
Preuve de la stabilité des lois binomiales propriétés 9.11 du cours chapitre 9

Montrons que :

si X_1 et X_2 sont deux var indépendantes, $X_1 \hookrightarrow \mathcal{B}(n_1, p)$ et $X_2 \hookrightarrow \mathcal{B}(n_2, p)$ alors

$X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$.

On a : $X_1 + X_2(\Omega) = [0, n_1 + n_2]$, on cherche $\forall k \in [0, n_1 + n_2], P(X_1 + X_2 = k)$.

En utilisant le SCE ($X_1 = i$) _{$i \in [0, n_1]$} et la FPT, la loi de la somme est donnée par : $\forall k \in [0, n_1 + n_2]$,

$$\begin{aligned}
 P(X_1 + X_2 = k) &= \sum_{i \in [0, n_1], j \in [0, n_2], k=i+j} P(X_1 = i \cap X_2 = j) \\
 &= \sum_{i \in [0, n_1], k-i \in [0, n_2]} P(X_1 = i \cap X_2 = k-i) \text{ donc contrainte sur } i \\
 &= \sum_{i=0}^k P(X_1 = i \cap X_2 = k-i) \text{ car } k-i \geq 0 \Leftrightarrow i \leq k \\
 &= \sum_{i=0}^k P(X_1 = i)P(X_2 = k-i) \text{ car } X_1, X_2 \text{ indépendantes} \\
 &= \sum_{i=0}^k \binom{n_1}{i} p^i q^{n_1-i} \binom{n_2}{k-i} p^{k-i} q^{n_2-k+i} \text{ car } X_1 \hookrightarrow \mathcal{B}(n_1, p), X_2 \hookrightarrow \mathcal{B}(n_2, p) \\
 &= p^k q^{n_1+n_2-k} \sum_{i=0}^k \binom{n_1}{i} \binom{n_2}{k-i}
 \end{aligned}$$

On utilise ensuite la formule de Vandermonde (voir à la suite pour la preuve de cette formule) :

$$\sum_{i=0}^k \binom{n_1}{i} \binom{n_2}{k-i} = \binom{n_1 + n_2}{k}.$$

Ainsi : $\forall k \in [0, n_1 + n_2], P(X_1 + X_2 = k) = \binom{n_1+n_2}{k} p^k q^{n_1+n_2-k}$.

Ce qui prouve que : $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$.

Preuve de la formule de Vandermonde

On va montrer par récurrence sur $n \in \mathbb{N}, \forall (m, p) \in \mathbb{N}^2, \sum_{k=0}^p \binom{m}{k} \binom{n}{p-k} = \binom{m+n}{p}$.

Pour $n \in \mathbb{N}$, on note $\mathcal{P}(n)$ l'assertion : $\forall (m, p) \in \mathbb{N}^2, \sum_{k=0}^p \binom{m}{k} \binom{n}{p-k} = \binom{m+n}{p}$.

NB : le coefficient binomial est nul si $0 \leq k \leq n$ n'est pas vérifiée.

Initialisation : $n = 0$, alors forcément $k = p$ et on a :

$$\binom{m}{p} \binom{0}{0} = \binom{m}{p} = \binom{m+0}{p}$$

$\mathcal{P}(0)$ est vraie.

Hérédité : on suppose que $\mathcal{P}(n)$ est vraie pour un certain entier n .

Soit $(m, p) \in \mathbb{N}^2$, alors si $p = 0$,

$$\sum_{k=0}^p \binom{m}{k} \binom{n+1}{p-k} = \binom{m}{0} \binom{n+1}{0} = 1 = \binom{m+n+1}{0}.$$

Sinon par la relation de Pascal :

$$\sum_{k=0}^p \binom{m}{k} \binom{n+1}{p-k} = \sum_{k=0}^p \binom{m}{k} \left[\binom{n}{p-k} + \binom{n}{p-k-1} \right] = \sum_{k=0}^p \binom{m}{k} \binom{n}{p-k} + \sum_{k=0}^p \binom{m}{k} \binom{n}{p-k-1}.$$

