Chapitre 19 : Couples de variables aléatoires discrètes

ECE3 Lycée Carnot

23 mai 2011

Dans ce dernier chapitre de probabilités de l'année, nous allons introduire l'étude de couples de variables aléatoires, c'est-à-dire l'étude simultanée de deux variables aléatoires. Rien de très nouveau au niveau des techniques utilisées, le but est de présenter un peu de vocabulaire et faire quelques calculs sur trois exemples détaillés.

1 Loi d'un couple de variables aléatoires

Définition 1. Un couple de variables aléatoires (X, Y) est la donnée de deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$. Une façon plus technique de voir les choses est de dire qu'un couple est une application $(X, Y) : \Omega \to \mathbb{R}^2$.

Exemple 1 : On lance simultanément deux dés (ça faisait longtemps), et on note X le plus grand des deux chiffres obtenus et Y le plus petit (au sens large).

Exemple 2 : Une urne contient 4 boules blanches, 3 boules vertes et 2 boules bleues. On tire 3 boules dans l'urne et on note X le nombre de boules blanches obtenues et Y le nombre de boules vertes.

Exemple 3 : On effectue une suite de lancers avec une pièce non équilibrée pour laquelle la probabilité d'obtenir Pile vaut $\frac{3}{4}$ (et celle d'obtenir face vaut donc $\frac{1}{4}$). On note X la longueur de la première chaîne obtenue, c'est-à-dire le nombre de tirages initiaux donnant le même résultat que le premier tirage. On note Y la longueur de la deuxième chaîne. Ainsi, si les premiers tirages donnent $PPPPFFP$ (peu importe la suite), on aura $X = 4$ et $Y = 2$.

Définition 2. La **loi conjointe** d'un couple de variables aléatoires (X, Y) est la donnée des probabilités $P((X = i) \cap (Y = j))$, pour $i \in X(\Omega)$ et $j \in Y(\Omega)$. On l'appelle aussi plus simplement loi du couple (X, Y).

Remarque 1. Certaines des probabilités $P((X = i) \cap (Y = j))$ peuvent être nulles, même si $P(X = i)$ et $P(Y = j)$ sont toutes les deux non nulles.

Remarque 2. Cette loi est souvent présentée sous forme d'un tableau à double entrée, les valeurs prises par X apparaissant par exemple en ligne et celles prises par Y en colonne.

Exemple 1 : On a $X(\Omega) = \{1; 2; 3; 4; 5; 6\}$; $Y(\Omega) = \{1; 2; 3; 4; 5; 6\}$, et la loi conjointe se calcule sans difficulté : $P((X = i) \cap (Y = j)) = 0$ si $i < j$ (le plus grand nombre ne peut pas être inférieur au plus petit), $P((X = i) \cap (Y = y_j)) = \frac{1}{36}$ si $i = j$ (le seul tirage favorable sur les 36 possibles est le tirage (i, i), et $P((X = i) \cap (Y = y_j)) = \frac{1}{18}$ si $i > j$ (les deux tirages (i, j) et (j, i) sont possibles), ce qu'on peut résumer par le tableau suivant (X en ligne, Y en colonne) :

$Y\backslash X$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
2	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
3	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
4	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$
6	0	0	0	0	0	$\frac{1}{36}$

Exemple 2 : On a ici $X(\Omega) = \{0;1;2;3\}$ et $Y(\Omega) = \{0;1;2;3\}$. On ne peut bien sûr avoir $X+Y > 3$ puisqu'on ne tire que trois boules. Lorsque cela a un sens, on a $P((X=i)\cap(Y=j)) = \dfrac{\binom{4}{i}\binom{3}{j}\binom{2}{3-i-j}}{\binom{9}{3}}$, ce qui donne le tableau suivant :

$Y\backslash X$	0	1	2	3
0	0	$\frac{4}{84}$	$\frac{12}{84}$	$\frac{4}{84}$
1	$\frac{3}{84}$	$\frac{24}{84}$	$\frac{18}{84}$	0
2	$\frac{6}{84}$	$\frac{12}{84}$	0	0
3	$\frac{1}{84}$	0	0	0

Remarque 3. On note parfois l'évenement $(X=i)\cap(Y=j)$ sous la forme $(X,Y) = (i,j)$.

