Couples de variables aléatoires réelles discrètes

Dans tout ce chapitre, on considèrera deux variables aléatoires discrètes X et Y définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

1 Loi d'une variable aléatoire sachant un événement

Soit A un événement de probabilité non-nulle.
La loi de X sachant A, appelée aussi loi conditionnelle de X sachant A, est la loi de X pour la probabilité P_A.
Autrement dit, c'est la donnée des probabilités $P_A(X = k)$, pour tout $k \in X(\Omega)$.

Remarque : il sera fréquent que certaines de ces probabilités soient nulles, car sachant A, X ne pourra pas toujours prendre toutes les valeurs.

Exemple : On dispose d'une pièce équilibrée, d'un dé équilibré, et d'un dé truqué dont la probabilité d'obtenir six vaut 1/3. On réalise l'expérience suivante. On lance la pièce : si on obtient pile, alors on lance le dé équilibré jusqu'à l'obtention d'un six, et si on obtient face, alors on lance le dé truqué jusqu'à l'obtention d'un six.
Soit X la variable aléatoire égale au nombre de lancers de dé effectués, et F l'événement "la pièce donne face".
Déterminer la loi de X sachant F, ainsi que la loi de X sachant $\overline{F}$. Déterminer alors la loi de X.

2 Loi d'un couple de variables aléatoires discrètes

Définition

> On appelle couple de variables aléatoires discrètes (X, Y) l'application $\Omega \to \mathbb{R}^2$
> $\omega \mapsto (X(\omega), Y(\omega))$.
> La loi du couple (X, Y) (ou loi conjointe) est caractérisée par la donnée de $X(\Omega)$, de $Y(\Omega)$ et des valeurs $P((X = i) \cap (Y = j))$ pour tout couple $(i, j) \in X(\Omega) \times Y(\Omega)$.

Remarque : La probabilité $P((X = i) \cap (Y = j))$ est aussi notée $P(X = i, Y = j)$.

Exemple :
On lance deux dés équilibrés à 4 faces distinctes et on note X_1 (resp. X_2) le résultat du premier lancer de dé (resp. du deuxième lancer). Soit $Y = max(X_1, X_2)$.
Déterminer la loi du couple (X_1, Y) . On pourra présenter les résultats dans un tableau à double entrée.

Remarques importantes :

1. Contrairement aux tableaux de loi d'une variable, $\boxed{\text{ce tableau contient des probabilités nulles}}$.
   En effet, le couple (X_1, Y) ne prend pas toutes les valeurs de $X_1(\Omega) \times Y(\Omega) = [\![1, 4]\!]^2$, puisque $Y \geq X_1$.
   En pratique, on ne cherchera pas à déterminer en amont l'ensemble $(X, Y)(\Omega)$ des valeurs prises par le couple. Comme dans l'exemple, on déterminera $X(\Omega)$ et $Y(\Omega)$ puis on étudiera la loi du couple (X, Y).
   A retenir : $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$ mais ces ensembles sont en général différents.
2. $\boxed{\text{Si on somme tous les coefficients du tableau de la loi conjointe on obtient 1}}$.
3. > Si la loi de Y sachant $(X = i)$ est connue, on utilisera la formule
   > $P((X = i) \cap (Y = j)) = P(X = i) \times P_{(X=i)}(Y = j)$ pour obtenir la loi du couple.
4. Si les variables X et Y sont $\boxed{\text{indépendantes}}$ (cf rappel ci-dessous), alors pour tout $i, j \in X(\Omega) \times Y(\Omega)$, $P((X = i) \cap (Y = j)) = P(X = i)P(Y = j)$ donc la loi du couple (X, Y) est connue dès que l'on connait les lois de chacune des variables.

Propriétés importantes d'une loi de couple :

1. pour tout $(i, j) \in X(\Omega) \times Y(\Omega)$, $P((X = i) \cap (Y = j)) \geq 0$.
2. $\displaystyle\sum_{(i,j)\in X(\Omega)\times Y(\Omega)} P((X = i) \cap (Y = j)) = 1$
   Les séries doubles n'étant pas à votre programme, on ne regardera cette propriété que dans le cas où X et Y sont finies ; il n'y aura donc pas de problème de convergence.

