CHAPITRE 10

Couples de va. aléatoires discrètes

Dans toute cette partie, on supposera données deux variables aléatoires discrètes X et Y, où on notera $Supp(X) \subset \{x_i, i \in \mathbb{N}\}$ et $Supp(Y) \subset \{y_j, j \in \mathbb{N}\}$, avec les x_i deux à deux distincts ou de probabilité nulle, de même que les y_j .

τ Loi

I-1 Loi conjointe



On appelle *vecteur aléatoire discret* toute *n*-liste (X_1, \ldots, X_n) de v.a.d. X_1, \ldots, X_n définies sur un même espace probabilisé.

En particulier, si n = 2, on appelle (X_1, X_2) un *couple* de v.a.d..

■ Exemple 1 :

On lance un dé une infinité de fois et on note respectivement X et Y les première et deuxième apparitions de 6, avec X=0 si 6 n'apparait jamais et Y=0 si 6 n'apparait qu'une fois maximum. Alors (X,Y) est un couple de var. al. discrètes.

Définition

Soit (X, Y) un couple de v.a.d. sur $(\Omega, \mathcal{P}(\Omega), p)$. On appelle *support* du couple (X, Y) et on note Supp((X, Y)) tout ensemble $S \subset (X, Y)(\Omega)$ tel que $P((X, Y) \in S) = 1$.

Remarque :

Trivialement, on a

$$Supp((X,Y)) = \{(X(w),Y(w)) \mid w \in \Omega\} \subset Supp(X) \times Supp(Y)$$

ou, de manière équivalente (en loi) pour les variables discrètes

$$Supp((X,Y)) = \{(x,y) \in Supp(X) \times Supp(Y) \mid P(X=x,Y=y) \neq 0\}$$

■ Exemple 2 :

On reprend l'exemple de X et Y resp. les première et deuxième apparitions de 6. Alors

$$Supp\big((X,Y)\big)\subset\{(x,y)\in\mathbb{N}^*\mid y>x\}\cup\underbrace{\{(x,0)\mid x\in\mathbb{N}\}}$$

le 6 apparaît au plus une fois

Définition

Soit (X,Y) un couple de v.a.d.. On appelle loi du couple (X,Y), ou loi conjointe des variables X et Y l'application

$$P_{(X,Y)}: Supp((X,Y)) \rightarrow [0;1]$$

 $(x,y) \mapsto p((X=x) \cap (Y=y))$

■ Exemple 3 :

Commençons à chercher la loi conjointe dans le cas de l'exemple précédent. ($1^{\text{ère}}$ et $2^{\text{ème}}$ apparition de 6.)

- $Supp((X,Y)) = \{(n,k) \in \mathbb{N} \mid 1 \le n < k\} \cup \mathbb{N} \times \{0\}.$
- Soient $n, k \in Supp((X, Y))$. On cherche p(X = n, Y = k):
- \star Si n=0, alors pour tout $k\in\mathbb{N}, p(X=n,Y=k)=0$:

En effet, on sait que $X \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$ et donc p(X=0)=0, d'où, comme $(X=n,Y=k)\subset (X=n)$, on a, pour tout $k\in\mathbb{N}$

$$0 \le p(X = 0, Y = k) \le p(X = 0) = 0$$
, d'où $p(X = n, Y = k) = 0$

* Si
$$1 \le n < k$$
, alors $p(X = n, Y = k) = q^{k-2}p^2$, où $p = 1/6$ et $q = 1 - p$:

En effet, posons $A_i =$ "le 6 apparaît au rang i." Alors

$$p(X = n, Y = k) = p(\overline{A_1} \cap \dots \overline{A_{n-1}} \cap A_n \cap \overline{A_{n+1}} \dots \overline{A_{k-1}} \cap A_k)$$

puis par indépendance des A_i et parce que $p(A_i) = \frac{1}{6}$, on a bien

$$p(X = n, Y = k) = \underbrace{q \dots q}_{n-1 \text{ fois } \text{de } n+1 \text{ à } k-1} p = q^{n-1} p q^{k-1-n} p = q^{k-2} p^2$$

$$\star$$
 Si $n\in\mathbb{N}^*,\,P(X=n,Y=0)=0$:

C'est l'événement "6 n'apparait qu'une seule fois et c'est au rang n". Avec le système complet $\{(Y=0)\} \cup \{(Y=n)\}_{k>2}$, on peut écrire

$$P(X = n) = P(X = n, Y = 0) + \sum_{k=2}^{+\infty} P(X = n, Y = k)$$

$$= P(X = n, Y = 0) + \sum_{k=n+1}^{+\infty} q^{k-2} p^{2}$$

$$= P(X = n, Y = 0) + pq^{n-1} \text{ (après calcul, laissé en exercice)}$$

D'où P(X = n) = P(X = n, Y = 0) + P(X = n) et donc P(X = n, Y = 0) = 0 (Ce résultat peut se rédémontrer un peu plus rapidement à l'aide des loi conditionnelles.)

• Conclusion :

1

$$Supp((X,Y)) = \{(n,k) \in \mathbb{N}^2 \mid 1 \le n < k\}$$
$$P(X = n, Y = k) = p^2 q^{k-2} \text{ si } 1 \le n < k$$

■ Exemple 4 :

Avec les variables de l'exemple précédent, calculons P("X + Y impair"):

Par la décomposition en parité différente, on a

$$P("X + Y \text{ impair"}) = P(\underbrace{"X \text{ pair et } Y \text{ impair"}}_{A}) + P("X \text{ impair et } Y \text{ pair"})$$

