

DM 5 - sujet A

THÈME : VECTEURS ALÉATOIRES

Exercice I

Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$ telles que :

$$\forall (i, j) \in \mathbb{N}^2, \quad \mathbf{P}([X = i] \cap [Y = j]) = \frac{1}{e^{2^{i+1}} j!}.$$

1. Déterminer les lois de X et Y .
2. Reconnaître la loi de $X + 1$. En déduire l'existence et les valeurs de $\mathbf{E}(X)$ et $\mathbf{V}(X)$.
3. Justifier l'existence et déterminer les valeurs de $\mathbf{E}(Y)$ et $\mathbf{V}(Y)$.
4. Les variables aléatoires X et Y sont-elles indépendantes?
5. Calculer $\mathbf{P}(X = Y)$.

Exercice II

Soit $p \in \mathbb{N}^*$. Considérons la fonction X suivante :

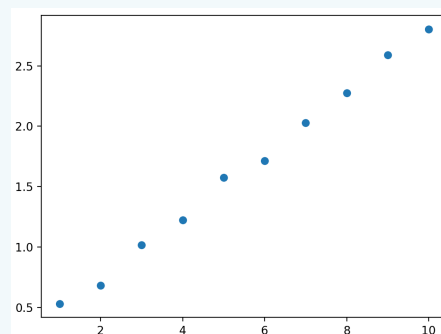
```
import numpy as np
import numpy.random as rd

def X(p):
    N=1+np.floor(p*rd.rand()) # floor(.) renvoie la partie entière
    x=0
    i=0
    while i < N:
        x+=(rd.random() < 0.5)
        i+=1
    return x
```

6. Reconnaître la loi suivie par la variable N ?
7. Calculer $\mathbf{E}(X)$.
8. Tester la cohérence du calcul grâce au code et résultat suivants :

```
import matplotlib.pyplot as plt

N=500
Y=np.zeros(10)
for p in range(1,11):
    s=0
    for i in range(N):
        s+=X(p)
    Y[p-1]=s/N
plt.plot(np.linspace(1,10,10),Y,'o')
plt.show()
```



DM 5 - sujet *

THÈME : VECTEURS ALÉATOIRES

Distance en variations totales

Toutes les variables aléatoires de ce problème sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans une partie de $\mathbb{N}$.

1. • *Préliminaires : Loi d'un couple*

Soit $p \in]0; 1/2[$. On considère deux variables aléatoires S, T respectivement à valeurs dans $\mathbb{N}$ et $\{0; 1\}$. On suppose que la loi du couple (S, T) est donnée par

$$\forall k \in \mathbb{N}, \quad \mathbf{P}([S = k] \cap [T = 0]) = \begin{cases} e^{-p} - p(1 - e^{-p}) & \text{si } k = 0 \\ 0 & \text{si } k = 1 \\ p^k e^{-p} / k! & \text{si } k \geq 2 \end{cases} \quad \text{et} \quad \mathbf{P}([S = k] \cap [T = 1]) = \begin{cases} p(1 - e^{-p}) & \text{si } k = 0 \\ p e^{-p} & \text{si } k = 1 \\ 0 & \text{si } k \geq 2. \end{cases}$$

- a) Vérifier que cela définit bien une loi de probabilité.
- b) Reconnaître les lois marginales de S et T .
- c) Exprimer $\mathbf{P}(S = T)$ en fonction de p .
- d) En déduire que $\mathbf{P}(S = T) \geq 1 - 2p^2$.

2. • *La distance en variation totale*

Pour X, Y deux variables aléatoires, on pose : $d(X, Y) = \sum_{n=0}^{+\infty} |\mathbf{P}([X = n]) - \mathbf{P}([Y = n])|$.

