Chapitre n°7

Couple de variable aléatoires discrètes 2

Dans tout la suite les variables aléatoires discrètes sont définies sur un même espace probabilisé $(\Omega,\ \mathcal{A},\ \mathrm{P})$

I. Fonctions de variables aléatoires discrètes

Proposition 1.1

Soit X et Y deux variables aléatoires. Soit f une fonction à valeurs réelles telle que $f(x, y)$ soit défini pour tout $x \in X(\Omega)$ et pour tout $y \in Y(\Omega)$.
Alors $f(X, Y)$ est une variable aléatoire discrète définie sur Ω.

Exercice 1

Dans chacun des cas suivants expliciter qu'elle la fonction f

- $\max(X, Y)$ est une vad on pose $f(.,.) =$
- $\min(X, Y)$ est une vad on pose $f(.,.) =$
- $X + Y$
- XY

Théorème 1.2 — Théorème de transfert

Soit $(X,\ Y)$ un couple de variables aléatoires discrètes et $g(x, y)$ une fonction définie sur l'ensemble $X(\Omega) \times Y(\Omega)$.
Alors

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

sous réserve que cette dernière série converge absolument.

Méthode :

- **Calculer la loi d'un maximum :**
  Pour calculer la loi de $\max(X, Y)$ on peut remarquer que pour tout réel x

  $$[\max(X, Y) \leqslant x] = [X \leqslant x] \cap [Y \leqslant x]$$

  Si de plus X et Y sont **indépendantes** alors $\mathrm{P}(\max(X, Y) \leqslant x) = \mathrm{P}(X \leqslant x)\,\mathrm{P}(Y \leqslant x)$.
  Ce qui permet de calculer la fonction de répartition.
- **Calculer la loi d'un minimum :**
  Pour calculer la loi de $\min(X, Y)$ on peut remarquer que pour tout réel x

  $$[\min(X, Y) > x] = [X > x] \cap [Y > x]$$

  Si de plus X et Y sont **indépendantes** alors $\mathrm{P}(\min(X, Y) > x) = \mathrm{P}(X > x)\,\mathrm{P}(Y > x)$.
  Et en utilisant $\mathrm{P}(\square \leqslant x) = 1 - \mathrm{P}(\square > x)$ on retrouve la fonction de répartition.

Exercice 2

Soit X et Y deux variables aléatoires discrètes indépendantes. suivant la loi uniforme sur $[\![1, n]\!]$. Calculer la loi du maximum et du minimum de X et Y

Méthode :

Pour calculer la loi d'une somme $X+Y$:

- On remarque que pour tout entier naturel

$$[X+Y=n] = \bigcup_{k=0}^{n} ([X=k]\cap[Y=n-k]) = \bigcup_{k=0}^{n} ([X=n-k]\cap[Y=k]) = \bigcup_{\substack{0\leqslant i\leqslant n\\ 0\leqslant j\leqslant n\\ i+j=n}} ([X=i]\cap[Y=j])$$

- Ces événements entre parenthèses sont incompatibles deux-à-deux.
- Si de plus X et Y sont **indépendantes** alors

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

Attention:

Il faut adapter les indices des $\sum$ précédentes au support de X et Y

Lemme 1.3 — Égalité de Vandermonde

Soit m,n et k trois entiers naturels alors

$$\sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j} = \binom{n+m}{k}$$

Démonstration. □

Théorème 1.4

- **Somme de deux lois binomiales « avec le même p »** : Soit n_1 et n_2 deux entiers naturels et $p \in]0,1[$. Soit $X \hookrightarrow \mathcal{B}(n_1, p)$ et $Y \hookrightarrow \mathcal{B}(n_2, p)$. On suppose de plus que X et Y sont indépendantes. Alors

$$X+Y \hookrightarrow \mathcal{B}(n_1+n_2, p)$$

- **Somme de deux lois de Poisson** : Soit λ_1 et λ_2 deux réels strictement positifs.
Soit $X \hookrightarrow \mathcal{P}(\lambda_1)$ et $Y \hookrightarrow \mathcal{P}(\lambda_2)$. On suppose de plus que X et Y sont indépendantes.
Alors