Or, comme P(Y=1)=0, on a $\{(Y=2j+1)\}_{j\in\mathbb{N}^*}$ sys. compl. de "Y imp.", d'où

$$P(A) = P\left(X \text{ pair } \cap \left(\bigcup_{j=1}^{+\infty} Y = 2j+1\right)\right) = \sum_{j=1}^{+\infty} P\left(\underbrace{X \text{ pair }, Y = 2j+1}_{A_j}\right)$$

$$= \sum_{j=1}^{+\infty} \left(\sum_{i=1}^{j} P(X = 2i, Y = 2j+1)\right) \quad \text{car } \{(X = 2i)\}_{i=1...j} \text{ sys. q-c. de } A_j.$$

$$= \sum_{j=1}^{+\infty} \left(\sum_{i=1}^{j} p^2 q^{2j+1-2}\right) = p^2 q \sum_{j=1}^{+\infty} j(q^2)^{j-1}$$

$$= p^2 q \frac{1}{(1-q^2)^2} = p^2 q \frac{1}{(1-q)^2(1+q)^2}$$

$$= \frac{q}{(1+q)^2} \quad \text{car } 1-q=p$$

et de même (laissé en exercice), on a

$$P("X \text{ impair}, Y \text{ pair"}) = \frac{1}{(1+q)^2}$$

D'où la probabilité finale :

$$P("X + Y \text{ impair"}) = \frac{q+1}{(1+q)^2} = \frac{1}{1+q}$$

I-2 Loi marginale

Définition

Soit (X,Y) un couple de v.a.d.. La loi de X est appelée première loi marginale de (X,Y) et la loi de Y est appelée deuxième loi marginale de (X,Y).

Commentaires:

On peut retrouver les lois marginales à partir de la loi conjointe :

Théorème 1

Soit
$$(X,Y)$$
 un couple de v.a.d. de loi $p_{(X,Y)}$. Alors
$$\forall x \in Supp(X), \qquad p(X=x) = \sum_{j=0}^{+\infty} p\big((X=x) \cap (Y=y_j)\big)$$

$$\forall y \in Supp(Y), \qquad p(Y=y) = \sum_{j=0}^{+\infty} p\big((X=x_j) \cap (Y=y)\big)$$

Démonstration :

On démontre ceci pour p(X=x) (le cas de p(Y=y) étant similaire.) Par définition de X, la famille $\{(Y=y_j)\}_{j\in\mathbb{N}}$ est un système complet de l'univers Ω . Alors

$$\begin{split} p(X=x) &= p\left((X=x)\cap\Omega\right) \\ &= p\left((X=x)\cap \bigsqcup_{j=0}^{+\infty}(Y=y_j)\right) \\ &= p\left(\bigsqcup_{j=0}^{+\infty}\left((X=x)\cap(Y=y_j)\right)\right) \\ &= \sum_{j=0}^{+\infty}p\left((X=x)\cap(Y=y_j)\right) \end{split}$$

2 Exercice 1

On pose X et Y les premières et deuxième apparition de 6 dans une infinité de lancers. Déterminer la loi marginale de Y.

Solution

Tout d'abord, on note que $Supp(Y) = \mathbb{N} - \{0,1\}$ car $Supp((X,Y)) \subset \mathbb{N}^* \times [2,+\infty[$. Ainsi, si $k \geq 2$,

$$p(Y = k) = \sum_{n=0}^{+\infty} p((Y = k) \cap (X = n))$$
$$= \sum_{n=1}^{k-1} q^{k-2} p^2 = (k-1)q^{k-2} p^2$$

[-3

Loi conditionnelle



Soit (X, Y) un couple de v.a.d.

• Pour tout $y \in Supp(Y)$ tel que $p(Y = y) \neq 0$, on appelle loi conditionnelle de X par-rapport à [Y = y], l'application

$$\begin{array}{ccc} Supp(X) & \to & [0;1] \\ x & \mapsto & p_{[Y=y]}(X=x) = \frac{p((X=x) \cap (Y=y))}{p(Y=y)} \end{array}$$

• Pour tout $x \in Supp(X)$ tel que $p(X = x) \neq 0$, on appelle *loi conditionnelle de Y* par-rapport à [X = x], l'application

$$\begin{array}{ccc} Supp(Y) & \rightarrow & [0;1] \\ y & \mapsto & p_{[X=x]}(Y=y) = \frac{p((X=x) \cap (Y=y))}{p(X=x)} \end{array}$$

■ Exemple 5 :

Reprenons l'exemple traité jusqu'ici.

• Soit $k\geqslant 2$. Loi conditionnelle de X par-rapport à [Y=k]: Si $k\geqslant 2$, on a $p(X=n)\neq 0$. La loi conditionnelle sachant [Y=k] est bien définie. Tout d'abord, si k fixé, alors $Supp\left(X_{/(Y=k)}\right)=[\![1;k-1]\!]$. Ensuite, si $n\in [\![1;k-1]\!]$, on a

$$p_{[Y=k]}(X=n) = \frac{p(Y=k, X=n)}{p(Y=k)} = \frac{q^{k-2}p^2}{(k-1)q^{k-2}p^2} = \frac{1}{k-1}$$

On a

$$X_{/(Y=k)} \hookrightarrow \mathcal{U}_{\lceil 1;k-1 \rceil}$$
 (intuitif?)

• Soit $n \ge 1$ <u>loi conditionnelle de Y par-rapport à [X = n]</u>: Si $n \ge 1$, on a $p(X = n) \ne 0$). La loi conditionnelle sachant [X = n] est bien définie. De plus,

$$Supp\left(Y_{/(X=n)}\right) = [n+1; +\infty[$$

et, pour tout $k \in [n+1; +\infty]$, on a

$$p_{[X=n]}(Y=k) = \frac{p(Y=k, X=n)}{p(X=n)} = \frac{q^{k-2}p^2}{q^{n-1}p} = pq^{k-n-1}$$

On se rend ainsi compte que

$$(Y-n)_{/(X=n)} \hookrightarrow \mathcal{G}(p)$$



Remarque:

Dans l'exemple précédent, tout se passe comme si on recommençait à zéro à partir de X=n (phénomène "sans mémoire"!)

Commentaires:

On note que dans beaucoup de cas, on n'a pas à calculer la loi conditionnelle, mais l'expérience amène à la trouver directement, comme dans l'exemple suivant :

■ Exemple 6 :

On tire une boule au hasard dans une urne contenant n boules identiques numérotées de 1 à n. On note X le résultat. On tire ensuite X fois de manière indépendante sur une cible avec une probabilité p de succès : "atteindre la cible". On note Y le nombre de succès.