- a) Vérifier que $d(X, Y)$ est bien défini et $d(X, Y) \leq 2$.
- b) Que dire de X et Y si $d(X, Y) = 0$?
- c) Soient X, Y et Z trois variables aléatoires. Montrer que $d(X, Y) = d(Y, X)$ et $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

3. • *Une majoration de la distance*

- a) Justifier que pour toute partie A de $\mathbb{N}$, on a $|\mathbf{P}(X \in A) - \mathbf{P}(Y \in A)| \leq \mathbf{P}([X \neq Y])$.
- b) On pose $A = \{n \in \mathbb{N} \mid \mathbf{P}([X = n]) \geq \mathbf{P}([Y = n])\}$, justifier que

$$d(X, Y) = 2 \left| \mathbf{P}([X \in A]) - \mathbf{P}([Y \in A]) \right| \quad \text{puis} \quad d(X, Y) \leq 2\mathbf{P}([X \neq Y]).$$

- c) On suppose dans cette question que X s'exprime sous la forme $X = \sum_{i=1}^n X_i$ où les variables $(X_1, \dots, X_n)$ sont indépendantes et de même loi. De même, on suppose que $Y = \sum_{i=1}^n Y_i$ où les variables $(Y_1, \dots, Y_n)$ sont indépendantes et de même loi. Montrer alors que

$$d(X, Y) \leq 2n - 2 \sum_{i=1}^n \mathbf{P}(X_i = Y_i).$$

Indication. On pourra remarquer que $[X \neq Y] = \bigcup_{i \in \{1, \dots, n\}} [X_i \neq Y_i]$.

4. • *Inégalité*

Soient $\lambda \in \mathbb{R}_+^*$ et $n \in \mathbb{N}^*$. Montrer que si $X \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ et $Y \hookrightarrow \mathcal{P}(\lambda)$, alors : $d(X, Y) \leq \frac{4\lambda^2}{n}$.

5. Quel théorème retrouve-t-on lorsque $n \rightarrow +\infty$?

DM 5 - éléments de solution

Sujet A

1. On a $X(\Omega) = \mathbb{N}$ et pour $i \in \mathbb{N}$

$$\begin{aligned} \mathbf{P}(X=i) &= \sum_{j=0}^{+\infty} \frac{1}{e 2^{i+1} j!} = \frac{1}{2^{i+1}} \sum_{j=0}^{+\infty} \frac{1}{j!} \\ &= \frac{1}{2^{i+1}}. \end{aligned}$$

Car on reconnaît une série exponentielle de paramètre 1.

On a aussi $Y(\Omega) = \mathbb{N}$ et pour $j \in \mathbb{N}$

$$\begin{aligned} \mathbf{P}(Y=j) &= \sum_{i=0}^{+\infty} \frac{1}{e 2^{i+1} j!} = \frac{1}{2e j!} \sum_{i=0}^{+\infty} \left(\frac{1}{2}\right)^i \\ &= \frac{1}{2e j!} \cdot \frac{1}{1 - \frac{1}{2}} = \frac{1}{e j!}. \end{aligned}$$

Car on reconnaît une série géométrique de raison $1/2$.

2. On vérifie que $(X+1)(\Omega) = \mathbb{N}^*$ et pour $i \in \mathbb{N}^*$

$$\begin{aligned} \mathbf{P}(X+1=i) &= \mathbf{P}(X=i-1) \\ &= \frac{1}{2^i} = \frac{1}{2} \cdot \left(\frac{1}{2}\right)^{i-1}. \end{aligned}$$

On constate que $X+1 \hookrightarrow \mathcal{G}\left(\frac{1}{2}\right)$.

On en déduit que

$$\begin{aligned} \mathbf{E}(X) &= \mathbf{E}(X+1) - 1 = 2 - 1 = 1 \\ \mathbf{V}(X) &= \mathbf{V}(X+1) = 2. \end{aligned}$$

3. Pour tout $(i, j) \in X(\Omega) \times Y(\Omega)$.

$$\mathbf{P}([X=i] \cap [Y=j]) = \mathbf{P}(X=i) \mathbf{P}(Y=j).$$

Les variables sont indépendantes.

4. Comme Y suit une loi de Poisson de paramètre 1,

$$\mathbf{E}(Y) = 1, \quad \mathbf{V}(Y) = 1.$$

5. Comme $([X=i])_{i \in \mathbb{N}}$ est un système complet d'événements, la formule des probabilités totales donne :

$$\begin{aligned} \mathbf{P}(X=Y) &= \sum_{i=0}^{+\infty} \mathbf{P}([X=Y] \cap [X=i]) \\ &= \sum_{i=0}^{+\infty} \mathbf{P}([Y=i] \cap [X=i]) \\ &= \sum_{i=0}^{+\infty} \frac{1}{e 2^{i+1} i!} \\ &= \frac{1}{2e} \sum_{i=0}^{+\infty} \frac{(1/2)^i}{i!} = \frac{1}{2e} e^{1/2} = \frac{1}{2\sqrt{e}}. \end{aligned}$$

6. Si U suit une loi uniforme continue sur $[0; 1]$, alors pour $p \in \mathbb{N}^*$,

$$N = 1 + \lfloor pU \rfloor$$

suit une loi uniforme discrète sur $\llbracket 1; p \rrbracket$.

