CHAMBRE DE COMMERCE ET D'INDUSTRIE DE PARIS
DIRECTION DE L'ENSEIGNEMENT
Direction des Admissions et Concours

ECOLE DES HAUTES ETUDES COMMERCIALES
ECOLE SUPERIEURE DE COMMERCE DE PARIS
ECOLE SUPERIEURE DE COMMERCE DE LYON

CONCOURS D'ADMISSION SUR CLASSES PREPARATOIRES

OPTION ECONOMIQUE

MATHEMATIQUES II

Samedi 15 Mai 1999, de 8h. à 12h.

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.

On désire tester si un algorithme générateur de nombres aléatoires est satisfaisant.
On étudie quelques aspects probabilistes permettant de répondre à cette question.
Les parties 1 et 2 sont indépendantes. La partie 3 utilise les notations et les résultats des parties précédentes.

Notation
Si X et Y sont deux variables aléatoires discrètes, on note $cov(X,Y)$ leur covariance, si celle-ci existe.

Partie 1

Soit n et s des entiers supérieurs ou égaux à 2. On considère une urne contenant des boules de couleurs C_1 , ..,C_s . Les boules de couleur C_i sont en proportion p_i . On a donc $\sum_{i=1}^{s} p_i = 1$ et on suppose que, pour tout i, $p_i > 0$.

On effectue n tirages successifs d'une boule avec remise.
Pour tout i de $\{1,..,s\}$, on note X_i la variable aléatoire égale au nombre de boules de couleur C_i obtenues à l'issue des n tirages (on remarque que la variable X_i dépend de n) . On définit la variable aléatoire U_n par : $U_n = \sum_{i=1}^{s} \frac{(X_i - np_i)^2}{np_i}$.

A. Etude des variables X_i.

1) Déterminer la loi de X_i , son espérance et sa variance.

2) Soit $(i,j) \in \{1,..,s\}^2$ tel que $i \neq j$. Déterminer la loi de $X_i + X_j$ et sa variance .
En déduire que $cov(X_i, X_j) = -np_i p_j$.

B. On suppose dans cette partie que $s = 2$.

1) Montrer que $U_n = Z_1^2$ où $Z_1 = \dfrac{X_1 - np_1}{\sqrt{np_1 p_2}}$. (On utilisera la relation : $X_1 + X_2 = n$)

2) Par quelle loi peut-on approcher la loi de Z_1 lorsque n est grand ?

C. On suppose dans cette partie que $s = 3$ et que $p_1 = p_2 = \dfrac{1}{4}$ et $p_3 = \dfrac{1}{2}$.

On pose $Z_1 = \dfrac{2}{\sqrt{n}}\left(X_3 - \dfrac{n}{2}\right)$ *et* $Z_2 = \sqrt{\dfrac{2}{n}}\left(X_1 - X_2\right)$.

1) Montrer que $U_n = Z_1^2 + Z_2^2$. (On utilisera la relation : $X_1 + X_2 + X_3 = n$)

2) Déterminer les espérances et variances de Z_1 et Z_2 et $cov(Z_1, Z_2)$.

3) Par quelle loi peut-on approcher celle de Z_1 lorsque n est grand?

4) Pour i élément de $\{1,..,n\}$, on définit la variable T_i par : $T_i = 1$ si au $i$ème tirage on a obtenu une boule de couleur C_1, $T_i = -1$ si au $i$ème tirage on a obtenu une boule de couleur C_2 , $T_i = 0$ si au $i$ème tirage on a obtenu une boule de couleur C_3 .
   a) Exprimer $X_1 - X_2$ à l'aide des variables T_i .
   b) En déduire que l'on peut approcher la loi de Z_2 , quand n est grand, par la loi normale centrée réduite.