$$X+Y \hookrightarrow \mathcal{P}(\lambda_1+\lambda_2)$$

- Soit $X_1, X_2, ..., X_n$ des variables aléatoires indépendantes suivant une loi de Bernoulli $\mathcal{B}(p)$, alors

$$\sum_{i=1}^{n} X_i \hookrightarrow$$

Démonstration. À savoir faire □

II. Covariance d'un couple de variables aléatoires

II. 1 Indépendance et conséquences

Définition 2.1 — Rappel : indépendance de deux variables aléatoires discrètes

Soit X et Y deux variables aléatoires discrètes définies sur un même espace. probabilisé $(\Omega, \mathcal{A}, \mathrm{P})$. On dit que X et Y sont *indépendantes* si et seulement si

$$\forall x \in X(\Omega) \quad \forall y \in Y(\Omega) \qquad \mathrm{P}([X = x] \cap [Y = y]) = \mathrm{P}([X = x])\,\mathrm{P}([Y = y])$$

Proposition 2.2 — Rappel espérance d'une somme : linéarité de l'espérance

Soit X et Y deux variables aléatoires discrètes [a] admettant une espérance alors $X + Y$ admet une espérance et

$$E(X + Y) = E(X) + E(Y)$$

Plus généralement si α et β sont des réels $\alpha X + \beta Y$ admet une espérance et

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

[a] **non nécessairement indépendantes**

Démonstration. À savoir faire □

Proposition 2.3 — Espérance d'un produit

Soit X et Y deux variables aléatoires discrètes **indépendantes** admettant une espérance telle que XY admet une espérance alors

$$E(XY) = E(X)E(Y)$$

Attention:

Quelle sont les différences entre les hypothèses de ces deux théorèmes ?

Démonstration. □

II. 2 Covariance

Définition 2.4 — Covariance

Soit X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2. On note alors

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

Exercice 3

Soit X et Y deux variables aléatoires dont la loi du couple est

X/Y	1	2	3	loi de X
1	1/6	1/6	1/6	1/2
2	1/6	1/6	1/6	1/2
loi de Y	1/3	1/3	1/3	1

Calculer la covariance.

Proposition 2.5 — Propriétés de la covariance

Soit X, Y et Z trois variables aléatoires admettant des moments d'ordre 2. Soit α et β deux réels.

1. **Symétrie** : $\text{Cov}\,(X,Y) = \text{Cov}\,(Y,X)$.
2. **La variance comme covariance** $\text{Cov}\,(X,X) = V(X)$
3. **Linéarité à gauche** $\text{Cov}\,(\alpha X + \beta Y, Z) = \dots$
4. **Linéarité à droite** $\text{Cov}\,(X, \alpha Y + \beta Z) = \dots$
5. Les deux dernières propriétés peuvent être regroupées en **bilinéarité**.

Exercice 4 — Calculs classiques

X et Y sont deux variables aléatoires discrètes admettant un moment d'ordre 2. Simplifier les expressions suivantes

1. $\text{Cov}\,(X+Y, X+Y)$
2. $\text{Cov}\,(X+Y, X-Y)$
3. $\text{Cov}\,(X-Y, X-Y)$

Proposition 2.6 — Variable presque certaine

Soit c une constante (ou une variable aléatoire presque certaine égale à c). Soit X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2.

- $\text{Cov}\,(X, c) = 0$
- $\text{Cov}\,(X+c, Y) = \text{Cov}\,(X,Y)$

Théorème 2.7 — Formule de Huygens

Soit X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2. Alors

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

Démonstration. □

Proposition 2.8 — Lien entre covariance et indépendance

Soit X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2. Si X et Y sont indépendantes alors

$$\text{Cov}\,(X,Y) = 0$$

Attention:

La réciproque est fausse !!

Exercice 5

Soit X et Y deux variables aléatoires dont la loi du couple est donnée ci-dessous. Calculer la covariance de (X,Y).

