

DM 4

THÈME : VECTEURS ALÉATOIRES, ESPÉRANCE CONDITIONNELLE
Sujet facultatif

Problème

Somme dont le nombre de termes est aléatoire

Problème 1. ♦♦♦♦

VeAp7

• Sommes de fonctions indicatrices

Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$.Soit A , une partie de $\mathbb{N}$, on pose $p(A)$ la probabilité $\mathbf{P}(X \in A)$ et on suppose que $0 < p(A) < 1$. On définit alors la variable aléatoire $\mathbf{1}_A(X)$ par :

$$\forall \omega \in \Omega, \quad \mathbf{1}_A(X)(\omega) = \begin{cases} 1 & \text{si } X(\omega) \in A \\ 0 & \text{si } X(\omega) \notin A. \end{cases}$$

1. a) Reconnaître la loi suivie par $\mathbf{1}_A(X)$.
- b) Soient $X_1, X_2, \dots, X_n$ des variables aléatoires suivant une même loi et indépendantes. Quelle est la loi suivie par la variable aléatoire $S_{A,n} = \sum_{k=1}^n \mathbf{1}_A(X_k)$?

Dans la suite, on considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires suivant une même loi et indépendantes. On considère par ailleurs une variable aléatoire N indépendante des variables X_k ($k \in \mathbb{N}^*$). Pour tout $n \in \mathbb{N}$, on pose $p_n = \mathbf{P}(N = n)$. On définit alors la variable aléatoire T_A par :

$$\forall \omega \in \Omega, \quad T_A(\omega) = \begin{cases} \sum_{k=1}^{N(\omega)} \mathbf{1}_A(X_k)(\omega) & \text{si } N(\omega) \geq 1 \\ 0 & \text{sinon.} \end{cases}$$

2. Soit $k \in \mathbb{N}$. Montrer que :

$$\mathbf{P}(T_A = k) = \sum_{n=k}^{+\infty} p_n \binom{n}{k} p(A)^k (1 - p(A))^{n-k}.$$

3. Montrer que

$$\forall (k, \ell) \in \mathbb{N}^2, \quad \mathbf{P}([T_A = k] \cap [T_{\bar{A}} = \ell]) = p_{k+\ell} \binom{k+\ell}{k} p(A)^k (1 - p(A))^\ell.$$

• Une caractérisation des lois de Poisson

4. Dans cette question, la variable N suit une loi de Poisson de paramètre λ .

a) Reconnaître la loi de T_A .b) Montrer que T_A et $T_{\bar{A}}$ sont indépendantes et déterminer la loi de $T_{\bar{A}}$.

5. On revient au cas général : on suppose que pour toute partie A de $\mathbb{N}$, les variables aléatoires T_A et $T_{\bar{A}}$ sont indépendantes. L'objectif est de montrer que la variable N suit une loi de Poisson. Pour cela, on pose

$$\forall t \in [-1; 1], \quad G(t) = \mathbf{E}(t^N).$$

- a) Soit A , une partie de $\mathbb{N}$. On note $p = p(A)$. Justifier que pour tout $s \in [-1; 1]$, $\mathbf{E}(s^{T_A}) = G((1-p) + ps)$. Adapter l'argument pour montrer de même que

$$\forall s, t \in [-1; 1], \quad \mathbf{E}(s^{T_A} t^{T_{\bar{A}}}) = G(ps + (1-p)t).$$

b) En déduire que pour tous $s, t \in [-1, 1]$,

$$G(ps + (1-p)t) = G((1-p) + ps)G(p + (1-p)t) \quad (\star)$$

c) On admet que la fonction G est dérivable sur $] -1; 1[$ et que la relation $(\star)$ est vraie pour tout $p \in]0; 1[$. Établir les égalités :

$$\forall p \in]0; 1[, \quad G(0) = G(p)G(1-p) \quad \text{et} \quad G'(0) = G'(p)G(1-p).$$

d) En déduire que la fonction G'/G est constante sur $]0; 1[$. Expliciter $G(x)$ pour $x \in [0; 1]$.

