Couples de v.a. discrètes (DS de l'année 2019-2020)

I. Exercices généraux

Exercice : EDHEC 2019 voie S

Partie 1 : fonction génératrice d'une v.a.r. discrète finie

Dans cette partie, n désigne un entier naturel non nul. On considère une variable aléatoire X prenant ses valeurs dans $[\![1, n]\!]$ et on appelle fonction génératrice de X, la fonction G définie par :

$$\forall t \in \mathbb{R}, \quad G(t) = \sum_{k=1}^{n} \mathbb{P}(\{X = k\})\, t^k$$

1. Calculer $G(1)$.

2. Exprimer l'espérance de X à l'aide de la fonction G.

3. Établir la relation : $\mathbb{V}(X) = G''(1) + G'(1) - \big(G'(1)\big)^2$.

Partie 2 : un résultat d'analyse

On pose, pour tout n de $\mathbb{N}^*$: $u_n = \sum\limits_{k=1}^{n} \dfrac{1}{k}$ et $h_n = \sum\limits_{k=1}^{n} \dfrac{1}{k^2}$.

4. ***a)*** Justifier que, pour tout entier naturel k non nul, on a : $\dfrac{1}{k+1} \leqslant \ln(k+1) - \ln(k) \leqslant \dfrac{1}{k}$.

b) Montrer alors : $\forall n \in \mathbb{N}^*,\ \ln(n) + \dfrac{1}{n} \leqslant u_n \leqslant \ln(n) + 1$.

c) En déduire un équivalent très simple de u_n lorsque n est au voisinage de $+\infty$.

5. Montrer que la suite $(h_n)_{n\in\mathbb{N}^*}$ est convergente.

Partie 3 : étude d'une expérience aléatoire

Dans cette partie, n désigne toujours un entier naturel non nul.

- On considère une urne contenant n jetons numérotés de 1 à n ($n \geqslant 2$).
  L'expérience aléatoire consiste à prélever tous ces jetons un par un, au hasard, et sans remise.
  Pour tout $i \in [\![1, n]\!]$, on note A_i la v.a.r. égale au numéro du jeton obtenu lors du $i^{\text{ème}}$ tirage.
- Pour tout $i \in [\![2, n]\!]$, on dit qu'il y a **record** à l'instant i si le $i^{\text{ème}}$ jeton tiré a un numéro plus grand que tous les numéros précédemment tirés.
  D'autre part, on convient qu'il y a systématiquement record à l'instant 1.
- Enfin, on note X_n la v.a.r. égale au nombre de records obtenus lorsque l'on procède à cette expérience dans une urne contenant n jetons.
  On note alors G_n la fonction génératrice de X_n, E_n son espérance et V_n sa variance.

6. Donner la loi de X_1.

7. ***a)*** Montrer : $X_n(\Omega) = [\![1, n]\!]$.

b) Déterminer $\mathbb{P}(\{X_n = 1\})$ et $\mathbb{P}(\{X_n = n\})$. En déduire les lois de X_2 et X_3.

c) En considérant le système complet d'événements $\big(\{A_n = n\}, \{A_n < n\}\big)$, montrer :

$$\forall n \geqslant 2,\ \forall j \in [\![2, n]\!], \quad \mathbb{P}(\{X_n = j\}) = \frac{1}{n}\ \mathbb{P}(\{X_{n-1} = j-1\}) + \frac{n-1}{n}\ \mathbb{P}(\{X_{n-1} = j\})$$

d) Donner la loi de X_4.

8. a) Vérifier que la formule obtenue à la question ***7.c)*** reste valable pour $j = 1$.

b) Établir la relation :

$$\forall n \geqslant 2,\ \forall t \in \mathbb{R}, \quad G_n(t) \ = \ \frac{t+n-1}{n}\ G_{n-1}(t) \qquad (\star)$$

c) En déduire :

$$\forall n \in \mathbb{N}^*,\ \forall t \in \mathbb{R}, \quad G_n(t) \ = \ \frac{1}{n!} \prod_{j=0}^{n-1} (t+j)$$

9. En dérivant la relation $(\star)$, trouver une relation entre E_n et E_{n-1} puis montrer :

$$\forall n \in \mathbb{N}^*,\ E_n = u_n$$

10. Recherche d'un équivalent de V_n.

a) En dérivant une deuxième fois la relation $(\star)$, montrer :

$$\forall n \geqslant 2,\ V_n - V_{n-1} = \frac{1}{n} - \frac{1}{n^2}$$

b) En déduire, pour tout entier naturel n non nul, V_n en fonction de u_n et h_n.

c) Montrer : $V_n \underset{n \to +\infty}{\sim} \ln(n)$.

Exercice : `ECRICOME 2018 - voie S`

Toutes les variables aléatoires dans ce problème sont supposées définies sur un même espace probabilisé noté $(\Omega, \mathscr{A}, \mathbb{P})$.

Partie 1 – Variables vérifiant une relation de Panjer

On dit qu'une variable aléatoire N, à valeurs dans $\mathbb{N}$ vérifie une relation de Panjer s'il existe un réel $a < 1$ et un réel b tels que :

$$\mathbb{P}(\{N = 0\}) \neq 1 \quad \text{et} \quad \forall k \in \mathbb{N}^*, \quad \mathbb{P}(\{N = k\}) = \left(a + \frac{b}{k}\right) \mathbb{P}(\{N = k-1\})$$

1. On suppose **dans cette question** que $a = 0$, et que b est un réel strictement positif.

a) Montrer :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(\{N = k\}) = \frac{b^k}{k!}\ \mathbb{P}(\{N = 0\})$$

b) Calculer $\sum\limits_{k=0}^{+\infty} \mathbb{P}(\{N = k\})$. En déduire que N suit une loi de Poisson de paramètre b.
Préciser son espérance et sa variance.

2. On suppose **dans cette question** que $a < 0$ et que $b = -2a$.

a) Montrer :
$$\forall k \geqslant 2, \quad \mathbb{P}(\{N = k\}) = 0$$

b) En déduire que N suit une loi de Bernoulli dont on précisera le paramètre en fonction de a.

3. On suppose **dans cette question** que Z suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0, 1[$.

a) Montrer :
$$\forall k \in [\![1, n]\!], \quad \mathbb{P}(\{Z = k\}) = \frac{p}{1-p} \times \frac{n-k+1}{k} \times \mathbb{P}(\{Z = k-1\})$$

b) En déduire que Z vérifie une relation de Panjer en précisant les valeurs de a et b correspondantes, en fonction de n et p.

4. On revient dans cette question au cas général : a est un réel vérifiant $a < 1$, b est un réel, et on suppose que N est une variable aléatoire, à valeurs dans $\mathbb{N}$, vérifiant la relation de Panjer.

a) Calculer $\mathbb{P}(\{N = 1\})$. En déduire : $a + b \geqslant 0$.

b) Montrer, pour tout entier $m \geqslant 1$:
$$\sum_{k=1}^{m} k\, \mathbb{P}(\{N = k\}) \; = \; a \sum_{k=0}^{m-1} (k+1)\, \mathbb{P}(\{N = k\}) + b \sum_{k=0}^{m-1} \mathbb{P}(\{N = k\})$$

c) En déduire que $\left((1-a) \sum\limits_{k=1}^{m} k\, \mathbb{P}(\{N = k\}) \right)_{m \geqslant 1}$ est majorée, puis que N admet une espérance. Préciser alors la valeur de $\mathbb{E}(N)$ en fonction de a et b.

d) Montrer que N admet un moment d'ordre 2 et :
$$\mathbb{E}\big(N^2\big) \; = \; \frac{(a+b)(a+b+1)}{(1-a)^2}$$

e) En déduire que N admet une variance et préciser la valeur de $\mathbb{V}(N)$ en fonction de a et b.

f) Montrer que $\mathbb{E}(N) = \mathbb{V}(N)$ si, et seulement si, N suit une loi de Poisson.

Partie 2 – Fonction génératrice

On notera dans la suite :
$$\forall k \in \mathbb{N}, \quad p_k = \mathbb{P}(\{N = k\})$$
où N est une variable aléatoire à valeurs dans $\mathbb{N}$.