X/Y	-1	1	loi de X
0	$1/3$	$1/3$	
1	0	$1/3$	
loi de Y			1

Proposition 2.9 — Variance d'une somme

Soit X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2. Soit α et β deux réels

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

$$V(\alpha X+\beta Y) = \alpha^2 V(X) + \beta^2 V(Y) + 2\alpha\beta\text{Cov}\,(X,Y)$$

Remarques :

R1 – On sait que si X et Y sont indépendantes, alors $\text{Cov}\,(X,Y) = 0$, ainsi :

R2 – On peut généraliser ce résultat. Si $X_1, ..., X_n$ sont n variables aléatoires indépendantes :

Exercice 6

Soient X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2.

1. Soit $t \in \mathbb{R}$. Montrer que $V(X+tY) = V(X) + 2\,\text{cov}(X,Y)t + V(Y)t^2$.
2. On note $P : t \mapsto V(X+tY)$. Justifier que, pour tout $t \in \mathbb{R}$, $P(t) \geq 0$. En déduire, à l'aide d'une considération sur un discriminant, que

$$\left|\frac{\text{Cov}\,(X,Y)}{\sigma(X)\sigma(Y)}\right| \leq 1$$

Remarque :

La deuxième question permet d'énoncer le résultat suivant, appelé *Inégalité de Cauchy-Schwarz* :

Définition 2.10 — Coefficient de corrélation linéaire

Soit X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2 et un écart-type non nul.
On définit le coefficient de corrélation linéaire

$$\rho(X,Y) = \frac{\text{Cov}\,(X,Y)}{\sigma_X \sigma_Y}$$

Où σ_X et σ_X sont les écarts types de X et Y.

Proposition 2.11 — Cauchy-Schwarz

Soit X et Y deux variable aléatoire admettant un moment d'ordre 2 alors

$$|\rho(X,Y)| \leqslant 1$$

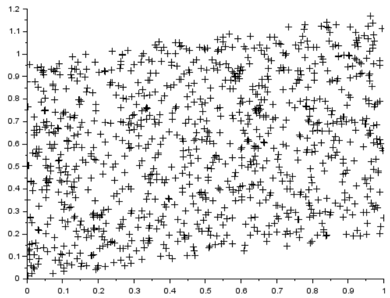
De plus il y' a égalité dans l'inégalité précédente si et seulement si il existe a et b deux constantes telles que

$$X = aY + b \qquad \text{ou } Y = aX + b$$

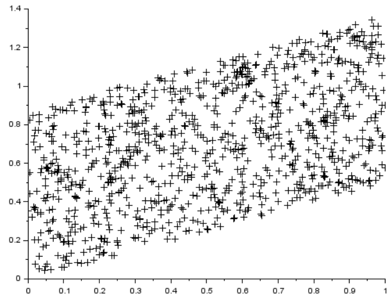
Exemple :

Soit X et Y deux variables indépendantes de loi uniforme sur $[0,1]$. On pose $Z = \rho X + \sqrt{1-\rho^2}Y$, $\rho \in]-1,1[$.

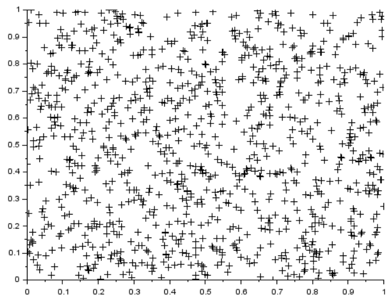
1. On calcule le coefficient de corrélation $\rho(X,Z)$.
2. Voici, selon les valeurs différentes valeurs de ρ, le nuage de points (X,Z).



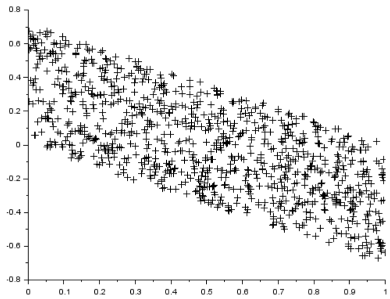
$\rho = 0,2$



$\rho = 0,5$



$\rho = 0$



$\rho = -0,7$

Remarque :

Autrement dit, le coefficient de corrélation linéaire de X et Y est un réel compris entre -1 et 1 : $-1 \leq \rho_{X,Y} \leq 1$

