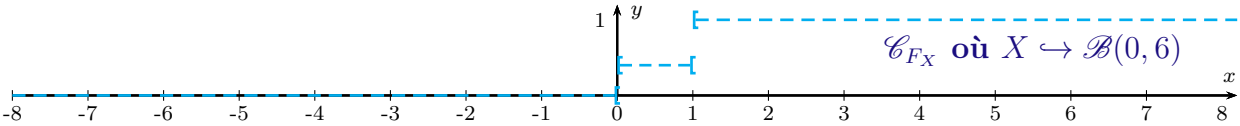
Chapitre 4 - couples et suites de variables aléatoires discrètes

Objectifs d'apprentissage - A la fin de ce chapitre, je sais :
- définir une fonction de répartition et savoir l'interpréter ☐
- définir et démontrer l'indépendance de plusieurs variables aléatoires ☐
- étudier un couple de variables aléatoires : loi de couple, lois marginales et conditionnelles ☐
- étudier une somme, un produit, un minimum ou un maximum de deux variables aléatoires ☐
- utiliser les propriétés de l'espérance et de la variance de plusieurs variables aléatoires, en particulier les cas d'indépendance et de suites de variables aléatoires ☐
- définir la covariance, le coefficient de corrélation de deux variables aléatoires et utiliser leurs propriétés ☐

Un complément sur les variables aléatoires : la fonction de répartition

Définitions et propriétés	**Exemples**
Définition : la **fonction de répartition** F d'une variable aléatoire X est la fonction : $x \mapsto P(X \leqslant x)$, définie sur $\mathbb{R}$ on la note également F_X	Remarque : F_X indique « où sont réparties » les valeurs de X « et avec quelles probabilités ». Exemple : si $X \hookrightarrow \mathscr{B}(1,p)$ $F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ p & \text{si} \quad 0 \leqslant x < 1 \\ 1 & \text{si} \quad 1 \leqslant 1 \end{cases}$
Propriété : la donnée de la fonction de répartition **caractérise la loi de la variable**.	Elle ne donne pas immédiatement l'ensemble $X(\Omega)$ mais $\forall k \in \mathbb{N}^*$, $F_X(k) - F_X(k-1) = P(X = k)$

Remarque : la fonction de répartition n'est pas continue dans le cas des variables discrètes, mais elle a de nombreuses propriétés :
- F_X est croissante sur $\mathbb{R}$ car si $x \leqslant y$ alors $[X \leqslant x] \subset [X \leqslant y]$
- $F_X(x) \xrightarrow[x \to -\infty]{} 0$ car quand $x \to -\infty$ la probabilité que X soit dans $]-\infty, x]$ tend vers 0
- $F_X(x) \xrightarrow[x \to +\infty]{} 1$ car quand $x \to +\infty$ la probabilité que X soit dans $]-\infty, x]$ tend vers 1



Etude de plusieurs variables aléatoires

Formule des probabilitées totales appliquées à des variables aléatoires

Il ne s'agit pas d'une nouvelle propriété mais d'un cas classique d'application de la formule des probabilités totales, lors de l'étude de deux variables aléatoires dépendantes.

Si X et Y sont deux variables à valeurs dans $\mathbb{N}$, alors $\{[X = j]; j \in X(\Omega)\}$ forme un système complet d'événements, donc :

$$\boxed{\forall i \in Y(\Omega), \quad P(Y = i) = \sum_{j \in X(\Omega)} P([Y = i] \cap [X = j]) = \sum_{j \in X(\Omega)} P_{[X=j]}(Y = i) P(X = j)}$$

Nota bene : la dernière égalité est valable si $\forall j \in X(\Omega), P(X = j) > 0$

