EXERCICE 3 Eml 2017

On considère une urne contenant initialement une boule bleue et deux boules rouges.

On effectue, dans cette urne, des tirages successifs de la façon suivante : on pioche une boule au hasard, on note sa couleur, puis on la replace dans l'urne en ajoutant une boule de la même couleur que celle qui vient d'être obtenue.

Pour tout k de $\mathbf{N}^*$, on note :

B_k l'événement : "on obtient une boule bleue au k-ième tirage"

R_k l'événement : "on obtient une boule rouge au k-ième tirage"

Partie I : Simulation informatique

1. Recopier et compléter la fonction suivante afin qu'elle simule l'expérience étudiée et renvoie le nombre de boules rouges obtenues lors des n premiers tirages, l'entier n étant entré en argument.

```
def EML(n) :
    b=1 # b désigne le nombre de boules bleues présentes dans l'urne
    r=2  # r désigne le nombre de boules rouges présentes dans l'urne
    s=0 # s désigne le nombre de boules rouges obtenues lors des n tirages
    for k in range(n) :
        x=np.random.uniform()
        if ...
            ...
        else
            ...
    return s
```

2. On exécute le programme suivant :

```
n=10
m=0
for i in range(1000):
    m+=EML(n)
print(m/1000)
```

On obtient `6.657`. Comment interpréter ce résultat ?

Partie II : Rang d'apparition de la première boule bleue et rang d'apparition de la première boule rouge

On définit la variable aléatoire Y égale au rang d'apparition de la première boule bleue et la variable aléatoire Z égale au rang d'apparition de la première boule rouge.

3. (a) Montrer : $\forall n \in \mathbf{N}^*, \mathbb{P}([Y = n]) = \frac{2}{(n+1)(n+2)}$.

   (b) La variable aléatoire Y admet-elle une espérance ? une variance ?

Partie III : Nombre de boules rouges obtenues au cours de n tirages

On définit, pour tout k de $\mathbf{N}^*$, la variable aléatoire X_k égale à 1 si on obtient une boule rouge au k-ième tirage et égale à 0 sinon.

On définit, pour tout n de $\mathbf{N}^*$, la variable aléatoire S_n égale au nombre de boules rouges au cours des n premiers tirages.

4. Donner, pour tout n de $\mathbf{N}^*$, une relation entre S_n et certaines variables aléatoires X_k pour $k \in \mathbf{N}^*$.
5. Déterminer la loi de X_1, son espérance et sa variance.
6. (a) Déterminer la loi du couple (X_1, X_2).
   (b) En déduire la loi de X_2.
   (c) Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?
7. Soit $n \in \mathbf{N}^*$ et $k \in [\![0; n]\!]$.
   (a) Calculer $\mathbb{P}\left(R_1 \cap \ldots \cap R_k \cap B_{k+1} \cap \ldots \cap B_n\right)$.
   (b) Justifier : $\mathbb{P}\left([S_n = k]\right) = \binom{n}{k}\mathbb{P}\left(R_1 \cap \ldots \cap R_k \cap B_{k+1} \cap \ldots \cap B_n\right)$,
   puis en déduire : $\mathbb{P}\left([S_n = k]\right) = \dfrac{2(k+1)}{(n+1)(n+2)}$
8. Montrer que, pour tout n de $\mathbf{N}^*$, S_n admet une espérance et : $\mathbf{E}\left[(\right] S_n) = \frac{2n}{3}$.
9. Soit $n \in \mathbf{N}^*$.
   (a) Montrer : $\forall k \in [\![0; n]\!], \quad \mathbb{P}_{[S_n=k]}\left([X_{n+1} = 1]\right) = \dfrac{k+2}{n+3}$.
   (b) En déduire : $\mathbb{P}\left([X_{n+1} = 1]\right) = \dfrac{\mathbf{E}\left[(\right] S_n) + 2}{n+3}$.
   (c) Déterminer alors la loi de la variable aléatoire X_{n+1}. Que remarque-t-on ?

Problème HEC 2019

Dans ce problème, on définit et on étudie les fonctions génératrices des moments et les fonctions génératrices des cumulants de variables aléatoires discrètes ou à densité.

Dans tout le problème :

— on note $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé et toutes les variables aléatoires introduites dans l'énoncé sont des variables aléatoires réelles définies sur $(\Omega, \mathcal{A})$;

— sous réserve d'existence, l'espérance et la variance d'une variable aléatoire X sont respectivement notées $\mathbf{E}[X]$ et $\mathbf{V}(X)$;

— pour toute variable aléatoire X et pour tout réel t pour lesquels la variable aléatoire e^{tX} admet une espérance, on pose :

$$M_X(t) = \mathbf{E}\left[e^{tX}\right] \quad \text{et} \quad K_X(t) = \ln(M_X(t)) \ ;$$

(les fonctions M_X et K_X sont respectivement appelées la *fonction génératrice des moments* et la *fonction génératrice des cumulants* de X)

— lorsque, pour un entier $p \in \mathbf{N}^*$, la fonction K_X est de classe $\mathcal{C}^p$ sur un intervalle ouvert contenant l'origine, on appelle *cumulant d'ordre p* de X, noté $Q_p(X)$, la valeur de la dérivée p-ème de K_X en 0 :

$$Q_p(X) = K_X^{(p)}(0).$$

Partie I. Fonction génératrice des moments de variables aléatoires discrètes

Dans toute cette partie :