7. Si on note $(X_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires toutes suivant une même loi de Bernoulli $\mathcal{B}(\frac{1}{2})$ alors X a la même loi que

$$\sum_{i=0}^N X_i$$

On en déduit que

$$\mathbf{E}(X) = \mathbf{E}\left(\sum_{i=0}^N X_i\right).$$

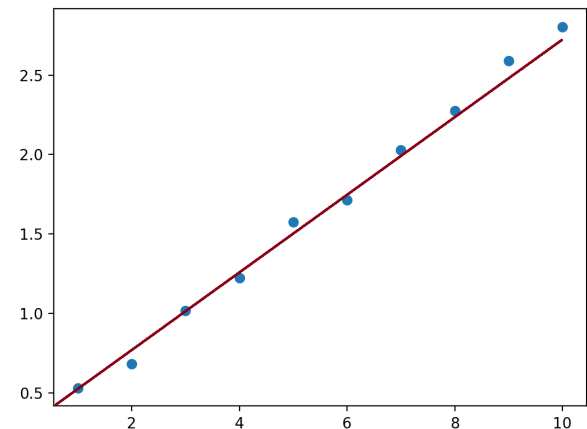
En reprenant la démonstration de l'exercice 75, on montre par la formule de l'espérance totale que

$$\mathbf{E}(X) = \mathbf{E}(N) \cdot \mathbf{E}(X_1)$$

Sachant que $N \hookrightarrow \mathcal{U}(\llbracket 1; p \rrbracket)$, $X_1 \hookrightarrow \mathcal{B}(1/2)$, il vient

$$\mathbf{E}(X) = \frac{p}{4}.$$

8. Le programme calcule pour chaque valeur de $p \in \llbracket 1; p \rrbracket$, une approximation de l'espérance de X . On vérifie graphiquement que la relation est linéaire et correspond à la formule précédente :



Sujet *

1.a) On vérifie que

$$\forall (k, \ell) \in \mathbb{N} \times \{0; 1\}, \quad \mathbf{P}([S = k] \cap [T = \ell]) \geq 0$$

$$\text{et} \quad \sum_{(k, \ell) \in \mathbb{N} \times \{0; 1\}} \mathbf{P}([S = k] \cap [T = \ell]) = 1 \quad (\bullet)$$

En effet, à partir des séries exponentielles

$$\begin{aligned} \sum_{k \in \mathbb{N}} \mathbf{P}([S = k] \cap [T = 0]) &= e^{-p} - p(1 - e^{-p}) + \sum_{k \geq 2} \frac{p^k e^{-k}}{k!} \\ &= -p + \sum_{k \geq 0} p^k e^{-k} / k! = 1 - p. \end{aligned}$$

puis

$$\sum_{k \in \mathbb{N}} \mathbf{P}([S = k] \cap [T = 1]) = p(1 - e^{-p}) + pe^{-p} = 0 = p.$$

On en déduit bien

$$\sum_{\ell \in \{0; 1\}} \sum_{k \in \mathbb{N}} \mathbf{P}([S = k] \cap [T = \ell]) = 1.$$

Comme les probabilités sont positives, cela justifie aussi qu'elles sont bien dans $[0; 1]$.

1.b) La variable S est à valeurs dans $\mathbb{N}$ avec pour $k \in \mathbb{N}$

$$\begin{aligned} \mathbf{P}(S = k) &= \mathbf{P}([S = k] \cap [T = 0]) + \mathbf{P}([S = k] \cap [T = 1]) \\ &= \begin{cases} e^{-p} & \text{si } k = 0 \\ pe^{-p} & \text{si } k = 1 \\ p^k e^{-k} / k! & \text{si } k \geq 2 \end{cases} \\ &= p^k e^{-k} / k!. \end{aligned}$$

On constate que S suit une loi de Poisson de paramètre p .

