ESSEC I 2011

Le sujet est composé de deux problèmes indépendants.

Problème 1 : Évolution des intentions de vote

Dans une élection à venir, deux candidats A et B se présentent.
Un groupe d'électeurs est composé de m individus, avec $m \geqslant 2$.

- Initialement, au jour appelé « jour 0 », le nombre d'individus préférant le candidat A vaut a (il y en a donc $m - a$ préférant le candidat B).
- Ensuite, chaque jour, un des individus au hasard dans le groupe en rencontre un autre, au hasard également, et il lui parle des élections. Si leurs intentions de vote diffèrent, il le convainc de voter comme lui.

Pour tout entier naturel n, on note X_n le nombre d'individus du groupe ayant l'intention de voter pour le candidat A le soir du $n^{\text{ème}}$ jour. Ainsi, X_n est une variable aléatoire à valeurs dans $[\![0, m]\!]$.
On remarque que X_0 est une variable aléatoire certaine : $\mathbb{P}\big([X_0 = a] \big) = 1$.

Partie I - Un cas particulier : $m = 4$

Dans cette partie, on étudie le cas d'un groupe formé de quatre électeurs.

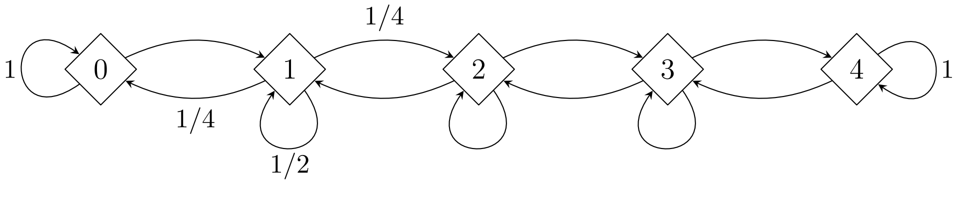
1. Soit i et j deux entiers dans $[\![0, 4]\!]$. On note $p_{i,j}$ la probabilité pour qu'il y ait exactement j personnes dans le groupe ayant l'intention de voter pour A un jour donné, sachant qu'il y en avait i la veille.

a) Justifier : $p_{0,0} = p_{4,4} = 1$.

b) Justifier : si i et j dans $[\![0, 4]\!]$ sont tels que $|i - j| \geqslant 2$, alors $p_{i,j} = 0$.

c) Établir : $p_{1,0} = p_{1,2} = \frac{1}{4}$ et $p_{1,1} = \frac{1}{2}$.

d) De la même façon, donner pour tout $(i, j) \in [\![0, 4]\!]^2$ la probabilité $p_{i,j}$.
On présentera les résultats sur le diagramme suivant, à reproduire et à compléter, et on justifiera quelques cas.



2. On définit la matrice $M = \begin{pmatrix} 1/2 & 1/3 & 0 \\ 1/4 & 1/3 & 1/4 \\ 0 & 1/3 & 1/2 \end{pmatrix}$,

et pour tout entier naturel n, la matrice colonne $U_n = \begin{pmatrix} \mathbb{P}\big([X_n = 1] \big) \\ \mathbb{P}\big([X_n = 2] \big) \\ \mathbb{P}\big([X_n = 3] \big) \end{pmatrix}$.

a) Pour tout entier naturel n, établir la relation : $U_{n+1} = M\, U_n$.
En déduire pour tout entier naturel n, l'égalité : $U_n = M^n\, U_0$.

b) Montrer que M admet trois valeurs propres distinctes α, β et γ, vérifiant $0 \leqslant \alpha < \beta < \gamma < 1$.

Justifier qu'il existe une matrice carrée P d'ordre 3 inversible, que l'on ne demande pas de préciser, et D une matrice diagonale d'ordre 3, à préciser, telles que $P^{-1} M P = D$.

c) En déduire que pour tout $k \in \{1, 2, 3\}$, la suite $\Big(\mathbb{P}\big([X_n = k]\big)\Big)_{n\in\mathbb{N}}$ est une combinaison linéaire des trois suites $\big(\alpha^n\big)_{n\in\mathbb{N}}$, $\big(\beta^n\big)_{n\in\mathbb{N}}$ et $\big(\gamma^n\big)_{n\in\mathbb{N}}$.

d) Montrer que pour tout $k \in \{1, 2, 3\}$, $\lim\limits_{n\to+\infty} \mathbb{P}\big([X_n = k]\big) = 0$.

3. Établir : $\lim\limits_{n\to+\infty} \Big(\mathbb{P}\big([X_n = 0]\big) + \mathbb{P}\big([X_n = 4]\big)\Big) = 1$.
Comment interpréter ce résultat ?