5. Montrer que, pour tout réel x de l'intervalle $[0, 1]$, la série $\sum\limits_{n \geqslant 0} p_n\, x^n$ est convergente.

On appelle alors **fonction génératrice de** N la fonction G définie sur $[0, 1]$ par :
$$\forall x \in [0, 1], \quad G(x) \; = \; \sum_{k=0}^{+\infty} p_k\, x^k \; = \; \sum_{k=0}^{+\infty} \mathbb{P}(\{N = k\})\, x^k$$

et on suppose dans cette partie que N vérifie une relation de Panjer avec $0 < a < 1$ et que $\dfrac{b}{a} > 0$. On pose : $\alpha = \dfrac{-(a+b)}{a}$.

On note enfin f la fonction définie par :
$$\forall x \in [0, 1], \quad f(x) = p_0\, (1 - a\, x)^{\alpha}$$

6. Montrer, pour tout $k \in \mathbb{N}$:

$$\forall x \in [0,1], \quad f^{(k)}(x) \ = \ k! \times p_k \, (1 - a\,x)^{\alpha - k}$$

7. Soit $x \in [0,1]$.

a) Pour tout entier $n \in \mathbb{N}$, montrer :

$$f(x) \ = \ \sum_{k=0}^{n} p_k \, x^k + (n+1)\, p_{n+1} \int_0^x (1 - a\,t)^{\alpha - n - 1} \, (x-t)^n \ dt$$

b) Vérifier, pour tout $t \in [0, x]$: $\dfrac{x-t}{1-at} \leqslant 1$. Puis montrer, pour tout $n \in \mathbb{N}$:

$$0 \ \leqslant \ \int_0^x (1-at)^{\alpha-n-1} \, (x-t)^n \ dt \ \leqslant \ \int_0^x (1-at)^{\alpha - 1} \ dt$$

c) En déduire :

$$G(x) \ = \ p_0 \, (1 - ax)^{\alpha}$$

En calculant $G(1)$, exprimer p_0 en fonction de a, b et α, et vérifier : $G'(1) = \mathbb{E}(N)$.

Partie 3 – formule de récursivité

On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires de même loi, à valeurs dans $\mathbb{N}$, mutuellement indépendantes et indépendantes de la variable N étudiée dans la question ***4.*** de la **Partie 1**.
On considère alors la variable aléatoire S définie par :

$$S \ = \ \begin{cases} 0 & \text{si } N = 0 \\ \sum\limits_{k=1}^{N} X_k & \text{si } N \geqslant 1 \end{cases}$$

Autrement dit :

$$\forall \omega \in \Omega, \quad S(\omega) = 0 \ \text{ si } \ N(\omega) = 0 \qquad \text{et} \qquad S(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega) \ \text{ sinon}$$

8. Calculer $\mathbb{P}(\{S = 0\})$ lorsque $a \in \,]0,1[$ à l'aide de la **Partie 2**.

9. Calculer $\mathbb{P}(\{S = 0\})$ lorsque N suit une loi de Poisson de paramètre λ.

10. Dans la suite du problème, on revient au cas général où N vérifie la relation de Panjer. On note toujours :

$$\forall k \in \mathbb{N}, \quad p_k = \mathbb{P}(\{N = k\})$$

et on notera également :

$$\forall k \in \mathbb{N}, \quad q_k = \mathbb{P}(\{X_1 = k\})$$

On considère pour tout entier $n \geqslant 1$, la variable aléatoire $S_n = \sum\limits_{k=1}^{n} X_k$, en convenant qu'on a $S_0 = 0$. Enfin, on admet le résultat suivant :

$$\forall n \in \mathbb{N}, \ \forall k \in \mathbb{N}^*, \quad \sum_{j=0}^{k} \left(a + \frac{b\,j}{k} \right) q_j \, \mathbb{P}(\{S_n = k - j\}) \ = \ \left(a + \frac{b}{n+1} \right) \mathbb{P}(\{S_{n+1} = k\})$$

Soit $k \in \mathbb{N}^*$.

a) Montrer :

$$\forall j \in [\![0, k]\!], \quad \mathbb{P}(\{S = k - j\}) \; = \; \sum_{n=0}^{+\infty} p_n \, \mathbb{P}(\{S_n = k - j\})$$

b) Montrer :

$$\sum_{j=0}^{k} \left(a + \frac{b\,j}{k}\right) q_j \, \mathbb{P}(\{S = k - j\}) \; = \; \sum_{n=0}^{+\infty} p_{n+1} \, \mathbb{P}(\{S_{n+1} = k\})$$

c) Justifier :

$$\mathbb{P}(\{S = k\}) \; = \; \sum_{n=0}^{+\infty} p_{n+1} \, \mathbb{P}(\{S_{n+1} = k\})$$

d) En déduire finalement :

$$\mathbb{P}(\{S = k\}) \; = \; \frac{1}{1 - a\,q_0} \sum_{j=1}^{k} \left(a + \frac{b\,j}{k}\right) q_j \, \mathbb{P}(\{S = k - j\})$$

Exercice : `HEC 2000`

Ce problème se compose de deux parties. Si le candidat ne parvient pas à établir un résultat demandé, il l'indiquera clairement, et il pourra pour la suite admettre ce résultat.
Dans tout le problème, n désigne un entier naturel non nul.
On considère une urne U_n contenant n boules numérotées de 1 à n.
On tire une boule au hasard dans U_n. On note k le numéro de cette boule.

- Si k est égal à 1, on arrête les tirages.
- Si k est supérieur ou égal à 2, on enlève de l'urne U_n les boules numérotées de k à n (il reste donc les boules numérotées de 1 à $k-1$), et on effectue à nouveau un tirage dans l'urne.

On répète ces tirages jusqu'à l'obtention de la boule numéro 1.
On note X_n la variable aléatoire égale au nombre de tirages nécessaires pour l'obtention de la boule numéro 1. On note Y_n la variable aléatoire égale à la somme des numéros des boules tirées.
On note $\mathbb{E}(X_n)$ et $\mathbb{V}(X_n)$ (respectivement $\mathbb{E}(Y_n)$ et $\mathbb{V}(Y_n)$) l'espérance et la variance de X_n (respectivement Y_n).

Dans tout le problème, $\mathbb{P}_B(A)$ désigne la probabilité conditionnelle de A sachant B, où B est un événement non négligeable.

Partie I

1. On pose : $h_n = \sum\limits_{k=1}^{n} \frac{1}{k} = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$.

a) Montrer, pour tout entier naturel k non nul, les inégalités :

$$\frac{1}{k+1} \; \leqslant \; \ln(k+1) - \ln(k) \; \leqslant \; \frac{1}{k}$$

où ln désigne le logarithme népérien.

b) En déduire les inégalités : $\ln(n+1) \leqslant h_n \leqslant 1 + \ln(n)$.

c) Déterminer un équivalent simple de h_n quand n tend vers l'infini.

2. On pose : $k_n = \sum\limits_{k=1}^{n} \frac{1}{k^2} = 1 + \frac{1}{2^2} + \cdots + \frac{1}{n^2}$.

a) Montrer, pour tout entier k supérieur ou égal à 2, l'inégalité :

$$\frac{1}{k^2} \; \leqslant \; \frac{1}{k-1} - \frac{1}{k}$$

b) En déduire la majoration : $k_n \leqslant 2$.

c) Déterminer un équivalent simple de $h_n - k_n$ quand n tend vers l'infini.

Partie II. Étude de la variable aléatoire X_n

On note I_n la variable aléatoire égale au numéro de la première boule tirée dans l'urne U_n.