5) On suppose avoir défini dans un programme Pascal :
   Type Tableau = Array[1..100] of integer;

   a) Ecrire une procédure *procedure Tirage(var C : Tableau);* permettant de simuler le tirage avec remise de 100 boules dans une urne contenant des boules de couleur C_1 ou C_2 ou C_3 en proportion respectivement $\dfrac{1}{4}, \dfrac{1}{4}, \dfrac{1}{2}$.
L'élément $C[i]$ vaut 1, 2 ou 3 et représente la couleur de la $i$ème boule tirée (C_1 , C_2 ou C_3). On utilisera la fonction *random* : *random*(4) retourne un entier aléatoire compris entre 0 et 3.
   b) Ecrire une fonction *Difference* de paramètre C qui retourne la valeur de $X_1 - X_2$.

D. On suppose $s = 4$ et $p_1 = p_2 = p_3 = p_4 = \dfrac{1}{4}$.

Soit A une matrice réelle à p lignes et m colonnes dont le coefficient en ligne i et colonne j est noté a_{ij}. On définit la matrice à m lignes et p colonnes appelée transposée de A et notée ${}^tA$ dont le coefficient en ligne i et colonne j vaut a_{ji}. On utilisera sans le démontrer que, pour toutes matrices A et B telles que le produit AB existe, ${}^t(A\,B) = {}^tB\ {}^tA$.

1) Pour i élément de $\{1, 2, 3, 4\}$, on pose $Y_i = \dfrac{X_i - np_i}{\sqrt{np_i}}$.
On note M la matrice carrée d'ordre 4 dont le coefficient en ligne i et colonne j vaut $cov(Y_i , Y_j)$. On définit N la matrice carrée d'ordre 4 dont tous les coefficients sont égaux à $\dfrac{1}{4}$.
Exprimer M en fonction de N et de I, où I désigne la matrice unité.

2) a) On définit 4 vecteurs de $\mathbb{R}^4$: $e_1 = \dfrac{1}{\sqrt{2}}(1,-1,0,0)$, $e_2 = \dfrac{1}{\sqrt{6}}(1,1,-2,0)$, $e_3 = \dfrac{1}{2\sqrt{3}}(1,1,1,-3)$, $e_4 = \dfrac{1}{2}(1,1,1,1)$.
Montrer que ces 4 vecteurs sont des vecteurs propres de N et qu'ils forment une base de $\mathbb{R}^4$.
   b) Soit Q la matrice de passage de la base canonique de $\mathbb{R}^4$ à la base (e_1, e_2, e_3, e_4) .
Montrer que ${}^tQ\,Q = I$.
Expliciter la matrice ${}^tQ\,M\,Q$.

3) On définit les variables Z_1, Z_2, Z_3 et Z_4 par : $\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{pmatrix} = {}^tQ \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}$.

a) Exprimer chaque Z_i en fonction de Y_1, Y_2, Y_3 et Y_4 et montrer que $Z_4 = 0$.

b) Montrer que, pour $i = 1, 2, 3$, les variables Z_i sont centrées.

c) Montrer que $U_n = Z_1^2 + Z_2^2 + Z_3^2$. (On pourra calculer $(Y_1, Y_2, Y_3, Y_4)\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}$)

Partie 2

1) Soit r un entier supérieur ou égal à 2.

Montrer la convergence de l'intégrale $\int_0^{+\infty} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}\, dx$. On note J_r sa valeur.

2) a) Pour tout réel t strictement positif, établir la convergence de l'intégrale $\int_t^{+\infty} x^{-\frac{1}{2}} e^{-\frac{x}{2}}\, dx$. On note $G(t)$ sa valeur.

b) Montrer à l'aide d'un changement de variable que, pour tout réel t strictement positif, $G(t) = 2\sqrt{2\pi}\left(1 - \Phi\left(\sqrt{t}\right)\right)$ où Φ désigne la fonction de répartition de la loi normale centrée réduite.

c) En déduire la convergence et la valeur J_1 de $\int_0^{+\infty} x^{-\frac{1}{2}} e^{-\frac{x}{2}}\, dx$.