e) Soit $\tilde{N} \mapsto \mathcal{P}(\mu)$ avec $\mu \in \mathbb{R}_*^+$. Expliciter $E(t^{\tilde{N}})$ pour $t \in [-1; 1]$. En déduire que N suit une loi de Poisson. Pour le second point, on pourra remarquer que pour tout $p \in \mathbb{N}$

$$E(t^N) = \sum_{k=0}^p \mathbf{P}(N = k) t^k + o_0(t^p).$$

Exercice Python

Exercice 2. ♦♦ Soit $p \in \mathbb{N}^*$. Le code suivant permet de simuler une variable aléatoire X :

VeA34

Editeur

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

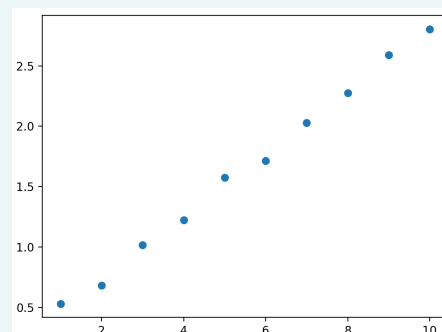
def X(p):
    N=1+np.floor(p*rd.random()) # floor(.) renvoie la partie entière
    x=0
    i=0
    while i < N:
        x+=(rd.random() < 0.5)
        i+=1
    return x
```

1. Reconnaître la loi suivie par la variable N ?
2. Calculer $E(X)$.
3. Tester la cohérence du calcul grâce au code et résultat suivants :

Editeur

```
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 4 - Solution

Exercice 1

1.a) On a $\mathbf{1}_A(X)(\Omega) = \{0; 1\}$. Donc $\mathbf{1}_A$ suit une loi de Bernoulli. Précisons le paramètre :

$$\mathbf{P}(\mathbf{1}_A(X) = 1) = \mathbf{P}(\{\omega \in \Omega \mid X(\omega) \in A\}) = \mathbf{P}(X \in A) = p(A).$$

Ainsi $\mathbf{1}_A(X) \hookrightarrow \mathcal{B}(p(A)).$

1.b) Par sommation de variables de Bernoulli, indépendantes et de même paramètre, on sait d'après le cours que

$$S_{A,n} \hookrightarrow \mathcal{B}(n, p(A)).$$

2. On a $T_A(\Omega) = \mathbb{N}$.

Soit $k \in \mathbb{N}^*$. Appliquons la formule des probabilités totales à l'aide du système complet d'événements $([N = n])_{n \in \mathbb{N}}$,

$$\begin{aligned} \mathbf{P}(T_A = k) &= \sum_{n \in \mathbb{N}} \mathbf{P}(N = n) \mathbf{P}_{[N=n]}(T_A = k) \\ &= \sum_{n \in \mathbb{N}} p_n \mathbf{P}_{[N=n]}(T_A = k) \\ &= \sum_{n=k}^{+\infty} p_n \mathbf{P}_{[N=n]}(T_A = k) \end{aligned}$$

car pour $k < n$, on a

$$\mathbf{P}_{[N=n]}(T_A = k) = 0.$$

Ensuite, par indépendance de N avec $S_{A,n}$ (lemme des coalitions)

$$\begin{aligned} \mathbf{P}_{[N=n]}(T_A = k) &= \mathbf{P}_{[N=n]}(S_{A,n} = k) = \mathbf{P}(S_{A,n} = k) \\ &= \binom{n}{k} p(A)^k (1 - p(A))^{n-k}. \end{aligned}$$

On en déduit que

$$\mathbf{P}(T_A = k) = \sum_{n=k}^{+\infty} p_n \binom{n}{k} p(A)^k (1 - p(A))^{n-k}.$$

L'égalité s'étend à $k = 0$, puisqu'on a aussi

$$\begin{aligned} \mathbf{P}(T_A = 0) &= \sum_{n=0}^{+\infty} p_n \mathbf{P}_{[N=n]}(T_A = 0) \\ &= p_0 \mathbf{P}_{[N=0]}(T_A = 0) + \sum_{n=0}^{+\infty} p_n \mathbf{P}_{[N=n]}(T_A = k). \end{aligned}$$