3. a) Quelle est la loi de I_n ?

b) Quelle est la loi conditionnelle de X_n sachant $\{I_n = 1\}$?

c) Si n est supérieur ou égal à 2, montrer :

$$\forall j \in \mathbb{N}^*, \forall k \in \{2, \ldots, n\}\,,\ \mathbb{P}_{\{I_n=k\}}(\{X_n = j\}) = \mathbb{P}(\{X_{k-1} = j - 1\})$$

4. a) Quelle est la loi de X_1 ?

b) Quel est l'événement $\{X_2 = 1\}$? Donner la loi de X_2, son espérance et sa variance.

c) Calculer $\mathbb{P}_{\{I_3=1\}}(\{X_3 = 2\})$, $\mathbb{P}_{\{I_3=2\}}(\{X_3 = 2\})$, $\mathbb{P}_{\{I_3=3\}}(\{X_3 = 2\})$.
Déterminer la loi de X_3, son espérance et sa variance.

d) Déterminer la loi du couple (I_3, X_3).
En déduire la covariance du couple (I_3, X_3).
Les variables aléatoires I_3 et X_3 sont-elles indépendantes ?

5. a) Montrer que X_n prend ses valeurs dans $\{1, 2, \ldots, n\}$.

b) Déterminer $\mathbb{P}(\{X_n = 1\})$ et $\mathbb{P}(\{X_n = n\})$.

c) Si n est supérieur ou égal à 2, montrer la relation :

$$\forall j \geqslant 2,\ \mathbb{P}(\{X_n = j\}) = \frac{1}{n} \sum_{k=1}^{n-1} \mathbb{P}(\{X_k = j - 1\})$$

d) Si n est supérieur ou égal à 3 et j supérieur ou égal à 2, calculer :

$$n\,\mathbb{P}(\{X_n = j\}) - (n-1)\,\mathbb{P}(\{X_{n-1} = j\})$$

En déduire, si n est un entier supérieur ou égal à 2 :

$$\forall j \geqslant 1,\ \mathbb{P}(\{X_n = j\}) = \frac{n-1}{n}\ \mathbb{P}(\{X_{n-1} = j\}) + \frac{1}{n}\ \mathbb{P}(\{X_{n-1} = j - 1\})$$

6. a) Si n est supérieur ou égal à 2, montrer, en utilisant ***5.d)*** :

$$\mathbb{E}(X_n) = \mathbb{E}(X_{n-1}) + \frac{1}{n}$$

b) En déduire $\mathbb{E}(X_n)$ et donner un équivalent simple de $\mathbb{E}(X_n)$ quand n tend vers l'infini.

7. a) Si n est supérieur ou égal à 2, calculer $\mathbb{E}(X_n^2)$ en fonction de $\mathbb{E}(X_{n-1}^2)$ et de $\mathbb{E}(X_{n-1})$.

b) En déduire : $\mathbb{V}(X_n) = h_n - k_n$ (en reprenant les notations de la **Partie I**).

c) Donner un équivalent de $\mathbb{V}(X_n)$ quand n tend vers l'infini.

8. Soit $(T_i)_{i \geqslant 1}$ une suite de variables aléatoires indépendantes telle que, pour tout i entier naturel non nul, T_i suit la loi de Bernoulli de paramètre $\frac{1}{i}$. On pose :

$$S_n \ = \ \sum_{i=1}^{n} T_i \ = \ T_1 + \cdots + T_n$$

a) Vérifier que X_1 et T_1 ont même loi.

b) Si n est supérieur ou égal à 2, montrer, pour tout entier naturel j non nul :

$$\mathbb{P}(\{S_n = j\}) \ = \ \frac{1}{n}\ \mathbb{P}(\{S_{n-1} = j-1\}) + \frac{n-1}{n}\ \mathbb{P}(\{S_{n-1} = j\})$$

En déduire que X_n et S_n ont même loi.

c) Retrouver ainsi $\mathbb{E}(X_n)$ et $\mathbb{V}(X_n)$.

Exercice : `ESCP 2004`

Dans tout le problème, r désigne un entier naturel vérifiant $1 \leqslant r \leqslant 10$. Une urne contient 10 boules distinctes B_1, B_2, ..., B_{10}. Une expérience aléatoire consiste à y effectuer une suite de tirages d'une boule **avec remise**, chaque boule ayant la même probabilité de sortir à chaque tirage.
Cette expérience est modélisée par un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.

Partie I : Étude du nombre de tirages nécessaires pour obtenir au moins une fois chacune des boules B_1, ... , B_r

On suppose que le nombre de tirages nécessaires pour obtenir au moins une fois chacune des boules B_1, ... , B_r définit une variable aléatoire Y_r sur $(\Omega, \mathscr{A}, \mathbb{P})$.

1. Cas particulier $r = 1$.
Montrer que la variable aléatoire Y_1 suit une loi géométrique ; préciser son paramètre, son espérance et sa variance.

2. On suppose que r est supérieur ou égal à 2.

a) Calculer la probabilité pour que les r boules B_1, B_2, ..., B_r sortent dans cet ordre aux r premiers tirages.

b) En déduire la probabilité $\mathbb{P}(\{Y_r = r\})$.

c) Préciser l'ensemble des valeurs que peut prendre la variable aléatoire Y_r.

3. On suppose encore que r est supérieur ou égal à 2. Pour tout entier i vérifiant $1 \leqslant i \leqslant r$, on désigne par W_i la variable aléatoire représentant le nombre de tirages nécessaires pour que, pour la première fois, i boules distinctes parmi les boules B_1, B_2, ... , B_r soient sorties (en particulier, on a : $W_r = Y_r$). On pose :

$$\left\{ \begin{array}{l} X_1 = W_1 \\ \forall i \in [\![2, r]\!],\ X_i = W_i - W_{i-1} \end{array} \right.$$

On admet que les variables aléatoires X_1, ..., X_r sont indépendantes.

a) Exprimer la variable aléatoire Y_r à l'aide des variables aléatoires X_1, ..., X_r.

b) Interpréter concrètement la variable aléatoire X_i pour tout i vérifiant $1 \leqslant i \leqslant r$.

c) Montrer que, pour tout i vérifiant $1 \leqslant i \leqslant r$, la variable aléatoire X_i suit une loi géométrique ; préciser son espérance et sa variance.

d) On pose : $S_1(r) = \sum\limits_{k=1}^{r} \dfrac{1}{k}$ et $S_2(r) = \sum\limits_{k=1}^{r} \dfrac{1}{k^2}$.
Exprimer l'espérance $\mathbb{E}(Y_r)$ et la variance $\mathbb{V}(Y_r)$ de Y_r à l'aide de $S_1(r)$ et de $S_2(r)$.

4. ***a)*** Si k est un entier naturel non nul, préciser le minimum et le maximum de la fonction $t \mapsto \dfrac{1}{t}$ sur l'intervalle $[k, k+1]$ et en déduire un encadrement de l'intégrale $\displaystyle\int_k^{k+1} \frac{1}{t}\ dt$.

b) Si r est supérieur ou égal à 2, donner un encadrement de $S_1(r)$ et en déduire la double inégalité :

$$10 \ln(r+1) \ \leqslant \ \mathbb{E}(Y_r) \ \leqslant \ 10\left(\ln(r)+1\right)$$

c) Si r supérieur ou égal à 2, établir par une méthode analogue à celle de la question précédente, la double inégalité :

$$1 - \frac{1}{r+1} \ \leqslant \ S_2(r) \ \leqslant \ 2 - \frac{1}{r}$$

En déduire un encadrement de $\mathbb{V}(Y_r)$.

Partie II : Étude du nombre de boules distinctes parmi les boules B_1, B_2, ..., B_r tirées au moins une fois au cours des n premiers tirages

Pour tout entier n supérieur ou égal à 1, on suppose que le nombre de boules distinctes parmi les boules B_1, B_2, ..., B_r tirées au moins une fois au cours des n premiers tirages, définit une variable aléatoire Z_n sur $(\Omega, \mathscr{A}, \mathbb{P})$. On note $\mathbb{E}(Z_n)$ l'espérance de Z_n et on pose $Z_0 = 0$.
Pour tout entier naturel n non nul et pour tout entier naturel k, on note $p_{n,k}$ la probabilité de l'événement $\{Z_n = k\}$ et on pose : $p_{n,-1} = 0$.