La variable T est à valeurs dans $\{0; 1\}$. En reprenant le calcul de la question précédente

$$\mathbf{P}(T = 0) = \sum_{k \in \mathbb{N}} \mathbf{P}([S = k] \cap [T = 0]) = 1 - p$$

$$\mathbf{P}(T = 1) = p.$$

Dès lors, T suit une loi de Bernoulli de paramètre p .

1.c) Comme T est à valeurs dans $\{0; 1\}$

$$\begin{aligned} \mathbf{P}(S = T) &= \mathbf{P}([S = 0] \cap [T = 0]) + \mathbf{P}([S = 1] \cap [T = 1]) \\ &= e^{-p} - p(1 - e^{-p}) + pe^{-p} \\ &= e^{-p}(1 + 2p) - p. \end{aligned}$$

1.d) À partir de l'inégalité de convexité

$$\forall x \in \mathbb{R}, \quad e^x \geq 1 + x.$$

On obtient pour $x = -p$.

$$\begin{aligned} \mathbf{P}(S = T) &\geq (1 - p)(1 + 2p) - p \\ &\geq 1 - 2p^2. \end{aligned}$$

2.a) Pour tout $n \in \mathbb{N}$

$$0 \leq |\mathbf{P}(X = n) - \mathbf{P}(Y = n)| \leq \mathbf{P}(X = n) + \mathbf{P}(Y = n).$$

Or les séries $\sum \mathbf{P}(X = n)$ et $\sum \mathbf{P}(Y = n)$ sont convergentes. Par le critère de comparaison, la série

$$\sum |\mathbf{P}(X = n) - \mathbf{P}(Y = n)|$$

est donc convergente. La quantité $d(X, Y)$ est bien définie avec

$$d(X, Y) \leq \sum_{n \in \mathbb{N}} \mathbf{P}(X = n) + \sum_{n \in \mathbb{N}} \mathbf{P}(Y = n) = 2.$$

2.b) Si $d(X, Y) = 0$ alors :

$$\forall n \in \mathbb{N}, \quad \mathbf{P}(X = n) = \mathbf{P}(Y = n).$$

Comme $X(\Omega) \subset \mathbb{N}$, $Y(\Omega) \subset \mathbb{N}$, on peut affirmer que X et Y ont même loi.

2.c) Comme

$$\forall n \in \mathbb{N}, \quad |\mathbf{P}(X = n) - \mathbf{P}(Y = n)| = |\mathbf{P}(Y = n) - \mathbf{P}(X = n)|.$$

on a bien $d(X, Y) = d(Y, X)$. De plus, par l'inégalité triangulaire, on a pour $n \in \mathbb{N}$

$$\begin{aligned} |\mathbf{P}(X = n) - \mathbf{P}(Y = n)| &= |\mathbf{P}(X = n) - \mathbf{P}(Z = n) + \mathbf{P}(Z = n) - \mathbf{P}(Y = n)| \\ &\leq |\mathbf{P}(X = n) - \mathbf{P}(Z = n)| + |\mathbf{P}(Z = n) - \mathbf{P}(Y = n)|. \end{aligned}$$

En sommant, on obtient bien

$$d(X, Y) \leq d(X, Z) + d(Z, X).$$

3.a) Les événements $[X = Y]$, $[X \neq Y]$ forment un système complet d'événements donc

$$\begin{aligned} \mathbf{P}([X \in A]) &= \mathbf{P}([X \in A] \cap [X = Y]) + \mathbf{P}([X \in A] \cap [X \neq Y]) \\ &= \mathbf{P}([Y \in A] \cap [X = Y]) + \mathbf{P}([X \in A] \cap [X \neq Y]) \\ &\leq \mathbf{P}([Y \in A]) + \mathbf{P}([X \neq Y]). \end{aligned}$$

Ce qui implique

$$\mathbf{P}([X \in A]) - \mathbf{P}([Y \in A]) \leq \mathbf{P}([X \neq Y]).$$

Par symétrie des rôles de X et Y

$$\mathbf{P}([Y \in A]) - \mathbf{P}([X \in A]) \leq \mathbf{P}([X \neq Y]).$$

D'où le résultat.