Indépendance (rappel, complément) : Les trois phrases suivantes sont équivalentes :

1. X et Y sont deux variables indépendantes
2. pour tout couple $(i, j) \in X(\Omega) \times Y(\Omega)$, $P((X = i) \cap (Y = j)) = P(X = i)P(Y = j)$.
3. pour tout couple $(i, j) \in X(\Omega) \times Y(\Omega)$, $P((X \leq i) \cap (Y \leq j)) = P(X \leq i)P(Y \leq j)$.

Lemme des coalitions

Soit X et Y deux variables indépendantes. Alors pour toute fonction f (resp. g) définie sur $X(\Omega)$ (resp. $Y(\Omega)$), les variables $f(X)$ et $g(Y)$ sont encore indépendantes.

3 Lois marginales

Les variables X et Y sont appelées variables marginales du couple (X, Y) et leur loi, la loi marginale de X (resp. loi marginale de Y).

Méthode pour trouver les lois marginales à partir de la loi du couple.
Supposons connue la loi du couple (X, Y) : on connait donc toutes les valeurs $P([X = i] \cap [Y = j])$ pour $(i, j) \in X(\Omega) \times Y(\Omega)$.
On cherche maintenant la loi de X donc les valeurs $P(X = i)$ pour $i \in X(\Omega)$.

La famille des événements $\{(Y = j),\ j \in Y(\Omega)\}$ forme un système complet d'événements, donc d'après la formule des probabilités totales appliquée à ce système complet d'événements, on obtient :

pour tout $i \in X(\Omega)$ fixé $\qquad P(X = i) = \sum_{j \in Y(\Omega)} P((X = i) \cap (Y = j))$

De même, la loi de Y s'obtient à l'aide de la formule des probabilités totales appliquée au système complet d'événements $\{(X = i),\ i \in X(\Omega)\}$:

pour tout $j \in Y(\Omega)$ fixé $\qquad P(Y = j) = \sum_{i \in X(\Omega)} P((X = i) \cap (Y = j))$.

Retour exemple :
Déterminer la loi marginale de X_1 (commenter !) puis celle de Y.

4 Fonctions d'un couple de variables aléatoires discrètes

4.1 Cas général

Proposition

Soit X, Y deux variable aléatoires discrètes définies sur $(\Omega, \mathcal{A})$ et soit $u : X(\Omega) \times Y(\Omega) \to \mathbb{R}$.
Alors la variable $u(X, Y)$ définie pour tout $w \in \Omega$ par $u(X, Y)(\omega) = u(X(\omega), Y(\omega))$ est une variable aléatoire discrète sur Ω.

Proposition Formule de transfert dans le cas où X et Y sont des variables finies :

Soit X et Y deux variables finies, et u une fonction définie sur $X(\Omega) \times Y(\Omega)$ et à valeurs réelles.
Alors $\boxed{\text{la variable aléatoire } u(X,Y) \text{ est finie, donc admet une espérance}}$ et
$E(u(X, Y)) = \sum_{(i,j) \in X(\Omega) \times Y(\Omega)} u(i, j) P((X = i) \cap (Y = j))$ où $\sum_{(i,j) \in X(\Omega) \times Y(\Omega)}$ cache une somme double.

On a donc : $\boxed{E(u(X, Y)) = \sum_{i \in X(\Omega)} \sum_{j \in Y(\Omega)} u(i, j) P((X = i) \cap (Y = j)) = \sum_{j \in Y(\Omega)} \sum_{i \in X(\Omega)} u(i, j) P((X = i) \cap (Y = j))}$.

L'étude théorique dans le cas général n'est pas au programme. Nous verrons dans les sections suivantes les cas $u : (x, y) \in \mathbb{R}^2 \mapsto max(x, y)$,$u : (x, y) \in \mathbb{R}^2 \mapsto min(x, y)$, $u : (x, y) \in \mathbb{R}^2 \mapsto x + y$ et $u : (x, y) \in \mathbb{R}^2 \mapsto x \times y$.

4.2 Somme de 2 variables aléatoires discrètes

Retour exemple : déterminer la loi de $X_1 + X_2$.

