**Université Ibn Tofail**
**ENSAK 2012**

**Contrôle de Probabilité, S4**

*La présentation, la lisibilité, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.*

## Problème

*Ce problème a pour objet principal la modélisation d'un processus aléatoire ponctuel (discret) représenté par une suit de variables aléatoires de Bernoulli. Ce modèle est ensuite approché par un modèle continu, et dans la dernière partie, on s'intéresse, dans un cas particulier, à l'adéquation de ce modèle continu au modèle discret initial. Dans tout le problème, $\lambda$ désigne un nombre réel de l'intervalle ouvert $]0,1[$.*

### Partie I : Modèle discret.

On suppose donnée une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires de Bernoulli, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Pour tout $n$ de $\mathbb{N}$, on note $p_n$ le paramètre de la variable aléatoire $X_n$.
On suppose que $p_0$ appartient à l'intervalle ouvert $]0,1[$ et que pour tout $n$ de $\mathbb{N}$, on a les probabilités conditionnelles suivantes :

$$P_{(X_n=1)}(X_{n+1}=1) = P(X_n=1) = p_n \quad \text{et} \quad P_{(X_n=0)}(X_{n+1}=1) = \lambda P(X_n=1) = \lambda p_n$$

*[On rappelle que la probabilité conditionnelle $P_A(B)$ peut aussi se noter $P(B/A)$]*

1. (a) Montrer que pour tout entier $n$ de $\mathbb{N}$, on a : $p_{n+1} = (1-\lambda)p_n^2 + \lambda p_n$.
   (b) En déduire que pour tout entier $n$ de $\mathbb{N}$, on a : $0 < p_n < 1$
2. (a) Montrer que la suite $(p_n)_{n\in\mathbb{N}}$ est convergent et déterminer sa limite.
   (b) On pose $a = (1-\lambda)p_0 + \lambda$. Établir, pour tout $n$ de $\mathbb{N}$, l'inégalité : $p_n \leqslant a^n$.
   En déduire que la série de terme général $p_n$ est convergente.
3. Pour tout $n$ de $N$, on définit la variable aléatoire $Y_n$ par : $Y_n = \sum_{k=0}^{n} X_k$ et on note $E(Y_n)$ son espérance.

   Justifier l'existence de la limite $L$ de la suite $(E(Y_n))_{n\in\mathbb{N}}$.
4. (a) Exprimer, pour tout $n$ de $\mathbb{N}$, la covariance $Cov(X_n, X_{n+1})$ de $X_n$ et $X_{n+1}$ en fonction de $p_n$ et $p_{n+1}$. Les variables $X_n$ et $X_{n+1}$ sont-elles indépendantes ?
   (b) Montrer que $\lim_{n\to\infty}\left(\frac{p_{n+1}}{p_n}\right) = \lambda$.
   (c) Pour tout $n$ de $\mathbb{N}$, on note $r_n$ le coefficient de corrélation linéaire entre $X_n$ et $X_{n+1}$ :

$$r_n = \frac{Cov(X_n, X_{n+1})}{\sqrt{V(X_n)V(X_{n+1})}} \quad \text{où } V \text{ désigne la variance}$$

   Exprimer $r_n$ en fonction de $p_n$ et $p_{n+1}$.
   Montrer que lorsque $n$ tend vers $+\infty$, $r_n$ est équivalent à $\frac{1-\lambda}{\sqrt{\lambda}}p_n$

## Partie II : Simulation.

*On rappelle que la fonction Pascal* `random` *simule une variable aléatoire suivant une loi uniforme sur l'intervalle* $[0,1]$.

Soit $N$ un entier naturel non nul et inférieur ou égal à 200.

On considère la suite finie des $N+1$ variables aléatoires $X_0, X_1, ..., X_N$ vérifiant les conditions de la partie I, modélisée par l'arbre pondéré suivant, et on note encore $Y_N = X_0 + \cdots + X_N$.

