ESSEC

M B A

CONCOURS D'ADMISSION

Option scientifique

MATHEMATIQUES II

Année 2001

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à **encadrer** dans la mesure du possible les résultats de leurs calculs.
Ils ne doivent faire usage d'aucun document; l'utilisation de toute calculatrice et de tout matériel électronique est interdite. Seule l'utilisation d'une règle graduée est autorisée.
Si au cours de l'épreuve un candidat repère ce qui lui semble une erreur d'énoncé, il le signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

Le but du problème est l'étude du coefficient de corrélation linéaire de deux variables aléatoires qu'on aborde d'abord de façon générale (**partie I**), puis dans un cas particulier (**partie II**).

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

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

1. <u>Covariance des variables aléatoires X et Y</u>

   (a) Exprimer $\operatorname{cov}(\lambda X + Y, \lambda X + Y)$ en fonction de $V(\lambda X + Y)$ et en déduire la formule suivante pour tout nombre réel λ :
   $$V(\lambda X + Y) = \lambda^2 V(X) + 2\lambda \operatorname{cov}(X,Y) + V(Y)$$

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

2. <u>Coefficient de corrélation linéaire des variables aléatoires X et Y</u>
   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 $\text{cov}(X,Y)$ 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 $\mathcal{N}(0,1)$ et que $Y = X^2$. Préciser les espérances et les variances de X et Y ainsi que la covariance et le coefficient de corrélation de X et Y. Etudier 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 :

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

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

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

On considère dans toute la suite de cette partie deux nombres entiers p et n tels que $1 \leqslant p \leqslant n$ et une urne contenant n jetons numérotés de 1 à n. On extrait de cette urne simultanément et au hasard p jetons et on désigne alors par

- X la variable aléatoire indiquant le plus petit des numéros des p jetons tirés.
- Y la variable aléatoire indiquant le plus grand des numéros des p jetons tirés.
  On note $E(X)$, $V(X)$ et $E(Y)$, $V(Y)$ les espérances et variances des variables aléatoires X et Y.

$\boxed{\textbf{A}}$ Dans cette partie **A**, on suppose que $p = 2$ (autrement dit, on extrait deux jetons de l'urne et X et Y sont les variables aléatoires indiquant le plus petit et le plus grand des 2 numéros tirés).

2. Lois des variables aléatoires X et Y

(a) Quel est le nombre de parties à 2 éléments d'un ensemble à j éléments? à n éléments ?
En déduire les probabilités $P(Y \leqslant j)$ et $P(Y = j)$ pour $2 \leqslant j \leqslant n$, puis, en raisonnant de même, les probabilités $P(X \geqslant i)$ 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 $i = n$.

(b) 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 $V(Y)$.

3. **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.

4. Covariance et coefficient de corrélation linéaire des variables aléatoires X et Y

(a) Montrer que la probabilité $P(X = i \cap Y = j)$ est égale à $\dfrac{2}{n(n-1)}$ pour $1 \leqslant i < j \leqslant n$.

(b) En déduire sous forme factorisée l'espérance $E(X(Y-2))$ et montrer que :

$$E(\mathbf{XY}) = \frac{(n+1)(3n+2)}{12}$$

(c) En déduire sous forme factorisée la covariance et le coefficient de corrélation de X et Y. *On remarquera que ce coefficient de corrélation linéaire de X et Y est indépendant de n.*

$\boxed{\mathbf{B}}$ Dans cette partie **B**, on revient au cas général et le nombre entier p vérifie donc $1 \leqslant p \leqslant n$.

5. Lois des variables aléatoires X et Y

(a) Déterminer $P(Y \leqslant j)$ et $P(Y = j)$ pour $p \leqslant j \leqslant n$, $P(X \geqslant i)$ et $P(X = i)$ pour $1 \leqslant i \leqslant n-p+1$.

(b) Comparer les lois de $n+1-X$ et Y et en déduire les expressions de $E(X)$ en fonction de $E(Y)$ et de $V(X)$ en fonction de $V(Y)$.

6. Espérances et variances des variables aléatoires X et Y

(a) Vérifier l'égalité $jC_{j-1}^{p-1} = pC_j^p$ et en déduire les espérances $E(Y)$ et $E(X)$ en fonction de n.

(b) Vérifier l'égalité $(j+1)jC_{j-1}^{p-1} = (p+1)pC_{j+1}^{p+1}$ et en déduire sous forme factorisée $E((Y+1)Y)$, puis $E(Y^2)$, $V(Y)$ et $V(X)$ en fonction de n.

7. Covariance et coefficient de corrélation linéaire des variables aléatoires X et Y

(a) Expliciter la probabilité $P(X = i \cap Y = j)$ pour $p \leqslant j \leqslant n$ et $1 \leqslant i \leqslant j-p+1$.

(b) En déduire sous forme factorisée l'espérance $E((Y+1)(Y-X))$ et montrer que :

$$E(XY) = \frac{(n+1)[(p+1)n+p]}{(p+2)(p+1)}$$

(c) En déduire sous forme factorisée la covariance et le coefficient de corrélation de X et Y. *On remarquera que ce coefficient de corrélation linéaire de X et Y est indépendant de n.*

$\boxed{\mathbf{C}}$ Dans cette partie **C**, on suppose à nouveau $p = 2$ et on se propose de retrouver les résultats du **IIA** par une autre méthode, en ne supposant connues que les probabilités $P(X = i \cap Y = j)$.

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

On désigne par G la fonction génératrice du couple de variables aléatoires (X, Y), définie par :

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

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

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

(b) Montrer, en posant $w = u + v + uv$, c'est à dire $1 + w = (1+u)(1+v)$, qu'on a 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 $\dfrac{\partial w}{\partial u}$ et $\dfrac{\partial w}{\partial v}$ et retrouver $E(X)$, $E(Y)$, $E(X^2)$ et $V(X)$, $E(Y^2)$ et $V(Y)$, $E(XY)$ et $\mathrm{cov}(X,Y)$, et enfin le coefficient de corrélation linéaire de X et Y.