1 Indépendance de variables aléatoires

Définition : deux variables aléatoires discrètes X, Y sont dites **indépendantes** si les événements $[X = x]$ et $[Y = y]$ le sont, pour tous x, y i.e. si $\forall (x, y) \in X(\Omega) \times Y(\Omega)$ $P([X = x] \cap [Y = y]) = P(X = x)P(Y = y)$	Exemple : si X et Y désignent les résultats de deux dés, $P([X = 3] \cap [Y = 5]) = P([X = 3]) \times P([Y = 5])$ $= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} = P([X = i] \cap [Y = j])$
Propriété : plus généralement : deux variables aléatoires réelles discrètes X et Y sont indépendantes si, et seulement si, pour tous intervalles I, J de $\mathbb{N}$: $P([X \in I] \cap [Y \in J]) = P(X \in I)P(Y \in J)$ En particulier, pour tous réels x et y : $P([X \leqslant x] \cap [Y \leqslant y]) = P(X \leqslant x)P(Y \leqslant y)$	Avec l'exemple précédent, $P([X \leqslant 2] \cap [Y > 3]) = P(X \leqslant 2)P(Y > 3)$ $= \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$
Définition (pour n variables) : $X_1, \ldots X_n$ sont dites **mutuellement indépendantes** si, pour tout $(x_1, \ldots x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega)$: $P\left([X_1 = x_1] \cap [X_2 = x_2] \cap \ldots [X_n = x_n]\right) = \prod_{k=1}^{n} P(X_k = x_k)$ Les propriétés précédentes (dans le cas de deux variables) s'appliquent également	Exemple : si $X_1, \ldots, X_n$ sont les résultats de n tirages avec remise, alors $X_1, \ldots, X_n$ sont mutuellement indépendantes.
Propriété : l'indépendance mutuelle implique l'indépendance deux à deux	⚠ la réciproque est fausse (cf. *exercice 11*) : trois variables X, Y, Z peuvent être deux-à-deux indépendantes sans qu'elles soient mutuellement indépendantes.

2 Couple de variables aléatoires

X et Y désignent des variables aléatoires réelles discrètes.

2.1 Lois d'un couple, loi marginale, loi conditionnelle

Définitions : la **loi d'un couple** (X, Y) est la donnée des probabilités $\mathbf{P}([X = x] \cap [Y = y])$ pour $x \in X(\Omega)$ et $y \in Y(\Omega)$ étant donnée la loi du couple (X, Y), **la loi marginale** de X est donnée par $\forall x \in X(\Omega)$: $P(X = x) = \sum_{y \in Y(\Omega)} P([X = x] \cap [Y = y])$ *on définit de même pour Y* la **loi conditionnelle de Y sachant** $[X = x]$ avec $x \in X(\Omega)$ fixé, est donnée par les valeurs $P_{[X=x]}(Y = y)$, où $y \in Y(\Omega)$ *de même pour X sachant $[Y = y]$*	Exemples de couples : taille et poids d'un individu, somme et produit des résultats du lancer de deux dés... • la loi marginale est une formule des probabilités totales appliquée avec le S.C.E. $(Y = y)_{y \in Y(\Omega)}$ • la loi conditionnelle peut être utilisée par exemple dans le cas d'une « succession de deux lois » (loi de Poisson puis loi binomiale). • lorsque $X(\Omega)$ et $Y(\Omega)$ sont finis, on peut dresser $\boxed{\text{le tableau de la loi du couple}}$, par exemple, avec le lancer de deux dés, où X est la plus petite valeur, Y la plus grande.

3 Opérations classiques sur les variables aléatoires

Les sections somme, minimum et maximum donnent la méthodologie sur ces cas classiques.

3.1 Sommes de variables aléatoires

Si X_1 et X_2 sont à valeurs dans $\mathbb{N}$, pour déterminer la loi de $X_1 + X_2$, on peut procéder ainsi :

• pour $i \in \mathbb{N}$, écrire $[X_1 + X_2 = i]$ comme une réunion d'événements disjoints en s'appuyant sur le S.C.E. assoocié à X_1 :

$$[X_1 + X_2 = i] = \bigcup_{j\in\mathbb{N}} \Big([X_1 = j] \cap [X_1 + X_2 = i]\Big) = \bigcup_{j\in\mathbb{N}} \Big([X_1 = j] \cap [X_2 = i - j]\Big)$$

• passer aux probabilité ; par incompatibilité :

$$P(X_1 + X_2 = i) = P\left(\bigcup_{j\in\mathbb{N}} \Big([X_1 = j] \cap [X_2 = i - j]\Big)\right) = \sum_{j=0}^{+\infty} P([X_1 = j] \cap [X_2 = i - j])$$

