



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 2 — Option économique

2001

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Sujet

Le but du problème est l'étude du coefficient de corrélation linéaire de deux variables aléatoires, abordée d'abord de façon générale dans la partie I, puis dans un cas particulier dans la partie II.

Partie I

On considère deux variables aléatoires X et Y définies sur un même espace probabilisé, 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 de X et Y , que ces variables soient discrètes ou à densité, est définie par

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

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 réel λ :

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

- b) En déduire que $(\text{Cov}(X, Y))^2 \leq V(X)V(Y)$.

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

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

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

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

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

- b) Donner la valeur de ρ lorsque les variables aléatoires X et Y sont indépendantes.
- c) On suppose enfin que X suit la loi normale centrée réduite $\mathcal{N}(0, 1)$ et que $Y = X^2$.

Préciser les espérances et les variances de X et Y , ainsi que leur covariance et leur coefficient de corrélation. Étudier alors la réciproque de la question I.2.b.

Partie II

1. Calculs préliminaires

- a) On considère deux entiers naturels q et n tels que $n \geq q$. En raisonnant par récurrence sur n , établir la formule

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

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

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

On considère dans toute la suite de cette partie un entier $n \geq 2$ et une urne contenant n jetons numérotés de 1 à n .

On extrait de cette urne successivement et sans remise deux jetons. On désigne par :

- N_1 le numéro du premier jeton tiré ;
- N_2 le numéro du second jeton tiré ;

- X le plus petit des numéros des deux jetons tirés ;
- Y le plus grand des numéros des deux jetons tirés.

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

2. Lois conjointe et marginales des variables aléatoires N_1 et N_2

- a) Déterminer $P(N_1 = i)$ pour $1 \leq i \leq n$ et $P(N_2 = j \mid N_1 = i)$ pour $1 \leq j \leq n, j \neq i$.

En déduire $P(N_2 = j)$ pour $1 \leq j \leq n$, puis comparer les lois de N_1 et de N_2 .

- b) Calculer les espérances $E(N_1)$ et $E(N_2)$ et les variances $V(N_1)$ et $V(N_2)$.

- c) Déterminer les probabilités $P([N_1 = i] \cap [N_2 = j])$ pour $1 \leq i, j \leq n$, en distinguant les cas $i = j$ et $i \neq j$, et en déduire que

$$E(N_1 N_2) = \frac{(n+1)(3n+2)}{12}.$$

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 $\frac{2}{n(n-1)}$ pour $1 \leq i < j \leq n$. Que valent-elles sinon ?

- b) En déduire les probabilités $P(Y = j)$ pour $2 \leq j \leq n$ et $P(X = i)$ pour $1 \leq i \leq n-1$.

On vérifiera que les formules donnant $P(Y = j)$ et $P(X = i)$ restent valables si $j = 1$ ou $i = n$.

- c) Déterminer $P(X = i \mid Y = j)$ et $P(Y = j \mid X = i)$ pour $1 \leq i < j \leq n$, puis reconnaître la loi de X conditionnée par $Y = j$ et la loi de Y conditionnée par $X = i$.
- d) Comparer les lois de $n + 1 - X$ et de Y , autrement dit les probabilités $P(n + 1 - X = j)$ et $P(Y = j)$ pour $2 \leq j \leq n$.

En déduire que $E(n + 1 - X) = E(Y)$ et $V(n + 1 - X) = V(Y)$, puis exprimer $E(X)$ en fonction de $E(Y)$ et $V(X)$ 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) Pour $n \geq 3$, en déduire le coefficient de corrélation de X et Y . On remarquera que ce coefficient est indépendant de n dans ce domaine.

Préciser ce qu'il en est pour $n = 2$.

6. Utilisation de la fonction génératrice des variables aléatoires X et Y

On se propose de retrouver les résultats précédents par une autre méthode, en ne supposant connues que les probabilités $P([X = i] \cap [Y = j])$ pour $1 \leq i, j \leq n$.

On désigne par G la fonction génératrice du couple (X, Y) , définie pour tous réels u, v

par le polynôme

$$G(u, v) = \sum_{j=1}^n \sum_{i=1}^n P([X = i] \cap [Y = j])(1+u)^i(1+v)^j.$$

a) Montrer que $\frac{\partial G}{\partial u}(0, 0) = E(X)$ et $\frac{\partial G}{\partial v}(0, 0) = E(Y)$.

Donner des égalités analogues pour $\frac{\partial^2 G}{\partial u^2}(0, 0)$, $\frac{\partial^2 G}{\partial v^2}(0, 0)$ et $\frac{\partial^2 G}{\partial u \partial v}(0, 0)$.

b) En posant $w = u + v + uv$, c'est-à-dire $1 + w = (1 + u)(1 + v)$, montrer que, pour $u, v, w \neq 0$,

$$G(u, v) = \frac{2(1+w)}{n(n-1)u} \left(\frac{(1+w)^n - 1}{w} - \frac{(1+v)^n - 1}{v} \right).$$

En développant ci-dessus $(1+w)^n$ et $(1+v)^n$, quelle expression de $G(u, v)$ en déduit-on ?

c) Préciser les deux dérivées partielles $\partial w / \partial u$ et $\partial w / \partial v$, puis retrouver sous forme factorisée les nombres $E(X)$, $E(Y)$, $E(X^2)$, $V(X)$, $E(Y^2)$, $V(Y)$, $E(XY)$ et $\text{Cov}(X, Y)$, et pour terminer le coefficient de corrélation de X et Y pour $n \geq 3$. Préciser le cas $n = 2$.