Proposition 1. Les événements $(X=i)\cap(Y=j)$ pour i parcourant $X(\Omega)$ et j parcourant $Y(\Omega)$ forment un système complet d'événements. On a donc $\displaystyle\sum_{i\in X(\Omega);j\in Y(\Omega)} P((X=i)\cap(Y=j)) = 1$.

Démonstration. La preuve est la même que dans le cas où on n'a qu'une variable aléatoire. Les événements sont manifestement disjoints, et leur union vaut Ω. $\square$

Définition 3. Si (X,Y) est un couple de variables aléatoires, les lois de X et de Y sont appelées **lois marginales** du couple.

Proposition 2. Si (X,Y) est un couple de variables aléatoires discrètes, on peut obtenir les lois marginales à partir de la loi conjointe : $\forall i \in X(\Omega)$, $P(X=i) = \displaystyle\sum_{j\in Y(\Omega)} P((X=i)\cap(Y=j))$ (et symétriquement pour la loi de Y).

Démonstration. Cela découle immédiatement du fait que les événements $Y=j$ forment un système complet d'événements (il suffit d'écrire la formule des probabilités totales). $\square$

Exemple 1 : Pour connaitre les lois marginales à partir du tableau précédemment établi, c'est très simple, il suffit de faire les sommes des lignes du tableau (pour la loi de Y) ou des colonnes (pour celle de X) :

$Y\backslash X$	1	2	3	4	5	6	$P(Y=j)$
1	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{11}{36}$
2	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{9}{36}$
3	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{7}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{5}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{3}{36}$
6	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
$P(X=i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	

Exemple 2 : De façon tout à fait similaire, on va retrouver des lois hypergéométriques pour X et Y :

$Y\backslash X$	0	1	2	3	$P(Y=j)$
0	0	$\frac{4}{84}$	$\frac{12}{84}$	$\frac{4}{84}$	$\frac{20}{84}$
1	$\frac{3}{84}$	$\frac{24}{84}$	$\frac{18}{84}$	0	$\frac{45}{84}$
2	$\frac{6}{84}$	$\frac{12}{84}$	0	0	$\frac{18}{84}$
3	$\frac{1}{84}$	0	0	0	$\frac{1}{84}$
$P(X=i)$	$\frac{10}{84}$	$\frac{40}{84}$	$\frac{30}{84}$	$\frac{4}{84}$	

Remarque 4. On vient de voir qu'on pouvait déduire les lois marginales de la loi conjointe. Le contraire n'est pas possible en général.

Définition 4. La **loi conditionnelle de X sachant** $Y = j$ est la loi de la variable Z définie par $\forall i \in X(\Omega)$, $P(Z = i) = P_{Y=j}(X = i)$. On définit de même les lois conditionnelle de Y sachant $X = i$.

Exemple 3 : Dans le cas de variables infinies, on ne peut naturellement plus écrire la loi sous forme de tableau, et les calculs de lois marginales ou conditionnelles sont un peu plus formels. La loi de X se calcule assez aisément : $X(\Omega) = \mathbb{N}^*$ et on aura $X = k$ si la suite de lancers commence par k Pile suivi d'une face ou par k face suivis d'un Pile, cas incompatibles qui donnent $P(X = k) = \left(\frac{3}{4}\right)^k \times \frac{1}{4} + \left(\frac{1}{4}\right)^k \times \frac{3}{4} = \frac{3^k + 3}{4^{k+1}}$.