— on note n un entier supérieur ou égal à 2 ;

— toutes les variables aléatoires considérées sont discrètes et à valeurs entières ;

— on note S une variable aléatoire à valeurs dans $\{-1, 1\}$ dont la loi est donnée par :

$$\mathbf{P}(S=-1) = \mathbf{P}(S=1) = \frac{1}{2}.$$

1. Soit X une variable aléatoire à valeurs dans $[[-n, n]]$.
   (a) Pour tout $t \in \mathbf{R}$, écrire $M_X(t)$ sous la forme d'une somme et en déduire que la fonction M_X est de classe $\mathcal{C}^\infty$ sur $\mathbf{R}$.
   (b) Justifier pour tout $p \in \mathbf{N}^*$, l'égalité : $M_X^{(p)}(0) = \mathbf{E}[X^p]$.
   (c) Soit Y une variable aléatoire à valeurs dans $[[-n, n]]$ dont la fonction génératrice des moments M_Y est la même que celle de X.
   On note G_X et G_Y les deux polynômes définis par :

$$\forall x \in \mathbf{R}, \quad \begin{cases} G_X(x) &= \displaystyle\sum_{k=0}^{2n} \mathbf{P}(X = k-n)\, x^k \\ G_Y(x) &= \displaystyle\sum_{k=0}^{2n} \mathbf{P}(Y = k-n)\, x^k. \end{cases}$$

   i. Vérifier pour tout $t \in \mathbf{R}$, l'égalité $G_X(e^t) = e^{nt} M_X(t)$.
   ii. Justifier la relation : $\forall t \in \mathbf{R}$, $G_X(e^t) = G_Y(e^t)$.
   iii. En déduire que la variable aléatoire Y suit la même loi que X.

2. Dans cette question, on note X_2 une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(2, \frac{1}{2}\right)$.
   On suppose que les variables aléatoire X_2 et S sont indépendantes et on pose $Y_2 = SX_2$.
   (a) i. Préciser l'ensemble des valeurs possibles de la variable aléatoire Y_2.
   ii. Calculer les probabilités $\mathbf{P}(Y_2 = y)$ attachées au diverses valeurs possibles y de Y_2.
   (b) Vérifier que la variable aléatoire $X_2 - (S+1)$ suit la même loi que Y_2.

3. Dans cette question, on note X_n une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(n, \frac{1}{2}\right)$.
   On suppose que les variables aléatoires X_n et S sont indépendantes et on pose $Y_n = SX_n$.
   (a) Justifier que la fonction M_{X_n} est définie sur $\mathbf{R}$ et calculer $M_{X_n}(t)$ pour tout $t \in \mathbf{R}$.
   (b) Montrer que la fonction M_{Y_n} est donnée par : $\forall t \in \mathbf{R}$,

$$M_{Y_n}(t) = \frac{1}{2^{n+1}}\left((1+e^t)^n + (1+e^{-t})^n\right).$$

(c) En utilisant l'égalité $(1+e^{-t})^n = e^{-nt}(1+e^t)^n$, montrer que Y_n suit la même loi que la différence $X_n - H_n$, où H_n est une variable aléatoire indépendante de X_n dont on précisera la loi.

Partie II. Propriétés générales des fonctions génératrices des cumulants et quelques exemples

1. Soit X une variable aléatoire et $\mathcal{D}_X$ le domaine de définition de la fonction K_X.
   (a) Donner la valeur de $K_X(0)$.
   (b) Soit $(a,b) \in \mathbf{R}^2$ et $Y = aX + b$. Justifier pour tout réel t pour lequel at appartient à $\mathcal{D}_X$, l'égalité :
   $$K_Y(t) = bt + K_X(at).$$
   (c) On suppose ici que les variables aléatoires X et $-X$ suivent la même loi.
   Que peut-on dire dans ce cas des cumulants d'ordre impair de la variable aléatoire X ?
2. Soit X et Y deux variables aléatoires indépendantes et $\mathcal{D}_X$ et $\mathcal{D}_Y$ les domaines de définition respectifs des fonctions K_X et K_Y.
   (a) Montrer que pour tout réel t appartenant à la fois à $\mathcal{D}_X$ et $\mathcal{D}_Y$, on a : $K_{X+Y}(t) = K_X(t) + K_Y(t)$.
   (b) En déduire une relation entre les cumulants des variables aléatoires X, Y et $X+Y$.
3. Soit un réel $\lambda > 0$ et soit T une variable aléatoire qui suit la loi de Poisson de paramètre λ.
   (a) Déterminer les fonctions M_T et K_T.
   (b) En déduire tous les cumulants de T.
4. Soit $(T_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires telles que, pour tout $n \in \mathbf{N}^*$, la variable aléatoire T_n suit la loi de Poisson de paramètre n. Pour tout $n \in \mathbf{N}^*$, on pose $W_n = \dfrac{T_n - n}{\sqrt{n}}$.
   (a) Justifier la convergence en loi de la suite de variables aléatoires $(W_n)_{n\in\mathbf{N}^*}$ vers une variable aléatoire W.
   (b) Déterminer la fonction K_{W_n}.
   (c) Montrer que pour tout $t \in \mathbf{R}$, on a : $\lim\limits_{n\to+\infty} K_{W_n}(t) = K_W(t)$.