3.b) Comme $A \cup \bar{A} = \mathbb{N}$

$$\begin{aligned} d(X, Y) &= \sum_{n \in A} |\mathbf{P}(X = n) - \mathbf{P}(Y = n)| + \sum_{n \in \bar{A}} |\mathbf{P}(X = n) - \mathbf{P}(Y = n)| \\ &= \sum_{n \in A} \mathbf{P}(X = n) - \mathbf{P}(Y = n) + \sum_{n \in \bar{A}} -\mathbf{P}(X = n) + \mathbf{P}(Y = n). \end{aligned}$$

Or on a aussi

$$\begin{cases} \sum_{n \in A} \mathbf{P}(X = n) + \sum_{n \in \bar{A}} \mathbf{P}(X = n) = 1 \\ \sum_{n \in A} \mathbf{P}(Y = n) + \sum_{n \in \bar{A}} \mathbf{P}(Y = n) = 1. \end{cases}$$

D'où

$$\begin{aligned} d(X, Y) &= 2 \sum_{n \in A} \mathbf{P}(X = n) - 1 - \left(2 \sum_{n \in \bar{A}} \mathbf{P}(Y = n) - 1 \right) \\ &= 2 \sum_{n \in A} \mathbf{P}(X = n) - \mathbf{P}(X = n) \\ &= 2(\mathbf{P}(X \in A) - \mathbf{P}(Y \in A)). \end{aligned}$$

La seconde inégalité est directe avec l'inégalité de la question 3.a).

3.c) On a

$$\begin{aligned} d(X, Y) &\leq 2\mathbf{P}(X \neq Y) \\ &\leq 2\mathbf{P}\left(\bigcup_{i=1}^n [X_i \neq Y_i]\right) \quad (\bullet) \\ &\leq 2 \sum_{i=1}^n \mathbf{P}(X_i \neq Y_i) \\ &\leq 2 \sum_{i=1}^n (1 - \mathbf{P}(X_i = Y_i)). \end{aligned}$$

D'où le résultat.

Remarque. La relation $(\bullet)$ s'obtient en montrant par récurrence que

$$\forall A_1, \dots, A_n \in \mathcal{A}, \quad \mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbf{P}(A_i)$$

en s'aidant de la formule du crible

$$\begin{aligned} \forall A, B \in \mathcal{A}, \quad \mathbf{P}(A \cup B) &= \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \\ &\leq \mathbf{P}(A) + \mathbf{P}(B). \end{aligned}$$

4. Considérons $p = \lambda/n$ et les couples

$$(S_1, T_1), \quad \dots, \quad (S_n, T_n)$$

dont les lois sont identiques et explicitées à la questions

1. On suppose aussi les variables indépendantes. D'après

1.b), on sait que pour tout $i \in \llbracket 1; n \rrbracket$

$$S_i \hookrightarrow \mathcal{B}(\lambda/n) \quad \text{et} \quad T_i \hookrightarrow \mathcal{P}(\lambda/n).$$

Par les propriétés de stabilité par sommes dans le cas d'indépendance, on sait que

$$\bar{S} = \sum_{i=1}^n S_i \hookrightarrow \mathcal{B}(n, \lambda/n), \quad \bar{T} = \sum_{i=1}^n T_i \hookrightarrow \mathcal{P}(\lambda).$$

En particulier $\bar{S}$ et X ont même loi. De même pour Y et $\bar{T}$. Enfin, par la question précédente

$$\begin{aligned} d(X, Y) &= d(\bar{S}, \bar{T}) \\ &\leq 2n - 2 \sum_{i=1}^n \mathbf{P}(S_i = T_i) \\ &\leq 2n - 2n\mathbf{P}(S = T). \end{aligned}$$

Si $p = \lambda/n > 1/2$, l'inégalité est automatiquement vérifiée puisque $d(X, Y) \leq 2$. On peut donc supposer $p \leq 1/2$ et utiliser la minoration de la question 1.d)

$$d(X, Y) \leq 2n - 2n \left(1 - 2 \left(\frac{\lambda}{n}\right)^2\right) \leq \frac{4\lambda^2}{n}.$$

Ce qui permet de conclure.

5. On retrouve le théorème de convergence des lois binomiales vers la loi de Poisson.