise). *Soit X et Y deux v.a.r admettant chacune une espérance et deux réels a et b. Alors $X + Y$ admet une espérance et :*

$$E(aX + bY) = aE(X) + bE(Y) \ .$$

En particulier $E(X + Y) = E(X) + E(Y)$.

La loi d'une somme est généralement difficile à déterminer. Certains cas particuliers peuvent cependant être traités explicitement.

Remarque 4.8. *On va utiliser dans la preuve ci-dessous la propriété suivante, énoncé dans le dernier paragraphe de ce chapitre : si X_1, ..., X_n sont des variables aléatoires de Bernoulli mutuellement indépendantes de même paramètre p, alors $\sum_{i=1}^{n} X_i$ est de loi binomiale de paramètres n et p.*

Propriété 4.9 (Stabilité par somme de la loi binomiale)**.** *On suppose que X_1 et X_2 sont deux v.a.r indépendantes suivants respectivement des lois binomiales $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$. Notons $S = X_1 + X_2$ leur somme. Alors S suit la loi binomiale $\mathcal{B}(n_1 + n_2, p)$.*

Démonstration :

Propriété 4.10 (Stabilité par somme de la loi de Poisson)**.** *On suppose que X_1 et X_2 sont deux v.a.r indépendantes suivants respectivement des lois de Poisson $\mathcal{P}(\lambda_1)$ et $\mathcal{P}(\lambda_2)$. Notons $S = X_1 + X_2$ leur somme. Alors S suit la loi de Poisson $\mathcal{P}(\lambda_1 + \lambda_2)$.*

Démonstration :

4.3 Maximum et minimum

Si X et Y sont deux variables aléatoires définies sur un espace probabilisé (Ω, P), alors V définie par $V(\omega) = \min(X(\omega), Y(\omega))$ ainsi que U définie par $U(\omega) = \max(X(\omega), Y(\omega))$ sont également des variables aléatoire (admis). On retiendra les méthodes suivantes pour calculer leurs lois respectives :

Méthode pour le max U: On détermine tout d'abord $U(\Omega)$. Pour $k \in U(\Omega)$, on remarque que l'on a $U \leq k$ si et seulement si $X \leq k$ et $Y \leq k$:

$$[U \leq k] = [X \leq k] \cap [Y \leq k] \ ;$$

et donc

$$P(U \leq k) = P\left([X \leq k] \cap [Y \leq k]\right) \ .$$

Si de plus X et Y sont indépendantes, on obtient :

$$P(U \leq k) = P\left(X \leq k\right) P\left(Y \leq k\right) \ .$$

Une fois déterminée la fonction de répartition de U, on détermine sa loi en remarquant que $[U \leq k] = [U \leq k-1] \cup [U = k]$ (union disjointe) et donc $P(U \leq k) = P(U = k) + P(U \leq k-1)$, ce qui donne :

$$P(U = k) = P(U \leq k) - P(U \leq k-1) \ .$$

Méthode pour le min V : On remarque dans ce cas que pour $k \in V(\Omega)$, on a $V \geq k$ si et seulement si $X \geq k$ et $Y \geq k$ et on applique ensuite une méthode similaire à celle ci-dessus.

5 Covariance et corrélation

Une question que l'on se pose souvent à propos de deux variables aléatoires est celle de leur indépendance. Lorsqu'elles ne le sont pas, on souhaite également préciser dans quelle mesure elles sont corrélées. Cette notion est très importante dans les études statistiques (essais de médicament, économie, sociologie ...). L'outil pour mesurer l'écart à l'indépendance de deux variables aléatoires ou autrement dit le degré de correlation entre ces deux variables est le coefficient de corrélation.

5.1 Covariance

On introduit tout d'abord la notion de covariance.

Définition 5.1 (Covariance)**.** *On suppose que les v.a.r X et Y admettent un moment d'ordre 2 (i.e. X^2 et Y^2 admettent une espérance). Alors les v.a.r X et Y admettent un moment d'ordre 1, ainsi que la v.a.r $(X - E(X))(Y - E(Y))$ (admis). On définit alors la covariance de X et Y par :*

$$Cov(X, Y) = E\left((X - E(X))(Y - E(Y))\right) \ .$$

Pour calculer en pratique la covariance, on utilisera plutôt la formule suivante :

Propriété 5.2 (Formule de Huygens)**.** *Sous les hypothèses de la définition,*

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

Démonstration :

Remarque 5.3. *Par définition,* $Cov(X,X) = E\left((X - E(X))(X - E(X))\right) = E\left((X - E(X))^2\right) = V(X)$.
Dans la propriété précédente, on retrouve la formule : $V(X) = E(X^2) - (E(X))^2$.