Formule générale, appelée produit de convolution discret :
Soit X et Y deux variables aléatoires discrètes définies sur un même espace et soit $S = X + Y$.
Alors, en appliquant la formule des probabilités totales au système complet d'événements $\{(X = i), i \in X(\Omega)\}$ on obtient les expressions suivantes :

$$\forall n \in S(\Omega),\ P(S = n) = \sum_{i \in X(\Omega)} P((X = i) \cap (S = n)) = \sum_{i \in X(\Omega)} P((X = i) \cap (Y = n - i)) = \sum_{\substack{i \in X(\Omega) \\ \text{tel que } n - i \in Y(\Omega)}} P((X = i) \cap (Y = n - i))$$

si de plus, X et Y sont indépendantes :

$$\forall n \in S(\Omega), \quad P(S = n) = \sum_{i \in X(\Omega) /\ n - i \in Y(\Omega)} P(X = i) P(Y = n - i)$$

$\boxed{\text{Ce raisonnement sera à refaire sur votre copie sauf dans le cas où } X \text{ et } Y \text{ sont à valeurs dans } \mathbb{N}}$.

Dans ce cas, l'ensemble des $i \in X(\Omega)$ tel que $n - i \in Y(\Omega)$ est inclus dans $[\![0, n]\!]$ d'où le résultat suivant :

> `Proposition`
> Soit X et Y deux variables discrètes à valeurs dans $\mathbb{N}$, et soit $S = X + Y$. Alors S est une variable discrète à valeurs dans $\mathbb{N}$, et pour tout $n \in S(\Omega)$, $P(S = n) = \sum_{i=0}^{n} P((X = i) \cap (Y = n - i))$.

Application :
Soit $X \hookrightarrow \mathcal{P}(\lambda)$ et $Y \hookrightarrow \mathcal{P}(\mu)$, avec X et Y indépendantes. Alors $\boxed{X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)}$.

$\longrightarrow$ `preuve`

Généralisation :
Soit une famille $(X_k)_{1 \leq k \leq n}$ de variables indépendantes telles que pour tout $k \in [\![1, n]\!]$, $X_k \hookrightarrow \mathcal{P}(\lambda_k)$.
Alors $X_1 + X_2 + ... + X_n \hookrightarrow \mathcal{P}(\lambda_1 + \lambda_2 + ... + \lambda_n)$.

Exercice :
Soit X et Y deux variables indépendantes de même loi $\mathcal{U}([\![1, N]\!])$. Déterminer la loi de $X + Y$.

Existence et calcul de l'espérance de la somme $X + Y$:

1. Si X et Y admettent une espérance, alors par linéarité, $X + Y$ admet une espérance et $E(X + Y) = E(X) + E(Y)$.
2. Si on connait la loi de $S = X + Y$, on peut étudier l'espérance en utilisant la définition de l'espérance. Sous réserve d'existence, $E(S) = \sum_{k \in S(\Omega)} kP(S = k)$.
3. Si on ne connaît que la loi du couple (X, Y), et que X et Y sont des variables finies, on utilise la formule de transfert : $X + Y$ admet une espérance et $E(X + Y) = \sum_{(i,j) \in X(\Omega) \times Y(\Omega)} (i + j)\ P((X = i) \cap (Y = j))$.

4.3 Produit de 2 variables aléatoires discrètes

Exemple : soit X et Y deux variables indépendantes suivant une loi de Bernoulli de paramètre respectif p et q. Déterminer la loi de $X \cdot Y$.

Remarque : La loi du produit sera rarement demandée ; il n'y a pas de formule générale simple. En revanche, l'espérance $E(XY)$ apparait dans la formule de covariance (cf section suivante) donc devra être calculée.

Existence et calcul de l'espérance du produit $X\dot{Y}$:

1. si on connaît la loi de $X \cdot Y$ (cas rare), on étudie son espérance à l'aide de la définition.
2. si on ne connaît que la loi du couple (X, Y), et que X et Y sont deux variables finies, on utilise la formule de transfert (toujours la même !) :
   $X \cdot Y$ admet une espérance et $E(X \cdot Y) = \sum_{(i,j) \in X(\Omega) \times Y(\Omega)} ij \times P((X = i) \cap (Y = j))$
3. Cas particulier important :
   `Proposition admise`
   Soit X et Y deux variables discrètes indépendantes, et admettant une espérance.
   Alors la variable XY admet une espérance et $E(XY) = E(X)E(Y)$.

   $\longrightarrow$ `preuve dans le cas où` X `et` Y `sont des variables finies`

4.4 Maximum-minimum de 2 variables aléatoires discrètes

Soit X, Y deux variables aléatoires définies sur $(\Omega, \mathcal{A}, P)$. On pose $Z = max(X, Y)$ et $T = min(X, Y)$.