5. Étude des cas particuliers $n = 1$ et $n = 2$.

a) Déterminer la loi de Z_1 et donner son espérance.

b) On suppose, dans cette question, que r est supérieur ou égal à 2.
Déterminer la loi de Z_2 et montrer que son espérance est donnée par : $\mathbb{E}(Z_2) = \dfrac{19\,r}{100}$

6. Établir, pour tout entier naturel n non nul et pour tout entier naturel k au plus égal à r, l'égalité :

$$10\,p_{n,k} \ = \ (10-r+k)\,p_{n-1,k} + (r+1-k)\,p_{n-1,k-1} \qquad (*)$$

Vérifier que cette égalité reste vraie dans le cas où k est supérieur ou égal à $r+1$.

7. Pour tout entier naturel non nul n, on définit le polynôme Q_n par :

$$\left\{ \begin{array}{lcl} Q_0(X) & = & 1 \\ Q_n(X) & = & \sum\limits_{k=0}^{n} p_{n,k}\, x^k \end{array} \right.$$

a) Préciser les polynômes Q_1 et Q_2.

b) Calculer $Q_n(1)$ et exprimer $Q_n'(1)$ en fonction de $\mathbb{E}(Z_n)$, où Q_n' désigne la dérivée du polynôme Q_n.

c) En utilisant l'égalité $(*)$, établir, pour tout réel x et pour tout entier naturel n non nul, la relation suivante :

$$10\,Q_n(x) \ = \ (10-r+rx)\,Q_{n-1}(x) + x(1-x)\,Q_{n-1}'(x) \qquad (**)$$

d) En dérivant membre à membre l'égalité $(**)$, former, pour tout entier naturel n non nul, une relation entre les espérances $\mathbb{E}(Z_n)$ et $\mathbb{E}(Z_{n-1})$.
En déduire, pour tout entier naturel n, la valeur de $\mathbb{E}(Z_n)$ en fonction de n et de r.

8. ***a)*** Pour tout entier naturel n, le polynôme Q_n'' désigne la dérivée du polynôme Q_n'.
En utilisant une méthode semblable à celle de la question précédente, trouver pour tout entier naturel n non nul, une relation entre $Q_n''(1)$ et $Q_{n-1}''(1)$.
En déduire, pour tout entier naturel n non nul, l'égalité suivante :

$$Q_n''(1) \ = \ r(r-1)\left(1 + \left(\frac{8}{10}\right)^n - 2\left(\frac{9}{10}\right)^n\right)$$

b) Calculer, pour tout entier naturel n, la variance de la variable aléatoire Z_n en fonction de n et de r.

II. Couples de v.a.r. discrètes

Exercice : EML 97

On dispose d'un dé équilibré à 6 faces et d'une pièce truquée telle que la probabilité d'apparition de « pile » soit égale à p, $p \in]0,1[$.
On pourra noter $q = 1 - p$.
Soit N un entier naturel non nul fixé.
On effectue N lancers du dé ; si n est le nombre de « 6 » obtenus, on lance alors n fois la pièce.

On définit trois variables aléatoires X, Y, Z de la manière suivante :
- × Z indique le nombre de « 6 » obtenus aux lancers du dé,
- × X indique le nombre de « piles » obtenus aux lancers de la pièce,
- × Y indique le nombre de « faces » obtenues aux lancers de la pièce.

Ainsi, $X + Y = Z$ et, si Z prend la valeur 0, alors X et Y prennent la valeur 0.

1. Préciser la loi de Z, son espérance et sa variance.

2. Pour $k \in \mathbb{N}$, $n \in \mathbb{N}$, déterminer la probabilité conditionnelle $\mathbb{P}_{\{Z=n\}}(\{X = k\})$.
On distinguera les cas : $k \leqslant n$ et $k > n$.

3. Montrer, pour tout couple d'entiers naturels (k, n) :
- × si $0 \leqslant k \leqslant n \leqslant N$ alors $\mathbb{P}(\{X = k\} \cap \{Z = n\}) = \binom{n}{k}\binom{N}{n} p^k (1-p)^{n-k} \left(\frac{5}{6}\right)^{N-n} \left(\frac{1}{6}\right)^n$.
- × si $n > N$ ou $k > n$ alors $\mathbb{P}(\{X = k\} \cap \{Z = n\}) = 0$.

4. Calculer la probabilité $\mathbb{P}(\{X = 0\})$.

5. Montrer pour tout couple d'entiers naturels (k, n) tel que $0 \leqslant k \leqslant n \leqslant N$:
$$\binom{n}{k}\binom{N}{n} = \binom{N}{k}\binom{N-k}{n-k}$$
En déduire la probabilité $\mathbb{P}(\{X = k\})$.

6. Montrer que la variable aléatoire X suit une loi binomiale de paramètre $(N, \frac{p}{6})$.
Quelle est la loi de la variable aléatoire Y ?

7. Est-ce que les variables aléatoires X et Y sont indépendantes ?
Déterminer la loi du couple (X, Y).

8. En comparant les variances de Z et de $X + Y$, déterminer la covariance du couple (X, Y).

Exercice : ESCP 2001

Préliminaire

1. Montrer, pour tout entier naturel non nul n, l'égalité : $\sum\limits_{k=1}^{n} k^3 = \dfrac{n^2(n+1)^2}{4}$.

Partie I

Soit N un entier supérieur ou égal à 2.
Une urne contient N boules dont $N-2$ sont blanches et 2 sont noires.
On tire au hasard, successivement et *sans remise*, les N boules de cette urne.
Les tirages étant numérotés de 1 à N, on note X_1 la variable aléatoire égale au numéro du tirage qui a fourni, pour la première fois, une boule noire et X_2 la variable aléatoire égale au numéro du tirage qui a fourni, pour la deuxième fois, une boule noire.

2. Préciser l'espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ que l'on peut utiliser pour modéliser cette expérience aléatoire.

3. Soit i et j deux entiers de l'intervalle $[\![1, N]\!]$. Montrer :

$$\mathbb{P}(\{X_1 = i\} \cap \{X_2 = j\}) = \begin{cases} 0 & \text{si } 1 \leqslant j \leqslant i \leqslant N \\ \dfrac{2}{N(N-1)} & \text{si } 1 \leqslant i < j \leqslant N \end{cases}$$

4. Déterminer les lois de probabilité des variables X_1 et X_2. Ces variables sont-elles indépendantes ?

5. ***a)*** Démontrer que la variable $N+1-X_2$ a même loi que X_1.

b) Déterminer la loi de la variable $X_2 - X_1$ et la comparer à celle de X_1.

6. À l'aide des résultats de la question ***5*** :

a) Calculer les espérances $\mathbb{E}(X_1)$ et $\mathbb{E}(X_2)$.

b) Montrer l'égalité des variances $\mathbb{V}(X_1)$ et $\mathbb{V}(X_2)$.

c) Établir la relation : $2\,\mathrm{Cov}(X_1, X_2) = \mathbb{V}(X_1)$
(où $\mathrm{Cov}(X_1, X_2)$ désigne la covariance des variables X_1 et X_2).

7. Calculer $\mathbb{V}(X_1)$. En déduire $\mathbb{V}(X_2)$ et $\mathrm{Cov}(X_1, X_2)$.

Partie II

On suppose que A et B sont deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$, indépendantes, suivant la même loi uniforme sur l'ensemble $\{1, 2, \ldots, N\}$ et on désigne par D l'événement : « A ne prend pas la même valeur que B ».

8. Montrer que la probabilité de l'événement D est $\dfrac{N-1}{N}$.

9. Soit Y_1 et Y_2 les variables aléatoires définies par : $\begin{cases} Y_1 = \min(A, B) \\ Y_2 = \max(A, B) \end{cases}$

Calculer, pour tout couple (i, j) d'éléments de $\{1, 2, \ldots, N\}$, la probabilité conditionnelle :

$$\mathbb{P}_D(\{Y_1 = i\} \cap \{Y_2 = j\})$$

Exercice : `ESSEC 2001`

Le but du problème est l'étude du coefficient de corrélation linéaire de deux variables aléatoires qu'on aborde d'abord dans un cas particulier (**Partie I**), puis de façon générale (**Partie II**).