Exemple 5.4. *On reprend les données de l'exemple 1.3. Calculer* $Cov(X,Y)$.

Propriété 5.5 (Propriétés de la covariance)**.** *Les v.a.r utilisées sont supposées vérifier les hypothèse de la définition.*

1. *La covariance est symétrique : si* X *et* Y *sont deux v.a.r,* $Cov(X,Y) = Cov(Y,X)$.
2. *La covariance est linéaire à gauche : si* X_1, X_2 *et* Y *sont des v.a.r,* λ_1 *et* λ_2 *deux réels, alors*

$$Cov(\lambda_1\ X_1 + \lambda_2\ X_2, Y) = \lambda_1\ Cov(X_1,Y) + \lambda_2\ Cov(X_2,Y)\ .$$

   Elle est également linéaire à droite.
3. *Si* X *est une v.a.r et* A *est une v.a.r constante (presque sûrement), alors* $Cov(X,A) = 0$.

Démonstration :

Propriété 5.6 (Variance d'une somme)**.** *Soit X, Y deux v.a.r admettant des moments d'ordre 2. Alors $X+Y$ admet un moment d'ordre 2 et*
$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)\ .$$

Démonstration :

5.2 Cas de variables indépendantes

Propriété 5.7 (Espérance du produit de deux v.a.r indépendantes)**.** *Soit X, Y deux v.a.r indépendantes admettant des moments d'ordre 2. Alors*
$$E(XY) = E(X)E(Y)\ .$$

Démonstration : (dans le cas fini)

Corollaire 5.8. *Soit X, Y deux v.a.r indépendantes admettant des moments d'ordre 2. Alors*

- $Cov(X,Y) = 0$;
- $V(X+Y) = V(X) + V(Y)$.

Démonstration :

Remarque 5.9. *La réciproque du premier point est fausse (voir TD).*

Exemple 5.10. *On dispose d'une urne contenant* 1 *boule noire et* 3 *boules blanches. On pioche deux fois dans cette urne avec remise et l'on note* N *la v.a.r représentant le nombre de boules noires obtenues. Déterminer la variance de* N *(sans utiliser la connaissance de la loi de* N*).*
Indication : on utilisera la v.a. X_1 *égale à* 1 *si on a obtenu une boule noire au premier tirage et* 0 *sinon ainsi que la v.a.* X_2 *égale à* 1 *si on a obtenu une boule noire au second tirage et* 0 *sinon.*

5.3 Coefficient de corrélation

Dans ce paragraphe, on suppose que X et Y admettent des moments d'ordre 2.
Le coefficient de corrélation est une version "normalisée" de la covariance (qui ne varie pas si l'on multiplie une des v.a.r) par une constante).

Définition 5.11 (coefficient de corrélation linéaire)**.** *On suppose que* X *et* Y *ne sont pas constantes (de sorte que leur variance n'est pas nulle). On appelle coefficient de corrélation de* X *et* Y*, noté* $\rho(X,Y)$ *le réel*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} .$$

Remarque 5.12. *Si* a *est un réel positif,*

$$\rho(aX,Y) = \frac{Cov(aX,Y)}{\sqrt{V(aX)V(Y)}} = \frac{a\,Cov(X,Y)}{\sqrt{a^2V(X)V(Y)}} = \frac{a\,Cov(X,Y)}{a\sqrt{V(X)V(Y)}} = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} = \rho(X,Y) .$$

Propriété 5.13. *(admise) Sous les hypothèses de la définition, on a*

$$-1 \leq \rho(X,Y) \leq 1 .$$

De plus, on a $|\rho(X,Y)| = 1$ *si et seulement si il existe* a *et* b *tels que* $Y = aX + b$*.*

