Bordeaux I, 2011-2012. M1 Enseignement des Mathématiques.

Probabilités et Statistiques

Examen du 15 décembre 2011 (14h-18h)

Exercice 1

Un message composé de n symboles binaires '0' ou '1'. Lors de sa transmission à un destinataire, chaque symbole est perturbé avec la probabilité p et se transforme alors en symbole opposé. On fait l'hypothèse que les erreurs de transmission sont indépendantes.

1) *Envoi simple.* On envoie le message une fois à son destinataire. Déterminer la probabilité pour que le message reçu ne coïncide pas avec le message d'origine.

2) *Envoi doublé.* Par précaution, on choisit d'envoyer le même message deux fois.

a) Avec quelle probabilité le i-ième symbole du premier message reçu est-il identique au i-ième symbole du deuxième message reçu ?

b) Avec quelle probabilité les deux messages reçus sont-ils identiques ?

c) Sachant que les deux messages reçus sont identiques, déterminer la probabilité pour que le message reçu ne coïncide pas avec le message d'origine.

d) On suppose $n = 100$ et $p = 0.001$. Que valent les probabilités trouvées aux questions 1 et 2.c ? Commenter.

Exercice 2

1. On considère un point aléatoire (U, V) de loi uniforme sur le carré $[0,1] \times [0,1]$.

a) Pour x choisi arbitrairement dans $[0,1]$, représenter graphiquement l'ensemble : $\{(u,v) \in [0,1] \times [0,1] \ / \ \mid u - v \mid \leq x\}$.

b) Déterminer la densité de $\mid U - V \mid$.

c) Montrer que $\mid U - V \mid$ a pour espérance $1/3$ et pour variance $1/18$.

2. Une expérience consiste à tirer deux points du segment $[0,1]$ de manière indépendante suivant la loi uniforme sur $[0,1]$, puis à relever la longueur du segment déterminé par ces deux points.

On répète plusieurs fois cette expérience. A chaque étape n, on calcule la moyenne arithmétique M_n des longueurs des n premiers segments obtenus.

a) Le théorème central-limite apporte-t-il des renseignements sur la suite des variables aléatoires M_n ?

b) A l'aide d'un tableur, un élève a engendré 1000 segments et obtenu comme valeur moyenne de leurs longueurs 0.3398 . Au vu de ce résultat, le professeur doit-il suspecter une erreur de programmation ? *(On argumentera soigneusement la réponse.)*

Exercice 3

On considère la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ sur l'ensemble d'états $\{1,2,3\}$ et dont la matrice de transition est

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1-a & 0 & a \\ a & 1-a & 0 \end{pmatrix}$$

avec $0 \leq a \leq 1$.

1) Déterminer pour quelles valeurs de a la chaîne est irréductible. Déterminer alors sa période.

2) On suppose dans cette question $0 < a < 1$.

Sachant qu'à l'instant initial, la chaîne est dans l'état 1, la suite des distributions de probabilités $(\, P(X_n = 1)\ P(X_n = 2)\ P(X_n = 3)\,)_{n\in\mathbb{N}}$ est-elle convergente ? Si oui, en déterminer la limite.

2) Peut-on trouver une valeur de a $(0 \leq a \leq 1)$ pour laquelle les trois propriétés suivantes sont remplies simultanément :

- la chaîne est irréductible,

- la chaîne possède une et une seule distribution de probabilité invariante $\pi = (\pi_1\ \pi_2\ \pi_3)$,

- sachant qu'à l'instant initial, la chaîne est dans l'état 1, la suite $(\, P(X_n = 1)\ P(X_n = 2)\ P(X_n = 3)\,)_{n\in\mathbb{N}}$ n'est pas convergente ?

Problème

Toutes les variables aléatoires qui apparaissent dans le problème sont supposées définies sur un même espace de probabilité $(\Omega, \mathcal{A}, P)$.

On considère dans tout le problème une variable aléatoire X de fonction de répartition F_X et admettant une densité f_X.

On appelle *médiane théorique* de X une solution de l'équation $F_X(x) = 1/2$.

Pour n dans $\mathbb{N}^*$, on considère un échantillon $X_1, X_2, ..., X_n$ de variables aléatoires indépendantes de même loi que X , et $\overline{X} = (\sum_{i=1}^{n} X_i)/n$ désigne la *moyenne empirique de l'échantillon.*

On admet l'existence de variables aléatoires $Y_1, Y_2, ..., Y_n$ telles que, quel que soit ω dans Ω, $Y_1(\omega), Y_2(\omega), ..., Y_n(\omega)$ constituent un réarrangement croissant de

$X_1(\omega), X_2(\omega), ..., X_n(\omega)$, de telle sorte que quel que soit ω dans Ω, $Y_1(\omega) \leq Y_2(\omega) \leq ... \leq Y_n(\omega)$.

En particulier, $Y_1 = \min(X_1, X_2, ..., X_n)$ et $Y_n = \max(X_1, X_2, ..., X_n)$.

Si n est impair, et si on pose $n = 2l + 1$, la variable aléatoire Y_{l+1} est appelée la *médiane empirique de l'échantillon* $X_1, X_2, ..., X_n$.