Déterminer la loi conditionnelle de Y sachant (X=i) pour $i\in \llbracket 1,n \rrbracket$:

Pour i fixé, sachant que l'événement (X=i) est réalisé, on observe que $Y/_{X=i}$ compte le nombre de succès dans une série d'épreuves de Bernoulli indépendantes de succès "atteindre son objectif" de probabilité p. Ainsi :

$$Y/_{X=i} \hookrightarrow \mathcal{B}(i,p)$$

II Corrélation

II-1

Espérances

Dans cette section, on se donnera un couple de v.a. discrètes (X,Y). Dans le cadre de l'étude de la corrélation, on aura à étudier l'existence et la valeur de $\mathbb{E}[XY]$. Il existe pour nous deux façons d'aborder cette question :

 \star La première consiste à considérer la variable Z=XY comme une variable habituelle, en cherchant

$$Supp(Z) = \{z_k \mid k \in \mathbb{N}\}$$
 et $P(Z = z)$ pour tout $z \in Supp(Z)$ puis valider l'existence et calculer $\mathbb{E}[Z]$ de la manière habituelle.

 \star La deuxième consiste à utiliser un résultat calculable directement à partie de la loi conjointe :

Théorème 2 de transfert

Soient

- $lacktriangleq X, Y \ deux \ variables \ finies \ de \ support \ index\'e \ resp. \ par \ [\![0,N]\!] \ et \ [\![0,M]\!]$
- lacksquare D un sous ensemble de \mathbb{R}^2 contenant Supp(X,Y),
- $g: D \to \mathbb{R}$ une fonction

alors

$$E(g(X,Y)) = \sum_{i=0}^{N} \sum_{j=0}^{M} g(x_i, y_j) P(X = x_i, Y = y_j) = \sum_{j=0}^{M} \sum_{i=0}^{N} g(x_i, y_j) P(X = x_i, Y = y_j)$$

■ Exemple 7:

On tire une boule au hasard dans une urne contenant n boules identiques numérotées de 1 à n. On note X le résultat. On tire ensuite X fois de manière indépendante sur une cible avec une probabilité p d'atteindre la cible. On note Y le nombre de succès. Calculons $\mathbb{E}[XY^2]$ (On admettra que $\sum_{k=1}^{n} k^3 = \frac{(n(n+1))^2}{4}$):

D'après le théorème de transfert, on sait que
$$\mathbb{E}[XY^2] = \sum_{i=1}^n \sum_{j=0}^n ij^2 P(X=i,Y=j) = \sum_{i=1}^n \sum_{j=0}^i ij^2 P(X=i,Y=j)$$

$$= \sum_{i=1}^n \sum_{j=0}^n ij^2 P(X=i) P_{X=i}(Y=j) = \frac{1}{n} \sum_{i=1}^n i \underbrace{\left(\sum_{j=0}^i j^2 P_{X=i}(Y=j)\right)}_{\mathbb{E}[(Y/X=i)^2]}$$
 Or, on a $T = Y_{/(X=i)} \hookrightarrow \mathcal{B}(i,p)$, et donc, $\mathbb{E}[T^2] = V_n(T) + \mathbb{E}[T]^2 = ip(1-p) + i^2p^2$, d'où
$$\mathbb{E}[XY^2] = \frac{p(1-p)}{n} \sum_{i=1}^n i^2 + \frac{p^2}{n} \sum_{i=0}^n i^3$$
 et donc
$$\mathbb{E}[XY^2] = p(n+1) \left(\frac{(1-p)(2n+1)}{6} + \frac{p n(n+1)}{4}\right)$$

? Exercice 2

En utilisant le théorème de transfert, montrer que dans l'exemple précédent,

$$\mathbb{E}[Y] = \frac{n+1}{2}p$$
 ainsi que $\mathbb{E}[Y^2] = \frac{p(n+1)}{6}(3+2pn-2p)$.

(On pourra remarquer que $Y = X^0 Y$.) En déduire V(Y).

II-2 Covariance

Définition

Soit (X, Y) un couple de v.a.d. Si la variable aléatoire (X - E(X))(Y - E(Y)) admet une espérance, on appelle *covariance du couple* (X, Y) la valeur

$$Cov(X,Y) = \mathbb{E}[(X - E(X))(Y - E(Y))]$$

En général, on n'utilise pas cette formule pour trouver la covariance en pratique. En effet, la formule peut nettement se simplifier (et peut même faire office de définition)

Théorème 3 (Formule de Huyghens-Koenig)

Soit (X,Y) un couple de v.a.d. Si $\mathbb{E}[X], \mathbb{E}[Y]$ et $\mathbb{E}[XY]$ existent, alors le couple (X,Y) admet une covariance qui peut être calculée grâce à la formule

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

Démonstration :

 \star Existence de Cov(X,Y):

Il s'agit de montrer que $\mathbb{E}\big[(X-E(X))(Y-E(Y))\big]$ existe :

$$|(X - E(X))(Y - E(Y))| = |XY + \mathbb{E}[XY] - X\mathbb{E}[X] - Y\mathbb{E}[X]|$$

$$\leq |XY| + \mathbb{E}[|XY|] + |X|\mathbb{E}[|X|] + |Y|\mathbb{E}[|X|]$$

Or, par hypothèse, l'espérance de |XY|, X et Y existent, ce qui signifie que l'espérance de (X-E(X))(Y-E(Y)) existe.

 \star Valeur de Cov(X,Y) :

Il suffit de développer l'espérance par linéarité de l'espérance.

On peut même encore un peu simplifier les hypothèses de la formule de H-K dans certains cas :

Propriété 4

Soit (X,Y) un couple de v.a. discrètes. Si les variables X et Y admettent un moment d'ordre 2, alors l'espérance $\mathbb{E}[XY]$ existe.