Remarque 5.14. *Dans le cas on a* $|\rho(X,Y)| = 1$*, les v.a.r dépendent de manière affine l'une de l'autre. Lorsque* $|\rho(X,Y)|$ *est proche de* 1*, on dit que* X *et* Y *sont fortement corrélées. Lorsque* $|\rho(X,Y)|$ *est proche de* 0*, on dit que* X *et* Y *sont faiblement corrélées ; lorsque* $|\rho(X,Y)| = 0$*, qu'elles sont décorrélées. C'est le cas lorsqu'elles sont indépendantes (mais ce ne sont pas les seuls cas).*

6 Suites de variables aléatoire

Cette partie généralise le chapitre sur les couples de variables aléatoires discrètes que nous avons vu au cas de suites de variables aléatoires discrètes.
On se place dans tout ce chapitre dans un espace probabilisé $(\Omega, \mathcal{A}, P)$, sur lequel sont définies les variables aléatoires considérées.

6.1 Définitions

Définition 6.1. *On appelle vecteur aléatoire discret la donnée d'un n-uplet $(X_1, ..., X_n)$ de variables aléatoires discrètes pour un $n \in \mathbb{N}^*$.*
On appelle suite infinie de variables aléatoires d'une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indexée par $\mathbb{N}$.

Exemple 6.2. *Si on lance 10 fois un dé équilibré on peut considérer le vecteur aléatoire $(X_1, ..., X_{10})$ des résultats successifs obtenus.*

Propriété 6.3. *(admise) Soit $(X_1, ..., X_n)$ un vecteur aléatoire sur $(\Omega, \mathcal{A}, P)$ et soit $f : \mathbb{R}^n \to \mathbb{R}$ une application. Alors l'application :*

$$\begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & f(X_1(\omega), ..., X_n(\omega)) \end{array}$$

définit encore une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, notée $f(X_1, ..., X_n)$.

6.2 Indépendance

On généralise la notion vue pour un couple de variables aléatoires.

Définition 6.4 (Indépendance)**.** *On dit que des variables aléatoires $X_1, ..., X_n$ sont mutuellement indépendantes si pour tout $(x_1, ..., x_n) \in X_1(\Omega), ..., X_n(\Omega)$ on a :*

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

On dit qu'une suite infinie de variables aléatoires sont indépendantes si toute sous-suite finie est constituée de variables aléatoiresmutuellement indépendantes.

Remarque 6.5. *Attention à ne pas confondre avec l'indépendance deux à deux des X_i.*

Propriété 6.6 (Lemme des coalitions, admis)**.** *(admis) Soit $(X_1, ..., X_n)$ un vecteur aléatoire formé de variables aléatoires mutuellement indépendantes. Soient I_1 et I_2 deux parties de $\{1, ..., n\}$ telles que $I_1 \cap I_2 = \emptyset$. Alors toute variable aléatoire fonction des variables aléatoires X_i où $i \in I_1$ est indépendante de toute fonction des variables aléatoires X_j où $j \in I_2$.*

Exemple 6.7. *Soit (X_n) une suite de variables aléatoires indépendantes. Alors les variables aléatoires $f(X_1, ..., X_p)$ et $g(X_{p+1}, ..., X_n)$ sont indépendantes pour toutes fonctions f et g.*

Exemple 6.8. *On lance une pièce 100 fois et on note X le nombre de Pile moins le nombre de Face dans les 50 premiers lancers, et Y le nombre de Pile moins le nombre de Face dans les 50 derniers lancers. Les variables aléatoires X et Y sont alors indépendantes d'après le lemme des coalitions.*

6.3 Somme de variables aléatoires

On étend tout d'abord les propriétés de l'espérance et de la variance d'une somme de variables aléatoires.

Propriété 6.9. *(admis) Soit (X_i) une suite de variables aléatoires. Alors pour tout $n \in \mathbb{N}$ on a :*

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

Propriété 6.10. *(admis) Soit (X_i) une suite de variables aléatoires indépendantes deux à deux (ou mutuellement indépendantes, a fortiori). Alors pour tout $n \in \mathbb{N}$ on a :*

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$$

On termine par le cas particulier d'une somme de variables aléatoires de Bernoulli indépendantes.

Propriété 6.11. *Soient X_1, ..., X_n des variables aléatoires de Bernoulli mutuellement indépendantes de même paramètre p. Alors $\sum_{i=1}^{n} X_i$ est de loi binomiale de paramètres n et p.*
Conséquence : l'espérance d'une v.a. de loi binomiale de paramètres n et p est np et sa variance $np(1-p)$.