I. Pour tout réel x et tout entier k tel que $1 \leq k \leq n$, on note $J_k(x)$ la variable aléatoire définie par

$$J_k(x) = \begin{cases} 1 & si \quad \{X_k \leq x\} \\ 0 & si \quad \{X_k > x\} \end{cases}$$

et on pose $C_n(x) = \sum_{k=1}^{n} J_k(x)$.

1. a) Montrer que Y_1 et Y_n ont pour densités respectives les fonctions f_{Y_1} et f_{Y_n}, définies pour tout réel x par

$$f_{Y_1}(x) = n(1 - F_X(x))^{n-1} f_X(x) \qquad f_{Y_n}(x) = n(F_X(x))^{n-1} f_X(x)$$

b) Quelle est la loi de probabilité de $C_n(x)$?

c) Justifier l'égalité $\{Y_k \leq x\} = \{C_n(x) \geq k\}$.

d) Montrer que pour tout k tel que $1 \leq k \leq n$ et tout réel x,

$$F_{Y_k}(x) = \sum_{j=k}^{n} \begin{pmatrix} n \\ j \end{pmatrix} (F_X(x))^j (1 - F_X(x))^{n-j}$$

e) Montrer que pour j entier compris entre 1 et n,

$$j \begin{pmatrix} n \\ j \end{pmatrix} = (n - j + 1) \begin{pmatrix} n \\ j-1 \end{pmatrix}$$

f) Montrer que f_{Y_k} est définie par

$$f_{Y_k}(x) = k \begin{pmatrix} n \\ k \end{pmatrix} (F_X(x))^{k-1} (1 - F_X(x))^{n-k} f_X(x)$$

g) Montrer que si X admet une espérance, alors pour tout k tel que $1 \leq k \leq n$, Y_k admet une espérance, et que si X admet une variance, alors Y_k admet une variance.

II. Dans cette deuxième partie, on suppose que la fonction de répartition F_X est définie par

$$F_X(x) = \begin{cases} 1 - \dfrac{1}{\sqrt{x}} & si \quad x \geq 1 \\ 0 & si \quad x < 1 \end{cases}$$

1. a) Représenter le graphe de F_X. Justifier que X est une variable aléatoire à densité. Calculer la densité f_X.

b) La variable aléatoire X a-t-elle une espérance ? A-t-elle une variance ?

c) Montrer l'existence et l'unicité de la médiane théorique de X et la calculer.

d) Pour k tel que $1 \leq k \leq n$, expliciter f_{Y_k}, et en déduire un équivalent de $f_{Y_k}(x)$ lorsque x tend vers $+\infty$.

2. On suppose dans cette question $n \geq 3$.

a) Montrer que pour tout k tel que $1 \leq k \leq n-2$, Y_k admet une espérance.

b) En utilisant le changement de variable $t = 1/\sqrt{x}$, montrer que pour tout k tel que $1 \leq k \leq n-2$,

$$E(Y_k) = k \begin{pmatrix} n \\ k \end{pmatrix} \int_0^1 t^{n-k-2}(1-t)^{k-1} dt$$

c) On admet que pour tout couple d'entiers positifs ou nuls (r, s),

$$\int_0^1 t^r (1-t)^s dt = \frac{r!\, s!}{(r+s+1)!}$$

En déduire l'expression de $E(Y_k)$ pour k tel que $1 \leq k \leq n-2$.

3. On suppose dans cette question n impair et supérieur ou égal à 5. On pose $n = 2l+1$.

a) Etablir l'égalité $E(Y_{l+1}) = 4 + \dfrac{6}{l-1}$.

b) On suppose de plus $n \geq 9$. Justifier l'existence de la variance de Y_{l+1}. On admet sa valeur :

$$var(Y_{l+1}) = \frac{4(2l+1)(l+1)(4l-7)}{(l-1)^2(l-2)(l-3)}$$

b) On suppose dans cette question $l = 100$. Utiliser l'inégalité de Bienaymé-Tchébytchev pour déterminer un intervalle borné I tel que $P(Y_{l+1} \in I) \geq 0.90$.

III. Dans cette partie, on suppose que la loi de X est normale, d'espérance μ et de variance 1.

1) Justifier l'existence et l'unicité de la médiane théorique m_X de X, et l'exprimer en fonction de μ.

2) Pour n dans $\mathbb{N}$, on pose $k(n) = l+1$, où l désigne la partie entière de $n/2$. On admet que la suite $\sqrt{\dfrac{2n}{\pi}}\,(Y_{k(n)} - m_X)$ converge en loi vers la loi normale centrée réduite $\mathcal{N}(0,1)$. Utiliser ce résultat pour déterminer un intervalle de confiance asymptotique pour la médiane théorique de X au niveau de confiance de 0.95.

3) On considère l'estimateur $\overline{X}$ de μ.

a) Quelle est sa loi ?

b) Utiliser cet estimateur pour déterminer un intervalle de confiance pour m_X au niveau de confiance de 0.95.

4) Comparer les deux intervalles de confiance obtenus et commenter le résultat.