AP 5

Variables aléatoires discrètes

Exercice 1 (ECRICOME)

Soit n un entier naturel non nul.
Dans une fête foraine, un stand propose le jeu suivant : le joueur lance n fois une pièce et compte le nombre de Pile obtenus. Si ce nombre est pair, le joueur est déclaré vainqueur, et s'il est impair, il est déclaré perdant.
Si le joueur est déclaré vainqueur, il gagne 10 euros pour chaque Pile obtenu, mais s'il a perdu, il doit payer 10 euros pour chaque Pile obtenu.
En particulier, s'il n'obtient aucun Pile, il est déclaré vainqueur, mais ne remporte rien. La pièce est truquée, et à chaque lancer, la probabilité d'obtenir Pile est égale à p ($p \in]0,1[$), et celle d'obtenir Face est de $1-p$.
On notera X la variable aléatoire égale au nombre de Pile obtenus, et G la variable aléatoire égale au gain algébrique du joueur.
Enfin, on notera A l'événement : "le joueur est déclaré vainqueur" et on dira que le jeu est favorable au joueur si l'espérance mathématique de la variable aléatoire G est positive.

Partie I

Dans cette partie, on suppose que $n = 3$ et $p = \dfrac{2}{3}$.

1. Reconnaître la loi de X et vérifier que : $\quad P(A) = \dfrac{13}{27}$.
2. Définir en `Python` une fonction d'entête `def simulX()` qui effectue trois lancers de la pièce et compte le nombre de piles obtenus.

   Cette fonction fera appelle à la commande `rd.random()` (de la bibliothèque `numpy.randon`, importée avec l'abréviation `rd`) qui retourne un nombre choisi au hasard de l'intervalle $]0,1[$.
3. Montrer que : $G(\Omega) = \{-30, -10, 0, 20\}$, puis expliciter la loi de G.
4. Calculer l'espérance de G. Le jeu est-il favorable au joueur ?
5. Définir en `Python` une fonction d'entête `def simulG()` qui simule la variable aléatoire G. Cette fonction fera appelle à la fonction `simulX` définie précédemment.

Partie II

Dans cette partie, on revient au cas général, où n est entier naturel non nul et $p \in]0,1[$.
Celui qui tient le stand souhaite rendre le jeu plus attractif en affichant "À ce jeu, il y a plus de gagnants que de perdants !", et cherche donc les conditions nécessaires sur p et n pour que son affichage ne soit pas mensonger.
Soit Y la variable aléatoire définie par : $Y = (-1)^X$.
Autrement dit, Y prend la valeur 1 lorsque X prend une valeur paire, et Y prend la valeur -1 lorsque X prend une valeur impaire.

6. (a) On note $Z = \dfrac{Y+1}{2}$.
   Déterminer $Z(\Omega)$, puis montrer que Z suit une loi de Bernoulli de paramètre $P(A)$.

   (b) Démontrer que : $E(Y) = 2P(A) - 1$.
7. (a) Donner la loi de X.

(b) En déduire que l'on a également : $E(Y) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} p^k (1-p)^{n-k}$,
puis que : $E(Y) = (1-2p)^n$.

8. Exprimer alors la valeur de $P(A)$ en fonction de n et p.

9. Démontrer que
$$P(A) \geq \frac{1}{2} \Longleftrightarrow \left[p \leq \frac{1}{2} \text{ OU ''}n \text{ est pair''} \right]$$

Partie III

Le concepteur du jeu souhaite cependant vérifier que, tout en laissant son jeu attractif (c'est à dire en faisant en sorte que $P(A) \geq \frac{1}{2}$), son activité soit rentable pour lui, autrement dit que le jeu soit défavorable au joueur (c'est à dire que $E(G) \leq 0$).

10. Exprimer G en fonction de X et Y. En déduire que : $E(G) = 10 \sum_{k=0}^{n} (-1)^k k P(X = k)$.

11. Démontrer que : $\forall k \in [\![1, n]\!]$, $k \binom{n}{k} = n \binom{n-1}{k-1}$.

12. Montrer que : $E(G) = -10np(1-2p)^{n-1}$.

13. Démontrer alors que :
$$\begin{cases} P(A) \geq \frac{1}{2} \\ \\ E(G) \leq 0 \end{cases} \Longleftrightarrow p \leq \frac{1}{2}$$

14. (a) Étudier la fonction f définie sur $\left[0, \frac{1}{2}\right]$ par : $\forall x \in \left[0, \frac{1}{2}\right]$, $f(x) = x(1-2x)^{n-1}$.