Partie II - Le cas général

On revient dans cette partie au cas général d'un groupe de m électeurs.
On note $\pi_{n,k} = \mathbb{P}\big([X_n = k]\big)$, la probabilité pour qu'il y ait exactement k électeurs envisageant de voter pour A à l'issue du $n^{\text{ème}}$ jour.

4. Soit n un entier naturel.

a) Établir les trois relations :

$$\forall k \in [\![0, m-1]\!], \quad \mathbb{P}_{[X_n=k]}\big([X_{n+1} = k+1]\big) = \frac{k\,(m-k)}{m(m-1)}$$

$$\forall k \in [\![1, m]\!], \quad \mathbb{P}_{[X_n=k]}\big([X_{n+1} = k-1]\big) = \frac{k\,(m-k)}{m(m-1)}$$

$$\forall k \in [\![1, m-1]\!], \quad \mathbb{P}_{[X_n=k]}\big([X_{n+1} = k]\big) = 1 - \frac{2k\,(m-k)}{m(m-1)}$$

b) En déduire la relation, si $k \in [\![1, m-1]\!]$:

$$\pi_{n+1,k} = \frac{(k-1)(m+1-k)\ \pi_{n,k-1} + \big(m(m-1) - 2k(m-k)\big)\ \pi_{n,k} + (k+1)(m-1-k)\ \pi_{n,k+1}}{m(m-1)}$$

5. *a)* Montrer par récurrence que pour tout entier naturel n et pour tout $k \in [\![1, m-1]\!]$:

$$\pi_{n,k} \leqslant \left(\frac{m(m-1) - 2}{m(m-1)}\right)^n$$

b) En déduire, pour tout $k \in [\![1, m-1]\!]$, la limite de $\pi_{n,k}$ lorsque n tend vers $+\infty$.

6. On définit l'évènement V_A (respectivement V_B) suivant : « au bout d'un certain nombre de jours, tous les individus du groupe ont l'intention de voter pour A (respectivement pour B) ».

a) Montrer : $\mathbb{P}(V_A) = \lim\limits_{n\to+\infty} \mathbb{P}\big([X_n = m]\big)$ et $\mathbb{P}(V_B) = \lim\limits_{n\to+\infty} \mathbb{P}\big([X_n = 0]\big)$.

b) Montrer : $\mathbb{P}(V_A) + \mathbb{P}(V_B) = 1$.
Que signifie ce résultat ?

7. Pour tout entier naturel n, on pose $Z_n = X_{n+1} - X_n$.

a) Justifier : $Z_n(\Omega) = \{-1, 0, 1\}$.

b) Exprimer $\mathbb{P}\big([Z_n = 1]\big)$ en fonction des probabilités $\pi_{n,k}$ avec $k \in [\![1, m-1]\!]$.

c) Comparer $\mathbb{P}\big([Z_n = -1]\big)$ et $\mathbb{P}\big([Z_n = 1]\big)$.

d) En déduire que $\mathbb{E}(Z_n) = 0$.

e) Montrer que la suite $\big(\mathbb{E}(X_n)\big)_{n\in\mathbb{N}}$ est constante et déterminer cette constante en fonction de a.

8. Montrer que $\mathbb{P}(V_A) = \dfrac{a}{m}$ et interpréter ce résultat.

Problème 2 - Une propriété limite des lois de Pareto

Question préliminaire

Soit g une fonction continue sur un intervalle I, à valeurs réelles.

1. a) Montrer que pour tout α et β dans I tels que $\alpha < \beta$:

$$\frac{1}{\beta-\alpha}\int_{\alpha}^{\beta} g(t)\ dt = \int_0^1 g\big(\alpha+(\beta-\alpha)x\big)\ dx$$

b) Soit a, b, c, d dans I tels que $a < c < d < b$.
On suppose g décroissante sur I, établir l'encadrement :

$$\frac{1}{b-c}\int_c^b g(t)\ dt \leqslant \frac{1}{d-c}\int_c^d g(t)\ dt \leqslant \frac{1}{d-a}\int_a^d g(t)\ dt$$

Partie I - Partie fractionnaire d'une variable à densité

- Pour tout réel x positif ou nul :
  - on note $[x]$ la *partie entière* de x. On rappelle qu'il s'agit de l'unique entier naturel n qui vérifie l'encadrement : $n \leqslant x < n+1$.
  - on note $\big\{x\big\} = x - [x]$, que l'on appelle la *partie fractionnaire* de x.