3) Soit r un entier naturel non nul. On définit la fonction f_r par :

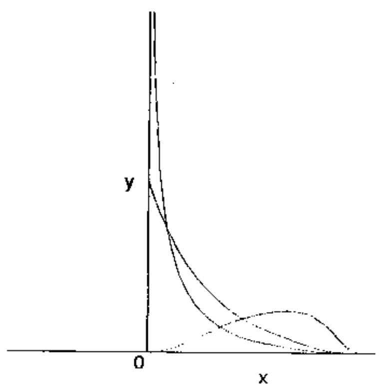
$$\forall x > 0,\quad f_r(x) = \frac{1}{J_r} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} \qquad \forall x \leq 0,\quad f_r(x) = 0.$$

a) Montrer que f_r est une densité de probabilité.

On dit qu'une variable aléatoire X suit la loi du χ^2 (lire khi-deux) à r degrés de liberté si et seulement si X admet f_r pour densité.

b) Quelle loi reconnaît-on pour $r = 2$?

c) Recopier le graphique ci-dessous en identifiant les courbes représentatives des trois fonctions f_1, f_2, f_7 en justifiant avec précision la réponse.



Partie 3

On reprend les notations de la partie 1, avec s entier quelconque supérieur ou égal à 2.

On admet que, n étant grand, la variable U_n suit la loi du χ^2 à s-1 degrés de liberté.

On considère un algorithme générateur de nombres entiers aléatoires compris entre 1 et 4.
On utilise cet algorithme pour créer un échantillon de 10000 nombres compris entre 1 et 4. On obtient alors :
2602 fois le nombre 1
2534 fois le nombre 2
2422 fois le nombre 3
2442 fois le nombre 4.

On se propose de tester la fiabilité de cet algorithme par l'examen de l'échantillon précédent.
On fait l'hypothèse que le nombre (aléatoire) fourni par le générateur suit une loi uniforme sur l'ensemble $\{1, 2, 3, 4\}$.

On donne les valeurs numériques suivantes :

- $0{,}95 = F_3(7{,}81)$, où F_3 désigne la fonction de répartition de la loi du χ^2 à 3 degrés de liberté,
- $\frac{1}{2500}\left(102^2 + 34^2 + 78^2 + 58^2\right) = 8{,}4032$.

En introduisant des variables convenables X_1, X_2, X_3, X_4 et la variable U_n associée, justifier le rejet de l'hypothèse de répartition uniforme des nombres fournis par le générateur.

ANNALES DE MATHEMATIQUES 1999

HEC, ESCP-EAP, EM-LYON : MATH II

CORRIGE

PARTIE–I

A. Etude des variables X_i

QUESTION–1

On effectue n tirages indépendants, avec remise ; à chaque tirage il y a deux issues possibles :

le succès, tirer une boule d'une couleur C_i avec la probabilité p_i,

l'échec, tirer une boule d'une autre couleur avec la probabilité $1 - p_i = q_i$;

X_i indique le nombre de succès, donc

X_i suit la loi binomiale $\mathcal{B}(n, p_i)$

$$\forall k \in [\![0; n]\!],\ p(X_i = k) = \binom{n}{k} p_i^k q_i^{n-k}.\ E(X_i) = np_i \text{ et } V(X_i) = np_iq_i.$$

QUESTION–2

Dans cette question, le schéma est le même que dans la question précédente ; le succès consiste à tirer une boule de couleur C_i ou C_j : la probabilité de succès est alors $p_i + p_j$

$X_i + X_j$ suit la binomiale $\mathcal{B}(n, p_i + p_j)$;
$V(X_i + X_j) = n(p_i + p_j)(1 - p_i - p_j)$.