(b) Pour une valeur de n fixée, comment le concepteur du jeu doit-il truquer sa pièce (c'est à dire quelle valeur doit-il donner à $p \in \left[0, \frac{1}{2}\right]$) pour optimiser la rentabilité de son activité ?

15. Compléter la fonction suivante qui prend en entrée n et p et renvoie en sortie un vecteur ligne contenant 200 simulations de la variable aléatoire G :

```python
def simulationsG(n,p):
    Y = np.zeros(200)
    for k in range(200):
        Y = .....
    return( ..... )
```

Partie IV

Le forain décide de fixer $n = 2$ et $p = \frac{1}{4}$. En période estivale, il pense pouvoir compter sur la participation de 200 clients dans la journée. Avant de se décider à installer son stand, il voudrait être certain, avec un risque d'erreur inférieur à 10%, qu'il gagnera plus de 100 euros dans la journée.
Pour tout entier i compris entre 1 et 200, on note alors G_i le gain algébrique du i-ème joueur.
On note aussi J la variable aléatoire égale au gain du forain sur toute la journée.

16. Pour tout entier $i \in [\![1, 200]\!]$, donner la loi de G_i et calculer son espérance et sa variance.

17. Exprimer la variable aléatoire J en fonction des variables aléatoires G_i.
Démontrer alors que $E(J) = 500$ et $V(J) = 11250$.

18. Justifier que : $P(J \leq 100) \leq P(|J - 500| \geq 400)$.

19. *On appelle "inégalité de Bienaymé-Tchébychev", l'inégalité suivante :*

$$\forall \varepsilon > 0,\ P(|J - E(J)| \geq \varepsilon) \leq \frac{V(J)}{\varepsilon^2}.$$

Utiliser l'inégalité de Bienaymé-Tchébychev pour montrer que : $P(J \leq 100) \leq \dfrac{9}{128}$.

20. Compte tenu de ses exigences de rentabilité, le forain peut-il installer son stand ?

21. Compléter le programme suivant pour qu'il affiche une valeur approchée de $P(J \leq 100)$ (obtenue après 10000 simulations de la variable J) :

```python
n = 2
p = 1/4
res = 0
for k in range(10000):
    g = simulationsG(n,p)
    J = -np.sum(g)
    if ..... :
        res = res+1
f = .....
print(f)
```

Exercice 2 (EDHEC)

Un mobile se déplace sur les points à coordonnées entières d'un axe d'origine O.
Au départ, le mobile est à l'origine.
Le mobile se déplace selon la règle suivante : s'il est sur le point d'abscisse k à l'instant n, alors, à l'instant $(n+1)$ il sera sur le point d'abscisse $(k+1)$ avec la probabilité p $(0 < p < 1)$ ou sur le point d'abscisse 0 avec la probabilité $1 - p$.
Pour tout $n \in \mathbb{N}$, on note X_n l'abscisse de ce point à l'instant n et l'on a donc $X_0 = 0$.
On admet que, pour tout $n \in \mathbb{N}$, X_n est définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$.
Par ailleurs, on note T l'instant auquel le mobile se trouve pour la première fois à l'origine (sans compter son positionnement au départ).
Par exemple, si les abscisses successives du mobile après son départ sont 0, 0, 1, 2, 0, 0, 1, alors on a $T = 1$. Si les abscisses successives sont : 1, 2, 3, 0, 0, 1, alors on a $T = 4$.
On admet que T est une variable aléatoire définie sur $(\Omega, \mathscr{A}, P)$.

1. (a) Pour tout k de $\mathbb{N}^*$, exprimer l'évènement $(T = k)$ en fonction d'évènements mettant en jeu certaines des variables X_i.
   (b) Donner la loi de X_1.
   (c) En déduire $P(T = k)$ pour tout k de $\mathbb{N}^*$, puis reconnaître la loi de T.

2. (a) Montrer par récurrence que, pour tout entier naturel n, $X_n(\Omega) = [\![0, n]\!]$.
   (b) Pour tout n de $\mathbb{N}^*$, utiliser le système complet d'évènements $(X_{n-1} = k)_{0 \leq k \leq n-1}$ pour montrer que : $P(X_n = 0) = 1 - p$.

3. (a) Établir que : $\forall n \in \mathbb{N},\ \forall k \in \{1, 2, \dots n+1\},\ P(X_{n+1} = k) = pP(X_n = k - 1)$.

(b) En déduire que : $\forall n \in \mathbb{N}^*,\ \forall k \in \{0, 1, 2 \ldots, n-1\},\ P(X_n = k) = p^k(1-p)$.
En déduire également la valeur de $P(X_n = n)$. Donner une explication probabiliste de ce dernier résultat.