On cherche à étudier cette situation à l'aide du programme suivant :

```
Program evaluation;
   var lambda,p0 : real;

function bernoulli(p:real):integer;
  begin
     if random <= p then bernoulli :=1 else bernoulli :=0;
  end;

function simulation(N:integer):integer;
  var c,i,x : integer; a,p,q :real;
  begin
    p:=p0; x:=bernoulli(p); c:=x;
    for i:=1 to N do
      begin
        q:=p;
        if x=0 then q:=p*lambda;
        x:=bernoulli(q); c:=c+x; p:= (1-lambda)*p*p + lambda * p;
      end;
    simulation:=c;
  end;

var y,k, N :integer ; T: array[0..200] of integer; begin
  readln(lambda);readln(p0);readln(N);randomize;
  for k:=0 to N do T[k]:=0;
  for k:=1 to 10000 do
    begin
      y:=simulation(N); T[y] := T[y]+1;
    end;
  for k:=0 to N do
     begin
       write(T[k]); write(' ');
     end;
  readln;
end.
```

1. Expliquer le résultat rendu par la fonction `bernoulli`.

2. Expliquer le fonctionnement de la fonction `simulation` et donner en particulier la signification du résultat rendu.

3. Le programme `evaluation` permet de simuler une variable aléatoire. En se référant à la loi faible des grands nombres, quelle loi de probabilité peut-on simuler grâce à ce programme?

## Partie III : Modèle continu.

soit $\ell$ tel que $0 < \ell < 1$ et soit $T$ un réel strictement positif. Pour tout $t$ de $[0,T]$, on définit une variable aléatoire $X(t)$ sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ qui suit une loi de Bernoulli de paramètre $p(t)$, c'est à dire que

: $p(t) = P(X(t) = 1)$. On suppose que la fonction $p$ est définie et dérivable sur $[0, T]$, de dérivée $p'$, et vérifie la relation :

$$\forall t \in [0, T] \qquad p'(t) = (1-\ell)p(t)(p(t)-1)$$

On note $p(0) = p_0$ et on suppose que $p_0$ appartient à l'intervalle ouvert $]0, 1[$.

1. Soit $f$ la fonction définie sur $[0, T]$ par $f(t) = p(t) \times e^{(1-\ell)t}$. Montrer que $f$ est croissante sur $[0, T]$ et en déduire que la fonction $p$ ne s'annule pas sur $[0, T]$.

2. (a) Soit $g$ la fonction définie sur $[0, T]$ par : $g(t) = \dfrac{e^{-(1-\ell)t}}{p(t)}$. Exprimer $g'(t)$ en fonction de $\ell$ et $t$ et en déduire qu'il existe une constante $k$ telle que, pour tout $t$ de $[0, T]$, $g(t) = k + e^{(\ell-1)t}$.

   (b) Montrer que, pour tout $t$ de $[0, T]$, on a : $p(t) = \dfrac{p_0}{p_0 + (1-p_0)e^{(1-\ell)t}}$.

   (c) Dresser le tableau de variations de $p$ sur $[0, T]$. Soit $(C)$ la courbe représentative de $p$ dans le plan rapporté à un repère orthogonal. À quelle condition, portant sur $p_0$, la courbe $(C)$ présente-t-elle un point d'inflexion? Quelles sont alors les coordonnées de ce point?

3. Pour tout $n \in \mathbb{N}^*$, on note $\delta = \dfrac{T}{n}$ et pour tout $k \in [\![0, n]\!]$, $t_k = k\delta$.

   Pour tout $n \in \mathbb{N}^*$, on définit la variable aléatoire $Z_n$ par : $Z_n = \sum\limits_{k=0}^{n} X(t_k)$, d'espérance $E(Z_n)$.

   (a) Montrer que la suite $\left(\dfrac{E(Z_n)}{n}\right)_{n\in\mathbb{N}^\times}$ est convergente et de limite $\dfrac{1}{T}\int\limits_0^T p(t)\, dt$. Cette limite sera notée $m(T)$ dans la suite de cette partie.