Démonstration :

XY est une variable aléatoire définie sur le même espace probabilisé que X et Y. Or, On rappelle que, pour tout $x, y \in \mathbb{R}$, on a

$$x^{2} + y^{2} - 2xy = \underbrace{(x - y)^{2}}_{\geqslant 0}$$

ce qui donne

$$xy \leqslant \frac{1}{2} \left(x^2 + y^2 \right)$$

Appliqué à |X| et |Y|, on obtient

$$|XY| \leqslant \frac{1}{2} \left(X^2 + Y^2 \right)$$

Par hypothèse, X et Y admettent un moment d'ordre 2, ainsi, $\mathbb{E}[X^2]$ et $\mathbb{E}[Y^2]$ existent, d'où $\mathbb{E}[X^2 + Y^2]$ existe.

Par comparaison de variables, l'espérance de XY existe. \square

■ Exemple 8 :

On effectue une série infinie de lancers indépendants d'un dé équilibré et on note A le rang d'apparition du premier 1, et B le rang d'apparition du premier 2. Montrons que $\mathbb{E}[AB]$ existe.

On sait que $A, B \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$. On en déduit que A et B admettent des moments d'ordre 2, ainsi on peut dire que $\mathbb{E}[AB]$ existe.

(Mais on n'a pas la valeur pour autant... Le calcul est HP mais pour information, on $a \mathbb{E}[AB] = \frac{6-2p}{p^2}$)

Corollaire

Soit (X,Y) un couple de v.a.d. Si les variables X et Y admettent un moment d'ordre Z, alors le couple (X,Y) admet une covariance et

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

Démonstration :

Les hypothèses permettent d'appliquer le théorème qui précise que l'espérance de XY existe. Ensuite, le théorème de Koenig Huyghens permet de conclure. \square

■ Exemple 9 :

Si A et B sont les moments de première apparition respectivement de 1 et de 2 dans une série infinie de lancers, cherchons Cov(A,B) en admettant que l'on connait $\mathbb{E}[AB]$:

D'après le début de l'exemple déja traité précédemment, on sait que $A, B \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$ et admettent chacune un moment d'ordre 2. Ainsi, $\mathbb{E}[AB]$, Cov(A, B), $\mathbb{E}[A]$ et $\mathbb{E}[B]$ existent, avec comme formule

$$Cov(A, B) = \mathbb{E}[AB] - \mathbb{E}[A] \mathbb{E}[AB]$$

Après calcul de $\mathbb{E}[AB]$ (admis) on trouverait (avec $p = \frac{1}{6}$)

$$\mathbb{E}[XY] = \frac{6 - 2p}{p^2}$$

De plus, on sait que $\mathbb{E}[A] = \mathbb{E}[B] = \frac{1}{p} = 6$. D'où

$$Cov(A, B) = \frac{6 - 2p}{p^2} - \frac{1}{p^2} = \frac{5 - 2p}{p^2} = 132$$



Remarque :

Si les variables sont finies, alors la covariance existe toujours.

2 Exercice 3

On tire une boule au hasard dans une urne contenant n boules identiques numérotées de 1 à n. On note X le résultat. On tire ensuite X fois de manière indépendante sur une cible avec une probabilité d'atteindre son objetif de p. On note Y le nombre de succès. Déterminer Cov(X,Y).

Solution

On sait d'après la formule de H-K, que

$$Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

D'après le théorème de transfert, on sait que

$$\mathbb{E}[XY] = \sum_{i=1}^{n} \sum_{j=0}^{n} ij P(X=i, Y=j) = \sum_{i=1}^{n} \sum_{j=0}^{i} ij P(X=i, Y=j)$$

$$= \sum_{i=1}^{n} \sum_{j=0}^{n} ij P(X=i) P_{X=i}(Y=j) = \frac{1}{n} \sum_{i=1}^{n} i \underbrace{\left(\sum_{j=0}^{i} j \binom{j}{i} p^{j} (1-p)^{i-j}\right)}_{\mathbb{E}[Y:Y:Y]}$$

Or,

$$\mathbb{E}\left[Y_{/X=i}\right] = ip$$

d'où

$$\mathbb{E}[XY] = \frac{1}{n} \sum_{i=1}^{n} i^2 p = \frac{p}{n} \frac{n(n+1)(2n+1)}{6} = p \frac{(n+1)(2n+1)}{6}$$

Ainsi

$$Cov(X,Y) = p \frac{(n+1)(2n+1)}{6} - \frac{n+1}{2} \mathbb{E}[Y] = p \frac{(n+1)(2n+1)}{6} - p \left(\frac{n+1}{2}\right)^2$$
$$= p \frac{(n+1)}{12} \left(2(2n+1) - 3(n+1)\right)$$

i.e.

$$Cov(X,Y) = p\frac{(n+1)(n-1)}{12}$$

Propriété 5

Soient X,Y des v.a.d. définies sur un même espace probabilisé admettant chacune un moment d'ordre 2. On a

$$Cov(X,X) = V(X)$$

$$Cov(X,Y) = Cov(Y,X)$$

Propriété 6 (bilinéarité de la covariance)

Soient X,Y des v.a.d définies sur un même espace probabilisé admettant chacune un moment d'ordre 2. On a

- $Cov(\centerdot, Y) : X \mapsto Cov(X, Y)$ est linéaire au sens suivant : Si X, X' sont deux v.a.d. telles que $\mathbb{E}[XY]$ et $\mathbb{E}[X'Y]$ existent, alors $\forall \alpha, \beta \in \mathbb{R}, \ Cov(\alpha X + \beta X', Y) = \alpha \ Cov(X, Y) + \beta \ Cov(X', Y)$
- $Cov(X, •): Y \mapsto Cov(X, Y)$ est linéaire au sens suivant : Si Y, Y' sont deux v.a.d. telles que $\mathbb{E}[XY]$ et $\mathbb{E}[XY']$ existent, alors $\forall \alpha, \beta \in \mathbb{R}, \ Cov(X, \alpha Y + \beta Y') = \alpha \ Cov(X, Y) + \beta \ Cov(X, Y')$

■ Exemple 10 :

Soient X,Y deux variables discrètes admettant la même variance. Montrons que Cov(X+Y,X-Y)=0 :

Par bilinéarité de la covariance, on a

$$\begin{array}{lcl} Cov(X+Y,X-Y) & = & Cov(X,X-Y) + Cov(Y,X-Y) \\ & = & \underbrace{Cov(X,X)}_{V(X)} - Cov(X,Y) + Cov(Y,X) - \underbrace{Cov(Y,Y)}_{V(Y)} \\ & = & V(X) - V(Y) \quad \text{par symétrie de la covariance} \\ & = & 0 \end{array}$$

Théorème 7

Soient X,Y des v.a.d. définies sur un même espace probabilisé admettant chacune un moment d'ordre 2. On a

$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$$

Démonstration :

On applique les définitions et la linéarité de l'espérance. \Box

♣ R

Remarque :

Ce théorème permet aussi dans certains cas de calculer Cov(X,Y) avec la formule :

$$Cov(X,Y) = \frac{1}{2} \left(V(X+Y) - V(X) - V(Y) \right)$$

■ Exemple 11 :

On effectue n lancers indépendants d'une même dé équilibré. On note X le nombre de fois où on obtient 1 et Y le nombre de fois où on obtient 2. On cherche à trouver Cov(X,Y).