Or

$$p_0 \mathbf{P}_{[N=0]}(T_A = 0) = p_0 \cdot 1 = p_0$$

et d'autre part pour $n \in \mathbb{N}^*$, on a toujours

$$\begin{aligned} \mathbf{P}_{[N=n]}(T_A = 0) &= \mathbf{P}_{[N=n]}(S_{A,n} = 0) = \mathbf{P}(S_{A,n} = 0) \\ &= \binom{n}{0} p(A)^0 (1 - p(A))^{n-0}. \end{aligned}$$

3. Notons que $T_A + T_{\bar{A}} = N$. Ainsi

$$\begin{aligned} [T_A = k] \cap [T_{\bar{A}} = \ell] &= [T_A = k] \cap [N - T_A = \ell] \\ &= [T_A = k] \cap [N = \ell + k]. \end{aligned}$$

On en déduit

$$\begin{aligned} \mathbf{P}([T_A = k] \cap [T_{\bar{A}} = \ell]) &= \mathbf{P}([T_A = k] \cap [N = \ell + k]) \\ &= \mathbf{P}(N = \ell + k) \mathbf{P}_{[N=\ell+k]}(T_A = k) \\ &= \mathbf{P}(N = \ell + k) \mathbf{P}_{[N=\ell+k]}(S_{A,\ell+k} = k) \\ \mathbf{P}([T_A = k] \cap [T_{\bar{A}} = \ell]) &= \mathbf{P}(N = \ell + k) \mathbf{P}(S_{A,\ell+k} = k) \end{aligned}$$

par indépendance entre $S_{A,\ell+k}$ et N . Le résultat s'en déduit.

4.a) La variable T_A est à valeurs dans $\mathbb{N}$. On a en reprenant la question 2, pour tout $k \in \mathbb{N}$

$$\begin{aligned} \mathbf{P}(T_A = k) &= \sum_{n=k}^{+\infty} \frac{e^{-\lambda} \lambda^n}{n!} \binom{n}{k} p(A)^k (1 - p(A))^{n-k} \\ &= e^{-\lambda} \frac{p(A)^k}{k!} \sum_{n=k}^{+\infty} \frac{(1 - p(A))^{n-k}}{(n-k)!} \lambda^n \\ \text{car } \frac{1}{n!} \binom{n}{k} &= \frac{1}{n!} \cdot \frac{n!}{k!(n-k)!} = \frac{1}{k!(n-k)!}. \end{aligned}$$

On effectue ensuite le changement d'indice $i = n - k$.

$$\begin{aligned} \mathbf{P}(T_A = k) &= \frac{e^{-\lambda} (\lambda p(A))^k}{k!} \sum_{i=0}^{+\infty} \frac{(\lambda(1 - p(A)))^i}{i!} \\ &= \frac{e^{-\lambda} (\lambda p(A))^k}{k!} e^{\lambda(1 - p(A))} \\ &= \frac{(\lambda p(A))^k}{k!} e^{-\lambda p(A)}. \end{aligned}$$

Dans ce cas, T_A suit une loi de Poisson

$$T_A \hookrightarrow \mathcal{P}(\lambda p(A)).$$

4.b) De même avec $\bar{A}$ de probabilité $1 - p(A)$

$$T_{\bar{A}} \hookrightarrow \mathcal{P}(\lambda(1 - p(A))).$$

On peut maintenant tester l'indépendance.
Soient $k, \ell \in \mathbb{N}$.

$$\begin{aligned} & \mathbf{P}(T_A = k) \mathbf{P}(T_{\bar{A}} = \ell) \\ &= \frac{(\lambda p(A))^k}{k!} e^{-\lambda p(A)} \cdot \frac{(\lambda(1-p(A)))^\ell}{\ell!} e^{-\lambda(1-p(A))} \\ &= \frac{\lambda^{k+\ell} e^{-\lambda}}{k! \ell!} \cdot p(A)^k (1-p(A))^\ell \\ &= \frac{\lambda^{k+\ell} e^{-\lambda}}{(k+\ell)!} \cdot \frac{(k+\ell)!}{k! \ell!} \cdot p(A)^k (1-p(A))^\ell \\ &= p_{k+\ell} \cdot \binom{k+\ell}{k} \cdot p(A)^k (1-p(A))^\ell. \end{aligned}$$

En reprenant le résultat de la question 3, il vient

$$\mathbf{P}(T_A = k) \mathbf{P}(T_{\bar{A}} = \ell) = \mathbf{P}([T_A = k] \cap [T_{\bar{A}} = \ell]).$$

Ce qui conclut bien sur l'indépendance des deux variables.