- Il est égal à 1 dans le cas où l'une des variables est fonction affine croissante de l'autre variable.
- Il est égal à -1 dans le cas où l'une des variables est fonction affine décroissante de l'autre variable.
- Les valeurs intermédiaires renseignent sur le degré de dépendance linéaire entre les deux variables.
- Plus le coefficient $\rho_{X,Y}$ est proche des valeurs extrêmes -1 et 1 , plus la corrélation entre les variables est forte.
- Deux variables aléatoires dont la covariance est nulle (et donc le coefficient de corrélation linéaire est nul) sont dites **non corrélées.**
- Une corrélation positive ($\rho_{X,Y} > 0$) indique que les variables X et Y varient dans le même sens.
- Une corrélation négative ($\rho_{X,Y} < 0$) indique que les variables X et Y varient en sens inverse.

III. Suites de variables aléatoires discrètes.

Dans cette partie toutes les variables aléatoires sont discrètes et définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathrm{P})$

III. 1 Indépendance

Définition 3.1 — indépendance mutuelle de n variables aléatoires discrètes

Soit X_1, ..., X_n n variables aléatoires. On dit qu'elles sont **mutuellement indépendantes** si et seulement si

$$\forall x_1 \in X_1(\Omega),\ \forall \qquad \mathrm{P}\left(\bigcap [X_i = x_i]\right) =$$

Définition 3.2

Soit $(X_i)_{i\in\mathbb{N}}$ une suite infinie de variables aléatoires discrètes. On dit que ces variables aléatoires sont mutuellement indépendantes si et seulement si pour tout partie $I \subset \mathbb{N}$ finie les variables $(X_i)_{i\in I}$ sont mutuellement indépendantes au sens précédent.

Proposition 3.3 — Lemme des coalitions

- Si X et Y sont deux variables aléatoires discrètes indépendante et si f et g sont deux fonctions numériques définies respectivement sur $X(\Omega)$ et $Y(\Omega)$ alors $f(X)$ et $g(Y)$ sont indépendantes.
- Soient $X_1,\dots,X_n$ mutuellement indépendantes et soit $p \in [\![2, n-1]\!]$. Alors toute variable aléatoire fonction des variables $X_1,\cdots,X_p$ est indépendante de toute variable aléatoire fonction des variables $X_{p+1},\cdots,X_n$.

Exemple :

Si X_1 X_2 et X_3 sont mutuellement indépendantes alors X_1 est indépendante de $\max(X_2, X_3)$

Proposition 3.4 — Espérance et variance d'une somme

Soit X_1, X_2, … X_n, n variables aléatoires qui admettent des espérances . Alors $(X_1+X_2+\cdots+X_n)$ admet une espérance et

$$E(X_1+X_2+\cdots+X_n) =$$

Si de plus ces variables aléatoires sont **mutuellement indépendantes** et admettent des moment d'ordre 2, alors $X_1+X_2+\cdots+X_n$ admet un moment d'ordre 2 et :

$$V(X_1+X_2+\cdots+X_n) =$$

Proposition 3.5 — variance d'une somme : cas général

Soit X_1, X_2, … X_n, n variables aléatoires qui admettent des moment d'ordre 2

$$V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k) + 2\sum_{1\leqslant i<j\leqslant n} \operatorname{Cov}(X_i, X_j)$$

Et on sait que la somme $\displaystyle\sum_{1\leqslant i<j\leqslant n} \operatorname{Cov}(X_i, X_j)$ comporte $\displaystyle\binom{n}{2} = \frac{n(n+1)}{2}$ termes.

On peut également généraliser les résultats de stabilité vus précédemment :

- Si $X_1, X_2, ..., X_n$ sont n variables indépendantes suivant une loi de Poisson, $X_i \hookrightarrow \mathcal{P}(\lambda_i)$, alors

$$\sum_{i=1}^{n} X_i \hookrightarrow$$

- Si $X_1, X_2, ..., X_n$ sont n variables indépendantes suivant une loi Binomiale, $X_i \hookrightarrow \mathcal{B}(n_i, p)$, alors

$$\sum_{i=1}^{n} X_i \hookrightarrow$$