Partie I

1. Calculs préliminaires

a) On considère deux nombres entiers naturels q et n tels que $n \geqslant q$. En raisonnant par récurrence sur n, établir la formule suivante :

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

b) En faisant $q = 1,\ 2\ ,\ 3$, en déduire une expression factorisée des quatre sommes suivantes :

$$\sum_{k=1}^{n} k \quad ; \quad \sum_{k=2}^{n} k(k-1) \quad ; \quad \sum_{k=1}^{n} k^2 \quad \text{et} \quad \sum_{k=3}^{n} k(k-1)(k-2)$$

On considère dans toute la suite de cette partie un nombre entier $n \geqslant 2$ et une urne contenant n jetons numérotés de 1 à n.
On extrait de cette urne successivement et sans remise 2 jetons et on désigne alors par :

- N_1 la variable aléatoire indiquant le numéro du premier jeton tiré,
- N_2 la variable aléatoire indiquant le numéro du second jeton tiré,
- X la variable aléatoire indiquant le plus petit des numéros des 2 jetons tirés,
- Y la variable aléatoire indiquant le plus grand des numéros des 2 jetons tirés.

On note $\mathbb{E}(N_1)$ et $\mathbb{V}(N_1)$, $\mathbb{E}(N_2)$ et $\mathbb{V}(N_2)$, $\mathbb{E}(X)$ et $\mathbb{V}(X)$, $\mathbb{E}(Y)$ et $\mathbb{V}(Y)$ les espérances et variances des quatre variables aléatoires N_1, N_2, X, Y.

2. Lois conjointe et marginales des variables aléatoires N_1 et N_2.

a) Déterminer les probabilités $\mathbb{P}(\{N_1 = i\})$ pour tout $1 \leqslant i \leqslant n$, et $\mathbb{P}_{\{N_1=i\}}(\{N_2 = j\})$ pour tout $1 \leqslant j \leqslant n$, $j \neq i$.
En déduire : $\forall j \in [\![1, n]\!]$, $\mathbb{P}(\{N_2 = j\}) = \dfrac{1}{n}$. Puis comparer les lois de N_1 et N_2.

b) Calculer les espérances $\mathbb{E}(N_1)$ et $\mathbb{E}(N_2)$, les variances $\mathbb{V}(N_1)$ et $\mathbb{V}(N_2)$.

c) Montrer, pour tout $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant n$:

$$\mathbb{P}(\{N_1 = i\} \cap \{N_2 = j\}) = \begin{cases} \dfrac{1}{n(n-1)} & \text{si } i \neq j \\ 0 & \text{sinon} \end{cases}$$

et en déduire :

$$\mathbb{E}(N_1\, N_2) = \frac{(n+1)(3n+2)}{12}$$

En déduire la covariance et le coefficient de corrélation linéaire de N_1 et N_2.

d) Exprimer enfin sous forme factorisée la variance $\mathbb{V}(N_1 + N_2)$.

3. Lois conjointe, marginales et conditionnelles des variables aléatoires X et Y

a) Montrer, pour tout $(i, j) \in [\![1, n]\!]^2$ tel que $1 \leqslant i < j \leqslant n$: $\mathbb{P}(\{X = i\} \cap \{Y = j\}) = \dfrac{2}{n(n-1)}$.
Que valent ces probabilités sinon ?

b) En déduire les probabilités $\mathbb{P}(\{Y = j\})$ pour $2 \leqslant j \leqslant n$ et $\mathbb{P}(\{X = i\})$ pour $1 \leqslant i \leqslant n-1$.
(On vérifiera que les formules donnant $\mathbb{P}(\{Y = j\})$ et $\mathbb{P}(\{X = i\})$ restent valables si $j = 1$ ou $i = n$).

c) Déterminer les probabilités $\mathbb{P}_{\{Y=j\}}(\{X = i\})$ et $\mathbb{P}_{\{X=i\}}(\{Y = j\})$ pour $1 \leqslant i < j \leqslant n$, puis reconnaître la loi de X conditionnellement à $\{Y = j\}$ et la loi de Y conditionnellement à $\{X = i\}$.

d) Comparer les lois des variables aléatoires $n+1-X$ et Y.
En déduire que $\mathbb{E}(n+1-X) = \mathbb{E}(Y)$ et $\mathbb{V}(n+1-X) = \mathbb{V}(Y)$, puis en déduire les expressions de $\mathbb{E}(X)$ en fonction de $\mathbb{E}(Y)$ et de $\mathbb{V}(X)$ en fonction de $\mathbb{V}(Y)$.

4. **Espérances et variances des variables aléatoires X et Y**

a) Exprimer les espérances $\mathbb{E}(Y)$ et $\mathbb{E}(X)$ en fonction de n.

b) Exprimer sous forme factorisée $\mathbb{E}\big(Y(Y-2)\big)$, puis $\mathbb{E}\left(Y^2\right)$, $\mathbb{V}(Y)$ et $\mathbb{V}(X)$ en fonction de n.

5. **Covariance et coefficient de corrélation linéaire des variables aléatoires X et Y**

a) Vérifier : $X+Y = N_1+N_2$.
En déduire sous forme factorisée la variance de $X+Y$ et la covariance de X et Y.

b) En déduire le coefficient de corrélation de X et Y.
On remarquera que le coefficient de corrélation linéaire de X et Y est indépendant de n.

Partie II

On considère deux variables aléatoires X et Y définies sur un même espace probabilisé et admettant des espérances $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ et des variances $\mathbb{V}(X)$ et $\mathbb{V}(Y)$ et on suppose $\mathbb{V}(X) > 0$ (on rappelle que $\mathbb{V}(X) = 0$ si et seulement si, avec une probabilité égale à 1, X est constante). La covariance des deux variables aléatoires X et Y (que celles-ci soient discrètes ou à densité) est alors le nombre réel défini par :

$$\mathrm{Cov}(X,Y) \;=\; \mathbb{E}\big((X-\mathbb{E}(X))\,(Y-\mathbb{E}(Y))\big) \quad \text{ou encore} \quad \mathrm{Cov}(X,Y) \;=\; \mathbb{E}(X\,Y) - \mathbb{E}(X)\,\mathbb{E}(Y)$$

6. **Covariance des variables aléatoires X et Y**

a) Exprimer $\mathrm{Cov}(\lambda X+Y, \lambda X+Y)$ en fonction de $\mathbb{V}(\lambda X+Y)$ et en déduire la formule suivante pour tout nombre réel λ :

$$\mathbb{V}(\lambda X+Y) \;=\; \lambda^2\,\mathbb{V}(X) + 2\lambda\,\mathrm{Cov}(X,Y) + \mathbb{V}(Y)$$

b) En déduire : $\big(\mathrm{Cov}(X,Y)\big)^2 \;\leqslant\; \mathbb{V}(X)\,\mathbb{V}(Y)$.
A quelle condition nécessaire et suffisante a-t-on l'égalité : $\big(\mathrm{Cov}(X,Y)\big)^2 \;=\; \mathbb{V}(X)\,\mathbb{V}(Y)$?

7. **Coefficient de corrélation linéaire des variables aléatoires X et Y**
On suppose dans cette question les variances $\mathbb{V}(X)$ et $\mathbb{V}(Y)$ de X et Y strictement positives.

a) Exprimer le coefficient de corrélation linéaire ρ des variables aléatoires X et Y en fonction de $\mathrm{Cov}(X,Y)$ et des écarts-types $\sigma(X)$ et $\sigma(Y)$ des variables aléatoires X et Y et montrer que ρ appartient à $[-1,+1]$.
Préciser de plus à quelle condition nécessaire et suffisante ρ est égal à -1 ou $+1$.

b) Donner la valeur de ρ lorsque les variables aléatoires X et Y sont indépendantes.

Exercice : `EDHEC 2015`

Partie I

Dans cette partie, x désigne un réel de $[0,1[$.

1. *a)* Montrer : $\forall m \in \mathbb{N},\ 0 \;\leqslant\; \displaystyle\int_0^x \frac{t^m}{1-t^2}\,dt \;\leqslant\; \frac{1}{1-x^2} \times \frac{1}{m+1}$.

b) En déduire que : $\displaystyle\lim_{m\to+\infty} \int_0^x \frac{t^m}{1-t^2} = 0$.