5. Remarque. Notons que :

$$\forall t \in [-1; 1], \quad |t^N| \leq 1.$$

Par le théorème de domination, $\mathbf{E}(t^N)$ existe. La fonction G est bien posée sur $[-1; 1]$.

5.a) Soient $s \in [-1; 1]$, $n \in \mathbb{N}$, on a

$$\begin{aligned} \mathbf{E}(s^{T_A} \mid [N = n]) &= \mathbf{E}\left(s^{\sum_{k=1}^n \mathbf{1}_A(X_k)} \mid [N = n]\right) \\ &= \mathbf{E}\left(s^{\sum_{k=1}^n \mathbf{1}_A(X_k)}\right) \\ &\quad (\text{indépendance de } s^{\sum_{k=1}^n \mathbf{1}_A(X_k)} \text{ et } N, \text{ lemme des coalitions}) \\ &= \mathbf{E}\left(\prod_{k=1}^n s^{\mathbf{1}_A(X_k)}\right) \\ &= \prod_{k=1}^n \mathbf{E}(s^{\mathbf{1}_A(X_k)}) \\ &\quad (\text{indépendance par le lemme des coalitions}) \end{aligned}$$

$$\mathbf{E}(s^{T_A} \mid [N = n]) = \mathbf{E}(s^{\mathbf{1}_A(X_1)})^n \quad (\text{égalité en loi})$$

Or on a vu que $\mathbf{1}_A(X_1) \hookrightarrow \mathcal{B}(p(A))$. D'après la formule de transfert

$$\begin{aligned} \mathbf{E}(s^{\mathbf{1}_A(X_1)}) &= s^0 p(\mathbf{1}_A(X_1) = 0) + s^1 p(\mathbf{1}_A(X_1) = 1) \\ &= 1 - p(A) + p(A)s. \end{aligned}$$

$$\text{Il vient} \quad \mathbf{E}(s^{T_A} \mid [N = n]) = (1 - p(A) + p(A)s)^n.$$

Notons que $1 - p(A) + sp(A) \in [-1; 1]$ et

$$0 \leq \mathbf{P}(N = n) \mathbf{E}(|S|^{T_A} \mid [N = n]) \leq \mathbf{P}(N = n).$$

Par le critère de comparaison, la série

$$\sum \mathbf{P}(N = n) \mathbf{E}(s^{T_A} \mid [N = n])$$

converge absolument. Avec le système complet d'événements $([N = n])_{n \in \mathbb{N}}$, le théorème de l'espérance totale s'applique. Ainsi $\mathbf{E}(s^{T_A})$ converge avec

$$\begin{aligned} \mathbf{E}(s^{T_A}) &= \sum_{n=0}^{+\infty} \mathbf{P}(N = n) \mathbf{E}(s^{T_A} \mid [N = n]) \\ &= \sum_{n=0}^{+\infty} \mathbf{P}(N = n) (1 - p + ps)^n \quad (p = p(A)) \\ &= \mathbf{E}((1 - p + ps)^N) \\ \mathbf{E}(s^{T_A}) &= G(1 - p + ps) \end{aligned}$$

par la formule de transfert. Pour la seconde égalité, il suffit de procéder de même avec

$$\mathbf{E}(s^{T_A} t^{T_{\bar{A}}} \mid [N = n]) = (ps + (1-p)t)^n \quad (\bullet)$$

5.b) Comme T_A et $T_{\bar{A}}$ sont indépendantes, les variables

$$s^{T_A} \quad \text{et} \quad t^{T_{\bar{A}}}$$

sont aussi indépendantes (lemme des coalitions). En particulier

$$\mathbf{E}(s^{T_A} t^{T_{\bar{A}}}) = \mathbf{E}(s^{T_A}) \cdot \mathbf{E}(t^{T_{\bar{A}}})$$

En reprenant le raisonnement de la question précédente ou en prenant $s = 1$ dans la relation $(\bullet)$, on obtient

$$\mathbf{E}(t^{T_{\bar{A}}}) = G(p + (1-t)p).$$

Or on a vu que

$$\mathbf{E}(s^{T_A}) = G(ps + (1-p)) \quad \text{et} \quad \mathbf{E}(s^{T_A} t^{T_{\bar{A}}}) = G(ps + (1-p)t).$$

D'où le résultat.

