ECE 2 - Année 2018-2019
Lycée français de Vienne
Mathématiques - F. Gaunard
http://frederic.gaunard.com

Devoir Libre n°2

Couples de v.a. discrètes
Automne 2018.

Exercice 1

On considère des variables aléatoires $X_1, X_2, \dots, X_n$ suivant chacune la même loi de Bernoulli de paramètre p avec $0 < p < 1$, c'est à dire : $\forall k \in [\![1;n]\!],\ P([X_k = 1]) = p$ et $P([X_k = 0]) = 1 - p$.

On suppose que pour tout couple $(k, \ell) \in [\![1;n]\!]^2$ avec $k \neq \ell$, le coefficient de corrélation linéaire des variables X_k et X_ℓ est le même; on note r ce coefficient. On a donc :

$$\forall (k,\ell) \in [\![1;n]\!]^2, \qquad \frac{\mathrm{Cov}(X_k, X_\ell)}{\sqrt{V(X_k)V(X_\ell)}} = \begin{cases} 1 & \text{si } k = \ell \\ r & \text{si } k \neq \ell \end{cases}$$

(1) (a) Dans les cas (i) et (ii) suivants, calculer la valeur de r et exprimer la variance de la variable aléatoire $\displaystyle\sum_{k=1}^{n} X_k$ en fonction de n et p.

(i) Les variables aléatoires $X_1, X_2, \dots, X_n$ sont mutuellement indépendantes.
(ii) Les variables aléatoires $X_1, X_2, \dots, X_n$ sont toutes égales.

De plus, préciser la loi de $\displaystyle\sum_{k=1}^{n} X_k$ dans chacun des deux cas précédents.

(b) Montrer que pour tout $k \in [\![1;n]\!]$, la variance de la variable aléatoire $\displaystyle\sum_{i=1}^{n} X_i$ est donnée par la formule :

$$V\left(\sum_{i=1}^{n} X_i\right) = kp(1-p)(1 + (k-1)r)$$

(c) En déduire que le coefficient r est au moins égal à $-\dfrac{1}{n-1}$.

(2) On suppose dans cette question que n est au moins égal à 2.

(a) Montrer que r est égal à -1 si et seulement si on a : $P([X_1 = 1] \cap [X_2 = 1]) = p(2p - 1)$.
(b) Que vaut alors $P([X_1 = 0] \cap [X_2 = 0])$?
(c) En déduire que le coefficient r ne peut-être égal à -1 que lorsque

$$p = \frac{1}{2} \quad \text{et} \quad P([X_1 + X_2 = 1]) = 1.$$

(3) On suppose dans cette question que n est supérieur ou égal à 3 et que

$$P\left(\left[\sum_{k=1}^{n} X_k = 1\right]\right) = 1.$$

(a) Exprimer les valeurs de p et r en fonction de n.
(b) Déterminer les n-uplets $(x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ pour lesquels la probabilité

$$P\left(\bigcap_{k=1}^{n} [X_k = x_k]\right)$$

est strictement positive et la calculer.

Exercice 2

Partie I

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 $V(Y)$ et on suppose $V(X) > 0$ (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:

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

(1) **Covariance des variables aléatoires X et Y.**

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

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

(b) En déduire que $(\mathrm{cov}(X,Y))^2 \leqslant V(X)V(Y)$.
A quelle condition nécessaire et suffisante a-t-on l'égalité $((\mathrm{cov}(X,Y))^2 = V(X)V(Y)$?

(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 $\mathrm{cov}(X,Y)$ 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.
(c) On suppose enfin que X suit une loi normale centrée réduite $N(0,1)$ et $Y = X^2$.
Préciser les espérances et les variances de X et Y ainsi que la covariance et le coefficient de corrélation de X et Y. Étudier alors la réciproque de la question **2.b)**.