On utilise alors l'hypothèse de récurrence pour le triplet (n, m, p) et pour le triplet $(n, m, p-1)$:

$$\sum_{k=0}^p \binom{m}{k} \binom{n+1}{p-k} = \binom{m+n}{p} + \binom{m+n}{p-1}.$$

Puis de nouveau par la relation de Pascal :

$$\sum_{k=0}^p \binom{m}{k} \binom{n+1}{p-k} = \binom{m+n+1}{p}.$$

D'où $\mathcal{P}(n+1)$ est vraie.

Conclusion : par principe de récurrence, $\mathcal{P}(n)$ est vraie pour tout entier n .

**Preuve des 6 et 7 de la propriété 9.19 et preuve des théorèmes 9.21 et 9.22 du cours
chapitre 9**

On considère deux variables aléatoires X et Y définies sur un même espace probabilisé et admettant des espérances $E(X)$ et $E(Y)$ et des variances $V(X)$ et $V(Y)$ (on rappelle que $V(X) = 0$ si et seulement si, avec une probabilité égale à 1, X est constante). La covariance des deux variables aléatoires X et Y (que celles-ci soient discrètes ou à densité) est alors le nombre réel défini par :

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))], \text{ ou encore } E(XY) - E(X)E(Y).$$

1. Covariance des variables aléatoires X et Y

- (a) Exprimer $\text{Cov}(\lambda X + Y, \lambda X + Y)$ en fonction de $V(\lambda X + Y)$ et en déduire la formule suivante pour tout nombre réel λ :

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

- (b) En déduire que $(\text{Cov}(X, Y))^2 \leq V(X)V(Y)$, puis $|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$ (INÉGALITÉ DE CAUCHY-SCHAWRZ).

A quelle condition nécessaire et suffisante a-t-on l'égalité $(\text{Cov}(X, Y))^2 = V(X)V(Y)$?

- (c) En déduire L'INÉGALITÉ DE MINKOWSKI : $\sigma(X + Y) \leq \sigma(X) + \sigma(Y)$.

2. Coefficient de corrélation linéaire des variables aléatoires X et Y .

On suppose dans cette question les variances $V(X)$ et $V(Y)$ de X et Y strictement positives.

- (a) Exprimer le coefficient de corrélation linéaire ρ des variables aléatoires X et Y en fonction de $\text{Cov}(X, Y)$ et des écarts-types $\sigma(X)$ et $\sigma(Y)$ des variables aléatoires X et Y et montrer que ρ appartient à $[-1, +1]$.

Préciser de plus à quelle condition nécessaire et suffisante ρ est égal à -1 ou $+1$.

- (b) Donner la valeur de ρ lorsque les variables aléatoires X et Y sont indépendantes.

1. Covariance des variables aléatoires X et Y

- (a) Pour tout X ayant une variance on a : $\text{Cov}(X, X) = E(X^2) - E(X)^2 = V(X)$
donc $\text{Cov}(\lambda X + Y, \lambda X + Y) = V(\lambda X + Y)$ et

$$\begin{aligned} V(\lambda X + Y) &= E((\lambda X + Y)^2) - E(\lambda X + Y)^2 \\ &= E(\lambda^2 X^2 + 2\lambda XY + Y^2) - [\lambda E(X) + E(Y)]^2 \\ &= \lambda^2 E(X^2) + 2\lambda E(XY) + E(Y^2) \\ &\quad - \lambda^2 E(X)^2 - 2\lambda E(X)E(Y) - E(Y)^2 \\ &= \lambda^2 V(X) + 2\lambda \text{Cov}(X, Y) + V(Y) \end{aligned}$$

- (b) Comme une variance est positive ou nulle, le polynôme du second degré $P(\lambda) = \lambda^2 V(X) + 2\lambda \text{Cov}(X, Y) + V(Y)$ a un discriminant négatif ou nul.

$$\begin{aligned} \Delta &= 4\text{Cov}(X, Y)^2 - 4V(X)V(Y) \\ &= 4[\text{Cov}(X, Y)^2 - V(X)V(Y)] \leq 0 \end{aligned}$$

Donc $\text{Cov}(X, Y)^2 \leq V(X)V(Y)$, on en déduit immédiatement par stricte croissance de la fonction racine carrée alors que $\sqrt{\text{Cov}(X, Y)^2} \leq \sqrt{V(X)V(Y)}$ soit, $|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$.