2. a) Pour tout réel t de $[0,1[$ et pour tout k élément de $\mathbb{N}^*$, calculer $\sum_{j=0}^{k-1} t^{2j}$.

b) En déduire : $\sum_{j=0}^{k-1} \dfrac{x^{2j+1}}{2j+1} = \displaystyle\int_0^x \frac{1}{1-t^2}\,dt - \int_0^x \frac{t^{2k}}{1-t^2}\,dt$.

c) Utiliser la question ***1*** pour montrer que la série de terme général $\dfrac{x^{2j+1}}{2j+1}$ converge et exprimer $\sum_{j=0}^{+\infty} \dfrac{x^{2j+1}}{2j+1}$ sous forme d'une intégrale que l'on ne cherchera pas à calculer.

d) Conclure : $\forall k \in \mathbb{N}$, $\sum_{j=k}^{+\infty} \dfrac{x^{2j+1}}{2j+1} = \displaystyle\int_0^x \frac{t^{2k}}{1-t^2}\,dt$.

On admet sans démonstration que l'on a aussi : $\forall k \in \mathbb{N}$, $\sum_{j=k+1}^{+\infty} \dfrac{x^{2j}}{2j} = \displaystyle\int_0^x \frac{t^{2k+1}}{1-t^2}\,dt$.

Partie II

Un joueur réalise une suite de lancers indépendants d'une pièce.
Cette pièce donne pile avec la probabilité p $(0 < p < 1)$ et face avec la probabilité $q = 1 - p$.
On note N la variable aléatoire égale au rang d'apparition du premier pile.
Si N prend la valeur n, le joueur place n boules numérotées de 1 à n dans une urne, puis il extrait une boule au hasard de cette urne. On dit que le joueur a gagné si le numéro porté par la boule tirée est impair et on désigne par A l'événement : « le joueur gagne ».
On appelle X la variable aléatoire égale au numéro porté par la boule extraite de l'urne.

3. Reconnaître la loi de N.

4. a) Donner, pour tout entier naturel k supérieur ou égal à j, la valeur de $\mathbb{P}_{\{N=2j\}}\big(\{X = 2k+1\}\big)$.

b) Donner, pour tout entier naturel k supérieur ou égal à $j+1$, la valeur de $\mathbb{P}_{\{N=2j+1\}}\big(\{X = 2k+1\}\big)$.

c) Déterminer $\mathbb{P}_{\{N=2j\}}\big(\{X = 2k+1\}\big)$ lorsque k appartient à $[\![0, j-1]\!]$.

d) Déterminer $\mathbb{P}_{\{N=2j+1\}}\big(\{X = 2k+1\}\big)$ lorsque k appartient à $[\![0, j]\!]$.

5. a) Justifier : $\mathbb{P}\big(\{X = 2k+1\}\big) = \sum_{n=1}^{+\infty} \mathbb{P}\big(\{N = n\}\big)\,\mathbb{P}_{\{N=n\}}\big(\{X = 2k+1\}\big)$.

En admettant que l'on peut scinder la somme précédente selon la parité de n, montrer :

$$\forall k \in \mathbb{N},\ \mathbb{P}\big(\{X = 2k+1\}\big) = \frac{p}{q}\left(\sum_{j=k+1}^{+\infty} \frac{q^{2j}}{2j} + \sum_{j=k}^{+\infty} \frac{q^{2j+1}}{2j+1}\right)$$

b) En déduire : $\forall k \in \mathbb{N}$, $\mathbb{P}\big(\{X = 2k+1\}\big) = \dfrac{p}{q}\displaystyle\int_0^q \frac{t^{2k}}{1-t}\,dt$.

6. a) Montrer : $\lim\limits_{n\to+\infty} \displaystyle\int_0^q \frac{t^{2n+2}}{(1-t)^2(1+t)}\,dt = 0$.

b) Montrer :

$$\sum_{k=0}^{n} \mathbb{P}\big(\{X = 2k+1\}\big) = \frac{p}{q}\left(\int_0^q \frac{1}{(1-t)^2(1+t)}\,dt - \int_0^q \frac{t^{2n+2}}{(1-t)^2(1+t)}\,dt\right)$$

c) En déduire : $\mathbb{P}(A) = \dfrac{p}{q}\displaystyle\int_0^q \frac{1}{(1-t)^2(1+t)}\,dt$.

7. a) Trouver trois constantes réelles a, b et c telles que, pour tout t différent de 1 et de -1, on ait :

$$\frac{1}{(1-t)^2(1+t)} = \frac{a}{1-t} + \frac{b}{1+t} + \frac{c}{(1-t)^2}$$

b) Écrire $\mathbb{P}(A)$ explicitement en fonction de q.

c) En déduire : $\mathbb{P}(A) > \dfrac{1}{2}$.

Exercice : `ESSEC I 2011`

Problème 1 : Évolution des intentions de vote

Dans une élection à venir, deux candidats A et B se présentent.
Un groupe d'électeurs est composé de m individus, avec $m \geqslant 2$.

- Initialement, au jour appelé « jour 0 », le nombre d'individus préférant le candidat A vaut a (il y en a donc $m-a$ préférant le candidat B).
- Ensuite, chaque jour, un des individus au hasard dans le groupe en rencontre un autre, au hasard également, et il lui parle des élections. Si leurs intentions de vote diffèrent, il le convainc de voter comme lui.

Pour tout entier naturel n, on note X_n le nombre d'individus du groupe ayant l'intention de voter pour le candidat A le soir du $n^{\text{ème}}$ jour. Ainsi, X_n est une variable aléatoire à valeurs dans $[\![0, m]\!]$.
On remarque que X_0 est une variable aléatoire certaine : $\mathbb{P}\big(\{X_0 = a\}\big) = 1$.

Partie I - Un cas particulier : $m = 4$

Dans cette partie, on étudie le cas d'un groupe formé de quatre électeurs.

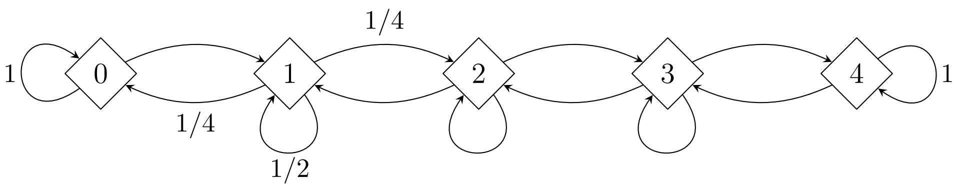
1. Soit i et j deux entiers dans $[\![0, 4]\!]$. On note $p_{i,j}$ la probabilité pour qu'il y ait exactement j personnes dans le groupe ayant l'intention de voter pour A un jour donné, sachant qu'il y en avait i la veille.

a) Justifier : $p_{0,0} = p_{4,4} = 1$.

b) Justifier : si i et j dans $[\![0, 4]\!]$ sont tels que $|i-j| \geqslant 2$, alors $p_{i,j} = 0$.

c) Établir : $p_{1,0} = p_{1,2} = \frac{1}{4}$ et $p_{1,1} = \frac{1}{2}$.

d) De la même façon, donner pour tout $(i,j) \in [\![0, 4]\!]^2$ la probabilité $p_{i,j}$.
On présentera les résultats sur le diagramme suivant, à reproduire et à compléter, et on justifiera quelques cas.



2. On définit la matrice $M = \begin{pmatrix} 1/2 & 1/3 & 0 \\ 1/4 & 1/3 & 1/4 \\ 0 & 1/3 & 1/2 \end{pmatrix}$,

et pour tout entier naturel n, la matrice colonne $U_n = \begin{pmatrix} \mathbb{P}\big(\{X_n = 1\}\big) \\ \mathbb{P}\big(\{X_n = 2\}\big) \\ \mathbb{P}\big(\{X_n = 3\}\big) \end{pmatrix}$.

a) Pour tout entier naturel n, établir la relation : $U_{n+1} = M\, U_n$.
En déduire pour tout entier naturel n, l'égalité : $U_n = M^n\, U_0$.

b) Montrer que M admet trois valeurs propres distinctes α, β et γ, vérifiant $0 \leqslant \alpha < \beta < \gamma < 1$.