   (b) Justifier la validité du changement de variable $u = e^{(1-\ell)t}$ dans l'intégrale $\int\limits_0^T p(t)\, dt$ et en déduire que l'on a :

   $$m(T) = \frac{1}{(1-\ell)T} \int\limits_1^{e^{(1-\ell)T}} \left(\frac{1}{u} - \frac{1-p_0}{p_0 + (1-p_0)u}\right) du$$

   (c) En déduire une expression de $m(T)$ en fonction de $p_0$, $\ell$ et $T$ et montrer que , lorsque $T$ tend vers $+\infty$, $p_0$ et $\ell$ étant fixés, $m(T)$ est équivalent à $-\dfrac{\ln(1-p_0)}{(1-\ell)T}$

## Partie IV : Retour au modèle discret.

Soit $n$ un entier naturel non fixé. Avec les notations des parties I et III, on suppose que $p_0 = \dfrac{1}{3}$, $\ell = \dfrac{1}{2}$ et $T = 2n(1-\lambda)$.

1. Montrer que la fonction $p$ définie dans la partie III est deux fois dérivable sur $[0, T]$, et montrer que pour tout $t$ de $[0, T]$: $\quad p''(t) = \dfrac{1}{4}(2p(t)-1)p(t)(p(t)-1) \quad$ où $p''$ désigne la dérivée seconde de $p$.

2. On rappelle que pour tout $k$ de $[\![0, n]\!]$, $\quad t_k = k\delta = k\dfrac{T}{n}$ et que $p_k$ a été défini dans la partie I.

   Pour tout $k$ de $[\![0, n]\!]$, on pose $\varepsilon_k = p(t_k) - p_k$.

   (a) Établir, pour tout $k$ de $[\![0, n-1]\!]$, l'inégalité suivante : $\quad |p(t_{k+1} - p(t_k) - \delta p'(t_k)| \leqslant \dfrac{\delta^2}{8}$.

   (b) Établir, pour tout $k$ de $[\![0, n-1]\!]$, l'égalité : $p(t_k) + \delta p'(t_k) - p_{k+1} = \varepsilon_k[1 - (1-\lambda)(1 - p(t_k) - p_k)]$.

   (c) En déduire, pour tout $k$ de $[\![0, n-1]\!]$, l'inégalité suivante : $|\varepsilon_{k+1}| \leqslant \dfrac{\delta^2}{8} + \dfrac{1}{3}(\lambda+2)|\varepsilon_k|$.

   (d) Établir, pour tout $k$ de $[\![0, n]\!]$, l'inégalité : $|\varepsilon_k| \leqslant 6(1-\lambda)$.

3. Pour tout réel $\alpha$ tel que $\alpha > 18(1-\lambda)$, on pose : $N(\alpha) = \dfrac{1}{1-\lambda} \ln\left(\dfrac{\alpha}{12(1-\lambda)} - \dfrac{1}{2}\right)$.

   (a) Vérifier que pour tout réel $\alpha > 18(1-\lambda)$, on a $N(\alpha) > 0$.

   (b) Montrer que si $n \leqslant N(\alpha)$, alors pour tout $k$ de $[\![0, n]\!]$, on a : $\left|\dfrac{p(t_k) - p_k}{p(t_k)}\right| \leqslant \alpha$.

   (c) Montrer que, pour $\alpha$ fixé, $\lim\limits_{\lambda \to 1} N(\alpha) = +\infty$

   (d) Conclure sur la qualité de l'approximation du modèle discret par le modèle continu, lorsque $\lambda$ se "rapproche" de 1.