• comme X_2 est à valeurs dans $\mathbb{N}$, $P(X_2 < 0) = 0$, donc $P(X_2 = i - j) = 0$ pour $j > i$ et :

$$P(X_1 + X_2 = i) = \sum_{j=0}^{i} P([X_1 = j] \cap [X_2 = i - j])$$

Cas particulier : si X_1 et X_2 sont indépendantes, on obtient :

$$P(X_1 + X_2 = i) = \sum_{j=0}^{i} P(X_1 = j)P(X_2 = i - j)$$

Remarques : on devra refaire à chaque fois le calcul, et l'adapter aux situations. ⚠ en particulier quand X commence à 1 (par exemple si X désigne le numéro d'un tirage) et de même la somme sur j peut s'arrêter à $i - 1$

3.2 Stabilité des lois binomiales et de Poisson

Propriété - somme de deux variables indépendantes suivant des lois binomiales : si X_1 et X_2 sont deux variables aléatoires indépendantes suivant les loi binomiales : $X_1 \hookrightarrow \mathscr{B}(n_1, p)$ et $X_2 \hookrightarrow \mathscr{B}(n_2, p)$, alors : $X_1 + X_2 \hookrightarrow \mathscr{B}(n_1 + n_2, p)$	Si X_1 donne le nombre de 6 réalisés lors de n_1 lancers d'un dé et X_2 le nombre de 3 lors de n_2 lancers, alors $X_1 \hookrightarrow \mathscr{B}(n_1, 1/6)$ et $X_2 \hookrightarrow \mathscr{B}(n_2, 1/6)$ et donc $X_1 + X_2$ le nombre total de 6 et de 3 réalisés lors de ces deux expériences suit la loi $\mathscr{B}(n_1 + n_2, 1/6)$
Propriété - somme de deux variables indépendantes suivant des lois de Poisson : si X_1 et X_2 sont deux variables aléatoires indépendantes suivant les loi de Poisson, $X_1 \hookrightarrow \mathscr{P}(\lambda_1)$ et $X_2 \hookrightarrow \mathscr{P}(\lambda_2)$, alors : $X_1 + X_2 \hookrightarrow \mathscr{P}(\lambda_1 + \lambda_2)$	Si X_1, le nombre de saumons, et X_2, le nombre d'aloses, qui passent dans une passe à poissons au cours d'une heure suivent les lois : $X_1 \hookrightarrow \mathscr{P}(5)$ et $X_2 \hookrightarrow \mathscr{P}(2)$ alors $X_1 + X_2$ le nombre total de poissons (saumons + aloses) suit la loi $\mathscr{P}(7)$

3.3 Loi d'un minimum, d'un maximum

L'idée de la méthode est de **passer par des inégalités** : pour que le maximum de deux valeurs soit inférieur ou égal à n, il faut et il suffit que les deux valeurs le soient (et supérieur pour le minimum).

La méthode s'appuie alors sur le résultat suivant qu'il faut pouvoir redémontrer :

Si X est une variable à valeurs dans $\mathbb{N}$, alors :

$$\forall k \in \mathbb{N}, \quad P(X = k) = P(X \leqslant k) - P(X \leqslant k - 1) = P(X > k - 1) - P(X > k)$$

En effet, $[X \leqslant k - 1] \cup [X = k] = [X \leqslant k]$ puis par incompatibilité des événements.

Remarque : ce résultat est le même que $F_X(k) - F_X(k-1) = P(X = k)$

Méthode : on suppose que X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ et on pose :