Justifier qu'il existe une matrice carrée P d'ordre 3 inversible, que l'on ne demande pas de préciser, et D une matrice diagonale d'ordre 3, à préciser, telles que $P^{-1}MP = D$.

c) En déduire que pour tout $k \in \{1, 2, 3\}$, la suite $\Big(\mathbb{P}\big(\{X_n = k\}\big)\Big)_{n\in\mathbb{N}}$ est une combinaison linéaire des trois suites $\big(\alpha^n\big)_{n\in\mathbb{N}}$, $\big(\beta^n\big)_{n\in\mathbb{N}}$ et $\big(\gamma^n\big)_{n\in\mathbb{N}}$.

d) Montrer que pour tout $k \in \{1, 2, 3\}$, $\lim\limits_{n\to+\infty} \mathbb{P}\big(\{X_n = k\}\big) = 0$.

3. Établir : $\lim\limits_{n\to+\infty} \Big(\mathbb{P}\big(\{X_n = 0\}\big) + \mathbb{P}\big(\{X_n = 4\}\big)\Big) = 1$.
Comment interpréter ce résultat ?

Partie II - Le cas général

On revient dans cette partie au cas général d'un groupe de m électeurs.
On note $\pi_{n,k} = \mathbb{P}\big(\{X_n = k\}\big)$, la probabilité pour qu'il y ait exactement k électeurs envisageant de voter pour A à l'issue du $n^{\text{ème}}$ jour.

4. Soit n un entier naturel.

a) Établir les trois relations :

$$\forall k \in [\![0, m-1]\!], \quad \mathbb{P}_{\{X_n=k\}}\big(\{X_{n+1} = k+1\}\big) = \frac{k\,(m-k)}{m(m-1)}$$

$$\forall k \in [\![1, m]\!], \quad \mathbb{P}_{\{X_n=k\}}\big(\{X_{n+1} = k-1\}\big) = \frac{k\,(m-k)}{m(m-1)}$$

$$\forall k \in [\![1, m-1]\!], \quad \mathbb{P}_{\{X_n=k\}}\big(\{X_{n+1} = k\}\big) = 1 - \frac{2k\,(m-k)}{m(m-1)}$$

b) En déduire la relation, si $k \in [\![1, m-1]\!]$:

$$\pi_{n+1,k} = \frac{(k-1)(m+1-k)\ \pi_{n,k-1} + \big(m(m-1) - 2k(m-k)\big)\ \pi_{n,k} + (k+1)(m-1-k)\ \pi_{n,k+1}}{m(m-1)}$$

5. *a)* Montrer par récurrence que pour tout entier naturel n et pour tout $k \in [\![1, m-1]\!]$:

$$\pi_{n,k} \leqslant \left(\frac{m(m-1)-2}{m(m-1)}\right)^n$$

b) En déduire, pour tout $k \in [\![1, m-1]\!]$, la limite de $\pi_{n,k}$ lorsque n tend vers $+\infty$.

6. On définit l'événement V_A (respectivement V_B) suivant : « au bout d'un certain nombre de jours, tous les individus du groupe ont l'intention de voter pour A (respectivement pour B) ».

a) Montrer : $\mathbb{P}(V_A) = \lim\limits_{n\to+\infty} \mathbb{P}\big(\{X_n = m\}\big)$ et $\mathbb{P}(V_B) = \lim\limits_{n\to+\infty} \mathbb{P}\big(\{X_n = 0\}\big)$.

b) Montrer : $\mathbb{P}(V_A) + \mathbb{P}(V_B) = 1$.
Que signifie ce résultat ?

7. Pour tout entier naturel n, on pose $Z_n = X_{n+1} - X_n$.

a) Justifier : $Z_n(\Omega) = \{-1, 0, 1\}$.

b) Exprimer $\mathbb{P}\big(\{Z_n = 1\}\big)$ en fonction des probabilités $\pi_{n,k}$ avec $k \in [\![1, m-1]\!]$.

c) Comparer $\mathbb{P}\big(\{Z_n = -1\}\big)$ et $\mathbb{P}\big(\{Z_n = 1\}\big)$.

d) En déduire que $\mathbb{E}(Z_n) = 0$.

e) Montrer que la suite $\big(\mathbb{E}(X_n)\big)_{n\in\mathbb{N}}$ est constante et déterminer cette constante en fonction de a.

8. Montrer que $\mathbb{P}(V_A) = \dfrac{a}{m}$ et interpréter ce résultat.

Exercice (ESSEC II - 2019)

Un modèle probabiliste d'une expérience aléatoire représente dans un certain sens le désordre qui intervient dans l'expérience et il est donc naturel que des outils soient introduits qui permettent de mesurer l'intensité de ce désordre. C'est le cas de la notion d'entropie qui fait l'objet du présent problème. On considèrera différentes situations et notamment la façon dont on mesure l'information que deux variables aléatoires s'apportent mutuellement.

- Dans la première partie on étudie le cas plus simple techniquement de variables dont la loi admet une densité. Les deuxièmes et troisièmes parties sont consacrées au cas discret.
- Dans la deuxième partie, on introduit les différentes notions d'entropie pour le cas de variables discrètes et dans la troisième partie, on examine comment on peut mesurer l'information apportée mutuellement par deux variables aléatoires.

Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. Pour toute variable aléatoire Y, on notera $\mathbb{E}(Y)$ son espérance lorsqu'elle existe.

Deuxième partie : Généralités sur l'entropie des variables discrètes

Soit A un ensemble fini non vide. On dit que X est une variable aléatoire dont la loi est à support A, si X est à valeurs dans A et si pour tout $x \in A$: $\mathbb{P}(\{X = x\}) > 0$.

6. Soit X une variable aléatoire de loi à support $\{0, 1, 2, \dots, n\}$ où n est un entier naturel. On appelle **entropie** de X le réel :

$$H(X) \ = \ -\sum_{k=0}^{n} \mathbb{P}(\{X = k\}) \log_2 \big(\mathbb{P}(\{X = k\})\big)$$

a) On définit la fonction $g : \{0, \dots, n\} \to \mathbb{R}$ en posant $g(k) = \log_2 \big(\mathbb{P}(\{X = k\})\big)$ pour k élément de $\{0, 1, \dots, n\}$. Montrer : $H(X) = -\mathbb{E}\big(g(X)\big)$.

b) Montrer : $H(X) \geqslant 0$.

c) Soit p un réel tel que $0 < p < 1$.
On suppose dans cette question que X suit la loi de Bernoulli $\mathcal{B}(p)$.

(i) Calculer $H(X)$ en fonction de p. On note ψ la fonction qui, à p, associe $H(X)$.

(ii) Montrer que ψ est concave sur $]0, 1[$.

(iii) Déterminer la valeur p_0 où ψ est maximale.

d) On suppose dans cette question que la loi de X est à support $\{0, 1, 2, 3\}$ avec les probabilités :

$$\mathbb{P}(\{X = 0\}) = \frac{1}{2} \quad ; \quad \mathbb{P}(\{X = 1\}) = \frac{1}{4} \quad ; \quad \mathbb{P}(\{X = 2\}) = \mathbb{P}(\{X = 3\}) = \frac{1}{8}$$

Calculer $H(X)$.

7. On commence par une inégalité générale, appelée **Inégalité de Jensen**.

a) Soit $N \geqslant 2$. Soit X une variable aléatoire de loi à support $\{x_1, x_2, \dots, x_N\}$ où les x_i sont des éléments distincts de $\mathbb{R}_+$. On pose $\mathbb{P}(\{X = x_i\}) = p_i$.
Montrer que, pour tout $1 \leqslant i \leqslant N$, on a : $p_i < 1$.