Pour déterminer la loi de Y, le plus simple est de passer par la loi de couple : on aura $(X, Y) = (i, j)$ si on débute par i Pile, puis j Face et à nouveau un Pile, ou bien i Face, j Pile et un Face, soit une probabilité de $P((X = i) \cap (Y = j)) = \left(\frac{3}{4}\right)^i \left(\frac{1}{4}\right)^j \times \frac{3}{4} + \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^j \times \frac{1}{4} = \frac{3^{i+1} + 3^j}{4^{i+j+1}}$. On utilise ensuite la formule des probabilités totales pour obtenir la loi de Y, avec le système complet d'évènements $(X = i)_{i\geqslant 1}$: $P(Y = j) = \sum_{i=1}^{+\infty} P(X = i) \times P_{X=i}(Y = j) = \sum_{i=1}^{+\infty} P((X = i) \cap (Y = j)) = \sum_{i=1}^{+\infty} \frac{3^{i+1} + 3^j}{4^{i+j+1}} = \frac{1}{4^j} \sum_{i=1}^{+\infty} \left(\frac{3}{4}\right)^{i+1} + \left(\frac{3}{4}\right)^j \sum_{i=1}^{+\infty} \frac{1}{4^{i+1}} = \frac{1}{4^j} \left(\frac{3}{4}\right)^2 \frac{1}{1 - \frac{3}{4}} + \left(\frac{3}{4}\right)^j \frac{1}{4^2} \frac{1}{1 - \frac{1}{4}} = \frac{3^2 + 3^{j-1}}{4^{j+1}}$.
On constate que la loi de Y n'est pas la même que celle de X (ce qui n'était pas vraiment évident a priori).

Si on souhaite voir à quoi ressemblent les lois conditionnelles, on a par exemple comme loi conditionnelle à $Y = j$ fixé : $P_{Y=i}(X = i) = \frac{P((X = i) \cap (Y = j))}{P(Y = j)} = \frac{3^{i+1} + 3^j}{4^{i+j+1}} \times \frac{4^{j+1}}{3^2 + 3^{j-1}} = \frac{3^{i+1} + 3^j}{4^i(3^2 + 3^{j-1})}$.

2 Indépendance de variables aléatoires

Définition 5. Deux variables aléatoires sont dites **indépendantes** si tous les couples d'événements $X = i$, $Y = j$ sont indépendants. Autrement dit, $\forall i \in X(\Omega)$, $\forall j \in Y(\Omega)$, $P((X = i) \cap (Y = j)) = P(X = i)P(Y = j)$.

Exemple : Un exemple idiot pour illustrer. Si on tire deux dés simultanément et qu'on note X le résultat du premier dé et Y celui du deuxième dé, X et Y sont des variables aléatoires indépendantes. On a $P((X = i) \cap (Y = j)) = \frac{1}{36}$ pour tout $1 \leqslant i \leqslant 6$ et $1 \leqslant j \leqslant 6$.

Exemple : Par contre, toujours dans le cas de l'inépuisable lancer de deux dés, si on prend pour X la somme des deux dés et pour Y leur produit, les deux variables ne sont pas indépendantes. On a par exemple $P(X = 8) = \frac{5}{36}$; $P(Y = 15) = \frac{1}{18}$, et $P((X = 8) \cap (Y = 15)) = \frac{1}{18}$.

Remarque 5. Deux variables aléatoires sont indépendantes si et seulement si toutes les lois conditionnelles de X sachant $Y = j$ sont identiques à la loi de X.

Remarque 6. Dans le cas où X et Y sont des variables aléatoires indépendantes, on peut obtenir la loi conjointe du couple (X,Y) à partir des deux lois marginales.