$$Z = \min(X, Y) \qquad T = \max(X, Y)$$

- loi de Z : elle s'obtient en déterminant $Z(\Omega)$ et en utilisant les relations :
  (a) $\forall n \in \mathbb{N}, \quad [Z > n] = [X > n] \cap [Y > n]$ *remarque préliminaire*
  (b) $\forall n \in \mathbb{N}, \quad P(Z = n) = P(Z > n-1) - P(Z > n)$ *résultat préliminaire*
  puis on utilise le résultat du **a.** pour les calculs de $P(Z > n-1)$ et $P(Z > n)$
- loi de T : elle s'obtient en déterminant $T(\Omega)$ et en utilisant les relations :
  (a) $\forall n \in \mathbb{N}, \quad [T \leqslant n] = [X \leqslant n] \cap [Y \leqslant n]$ *remarque préliminaire*
  (b) $\forall n \in \mathbb{N}, \quad P(T = n) = P(T \leqslant n) - P(T \leqslant n-1)$ *résultat préliminaire*
  puis on utilise le le résultat du **a.** pour les calculs de $P(T \leqslant n-1)$ et $P(T \leqslant n)$

Se référer à l'*exercice 14* (exercice type).

3.4 Produit

La loi d'un produit XY se détermine avec la loi d'un couple. Le seul résultat à connaître est le celui sur $E(XY)$ quand X et Y sont indépendantes (cf. plus bas).

3.5 Lemme des coalitions

Propriété : si $X_1, \ldots, X_n$ sont n variables aléatoires discrètes indépendantes, toute variable aléatoire fonction de $X_1, \ldots, X_p$ $(p < n)$ est indépendante de toute variable aléatoire fonction de $X_{p+1}, \ldots, X_n$	Exemple : si X, Y et Z sont des variables mutuellement indépendantes, alors $X^2 + Y^3$ est indépendante de $\ln(Z)$ (si elle existe).

4 Espérance, variance, covariance et coefficient de corrélation linéaire

Pour rappel, $E(X^r)$ si elle existe, soit d'après le théorème de transfert $\sum_{x \in X(\Omega)} x^r P(X = x)$ si elle converge absolument, est appelée moment d'ordre r de X ($r = 1$ ou $r = 2$ en général pour nous).

4.1 Espérance et variance

Propriété - linéarité de l'espérance : soit $(X_1, \ldots, X_n)$ des variables aléatoires et $(a_1, \ldots, a_n)$ des réels, alors : $$E\left(\sum_{k=1}^{n} a_k X_k\right) = \sum_{k=1}^{n} a_k E(X_k)$$	Remarque : c'est simplement un cas plus général (avec n variables) de la linéarité vue en première année
Théorème de transfert : soit X, Y deux variables aléatoires et f une fonction de deux variables définie sur un domaine contenant $X(\Omega) \times Y(\Omega)$ la variable aléatoire $Z = f(X, Y)$ admet une espérance si, et seulement si la somme $$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} f(x, y) P((X = x) \cap (Y = y))$$ est absolument convergente, et dans ce cas, la somme vaut $E(Z) = E(f(X, Y))$	Remarque : pas besoin donc dans ce cas de déterminer la loi de $f(X, Y)$ pour le calcul de son espérance, on « transfère » les lois de X et Y Exemple : si $Z = \min(X, Y)$ alors, sous réserve d'existence, $E(Z) = E(\min(X, Y))$ $= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \min(x, y) P((X = x) \cap (Y = y))$

Cas d'indépendance

Propriété : si X et Y sont **indépendantes** et admettent une espérance, alors XY admet une espérance et $$E(XY) = E(X)E(Y)$$ cette propriété s'étend à : $$E\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} E(X_k)$$ si $X_1, \ldots, X_n$ sont mutuellement indépendantes.	Si X_1 et X_2 donnent les résultats du lancer de deux dés alors, les deux résultats étant indépendants, $E(X_1X_2) = E(X_1)E(X_2) = \frac{7}{2}\frac{7}{2} = \frac{49}{4}$
Propriété : si X et Y sont **indépendantes** et admettent un moment d'ordre 2, alors $X+Y$ admet un moment d'ordre 2 et $$V(X+Y) = V(X) + V(Y)$$ cette propriété s'étend à : $$V\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} V(X_k)$$ si $X_1, \ldots, X_n$ admettent des moments d'ordre 2 et sont mutuellement indépendantes.	On le démontre avec la formule de Kœnig-Huygens (et la linéarité de l'espérance) $V(X+Y) = E\left[(X+Y)^2\right] - [E(X+Y)]^2$ $= E(X^2 + 2XY + Y^2) - [E(X) + E(Y)]^2$ $= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2$ $= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2 + 2(E(XY) - E(X)E(Y))$ soit $V(X+Y) = V(X) + V(Y)$ car par indépendance $E(XY) = E(X)E(Y)$

Variables centrées, réduites, centrées réduites

Pour rappel, si $E(X) = 0$, la variable est dite centrée et si $V(X) = 1$ la variable est dite réduite.