De plus, on a égalité si et seulement si $\Delta = 0$ donc si le polynôme s'annule, i.e. s'il existe λ_0 tel que $V(\lambda_0 X + Y) = 0$

Or (c'est le rappel de début) cela équivaut à $\lambda_0 X + Y$ constante presque sûrement.

Conclusion : Il y a égalité $\iff$ il existe $b \in \mathbb{R}$ tel que $Y = aX + b$ presque sûrement.

NB : on peut même être plus précis.

$\Delta = 0$, le polynôme s'annule, i.e. s'il existe λ_0 tel que $V(\lambda_0 X + Y) = 0$, alors on a $\lambda_0 = -\frac{\text{cov}(X, Y)}{V(X)} = -\frac{\sigma(Y)}{\sigma(X)}$ si $\text{cov}(X, Y) > 0$, ou $= \frac{\sigma(Y)}{\sigma(X)}$ si $\text{cov}(X, Y) < 0$, ou $= 0$ si $\text{cov}(X, Y) = 0$.

On a vu que $\lambda_0 X + Y$ constante presque sûrement donc il existe $b \in \mathbb{R}$ tel que : $\lambda_0 X + Y = b$ i.e. : $Y = -\lambda_0 X + b$. On posera $a = -\lambda_0$.

- (c) On a $(\sigma(X + Y))^2 - (\sigma(X) + \sigma(Y))^2 = V(X + Y) - V(X) - V(Y) - 2\sigma(X)\sigma(Y) = 2\text{cov}(X, Y) - 2\sigma(X)\sigma(Y) = 2(\text{cov}(X, Y) - \sigma(X)\sigma(Y))$.

Or d'après l'inégalité démontrée à la question précédente on a $\text{cov}(X, Y) - \sigma(X)\sigma(Y) \leq 0$ d'où : $(\sigma(X + Y))^2 - (\sigma(X) + \sigma(Y))^2 \leq 0$ c'est à dire que $\sigma(X + Y) \leq \sigma(X) + \sigma(Y)$.

2. Coefficient de corrélation linéaire des variables aléatoires X et Y .

On suppose dans cette question les variances $V(X)$ et $V(Y)$ de X et Y strictement positives.

- (a) Le coefficient de corrélation linéaire est $\rho = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$.

Et comme $(\text{Cov}(X, Y))^2 \leq V(X)V(Y)$ alors $\frac{(\text{Cov}(X, Y))^2}{V(X)V(Y)} \leq 1$ et en prenant la racine, $\left| \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)} \right| \leq 1$ (la covariance peut être négative!)

Conclusion : $\boxed{\rho \in [-1, 1]}$

Et l'on a $\rho = \pm 1 \iff \rho^2 = 1 \iff (\text{Cov}(X, Y))^2 = V(X)V(Y)$ donc à la condition nécessaire et suffisante qu'il existe une constante b telle que $Y = aX + b$ presque sûrement d'après la question 1b).

On a vu que $a = \frac{\sigma(Y)}{\sigma(X)}$ si $\text{cov}(X, Y) > 0$, donc $a > 0$ et alors $\rho = 1$.

$a = -\frac{\sigma(Y)}{\sigma(X)}$ si $\text{cov}(X, Y) < 0$ et $\rho = -1$.

- (b) Si X et Y sont indépendantes, alors $\text{Cov}(X, Y) = 0$ et donc $\rho = 0$. Attention, on peut montrer qu'il n'y a pas de réciproque!