**Université Iben Toufail**
**EN.S.A de Kénitra**
**Cycle Préparatoire-$S_4$**

# Statistiques et Probabilités

## EXERCICE 1 : PROBABILITÉS (LOI MONÔME) **8 points**

Soit $f_n$ la fonction définie par :

$$f_n(x) = \begin{cases} nx^{n-1} & \text{si } x \in [0,1] \\ 0 & \text{si non} \end{cases} \quad \text{où n désigne un entier naturel non nul.}$$

1. Montrer que $f_n$ est une densité de probabilité
2. On considère une variable aléatoire $X_n$ réelle dont une densité de probabilité est $f_n$ on dit alors que $X_n$ suit la loi monôme d'ordre $n$
   a. Reconnaitre la loi $X_1$
   b. Dans le cas où $n$ est supérieur ou égal à 2,déterminer la fonction de répartition $F_n$ de $x_n$ ainsi que son espérance $E(X_n)$ et sa variance $V(X_n)$.
3. On considère deux variables aléatoires $U_n$ et $V_n$ suivant la loi monôme d'ordre $n$ $(n, > 2)$, indépendantes.
   On pose $M_n = \sup(U_n, V_n)$ et $T_n = \inf(U_n, V_n)$.
   a. Pour tout réel $x$, écrire l'événement $(M_n \leqslant x)$ à l'aide des événements $(U_n \leqslant x)$ et $(V_n \leqslant x)$.
   b. En déduire une densité de $M_n$. Vérifier que $M_n$ suit une loi monôme dont on donnera l'ordre, puis déterminer sans calcule $E(M_n)$
   c. Exprimer $M_n + T_n$ en fonction de $U_n + V_n$ , puis déduire, sans calcule d'intégrale la valeur de $E(T_n)$
   d. la variable aléatoire $T_n$ suit-elle une loi monôme ?.

## EXERCICE 2 : STATISTIQUES À DEUX VARIABLES **6 points**

Le tableau ci-dessous donne en dirhams le montant des remboursements annuels $y_i$ effectués de 2007 à 2011 par un ménage, à la suite de divers emprunts :

| Année | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|
| Rang $x_i$ de l'année | 1 | 2 | 3 | 4 | 5 |
| $y_i$ | 6096 | 7602 | 9170 | 11155 | 15385 |

1. Construire le nuage de points $M_i(x_i\,;\,y_i)$, avec $i$ compris entre 1 et 5, associe à cette série statistique. On prendra comme unité graphique 2 cm pour 1 en abscisse et 1 cm pour 1 000 dirhams en ordonnée.
   On commencera les graduations au point de coordonnées (0 ; 6000).
2. On pose, pour $i$ variant de 1 à 5, $z_i = \ln y_i$.
   a. Calculer $z_i$ en arrondissant les valeurs à $10^{-3}$ près.

b. Déterminer une équation de la droite d'ajustement de $z$ en $x$, obtenue par la méthode des moindres carrés. Les coefficients obtenus seront arrondis au centième.

c. En déduire que l'on peut écrire une relation entre $y$ et $x$ sous la forme : $y = Ae^{Bx}$ avec A et B des réelles à déterminer

d. En supposant, que cet ajustement reste valable en 2012, estimer le montant des remboursements annuels de ce ménage en 2012, arrondi au dirhams.

3. Ce ménage disposait de 50 000 DH de revenu annuel en 2010. On estime que son revenu annuel augmente de 2 % par an.
La banque alerte ses clients lorsque le montant des remboursements des emprunts dépasse le tiers du montant des revenus.
En quelle année la banque alertera-t-elle ce ménage ? Justifier.

**EXERCICE 3 : STATISTIQUES À UNE SEULE VARIABLE** **6 points**

Des gendarmes ont effectués un contrôle de vitesse sur le bord d'une route nationale, et ont relevés les résultats suivants.

| Vitesse | [50 ; 70[ | [70 ; 90[ | [90 ; 110[ | [110 ; 130[ |
|---|---|---|---|---|
| Effectif | 20 | 60 | 15 | 5 |

1. Déterminer la population et le caractère étudié.
2. Calculer La vitesse moyenne des automobilistes contrôles.
3. a. Donner le tableau des fréquences cumulées.
   b. Calculer la médiane ; premier et troisième quartile de cette série.
   c. Construire la boite à moustache pour cette série .
4. la distribution étudiée est-elle asymétrique ?