Exemples : Dans nos deux premiers exemples, on constate sans difficulté et sans surprise que les variables X et Y ne sont pas indépendantes (le fait qu'on ait des 0 dans la tableau de la loi conjointe suffit à imposer qu'il n'y ait pas indépendance).
Le troisième exemple est moins intuitif. Le calcul des lois conditionnelles, qui sont distinctes de la loi marginale de X, prouve qu'il n'y a pas non plus indépendance. Une autre façon de voir les choses est de trouver un contre-exemple à l'indépendance des évènements $X = i$ et $Y = j$. On peut ici calculer simplement $P(X=1) = \frac{3+3}{4^2} = \frac{3}{8}$; $P(Y=1) = \frac{3^2+1}{4^2} = \frac{5}{8}$; et $P((X=1)\cap(Y=1)) = \frac{3^2+3}{4^3} = \frac{3}{16} \neq \frac{3}{8} \times \frac{5}{8}$. Les courageux vérifieront qu'en effectuant la même expérience avec une pièce équilibrée, on obtiendrait deux variables aléatoires indépendantes.
Un autre calcul intéressant sur cette expérience (quoique ne faisant pas vraiment intervenir le fait qu'on travaille sur un couple) est celui des espérances de X et de Y : on a (sous réserve d'existence) $E(X) = \sum_{k=1}^{+\infty} k \times \frac{3^k+3}{4^{k+1}} = \frac{3}{4^2}\sum_{k=1}^{+\infty} k\left(\frac{3}{4}\right)^{k-1} + \frac{3}{4^2}\sum_{k=1}^{+\infty} k\left(\frac{1}{4}\right)^{k-1} = \frac{3}{4^2}\left(\frac{1}{(1-\frac{3}{4})^2} + \frac{1}{(1-\frac{1}{4})^2}\right) = 3 + \frac{1}{3} = \frac{10}{3}$.
Pour Y, on a $E(Y) = \sum_{k=1}^{+\infty} k \times \frac{3^2+3^{k-1}}{4^{k+1}} = \frac{3^2}{4^2}\sum_{k=1}^{+\infty} k\left(\frac{1}{4}\right)^{k-1} + \frac{1}{4^2}\sum_{k=1}^{+\infty} k\left(\frac{3}{4}\right)^{k-1} = \frac{3^2}{4^2} \times \frac{1}{(1-\frac{1}{4})^2} + \frac{1}{4^2} \times \frac{1}{(1-\frac{3}{4})^2} = 1 + 1 = 2$. Encore une fois, les courageux pourront faire le calcul avec une pièce équilibrée et constater que dans ce cas les deux variables ont pour espérance 2.

Proposition 3. Si X et Y sont deux variables aléatoires indépendantes, on a $V(X+Y) = V(X) + V(Y)$.

Remarque 7. On peut calculer facilement la variance d'une loi binomiale à partir de la proposition précédente. En effet une loi binomiale de paramètre (n,p) est la somme de n variables aléatoires indépendantes de paramètre p, qui ont chacune pour variance $p(1-p)$. La variance de la binomiale vaut donc $np(1-p)$. Pour l'hypergéométrique, c'est plus compliqué, puisque la loi pour chaque tirage dépend des résultats des tirages précédents.

3 Opérations et variables aléatoires

Profitons de ce petit chapitre pour rappeler quelques cas particuliers de fonctions de deux variables aléatoires pour lesquelles on sait calculer facilement la loi. Commençons par le cas de la somme de deux variables :

Proposition 4. Si X et Y sont deux variables aléatoires indépendantes, la loi de leur somme $X+Y$ est donné par $P(X+Y=k) = \sum P(X=i)P(Y=k-i)$ (la somme portant sur les valeurs de i pour lesquelles $P(X=i)$ et $P(Y=k-i)$ sont toutes deux non nulles).

Démonstration. On a via la formule des probabilités totales $P(X+Y=k) = \sum P(X=i)_{i\in X(\Omega)}P_{X=i}(X+Y=k)$. Or, $P_{X=i}(X+Y=k) = P_{X=i}(Y=k-i)$ donc $P(X+Y=k) = \sum P_{i\in X(\Omega)}((X=i)\cap(Y=k-i))$. □

Exemple : Deux variables aléatoires X et Y suivant des lois uniformes respectivement sur $\{1;2;\dots;6\}$ et sur $\{1;2;3;4\}$ (par exemple les résultats d'un lancer de dés à six faces et à quatre faces). La loi de $X+Y$ se calcule aisément :

i	2	3	4	5	6	7	8	9	10
$P(X+Y=i)$	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{24}$

Proposition 5. Soient X et Y deux variables aléatoires indépendantes suivant des lois binomiales $\mathcal{B}_{m,p}$ et $\mathcal{B}_{n,p}$ alors $X+Y$ suit une loi binomiale de paramètre $\mathcal{B}_{m+n,p}$.