D'autre part, $V(X_i + X_j) = V(X_i) + V(X_j) + 2\,\mathrm{cov}(X_i, X_j)$, donc

$$\begin{aligned}
\mathrm{cov}(X_i, X_j) &= \frac{1}{2}\Big(V(X_i + X_j) - V(X_i) - V(X_j)\Big)\\
&= \frac{1}{2}\Big(n(p_i + p_j)(1 - p_i - p_j) - np_i(1 - p_i) - np_j(1 - p_j)\Big)\\
&= \frac{n}{2}\Big(p_i + p_j - p_i^2 - p_j^2 - 2p_ip_j - p_i + p_i^2 - p_j + p_j^2\Big)
\end{aligned}$$

$$\boxed{\mathrm{cov}(X_i, X_j) = -np_ip_j.}$$

 Jean MALLET et Michel MITERNIQUE

B. Dans cette partie, $s = 2$.

QUESTION–1

$$\begin{aligned}
U_n &= \frac{(X_1 - np_1)^2}{np_1} + \frac{(X_2 - np_2)^2}{np_2} \\
&= \frac{(X_1 - np_1)^2 p_2 + (X_2 - np_2)^2 p_1}{np_1 p_2} \\
&= \frac{(X_1 - np_1)^2 p_2 + (n - X_1 - n(1-p_1))^2 p_1}{np_1 p_2} \\
&\quad \text{car } X_1 + X_2 = n \text{ et } p_1 + p_2 = 1 \\
&= \frac{(X_1 - np_1)^2 p_2 + (-X_1 + np_1)^2 p_1}{np_1 p_2} \\
&= \frac{(X_1 - np_1)^2 (p_1 + p_2)}{np_1 p_2} \quad \text{car } (-X_1 + np_1)^2 = (X_1 - np_1)^2 \\
&= \frac{(X_1 - np_1)^2}{np_1 p_2} \quad \text{car } p_1 + p_2 = 1.
\end{aligned}$$

$$\boxed{\text{Donc, } U_n = Z_1^2, \text{ avec } Z_1 = \frac{X_1 - np_1}{\sqrt{np_1 p_2}}.}$$

QUESTION–2

La variable X_1 suit la loi binomiale de paramètres n, p_1 ; donc $E(X_1) = np_1$ et $V(X_1) = np_1(1 - p_1) = np_1 p_2$. Il s'ensuit que Z_1 **est la variable centrée réduite associée à** X_1.

D'après le théorème de la limite centrée, Z_1 **converge en loi vers une variable qui suit la loi normale centrée réduite.**

- Rappelons le théorème de la limite centrée, théorème qui n'est pas fréquemment utilisé :

Soit (Y_k) **pour** $k \geq 1$**, une suite de variables définies sur le même espace probabilisé** Ω**, indépendantes, de même loi, possédant une espérance et une variance non nulle. Alors la variable centrée réduite associée à la somme** $S_n = \sum_{k=1}^{n} Y_k$ **converge en loi vers une variable** Y **définie sur** Ω **qui suit la loi normale centrée réduite.**

Le théorème s'applique dans notre situation car X_1 peut être considérée comme la somme de n variables de Bernoulli, de paramètre p_1, indépendantes, admettant une espérance p_1 et une variance **non nulle** $p_1 q_1$.

C. On suppose ici que $s = 3$, $p_1 = p_2 = \frac{1}{4}$ et $p_3 = \frac{1}{2}$.

QUESTION–1

$$\begin{aligned}
U_n &= \frac{(X_1 - np_1)^2}{np_1} + \frac{(X_2 - np_2)^2}{np_2} + \frac{(X_3 - np_3)^2}{np_3} \\
&= \frac{4(X_1 - \frac{n}{4})^2}{n} + \frac{4(X_2 - \frac{n}{4})^2}{n} + \frac{2(X_3 - \frac{n}{2})^2}{n}.
\end{aligned}$$

 Jean MALLET et Michel MITERNIQUE

$$\begin{aligned} Z_1^2 + Z_2^2 &= \frac{4}{n}\Big(X_3 - \frac{n}{2}\Big)^2 + \frac{2}{n}\Big(X_1 - X_2\Big)^2 \\ &= \underbrace{\frac{2}{n}\Big(X_1 - X_2\Big)^2 + \frac{2}{n}\Big(X_3 - \frac{n}{2}\Big)^2}_{S_1} + \frac{2}{n}\Big(X_3 - \frac{n}{2}\Big)^2. \ \mathbf{(1)} \end{aligned}$$