  Par exemple, si $x = 12,34$, alors $[x] = 12$ et $\{x\} = 0,34$.
- Dans cette partie, X désigne une variable aléatoire à valeurs réelles admettant une densité f qui vérifie les propriétés :
  - f est nulle sur $]-\infty,0[$;
  - la restriction de f à $[0,+\infty[$ est continue et décroissante.

  On pose $M = f(0)$, c'est le maximum de f sur $\mathbb{R}$.

Soit $Y = \big\{X\big\} = X - \big[X\big]$, la variable aléatoire égale à la partie fractionnaire de X.
On note F_Y la fonction de répartition de Y.

2. Que vaut $F_Y(y)$ lorsque $y < 0$? Que vaut $F_Y(y)$ lorsque $y \geqslant 1$?
On justifiera les réponses.

3. Justifier l'égalité entre évènements : $[Y=0] = \displaystyle\bigcup_{n\in\mathbb{N}} [X=n]$.
En déduire : $F_Y(0) = 0$.

4. Soit y un réel de l'intervalle $]0,1[$.

a) Montrer l'égalité : $F_Y(y) = \displaystyle\sum_{n=0}^{+\infty} \int_n^{n+y} f(t)\ dt$.

b) Montrer, en utilisant la question préliminaire, les inégalités :
- Pour tout n entier naturel : $\displaystyle\int_n^{n+y} f(t)\ dt \geqslant y\int_n^{n+1} f(t)\ dt$.
- Pour tout n entier naturel non nul : $\displaystyle\int_n^{n+y} f(t)\ dt \leqslant y\int_{n-1+y}^{n+y} f(t)\ dt$.

c) En déduire : $y\displaystyle\int_0^{+\infty} f(t)\ dt \leqslant F_Y(y) \leqslant \int_0^y f(t)\ dt + y\int_y^{+\infty} f(t)\ dt$, puis l'encadrement :

$$y \leqslant F_Y(y) \leqslant y+M$$

Partie II - Premier chiffre significatif d'une variable de Pareto

Pour tout réel λ strictement positif, on définit la fonction g_λ sur $\mathbb{R}$ par $g_\lambda : x \mapsto \begin{cases} \dfrac{\lambda}{x^{\lambda+1}} & \text{si } x \geqslant 1 \\ 0 & \text{sinon} \end{cases}$.

5. Montrer que pour tout réel λ strictement positif, g_λ est une densité de probabilité sur $\mathbb{R}$ (loi dite *de Pareto*).

Dans toute la suite, on note Z_λ une variable aléatoire admettant g_λ pour densité.

6. Déterminer la fonction de répartition G_λ de Z_λ.

7. On note ln la fonction *logarithme népérien*, et log la fonction *logarithme décimal*.

Cette fonction est définie sur $]0, +\infty[$ par : $\log(x) = \dfrac{\ln(x)}{\ln(10)}$ pour tout réel x strictement positif.

On pose $X_\lambda = \log(Z_\lambda)$, et on note F_λ la fonction de répartition de X_λ.

a) Établir, pour tout réel x, l'égalité : $F_\lambda(x) = G_\lambda\big(10^x\big)$.

b) En déduire que X_λ suit une loi exponentielle dont on précisera le paramètre en fonction de λ.

8. On pose $Y_\lambda = \big\{X_\lambda\big\}$, la partie fractionnaire de X_λ.

Montrer, en utilisant les résultats de la partie I, que pour tout réel y de l'intervalle $]0, 1[$:

$$\lim_{\lambda \to 0^+} \mathbb{P}\big(\,[Y_\lambda \leqslant y]\,\big) = y$$

En déduire que, lorsque λ tend vers 0, Y_λ converge en loi vers une variable aléatoire suivant la loi uniforme sur l'intervalle $[0, 1]$.

9. Pour tout réel x supérieur ou égal à 1, on note $\alpha(x)$ le premier chiffre dans l'écriture décimale de x. C'est un entier de l'intervalle $[\![1, 9]\!]$.

Par exemple, $\alpha(50) = 5$ et $\alpha(213, 43) = 2$.

a) Pour tout $k \in [\![1, 9]\!]$, montrer l'équivalence :

$$\alpha(x) = k \;\Leftrightarrow\; \big\{\log(x)\big\} \in \big[\log(k), \log(k+1)\big[$$

b) On note $C_\lambda = \alpha\big(Z_\lambda\big)$ la variable aléatoire prenant comme valeur le premier chiffre de Z_λ.

Montrer, pour tout $k \in [\![1, 9]\!]$: $\displaystyle\lim_{\lambda \to 0^+} \mathbb{P}\big(\,[C_\lambda = k]\,\big) = \log\left(1 + \frac{1}{k}\right)$.

Cette loi limite obtenue pour le premier chiffre de Z_λ est appelée *loi de Benford*.