Démonstration. D'après la propriété précédente, on a $P(X+Y=k) = \sum_{i=0}^{i=k} P(X=i)P(Y=k-i) = \sum_{i=0}^{i=k} \binom{n}{i} p^i(1-p)^{n-i}\binom{m}{k-i}p^{k-i}(1-p)^{m-k+i} = p^k(1-p)^{n+m-k}\sum_{i=0}^{i=k}\binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}p^k(1-p)^{n+m-k}$ en utilisant la formule de Vandermonde. On reconnait bien une loi binomiale de paramètre $(n+m,p)$. □

Proposition 6. Soient X et Y deux variables indépendantes suivant des lois de Poisson de paramètres respectifs λ et μ, alors $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

Démonstration. C'est encore un calcul de somme : $P(X+Y=k) = \sum_{i=0}^{i=k} P(X=i)P(Y=k-i) = \sum_{i=0}^{i=k} \frac{e^{-\lambda}\lambda^i}{i!} \times \frac{e^{-\mu}\mu^{k-i}}{(k-i)!} = e^{-(\lambda+\mu)}\sum_{i=0}^{i=k}\frac{1}{i!(k-i)!}\lambda^i\mu^{k-i} = \frac{e^{-(\lambda+\mu)}}{k!}\sum_{i=0}^{i=k}\binom{k}{i}\lambda^i\mu^{k-i} = \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^k}{k!}$ via la formule du binôme de Newton. On reconnait bien une loi de Poisson de paramètre $\lambda+\mu$. □

Autre exemple où on sait calculer la loi, le cas du minimum (ou du maximum) de deux variables aléatoires, où il est plus facile de passer par la fonction de répartition :

Proposition 7. Soient X, Y deux variables aléatoires indépendantes de fonctions de répartitions respectives F_X et F_Y, alors la fonction de répartition de la variablé aléatoire $Z = \max(X,Y)$ est donnée par $F_Z = F_X F_Y$.

Démonstration. Par définition $F_Z(x) = P(Z \geqslant x)$. Mais dire que $\max(X,Y) \leqslant x$ est équivalent à dire que $X \leqslant x$ et que $Y \leqslant x$, donc $F_Z(x) = P((X \leqslant x) \cap (Y \leqslant x)) = P(X \leqslant x) \times P(Y \leqslant x) = F_X(x)F_Y(x)$. □

Exemple : Reprenons une nouvelle fois le lancer de deux dés. On note X le résultat du premier dé et Y celui du deuxième dé. Les deux fonctions de répartition sont F_X et F_Y sont les mêmes : sur $[-\infty;1[$, $F_X(x) = 0$; sur $[1;2[$, $F_X(x) = \frac{1}{6}$, sur $[2;3[$, $F_X(x) = \frac{2}{6}$ etc. Soit $Z = \max(X,Y)$. La fonction de répartition de Z est donc donnée par : $[1;2[$, $F_Z(x) = \frac{1}{36}$; $[2;3[$, $F_Z(x) = \frac{4}{36}$ etc. On peut ainsi retrouver la loi du minimum par une méthode différente de celle vue un peu plus haut dans ce même chapitre (rappelons au cas où que pour passer de la fonction de répartition à la loi, on utilise la formule $P(Z=i) = F_Z(i) - F_Z(i-1)$:

i	1	2	3	4	5	6
$P(Z=i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Remarque 8. Une fois connu le maximum de deux variables aléatoires, on peut en déduire le minimum en utilisant le fait que $X+Y = \min(X+Y) + \max(X+Y)$. On peut aussi utiliser le fait que si on pose $G_X(x) = 1 - F_X(x) = P(X > x)$ (une sorte d'anti-fonction de répartition), alors, si $U = \min(X,Y)$, on a $G_U(x) = G_X(x)G_Y(x)$ (ce qui se démontre de façon très similaire à ce qu'on vient de faire).