5.c) Il suffit de prendre $s = t = 0$ pour avoir la première relation. Ensuite, en dérivant par rapport à t dans $(\star)$, on a

$$\begin{aligned} (1-p)G'(ps + (1-p)t) \\ = G((1-p) + ps) \cdot (1-p)G'(p + (1-p)t). \end{aligned}$$

On évalue de nouveau en $s = t = 0$ et on divise par $1-p \neq 0$ pour obtenir la seconde relation.

5.d) Précisons que $G(p) \neq 0$ pour tout $p \in [0; 1]$.

$$\frac{G'(p)}{G(p)} = \frac{G'(0)}{G(1-p)} \cdot \frac{G(1-p)}{G(0)} = \frac{G'(0)}{G(0)}.$$

La fonction G'/G est bien constante. Notons que pour $p \in [0; 1]$

$$G(p) = \sum_{k=0}^{+\infty} \mathbf{P}(N = n) t^p \geq \mathbf{P}(N = n_0) > 0.$$

où n_0 est un indice, $n_0 \in \mathbb{N}(\Omega)$. Ensuite, si on pose $H(p) = \ln G(p)$, on a par composition la dérivée de H avec

$$H'(p) = \frac{G'(p)}{G(p)} = -\lambda \quad \text{où} \quad \lambda = -\frac{G'(0)}{G(0)}.$$

Précisons que G est décroissante, donc $G'(0) < 0$.
D'où le choix du signe $-$ dans la définition de λ , afin d'avoir un paramètre positif.

Par intégration

$$H(p) - H(1) = \int_1^p H'(t) dt = -\lambda(p-1) = \lambda(1-p).$$

Sachant que $G(1) = E(1) = 1$, $H(1) = 0$ et

$$H(p) = \ln(G(p)) = \lambda p \quad \text{puis} \quad G(p) = \exp(\lambda(1-p)).$$

5.e) D'après la formule de transfert (on a déjà prouvé la convergence), on a pour tout $t \in [-1; 1]$

$$\begin{aligned} E(t^{\tilde{N}}) &= \sum_{n=0}^{+\infty} P(\tilde{N} = n) t^n \\ &= \sum_{n=0}^{+\infty} e^{-\mu} \frac{\mu^n}{n!} t^n \\ &= e^{-\mu} \sum_{n=0}^{+\infty} \frac{(\mu t)^n}{n!} \\ E(t^{\tilde{N}}) &= e^{-\mu} e^{\mu t} = e^{\mu(1-t)}. \end{aligned}$$

On constate alors que pour $\mu = \lambda$:

$$\forall t \in [0; 1], \quad E(t^{\tilde{N}}) = E(t^N).$$

$$\text{D'où } \forall t \in [0; 1], \quad \sum_{n=0}^{+\infty} P(N = n) t^n = \sum_{n=0}^{+\infty} P(\tilde{N} = n) t^n.$$

Justifions que cela impose :

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

Pour cela, on montre que pour tout $p \in \mathbb{N}$

$$\sum_{n=0}^p P(N = n) t^n + o(t^p) = \sum_{n=0}^p P(\tilde{N} = n) t^n + o(t^p),$$

puis on conclut par unicité du développement limité.

Exercice 2

1. 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 $\llbracket 1; p \rrbracket$.

2. 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

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

À partir de la formule de l'espérance totale, on montre que

$$E(X) = E(N) \cdot E(X_1)$$

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

$$E(X) = \frac{1+p}{4}.$$

3. Le programme calcule pour chaque valeur de $p \in \llbracket 1; 10 \rrbracket$, une approximation de l'espérance de X . C'est une conséquence de la loi faible des grands nombres. On vérifie graphiquement que la relation est linéaire et correspond à la formule précédente :