On désire démontrer par récurrence la propriété suivante :

$\mathcal{P}(N)$: **Pour toute fonction φ convexe sur $\mathbb{R}_+$, si X est une variable aléatoire de loi à support $A \subset \mathbb{R}_+$ avec $\mathrm{Card}(A) = N$, on a :** $\mathbb{E}\big(\varphi(X)\big) \geqslant \varphi\big(\mathbb{E}(X)\big)$

b) Montrer que $\mathcal{P}(2)$ est vraie.

c) Soit $N \geqslant 3$. On suppose que $\mathcal{P}(N-1)$ est vérifiée. Soit X une variable aléatoire de loi à support $A = \{x_1, x_2, \dots, x_N\}$ où les x_i sont des éléments distincts de $\mathbb{R}_+$. On pose : $\mathbb{P}(\{X = x_i\}) = p_i$.
Pour i tel que $1 \leqslant i \leqslant N-1$, on pose : $p'_i = \dfrac{p_i}{1 - p_N}$.

(i) Montrer : $\sum\limits_{i=1}^{N-1} p'_i = 1$ et $\forall i \in [\![1, N-1]\!]$, $0 < p'_i < 1$.

(ii) Soit Y une variable aléatoire de loi à support $\{x_1, \dots, x_{N-1}\}$ telle que $\mathbb{P}(\{Y = x_i\}) = p'_i$ pour $1 \leqslant i \leqslant N-1$. Montrer : $\sum\limits_{i=1}^{N-1} p'_i\, \varphi(x_i) \ \geqslant\ \varphi\left(\sum\limits_{i=1}^{N-1} p'_i\, x_i\right)$.

(iii) Montrer : $\mathbb{E}\big(\varphi(X)\big) \ \geqslant\ \varphi\big(\mathbb{E}(X)\big)$.

d) Montrer que, si φ est *concave* sur $\mathbb{R}_+$, on a : $\mathbb{E}\big(\varphi(X)\big) \ \leqslant\ \varphi\big(\mathbb{E}(X)\big)$.

8. Soit X une variable aléatoire de loi à support $\{0, 1, \dots, n\}$.
On pose, pour k tel que $0 \leqslant k \leqslant n$, $p_k = \mathbb{P}(\{X = k\})$.

a) Montrer : $\sum\limits_{k=0}^{n} p_k \log_2\left(\dfrac{1}{(n+1)\, p_k}\right) \ \leqslant\ \log_2\left(\sum\limits_{k=0}^{n} \dfrac{p_k}{(n+1)\, p_k}\right) \ =\ 0$.

b) Montrer : $\sum\limits_{k=0}^{n} p_k \log_2\big((n+1)\, p_k\big) \ =\ \log_2(n+1) - H(X)$.

c) Montrer : $H(X) \leqslant \log_2(n+1)$.

d) On suppose que X suit la loi uniforme sur $\{0, 1, \dots, N\}$. Calculer $H(X)$.

9. Soient X et Y deux variables aléatoires *de même loi* à support $\{0, 1, \dots, n\}$.
On suppose en outre X et Y indépendantes.

a) Montrer : $\mathbb{P}(\{X = Y\}) \ =\ \sum\limits_{k=0}^{n} \big(\mathbb{P}(\{X = k\})\big)^2$.

b) On pose $v(k) = \mathbb{P}(\{X = k\})$ pour tout k élément de $\{0, 1, \dots, n\}$. Montrer :

$$2^{\mathbb{E}\left(\log_2\left(v(X)\right)\right)} \ \leqslant\ \mathbb{E}\left(2^{\log_2\left(v(X)\right)}\right) \ =\ \mathbb{E}\big(v(X)\big)$$

c) En déduire : $2^{-H(X)} \ \leqslant\ \mathbb{P}(\{X = Y\})$.

d) Donner un exemple de loi où l'inégalité précédente est une égalité.

Troisième partie : Entropie jointe et information mutuelle de deux variables discrètes

Soient X et Y deux variables aléatoires de lois à support $\{0, 1, \dots, n\}$.
On appelle **entropie jointe** de X et Y le réel :

$$H(X,Y) \;=\; -\sum_{k=0}^{n} \sum_{j=0}^{n} \mathbb{P}(\{X=k\} \cap \{Y=j\}) \log_2 \big(\mathbb{P}(\{X=k\} \cap \{Y=j\})\big)$$

avec la convention : $0 \times \log_2(0) = 0$.

11. ***a)*** On définit la fonction $g : \{0, 1, \dots, n\}^2 \to \mathbb{R} \cup \{-\infty\}$ en posant pour $(k, j) \in \{0, 1, \dots, n\}^2$:

$$g(k,j) \;=\; \log_2 \big(\mathbb{P}(\{X=k\} \cap \{Y=j\})\big)$$

Montrer : $H(X,Y) = -\mathbb{E}\big(g(X,Y)\big)$.

b) Montrer : $H(X,Y) = H(Y,X)$.

c) Pour tout k tel que $0 \leqslant k \leqslant n$, on pose :

$$H(Y \,|\, X=k) \;=\; -\sum_{j=0}^{n} \mathbb{P}_{\{X=k\}}(\{Y=j\}) \log_2 \big(\mathbb{P}_{\{X=k\}}(\{Y=j\})\big)$$

On appelle **entropie conditionnelle** de Y sachant X le réel :

$$H(Y \,|\, X) \;=\; \sum_{k=0}^{n} \mathbb{P}(\{X=k\})\, H(Y \,|\, X=k)$$

Montrer : $H(X,Y) \;=\; H(X) + H(Y \,|\, X)$.

d) Montrer que pour tout couple de variables aléatoires X et Y de lois à support $\{0, 1, \dots, n\}$, on a :

$$H(X) - H(X \,|\, Y) \;=\; H(Y) - H(Y \,|\, X)$$

12. On considère dans cette question deux variables aléatoires de lois à support $\{0, 1, 2, 3\}$. On suppose que la loi conjointe de (X, Y) est donnée par le tableau suivant :

$j \backslash k$	0	1	2	3
0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
3	$\frac{1}{4}$	0	0	0

(on lit dans la $k^{\text{ème}}$ colonne et la $j^{\text{ème}}$ ligne la valeur de $\mathbb{P}(\{X=k\} \cap \{Y=j\})$)

a) Déterminer la loi de X et montrer : $H(X) = \dfrac{7}{4}$.

b) Déterminer la loi de Y et calculer $H(Y)$.

c) Montrer : $H(X \mid Y) = \dfrac{11}{8}$.

d) Que vaut $H(Y \mid X)$?

e) Calculer $H(X,Y)$.

13. Soient X et Y deux variables aléatoires de lois à support $\{0,1,\ldots,n\}$. On appelle **information mutuelle** de X et de Y le réel :

$$I(X,Y) \;=\; \sum_{k=0}^{n} \sum_{j=0}^{n} \mathbb{P}(\{X=k\} \cap \{Y=j\}) \log_2\left(\frac{\mathbb{P}(\{X=k\} \cap \{Y=j\})}{\mathbb{P}(\{X=k\})\,\mathbb{P}(\{Y=j\})}\right)$$

a) Montrer : $I(X,Y) = I(Y,X)$.

b) Montrer : $I(X,Y) = H(X) - H(X \mid Y)$.

c) Montrer : $I(X,X) = H(X)$.

d) Que vaut $I(X,Y)$ si X et Y sont indépendantes ?

14. Soient X et Y deux variables aléatoires de lois à support $\{0,1,\ldots,n\}$. On fixe $0 \leqslant k \leqslant n$.

Pour $0 \leqslant j \leqslant n$, on pose : $p_j = \dfrac{\mathbb{P}(\{X=k\} \cap \{Y=j\})}{\mathbb{P}(\{X=k\})}$.

On suppose que $p_j > 0$ pour tout $0 \leqslant j \leqslant n$ et on pose : $x_j = \dfrac{\mathbb{P}(\{X=k\})\,\mathbb{P}(\{Y=j\})}{\mathbb{P}(\{X=k\} \cap \{Y=j\})}$.

a) Montrer : $\displaystyle\sum_{j=0}^{n} p_j = 1$.

b) Soit Z_k une variable aléatoire de loi à support $\{x_0,\ldots,x_n\}$ dont la loi est donnée par $\mathbb{P}(\{Z_k = x_j\}) = p_j$ pour $0 \leqslant j \leqslant n$. Montrer :

$$\mathbb{E}\big(\log_2(Z_k)\big) \;\leqslant\; 0$$

c) En déduire : $I(X,Y) \geqslant 0$.