Tout d'abord, on remarque que X et Y sont deux variables finies telles que

$$X, Y \hookrightarrow \mathcal{B}(n, p), \text{ avec } p = \frac{1}{6}$$

Ceci garantit l'existance des moments d'ordre 2 de X, Y et valide fonc la formule

$$Cov(X, Y) = \frac{1}{2} (V(X + Y) - V(X) - V(Y))$$

De plus, la variable X+Y correspond au nombre total de fois où on rencontre 1 ou 2. Or, comme on ne peut pas avoir en même temps 1 et 2, ceci est en réalité, le nombre de réalisation du succès "obtenir 1 ou 2" dans une série de n tentatives indépendantes. Ainsi, on sait que

$$X + Y \hookrightarrow \mathcal{B}(n, \alpha)$$
 avec $\alpha = \frac{1}{3} = 2p$

D'où

$$Cov(X,Y) = \frac{1}{2} (n\alpha(1-\alpha) - 2np(1-p)) = -np^2 = -\frac{n}{36}$$

Corollaire

Soient X,Y des v.a.d. définies sur un même espace probabilisé admettant chacune un moment d'ordre 2. On a

$$|Cov(X,Y)| \le \sigma(X)\sigma(Y)$$

Démonstration :

Observons le polynôme en $\lambda: V(Y+\lambda X)=V(Y)+\lambda^2 V(X)+2\lambda Cov(X,Y)$. La variance étant toujour positive, c'est un polynôme n'admettant qu'au maximum une racine, c'est-à-dire de discriminant négatif.

Ainsi,
$$0 \ge \Delta = 4Cov(X, Y)^2 - 4V(X)V(Y)$$
, D'où

$$Cov(X,Y)^2 \leqslant V(X)V(Y)$$

ce qui donne le résultat annoncé. \square

II-3

Coefficient de corrélation linéaire

Définition

Soient X,Y des v.a.d. sur un même espace probabilisé, admettant chacune un moment d'ordre 2 **non nul**. On appelle *coefficient de corrélation linéaire* de X et Y la valeur

$$r(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

■ Exemple 12 :

Revenons à l'exemple des tirs sur une cible. (On admettra que $V(X) = \frac{n^2 - 1}{12}$.) On a alors, grâce aux calculs effectués (dans les exemples et exercices précédents)

$$r(X,Y) = \frac{p\frac{(n+1)(n-1)}{12}}{\sqrt{\frac{n^2-1}{12}}\sqrt{\frac{p(n+1)}{12}(6+pn-7p)}} = \sqrt{\frac{p(n-1)}{pn+6-7p}}$$

Propriété 8

Soient X, Y des v.a.d.

- sur un même espace probabilisé,
- admettant chacune un moment d'ordre 2 non nul,

$$-1 \leqslant r(X, Y) \leqslant 1$$

Propriété 9

Soient X, Y des v.a.d.

- sur un même espace probabilisé,
- admettant chacune un moment d'ordre 2 non nul,

alors

$$|r(X,Y)|=1 \quad \Leftrightarrow \quad \text{il existe } a\in\mathbb{R}^*,\,b\in\mathbb{R} \text{ tels que } P(Y=aX+b)=1$$

Démonstration :

En reprenant la démonstration du corollaire précédent (partie 2a), on constate que

$$|r(X,Y)| = 1 \Leftrightarrow Cov(X,Y)^2 \leqslant V(X)V(Y) \Leftrightarrow \Delta = 0$$

Or, $\Delta=0$ équivaut exactement à dire que le polynôme en $\lambda:V(X+\lambda Y)=V(X)+\lambda^2V(Y)+2\lambda Cov(X,Y)$ admet une racine (double). On note λ_0 cette racine. Ainsi,

$$|r(X,Y)| = 1 \Leftrightarrow V(Y + \lambda_0 X) = 0$$

Or, une variable admet une variance nulle ssi elle est constante presque surement, c'est-à-dire s'il existe $b\in\mathbb{R}$ tel que

$$P(Y + \lambda_0 X = b) = 1$$

On note que $\lambda_0 \neq 0$ car sinon on aurait V(Y) = 0 ce qui est contraire à l'énoncé. En conclusion, avec les notations précédentes

$$|r(X,Y)| = 1 \quad \Leftrightarrow \quad P(Y = -\lambda_0 X + b) = 1$$

II-4

7

Indépendance

Définition

Étant donné un couple de variables aléatoires discrètes (X,Y) sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$, on dit que X et Y sont indépendantes si

$$p((X = x) \cap (Y = y)) = p(X = x) p(Y = y)$$
 $\forall x \in Supp(X), y \in Supp(Y)$

■ Contre-Exemple(s) :

On lance un dé une infinité de fois. On reprend les variables aléatoires X et Y donnant le rang de première et deuxième apparition de 6. Alors

$$P(X = 5 \cap Y = 3) = 0 \neq \underbrace{P(X = 5)}_{\neq 0} \underbrace{P(Y = 3)}_{\neq 0}$$

Les deux variables aléatoires ne peuvent donc pas être indépendantes.

2 Exercice 4

On lance une infinité de fois deux dés de manière indépendante. On note X le rang de première apparition de 6 sur le premier dé et Y le rang de première apparition d'un double. (Y vaut 0 s'il n'y a jamais de double.) Montrer que X et Y sont indépendantes.