Partie II

(1) **Calculs préliminaires**

(a) On considère deux nombres entiers naturels q et n tels que $n \geqslant q$. En raisonnant par récurrence sur n, établir la formule suivante :

$$\sum_{k=q}^{n} \binom{k}{q} = \binom{n+1}{q+1}$$

(b) En faisant $q = 1, 2, 3$, en déduire une expression factorisée des quatre sommes suivantes :

$$\sum_{k=1}^{n} k; \qquad \sum_{k=2}^{n} k(k-1) \quad et \quad \sum_{k=1}^{n} k^2; \qquad \sum_{k=3}^{n} k(k-1)(k-2).$$

On considère dans toute la suite de cette partie un nombre entier $n \geqslant 2$ et une urne contenant n jetons numérotés de 1 à n.
On extrait de cette urne successivement et sans remise 2 jetons et on désigne alors par :

- N_1 la variable aléatoire indiquant le numéro du premier jeton tiré
- N_2 la variable aléatoire indiquant le numéro du second jeton tiré.
- X la variable aléatoire indiquant le plus petit des numéros des 2 jetons tirés.
- Y la variable aléatoire indiquant le plus grand des numéros des 2 jetons tirés.

On note $E(N_1)$ et $V(N_1)$, $E(N_2)$ et $V(N_2)$, $E(X)$ et $V(X)$, $E(Y)$ et $V(Y)$ les espérances et variances des quatre variables aléatoires N_1, N_2, X, Y.

(2) **Lois conjointe et marginales des variables aléatoires N_1 et N_2.**

(a) Déterminer les probabilités $P(N_1 = i)$ pour $1 \leqslant i \leqslant n$ et $P(N_2 = j/N_1 = i)$ pour $1 \leqslant j \leqslant n, j \neq i$.
En déduire $P(N_2 = j)$ pour $1 \leqslant j \leqslant n$, puis comparer les lois de N_1 et N_2.
(b) Calculer les espérances $E(N_1)$ et $E(N_2)$, les variances $V(N_1)$ et $V(N_2)$.
(c) Déterminer les probabilités $P(N_1 = i \cap N_2 = j)$pour $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant n$ en distinguant les deux cas $i = j$ et $i \neq j$ et en déduire que :

$$E(N_1N_2) = \frac{(n+1)(3n+2)}{12}.$$

En déduire la covariance et le coefficient de corrélation linéaire de N_1 et N_2.
(d) Exprimer enfin sous forme factorisée la variance $V(N_1 + N_2)$.

(3) **Lois conjointe, marginales et conditionnelles des variables** aléatoires X et Y.

(a) Montrer que les probabilités $P(X = i \cap Y = j)$ sont égales à $\dfrac{2}{n(n-1)}$ pour $1 \leqslant i < j \leqslant n$.
Que valent-elles sinon ?
(b) En déduire les probabilités $P(Y = j)$ pour $2 \leqslant j \leqslant n$ et $P(X = i)$ pour $1 \leqslant i \leqslant n-1$.
(On vérifiera que les formules donnant $P(Y = j)$ et $P(X = i)$ restent valables si $j = 1$ ou $i = n$).
(c) Déterminer les probabilités $P_{Y=j}(X = i)$ et $P_{X=i}(Y = j)$ pour $1 \leqslant i < j \leqslant n$, puis reconnaître la loi de X conditionnée par $Y = j$ et la loi de Y conditionnée par $X = i$.
(d) Comparer les lois des variables aléatoires $n+1-X$ et Y, autrement dit les deux probabilités $P(n+1-X = j)$ et $P(Y = j)$ pour $2 \leqslant j \leqslant n$.
En déduire que $E(n+1-X) = E(Y)$ et $V(n+1-X) = V(Y)$, puis en déduire les expressions de $E(X)$ en fonction de $E(Y)$ et de $V(X)$ en fonction de $V(Y)$.

(4) **Espérances et variances des variables aléatoires X et Y.**

(a) Exprimer les espérances $E(Y)$ et $E(X)$ en fonction de n.
(b) Exprimer sous forme factorisée $E[(Y(Y-2)]$, puis $E(Y^2), V(Y)$ et $V(X)$ en fonction de n.

(5) **Covariance et coefficient de corrélation linéaire des variables aléatoires X et Y.**

(a) Vérifier que $X + Y = N_1 + N_2$, puis en déduire sous forme factorisée la variance de $X + Y$ et la covariance de X et Y.
(b) En déduire le coefficient de corrélation de X et Y.
On remarquera que ce coefficient de corrélation linéaire de X et Y est indépendant de n.