(c) Vérifier que : $\displaystyle\sum_{k=0}^{n} P(X_n = k) = 1$.

4. (a) Montrer que : $\forall n \geq 2,\ \displaystyle\sum_{k=1}^{n-1} kp^{k-1} = \frac{(n-1)\,p^n - n\,p^{n-1} + 1}{(1-p)^2}$.

(b) En déduire que : $E(X_n) = \dfrac{p(1-p^n)}{1-p}$.

5. (a) Montrer, en utilisant la question 3.(a), que : $\forall n \in \mathbb{N},\ E\left(X_{n+1}^2\right) = p\left(E\left(X_n^2\right) + 2E(X_n) + 1\right)$.

(b) Pour tout entier naturel n, on pose $u_n = E\left(X_n^2\right) + (2n-1)\dfrac{p^{n+1}}{1-p}$.

Montrer que : $u_{n+1} = pu_n + \dfrac{p(1+p)}{1-p}$.

(c) En déduire l'expression de u_n, puis celle de $E\left(X_n^2\right)$ en fonction de p et n.

(d) Montrer enfin que : $V(X_n) = \dfrac{p}{(1-p)^2}\left(1 - (2n+1)\,p^n(1-p) - p^{2n+1}\right)$.

6. Dans cette question et dans cette question seulement, on prend $p = \dfrac{1}{3}$.

(a) Compléter la fonction `Python` suivante qui prend n en entrée, simule l'expérience aléatoire étudiée, et renvoie en sortie la valeur prise par X_n.

```python
def X(n):
    x = 0
    for k in range(n):
        u = np.floor(rd.random()*3)
        if u==2 :
            x = ....
        else:
            x = ....
    return(x)
```

(b) Quelle instruction faut-il taper dans la console pour obtenir la valeur d'une simulation de la variable X_{10} ?

(c) On ajoute à la suite du programme précédent les instructions :

```python
p = 1/3
n = 10

x = np.zeros(10000)
for k in range(10000):
    x[k] = X(n)

print("m emp = ", np.mean(x))
print("m th = ", p*(1-p**n)/(1-p))

print("v emp = ", np.var(x))
print("v th = ", p/((1-p)^2)*(1-(2*n+1)*p^n*(1-p)-p^(2*n+1)))
```

Après exécution, la console affiche :

```
m emp =
   0.5108
m th =
   0.4999915
v emp=
   0.7548834
v th=
   0.7498222
```

Indiquer ce que fait ce programme et commenter les résultats obtenus.

Exercice 3 (ESSEC)

Partie 1 : Exemple introductif.

On effectue des lancers successifs (indépendants) d'un dé cubique équilibré, dont les faces sont numérotées de 1 à 6, et on note $X_1, X_2, \ldots X_n, \ldots$, les variables aléatoires donnant le numéro amené par le dé aux premier lancer, deuxième lancer, ...

Pour tout entier naturel n non nul, on note Y_n la somme des points obtenus aux n premiers lancers.

Enfin, pour tout entier naturel k non nul, la variable aléatoire T_k compte le nombre de celles des variables aléatoires $Y_1, Y_2, \ldots, Y_n, \ldots$ qui prennent une valeur inférieure ou égale à k.

Par exemple, si les cinq premiers numéros amenés par le dé sont, dans l'ordre : 3, 1, 2, 3, 6, alors les événements suivants sont réalisés : $(Y_1 = 3)$, $(Y_2 = 4)$, $(Y_3 = 6)$, $(Y_4 = 9)$, $(Y_5 = 15)$, et les variables aléatoires T_2, T_3, T_9 et T_{12} prennent respectivement pour valeurs 0, 1, 4 et 4.

1. **On s'intéresse dans cette question à la variable aléatoire T_{12}.**

   (a) Donner les valeurs prises par T_{12}. On explicitera un exemple de résultat correspondant à chacune des deux valeurs extrêmes.

   Quelle est la probabilité que T_{12} prenne la valeur 12 ?

   (b) Compléter le programme `Python` ci-dessous pour qu'il simule l'expérience aléatoire étudiée et affiche la valeur de T_{12}.

   On rappelle que, sur `Python`, `rd.random()` fournit un réel aléatoire appartenant à $]0, 1[$ et `np.floor` est la fonction partie entière.

```python
def simulT12():
    y = 0
    t = 0
    while ..... :
        x = np.floor(rd.random()*6) + 1
        y = .....
        t = .....
    return( ..... )
```

2. **On s'intéresse dans cette question à la variable aléatoire T_2.**

   (a) Déterminer la loi de probabilité de T_2.