Solution

X et Y suivent une loi géométrique, avec

$$X \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$$
 et $Y \hookrightarrow \mathcal{G}\left(\frac{6}{36}\right) = \mathcal{G}\left(\frac{1}{6}\right)$

Vérifions que

$$P(X = i \cap Y = j) = P(X = i)P(Y = j) \quad \forall i, j \in \mathbb{N}^*$$
:

il y a 25 couples qui ne contiennent ni un 6 sur le premier dé, ni un double. Ainsi,

· Si 0 < i < j:

$$P(X = i \cap Y = j) = \underbrace{\left(\frac{25}{36}\right)^{i-1}}_{\text{ni 6 ni double jusqu'à } i / 6 \text{ à i non double / pas de double jusqu'à }}_{\text{non double / pas de double jusqu'à } j - 1 / \text{double à } j$$

réorganisé autrement, on trouve bien

$$P(X = i \cap Y = j) = \left(\frac{5}{6}\right)^{i-1} \left(\frac{5}{6}\right)^{i-1} \frac{1}{6} \frac{5}{6} \left(\frac{5}{6}\right)^{j-i-1} \frac{1}{6} = \left(\frac{5}{6}\right)^{i-1} \frac{1}{6} \left(\frac{5}{6}\right)^{j-1} \frac{1}{6} = P(X = i)P(X = j)$$

· Si 0 < j < i:

$$P(X = i \cap Y = j) = \underbrace{\left(\frac{25}{36}\right)^{j-1}}_{\text{ni 6 ni double jusqu'à } j/\text{ double à } j, \text{ distinct de 6 } / \text{ pas de 6 jusqu'à } j - 1}_{\text{6 à } j} \underbrace{\left(\frac{5}{6}\right)^{j-i-1}}_{\text{6 à } j}$$

De même qu'avant, on trouve

$$P(X = i \cap Y = j) = P(X = i)P(X = j)$$

· Si 0 < i = j:

$$P(X = i \cap Y = j) = \underbrace{\left(\frac{25}{36}\right)^{j-1}}_{\text{ni 6 ni double jusqu'à } j / \text{ double 6 à } j /} = \left(\frac{5}{6}\right)^{j-1} \frac{1}{6} \left(\frac{5}{6}\right)^{j-1} \frac{1}{6}$$

Or, i = j, on retrouve donc bien

$$P(X = i \cap Y = j) = P(X = i)P(X = j)$$

Conclusion, X et Y sont bien indépendantes.

Propriété 10

X et Y sont indépendantes ssi l'une des assertions suivantes est vérifiée

- $X_{/Y=y}$ a même loi que X pour tout $y \in Supp(Y)$ de probabilité non nulle.
- la loi de $X_{/Y=y}$ est indépendante de y.

Démonstration :

Première assertion :

pour tout $x \in Supp(X)$, on a

$$P(X = x \cap Y = y) = P_{Y=y}(X = x)P(Y = y)$$

Ainsi, si les variables $X_{/Y=y}$ et X ont même loi, on obtient

$$P(X = x \cap Y = y) = P(X = x)P(Y = y)$$
, donc X et Y sont indépendantes

et si les variables X et Y sont indépendantes, alors par simplification de

$$P(X = x)P(Y = y) = P_{Y=y}(X = x)P(Y = y)$$

on obtient

$$P(X = x) = P_{Y=y}(X = x)$$
 donc $X_{/Y=y}$ et X ont même loi.

• Deuxième assertion :

Si X et Y sont indépendantes, alors clairement $P_{Y=y}(X=x) = \frac{P(X=x \cap Y=y)}{P(Y=y)} = P(X=x)$ est indépendante de la valeur y de Y.

Si $P_{Y=y}(X=x)$ est indépendant de y, notons $u_i = P_{Y=y}(X=x_i)$ pour tout i et $y \in Supp(Y)$. Alors, pour tout i, on a

$$P(X = x_i, Y = y_j) = u_i P(Y = y_j)$$

En somment sur $j \in \mathbb{N}$, on obtient par formule des probas totales :

$$P(X = x_i) = u_i$$

et donc finalement

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$
 i.e. Xet Y ind.

■ Exemple 13 :

Dans l'exemple de X et Y resp. les $1^{\text{ères}}$ et $2^{\text{èmes}}$ apparition du 6 dans une série infinie de lancers indépendants, on a vu que pour tout $n \in \mathbb{N}^*$, on a

$$X_{/(Y=n)} \hookrightarrow \mathcal{U}_{\{1,\dots,n\}}$$

La loi de $X_{/(Y=n)}$ dépend donc de la valeur de Y, ce qui signifie que les variables X et ne sont pas indépendantes. (On s'en doutait un peu!!).

■ Exemple 14 :

Si on a des variables X, Y telles que pour tout $n \in Supp(Y)$,

$$X_{/(Y=n)} \hookrightarrow \mathcal{G}\left(\frac{1}{2}\right)$$

Alors X et Y sont indépendantes et de surcroit, $X \hookrightarrow \mathcal{G}\left(\frac{1}{2}\right)$.



Remarque:

Le résultat précédent est évidemment également vrai en échangeant X et Y.

Lemme 11

Soit (X,Y) un couple de v.a.d. indépendantes, alors,

- \blacksquare pour tous $a, b \in \mathbb{R}$, X a et Y b sont indépendantes.
- \blacksquare mieux, pour toutes fonctions réelles g, h, les v.a.d. g(X) et h(Y) sont indépendantes.

Démonstration :

- la première est conséquence de la deuxième.
- On note $Supp(X) = \{x_i \mid i \in \mathbb{N}\}\$ et $Supp(Y) = \{y_i \mid j \in \mathbb{N}\}.$ Alors

$$Supp(q(X)) = \{q(x_i) \mid i \in \mathbb{N}\}$$
 $Supp(h(Y)) = \{h(y_i) \mid j \in \mathbb{N}\}$

Ainsi, pour tout $k \in Supp(g(X)), l \in Supp(h(Y))$, on a

$$\begin{array}{lcl} P(g(X) = k \cap h(Y) = l) & = & P(X \in g^{-1}(k) \cap Y \in h^{-1}(l)) \\ & = & P(X \in g^{-1}(k)) P(Y \in h^{-1}(l)) \\ & = & P(g(X) = k) P(h(Y) = l) \end{array}$$

Théorème 12

Soit (X,Y) un couple de v.a.d.