Propriété : avec X une variable aléatoire, on note $\mu = E(X)$ et $\sigma = \sigma(X)$ (où $\sigma(X) = \sqrt{V(X)}$ est l'écart-type), alors $X^* = \dfrac{X - \mu}{\sigma}$ est une **variable centrée réduite**.

4.2 Covariance

Définitions : la covariance d'un couple de variables aléatoires (X, Y) est, sous réserve d'existence : $\mathrm{Cov}(X,Y) = E\left((X - E(X))(Y - E(Y))\right)$ $= E(XY) - E(X)E(Y)$ si on note $\tilde{X}$ et $\tilde{Y}$ les variables centrées : $\mathrm{Cov}(X,Y) = E(\tilde{X}\tilde{Y})$	
Propriété : si X et Y admettent un moment d'ordre 2 alors $X, Y, X+Y$ admettent une variance, (X,Y) admet une covariance, et : $$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y)$$	
Propriété : si X et Y sont indépendantes, alors $\mathrm{Cov}(X,Y) = 0$ *Nota bene* : on retrouve alors $V(X+Y) = V(X) + V(Y)$	⚠ la réciproque est fausse.

Propriétés - opérations avec la covariance : pour trois variables X, Y, Z admettant un moment d'ordre 2 : • $V(X-Y) = V(X) + V(Y) - 2\,\mathrm{Cov}(X,Y)$ • $\mathrm{Cov}(X,X) = V(X)$ • $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$ • $\mathrm{Cov}(a,X) = \mathrm{Cov}(X,a) = 0$ pour $a \in \mathbb{R}$ • linéarité à gauche, pour a, b réels : $\mathrm{Cov}(aX+bY, Z) = a\,\mathrm{Cov}(X,Z) + b\,\mathrm{Cov}(Y,Z)$ • linéarité à droite, pour a, b réels : $\mathrm{Cov}(Z, aX+bY) = a\,\mathrm{Cov}(Z,X) + b\,\mathrm{Cov}(Z,Y)$	Quelques éléments de démonstration, on trouve ces résultats car : • • • • • •
Propriété - Inégalité de Schwarz : si X et Y admettent des moments d'ordre 2 $$\lvert\,\mathrm{Cov}(X,Y)\rvert \leqslant \sigma(X)\sigma(Y)$$	On peut interpréter cette propriété comme : « les variations couplées de X et Y ($\mathrm{Cov}(X,Y)$) sont moindres que le produit des variations ($\sqrt{V(X)}\sqrt{V(Y)}$) »

4.3 Coefficient de corrélation

Définition : avec X et Y deux variables aléatoires, le **coefficient de corrélation** du couple (X,Y) est : $$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$ pour X et Y de variances non nulles.	
Propriétés : • $-1 \leqslant \rho(X,Y) \leqslant 1$ ou $\lvert\rho(X,Y)\rvert \leqslant 1$ • $\lvert\rho(X,Y)\rvert = 1 \Leftrightarrow Y = aX + b$ avec $(a,b) \in \mathbb{R}^2$ • $\rho(X,Y) = \rho(Y,X)$ • $\rho(X,X) = 1$ • $\rho(aX+b, Y) = \rho(X,Y)$ pour $a > 0$ et $b \in \mathbb{R}$	
Définition et propriété : la **droite de régression linéaire** est la droite d'équation $y = ax + b$ où a et b sont choisis tels que $E((Y-(aX+b))^2)$ soit minimal. Cette droite a pour équation : $$\frac{y-\mu}{\sigma_Y} = \rho\frac{x-m}{\sigma_X}$$ où $m = E(X)$, $\mu = E(Y)$, $\rho = \rho(X,Y)$	