Alors $\boxed{\text{pour tout } x \in Z(\Omega),\ (Z \leq x) = (X \leq x) \cap (Y \leq x)}$.

En particulier, si X et Y sont indépendantes : $\boxed{\text{pour tout } x \in Z(\Omega),\ P(Z \leq x) = P(X \leq x)P(Y \leq x)}$.

$\longrightarrow$ dans ce cas, on passera par l'étude des fonctions de répartition de X et Y et non l'étude du couple.

De manière analogue :

$\boxed{\text{pour tout } x \in T(\Omega),\ (T > x) = (X > x) \cap (Y > x)}$.

En particulier, si X et Y sont indépendantes : $\boxed{\text{pour tout } x \in T(\Omega),\ P(T > x) = P(X > x)P(Y > x)}$.

Application 1 :
soit X, Y deux variables aléatoires indépendantes, et telles que $X \hookrightarrow \mathcal{G}(p)$, $Y \hookrightarrow \mathcal{G}(q)$, avec $p, q \in [0, 1]$.
Déterminer la loi de $Z = min(X, Y)$.
bonus : Reconnaître la loi de Z.

Application 2 :
soit X, Y deux variables aléatoires indépendantes et de même loi $\mathcal{U}([\![1, n]\!])$.
Déterminer la loi de $Z = max(X, Y)$

5 Covariance

Définition
Soit X, Y deux variables aléatoires discrètes admettant un moment d'ordre 2.
On appelle covariance du couple (X, Y) le réel $\boxed{Cov(X, Y) = E((X - E(X))(Y - E(Y)))}$.

Intuitivement, $Cov(X, Y)$ mesure la variation simultanée de X et Y.
Interprétation du signe de $Cov(X, Y)$: si $Cov(X, Y) \geq 0$, cela signifie que lorsque X augmente, en moyenne Y augmente aussi. En revanche, si $Cov(X, Y) \leq 0$, cela signifie que lorsque X augmente, en moyenne Y diminue.

`Proposition  Formule de Koenig-Huygens`
Soit X et Y deux variables admettant chacune un moment d'ordre 2.
Alors $Cov(X, Y) = E(XY) - E(X)E(Y)$.

Remarque
Si de plus X et Y sont $\boxed{\text{indépendantes}}$, alors $E(XY) = E(X)E(Y)$, et donc $\boxed{Cov(X, Y) = 0}$.

$\longrightarrow$ Attention, la réciproque est fausse !

Exemple :
Soit X et Y deux variables indépendantes suivant la même loi de Bernoulli de paramètre $p \in]0, 1[$.

1. Déterminer les lois de $X + Y$ et $X - Y$.
2. Montrer que $Cov(X + Y, X - Y) = 0$.
3. $X + Y$ et $X - Y$ sont-elles indépendantes ?

Propriétés de la covariance :

1. Si X admet une variance, alors $Cov(X, X) = V(X)$.
2. Si X est une variable constante alors, pour toute variable Y, $Cov(X, Y) = 0$
3. Si le couple (X, Y) admet une covariance alors le couple (Y, X) également, et $Cov(X, Y) = Cov(Y, X)$.
4. Il y a de plus des propriétés de linéarité par rapport à chacune des variables : autrement dit, sous réserve d'existence, $Cov(aX + Y, Z) = aCov(X, Z) + Cov(Y, Z)$ et $Cov(X, aY + Z) = aCov(X, Y) + Cov(X, Z)$.

$\longrightarrow$ `preuve`

Lien entre variance et covariance :
`Proposition`
$\boxed{\text{Si } X \text{ et } Y \text{ admettent une variance, alors } X + Y \text{ admet une variance et } V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)}$.

$\longrightarrow$ `preuve de la formule (existence admise)`

En particulier, si X et Y sont deux variables indépendantes admettant une variance,
alors $X + Y$ admet une variance et $V(X + Y) = V(X) + V(Y)$

Plus généralement, si $X_1, ...X_n$ sont n variables mutuellement indépendantes admettant une variance,
alors $\sum\limits_{k=1}^{n} X_k$ admet une variance et $V(\sum\limits_{k=1}^{n} X_k) = \sum\limits_{k=1}^{n} V(X_k)$.