   (b) Qu'obtient-on à l'affichage en exécutant le programme ci-dessous ?

```python
L = np.zeros(1,3)
for d1 in range(1,7):
    for d2 in range(1,7):
        if d1 > 2 :
            L[0] = L[0] + 1
        elif d1+d2 > 2 :
            L[1] = L[1] + 1
        else:
            L[2] = L[2] + 1
print(L/36)
```

Partie 2 : Cas général.

Dorénavant, on considère une suite $(X_i)_{i\geq 1}$ de variables aléatoires, définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$, mutuellement indépendantes, de même loi, à valeurs positives ou nulles.
Pour tout entier naturel n non nul, on pose alors : $Y_n = X_1 + X_2 + \cdots + X_n$ et on note F_n la fonction de répartition de la variable aléatoire Y_n.
On fixe un réel strictement positif x, et on s'intéresse au nombre T_x des variables aléatoires Y_n telles que l'événement $(Y_n \leq x)$ soit réalisé.

3. Démontrer que la suite $(F_n(x))_{n\geq 1}$ est décroissante.
4. Démontrer chacune des deux relations suivantes :
   - $P(T_x = 0) = 1 - F_1(x)$;
   - pour tout entier naturel n non nul, $P(T_x = n) = F_n(x) - F_{n+1}(x)$.
5. En déduire l'équivalence :

$$\sum_{n=0}^{+\infty} P(T_x = n) = 1 \quad \Leftrightarrow \quad \lim_{n\to+\infty} F_n(x) = 0.$$

   Autrement dit, T_x est une variable aléatoire si, et seulement si $\lim_{n\to+\infty} F_n(x) = 0$.

Partie 3 : Cas d'une loi géométrique.

Dans cette troisième partie, les variables aléatoires X_i, $i \in \mathbb{N}^*$, suivent la loi géométrique $\mathcal{G}(p)$ de paramètre $p \in]0,1[$ et on pose $q = 1 - p$. De plus, x est ici un entier naturel non nul fixé.
On rappelle que, par convention : $\binom{n}{m} = 0$ si n et m sont des entiers naturels tels que $m > n$.

6. Loi de Y_n (avec $n \in \mathbb{N}^*$).
   (a) Préciser $Y_n(\Omega)$.
   (b) Par un calcul de loi de somme, déterminer la loi de Y_2, puis celle de Y_3.
   (c) Démontrer que, pour tous entiers naturels m et n tels que $m \geq n$, on a l'égalité :

$$\sum_{k=n}^{m} \binom{k}{n} = \binom{m+1}{n+1}.$$

   (d) Démontrer par récurrence que, pour tout entier naturel n non nul :

$$P(Y_n = k) = \binom{k-1}{n-1} q^{k-n} p^n$$

   si k est un entier supérieur ou égal à n.

7. Calcul de $P(T_x = n)$.

   (a) Justifier que T_x est une variable aléatoire et préciser $T_x(\Omega)$.
   Calculer $P(T_x = 0)$.

   (b) Vérifier chacune des deux égalités suivantes :

$$\begin{aligned} F_n(x) &= p^n \sum_{k=n}^{x} \binom{k}{n} q^{k-n} - qp^n \sum_{k=n+1}^{x} \binom{k-1}{n} q^{k-n-1} \\ F_{n+1}(x) &= p^{n+1} \sum_{k=n+1}^{x} \binom{k-1}{n} q^{k-n-1} \end{aligned}$$

   En utilisant la question 4, en déduire le calcul de $P(T_x = n)$ pour n entier supérieur ou égal à 1.

   (c) Reconnaître la loi de T_x. Préciser son espérance et sa variance.

8. Sachant que les variables aléatoires X_1, $X_2 \ldots$ sont des temps d'attente, et en observant que la réalisation de n premiers succès équivaut à la réalisation du n-ième succès, donner une interprétation, soigneusement exposée, de chacune des variables aléatoires Y_n et T_x, et retrouver ainsi la loi de T_x.

Partie 4 : Cas d'une loi exponentielle.

Dans cette dernière partie, les variables aléatoires X_n suivent la loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda > 0$. On admettra qu'alors Y_n admet pour densité la fonction f_n définie sur $\mathbb{R}$ par :

$$f_n(t) = \begin{cases} 0 & \text{si } t < 0, \\ \dfrac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} & \text{si } t \geq 0. \end{cases}$$

9. À l'aide de la question 4, calculer $P(T_x = 0)$, puis $P(T_x = n)$ pour tout entier naturel n non nul.

10. Reconnaître la loi de T_x. Préciser son espérance et sa variance.