La somme S_1 des deux premiers termes vaut

$$\begin{aligned} S_1 &= \frac{2}{n}\Big(X_1 - X_2\Big)^2 + \frac{2}{n}\Big(n - X_1 - X_2 - \frac{n}{2}\Big)^2 \\ &\quad \text{car } X_1 + X_2 + X_3 = n \\ &= \frac{2}{n}\Big(X_1 - X_2\Big)^2 + \frac{2}{n}\Big(- X_1 - X_2 + \frac{n}{2}\Big)^2 \\ &= \frac{2}{n}\Big(X_1^2 + X_2^2 - 2X_1X_2 + X_1^2 + X_2^2 + \frac{n^2}{4} - nX_1 - nX_2 + 2X_1X_2\Big) \\ &= \frac{2}{n}\Big(2X_1^2 + 2X_2^2 + \frac{n^2}{4} - nX_1 - nX_2\Big) \\ &= \frac{4}{n}\Big(X_1^2 + X_2^2 + \frac{n^2}{8} - \frac{n}{2}X_1 - \frac{n}{2}X_2\Big) \\ &= \frac{4}{n}\Big(X_1^2 - \frac{n}{2}X_1 + \frac{n^2}{16} + X_2^2 + \frac{n^2}{16} - \frac{n}{2}X_2\Big) \\ &= \frac{4}{n}\Big((X_1 - \frac{n}{4})^2 + (X_2 - \frac{n}{4})^2\Big). \end{aligned}$$

Donc, en revenant à l'égalité **(1)**

$$Z_1^2 + Z_2^2 = \frac{4}{n}\Big(X_1 - \frac{n}{4}\Big)^2 + \frac{4}{n}\Big(X_2 - \frac{n}{4}\Big)^2 + \frac{2}{n}\Big(X_3 - \frac{n}{2}\Big)^2.$$

$$\boxed{\text{On a bien : } U_n = Z_1^2 + Z_2^2.}$$

QUESTION–2

- <u>Occupons-nous de Z_1.</u>

Par linéarité de l'espérance, on a :

$$\begin{aligned} E(Z_1) &= E\Big(\frac{2}{\sqrt{n}}\Big(X_3 - \frac{n}{2}\Big)\Big) \\ &= \frac{2}{\sqrt{n}}\Big(E(X_3) - \frac{n}{2}\Big) \\ &= \frac{2}{\sqrt{n}}\Big(\frac{n}{2} - \frac{n}{2}\Big) \quad \text{car } X_3 \text{ suit la binomiale } \mathcal{B}(n, \tfrac{1}{2}). \end{aligned}$$

$$\boxed{E(Z_1) = 0.}$$

$$\begin{aligned} V(Z_1) &= \frac{4}{n}V(X_3 - \frac{n}{2}) \quad \text{car } V(\lambda Z) = \lambda^2 V(Z) \\ &= \frac{4}{n}V(X_3) \quad \text{car } V(Z + a) = V(Z) \\ &= \frac{4}{n}n\frac{1}{2}\frac{1}{2}. \end{aligned}$$

$$\boxed{V(Z_1) = 1. \text{ La variable } Z_1 \text{ est centrée, réduite.}}$$

- <u>Occupons-nous de Z_2.</u>

$$\begin{aligned} V(Z_2) &= \sqrt{\frac{2}{n}}\Big(E(X_1) - E(X_2)\Big) \quad \text{par linéarité} \\ &= \sqrt{\frac{2}{n}}\Big(\frac{n}{4} - \frac{n}{4}\Big) \quad \text{puisque } X_1 \text{ et } X_2 \text{ suivent la binomiale } \mathcal{B}(n, \tfrac{1}{4}). \end{aligned}$$

$$\boxed{E(Z_2) = 0.}$$