- \blacksquare Si $\mathbb{E}[X]$ et $\mathbb{E}[Y]$ existent
- Si X et Y sont indépendantes,

alors $\mathbb{E}[XY]$ existe et

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Démonstration : admise.

LA RÉCIPROQUE DE L'INDÉPENDANCE EST FAUSSE!

Exemple : Soit X la variable aléatoire de loi uniforme sur $\{-1, 0, 1\}$.

$$p(X = -1) = p(X = 0) = p(X = 1) = \frac{1}{3}.$$

On pose $Y = X^2$, d'où

$$Supp(Y) = \{0, 1\}, \quad p(Y = 0) = \frac{1}{3}, \ p(Y = 1) = \frac{2}{3}$$

Y et X sont clairement non indépendants! (pour preuve, on peut par exemple vérifier que $p(X = 0 \cap Y = 0) \neq p(X = 0)p(Y = 0)$. Or, notons que XY = X, ainsi

$$\mathbb{E}[XY] = \mathbb{E}[X] = 0 = \mathbb{E}[X]\mathbb{E}[Y].$$

Corollaire

9

Soit (X,Y) un couple de v.a.d.

- \blacksquare Si $\mathbb{E}[X]$ et $\mathbb{E}[Y]$ existent
- Si X et Y sont indépendantes, alors Cov(X,Y) existe et

$$Cov(X, Y) = 0$$

Démonstration :

Comme $\mathbb{E}[X]$ et $\mathbb{E}[Y]$ existent, alors $\mathbb{E}[(X - \mathbb{E}[X])]$ et $\mathbb{E}[(Y - \mathbb{E}[Y])]$ existent.

De plus, comme X et Y sont indépendantes, alors $X - \mathbb{E}[X]$ et $Y - \mathbb{E}[Y]$ sont également indépendantes. (lemme).

D'après le théorème précédent, on sait donc que $\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$ existe, ce qui assure l'existence de la covariance.

Par simple calcul, on obtient encore une fois

$$Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

ce qui vaut 0 par indépendance de X et Y. \square

Corollaire

Soit (X,Y) un couple de v.a.d.

- Si X et Y admettent un moment d'ordre 2,
- $\blacksquare \ Si \ X \ et \ Y \ sont \ ind\'ependantes,$

alors V(X+Y) existe et

$$V(X+Y) = V(X) + V(Y)$$

Démonstration :

On utilise la décomposition de V(X+Y) grâce à la covariance :

$$V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$$

puis Cov(X, Y) = 0 par le corollaire précédent. \square

Corollaire

Soient X, Y des v.a.d.

- sur un même espace probabilisé,
- admettant chacune un moment d'ordre 2 non nul,
- \blacksquare indépendantes,

alors

$$r(X,Y) = 0.$$

Démonstration :

Si X et Y sont indépendantes, alors la covariance est nulle! \square

III Convolution

III-1

Généralités

Dans cette section, on se donne toujours X et Y deux variables aléatoires discrètes et on note

$$Supp(X) = \{x_i \mid i \in \mathbb{N}\} \quad \text{ et } \quad Supp(Y) = \{y_j \mid j \in \mathbb{N}\}$$

où les x_i (puis les y_i) sont deux à deux distincts de probabilité nulle.

But: Le but de ce paragraphe est de pouvoir déterminer la loi de la v.a.d. X + Y.

Stratégie : Utiliser la partition d'événements $\{(Y=y_j)\}_{j\in\mathbb{N}}$ ou $\{(X=x_i)\}_{i\in\mathbb{N}}$

En effet, si $x \in \mathbb{R}$, on peut écrire :

$$P(X+Y=x) = P\left(\bigsqcup_{i\in\mathbb{N}}(X+Y=x)\cap(X=x_i)\right) = P\left(\bigsqcup_{i\in\mathbb{N}}(x_i+Y=x)\cap(X=x_i)\right)$$
$$= P\left(\bigsqcup_{i\in\mathbb{N}}(Y=x-x_i)\cap(X=x_i)\right) = \sum_{i=0}^{+\infty}P\left((Y=x-x_i)\cap(X=x_i)\right)$$

Par exemple, si X et Y sont indépendantes, alors

$$P(X + Y = x) = \sum_{i=0}^{+\infty} P(Y = x - x_i) P(X = x_i)$$

Définition et Proposition

Soit (X,Y) un couple de v.a.d. de loi respectives $\mathcal{L}(X)$ et $\mathcal{L}(Y)$. On appelle *convolution* des lois de X et Y on note $\mathcal{L}(X) \star \mathcal{L}(Y) : \mathbb{R} \to \mathbb{R}$ l'application définie par

$$\mathcal{L}(X) \star \mathcal{L}(Y)(x) = \sum_{i=0}^{+\infty} P(X = x_i) P(Y = x - x_i)$$

où la série $\sum\limits_{i\in\mathbb{N}}P(X=x_i)P(Y=x-x_i)$ est convergente

Démonstration :

La série $\sum_{i\in\mathbb{N}} P(X=x_i)P(Y=x-x_i)$ est à termes positifs et on a

$$\forall i \in \mathbb{N}, \quad 0 \leqslant P(X = x_i)P(Y = x - x_i) \leqslant P(X = x_i)$$

où $P(X=x_i)$ est le terme général d'une série convergente. Ainsi, par le théorème de comparaison des séries à termes positifs, on a la convergence de $\sum_{i\in\mathbb{N}} P(X=x_i)$

$$x_i)P(Y = x - x_i)$$

(et même, de plus
$$0 \le \sum_{i=0}^{+\infty} P(X = x_i) P(Y = x - x_i) \le \sum_{i=0}^{+\infty} P(X = x_i) = 1.$$
)

Théorème 13

Soit (X,Y) un couple de v.a.d. indépendantes, alors la loi de X+Y est définie par

$$\mathcal{L}(X) \star \mathcal{L}(Y)$$
.

Démonstration :

La démonstration est le raisonnement effectué juste avant le théorème. \Box

Cas particuliers:

- Si X,Y sont des variables aléatoires discrètes indépendantes prenant des valeurs dans $\mathbb Z$: Alors :
- * $Supp(X+Y) \subset \mathbb{Z}$, d'où, si $x \in \mathbb{R} \mathbb{Z}$, on a P(X+Y=x) = 0.
- \star Si $k \in \mathbb{Z}$, on a

$$P(X+Y=k) = \sum_{j=-\infty}^{+\infty} P(X=k-j)P(Y=j)$$

- Si X,Y sont des variables aléatoires discrètes indépendantes prenant des valeurs dans $\mathbb N$: Alors :
- \star Supp $(X+Y) \subset \mathbb{Z}$, d'où, si $x \in \mathbb{R} \mathbb{N}$, on a P(X+Y=k) = 0
- \star Si $k \in \mathbb{N}$, on a

$$P(X + Y = k) = \sum_{j=0}^{+\infty} \underbrace{P(X = k - j)}_{=0 \text{ si } k - j < 0} P(Y = j)$$
$$= \sum_{j=0}^{k} P(X = x - j) P(Y = j)$$

III-2 Exemples

III.2-a) Convolution de Lois binomiales

Théorème 14

Soient $n, m \in \mathbb{N}$ et $p \in [0, 1]$. Alors $\mathcal{B}(n, p) \star \mathcal{B}(m, p) = \mathcal{B}(n + m, p)$.

Ou autrement dit

Théorème 15

Soient $n, m \in \mathbb{N}$ et $p \in [0; 1]$. Si X et Y sont deux v.a.d. indépendantes telles que

$$X \hookrightarrow \mathcal{B}(n,p)$$
 et $Y \hookrightarrow \mathcal{B}(m,p)$

alors

$$X + Y \hookrightarrow \mathcal{B}(n + m, p)$$

Démonstration :

- Par le calcul : On se donne $X \hookrightarrow \mathcal{B}(n,p)$ et $Y \hookrightarrow \mathcal{B}(m,p)$ deux v.a.d. indépendantes $Supp(X+Y) = \{0,1,\ldots,n+m\}$. Ainsi,
- * Si $x \notin Supp(X + Y)$, on a P(X + Y = x) = 0

* Si
$$x \in Supp(X+Y)$$
, on a $P(X+Y=x)$ =
$$\sum_{j=0}^{+\infty} \underbrace{P(X=x-j)}_{=0 \text{ si } x-j<0} P(Y=j)$$
$$= \sum_{j=0}^{x} P(X=x-j) P(Y=j)$$

On rappelle que si on n'a pas $0 \leq K \leq N$, la quantité $\binom{N}{K}$ est nulle. En utilisant cette convention, on peut écrire

$$P(X + Y = x) = \sum_{j=0}^{x} {n \choose x-j} p^{x-j} (1-p)^{n-x+j} {m \choose j} p^{j} (1-p)^{m-j}$$

$$= \sum_{j=0}^{x} {n \choose x-j} {m \choose j} p^{x} (1-p)^{n+m-x} = p^{x} (1-p)^{n+m-x} \sum_{j=0}^{x} {n \choose x-j} {m \choose j}$$

$$= {x \choose n+m} p^{x} (1-p)^{n+m-x}$$

$$= {n \choose n+m} p^{x} (1-p)^{n+m-x}$$
(Vandermonder)

• On aurait également pu établir ce résultat de manière en expliquant que $X \hookrightarrow \mathcal{B}(n,p)$ modélise par exemple n expériences de Bernoulli indépendantes, dont le succès a comme probabilité p.

De même, $Y \hookrightarrow \mathcal{B}(m,p)$ indépendante de X modélise les mêmes expériences de Bernoulli, mais m fois.

Ainsi, "en tout", X+Y modélise n+m expériences de Bernoulli de probabilité de succès p. \square

III.2-b) Convolution de lois de Poisson

Théorème 16

Soient $\lambda, \mu > 0$. Alors $\mathcal{P}(\lambda) \star \mathcal{P}(\mu) = \mathcal{P}(\lambda + \mu)$.

autrement dit

Corollaire

Soient $\lambda, \mu > 0$. Si X et Y sont deux v.a.d. indépedantes telles que

$$X \hookrightarrow \mathcal{P}(\lambda)$$
 et $Y \hookrightarrow \mathcal{P}(\mu)$

alors

$$X + Y \hookrightarrow P(\lambda + \mu)$$

Démonstration :

On se donne X et Y sont deux v.a.d. indépedantes telles que

$$X \hookrightarrow \mathcal{P}(\lambda)$$
 et $Y \hookrightarrow \mathcal{P}(\mu)$

alors

- $Supp(X + Y) = \mathbb{N}$.
- \star Si $x \notin Supp(X+Y) = \mathbb{N}$, on a

$$P(X+Y=x)=0$$

* Si $x \in Supp(X + Y) = \mathbb{N}$, on a

$$P(X + Y = x) = \sum_{j=0}^{+\infty} \underbrace{P(X = x - j)}_{=0 \text{ si } x - j < 0} P(Y = j)$$

$$= \sum_{j=0}^{x} P(X = x - j) P(Y = j)$$

$$= \sum_{j=0}^{x} e^{-\lambda} \frac{\lambda^{x - j}}{(x - j)!} e^{-\mu} \frac{\mu^{j}}{j!}$$

$$= e^{-\lambda - \mu} \sum_{j=0}^{x} \frac{\lambda^{x - j}}{(x - j)!} \frac{\mu^{j}}{j!}$$

$$= \frac{e^{-\lambda - \mu}}{x!} \sum_{j=0}^{x} {x \choose j} \lambda^{x - j} \mu^{j}$$

$$= \frac{e^{-(\lambda + \mu)}}{x!} (\lambda + \mu)^{x}$$

On a donc bien $X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$ \square