DS3 (vB) - Barème

Problème 1 (`HEC 2010, partie II`)

Soit p un réel de $]0,1[$ et $q = 1-p$. Soit X_1 et X_2 deux variables indépendantes de même loi géométrique de paramètre p (d'espérance $\frac{1}{p}$).
On pose : $Y = X_1 - X_2$, $T = \max(X_1, X_2)$ et $Z = \min(X_1, X_2)$.
On rappelle que $T + Z = X_1 + X_2$ et $T - Z = |X_1 - X_2| = |Y|$.

1. a) Rappeler sans démonstration les valeurs respectives de $\mathbb{V}(X_1)$ et de $\mathbb{P}([X_1 \leqslant k])$, pour tout k de $X_1(\Omega)$.

- **1 pt :** $\mathbb{V}(X_1) = \dfrac{1-p}{p^2}$
- **1 pt :** $\mathbb{P}([X_1 \leqslant k]) = 1 - \mathbb{P}([X_1 > k]) = 1 - (1-p)^k$

b) Calculer $\mathbb{E}(X_1 + X_2)$, $\mathbb{V}(X_1 + X_2)$, $\mathbb{E}(X_1 - X_2)$, $\mathbb{V}(X_1 - X_2)$.

- **1 pt : existence de tous les objets**
- **1 pt :** $\mathbb{E}(X_1 + X_2) = \dfrac{2}{p}$
- **1 pt :** $\mathbb{E}(X_1 - X_2) = 0$
- **1 pt : hypothèse d'indépendance pour les calculs de variance**
- **1 pt :** $\mathbb{V}(X_1 + X_2) = 2\dfrac{q}{p^2}$
- **1 pt :** $\mathbb{V}(X_1 - X_2) = 2\dfrac{q}{p^2}$

c) Écrire une fonction **Python** qui prend en entrée le paramètre p et qui renvoie 1 si l'événement $[X_1 = X_2]$ est réalisé et 0 sinon. On simulera les variables aléatoires X_1 et X_2 dans le programme.

```
def simul(p):
    X1 = rd.geometric(p)
    X2 = rd.geometric(p)
    if X1 == X2:
        return 1
    else:
        return 0
```

- **1 pt :** `X1 = rd.geometric(p)` **et** `X2 = rd.geometric(p)`
- **2 pts : structure conditionnelle bien écrite**
- **1 pt : tout est correct**

d) Établir la relation : $\mathbb{P}([X_1 = X_2]) = \dfrac{p}{1+q}$.

- **1 pt : la famille** $([X_2 = k])_{k \in \mathbb{N}^*}$ **est un système complet d'événements**
- **1 pt : FPT** $\mathbb{P}([X_1 - X_2 = 0]) = \sum\limits_{k=1}^{+\infty} \mathbb{P}([X_2 = k] \cap [X_1 - X_2 = 0])$
- **1 pt : indépendance de** X_1 **et** X_2
- **1 pt : décalage d'indice**

– **1 pt : somme d'une série géométrique de raison** q^2 **et résultat** ($\mathbb{P}([X_1 = X_2]) = \dfrac{p}{2-p} = \dfrac{p}{1+q}$)

2. a) Montrer que Z suit la loi géométrique de paramètre $1 - q^2$.
En déduire $\mathbb{E}(Z)$, $\mathbb{V}(Z)$ et $\mathbb{E}(T)$.

– **1 pt :** $Z(\Omega) \subset \mathbb{N}^*$
– **1 pt :** $[Z > k] \ = \ [X_1 > k] \cap [X_2 > k]$
– **1 pt : indépendance de** X_1 **et** X_2
– **1 pt : résultat**
– **1 pt :** $[Z \geqslant k] = [Z > k] \cup [Z = k]$
– **1 pt :** $[Z > k-1] = [Z = k] \cup [Z > k]$ **car** Z **est à valeurs entières**
– **1 pt : incompatibilité et conclusion** $Z \hookrightarrow \mathcal{G}\left(1 - q^2\right)$
– **1 pt :** $\mathbb{E}(Z) = \dfrac{1}{1-q^2}$
– **1 pt :** $\mathbb{V}(Z) = \dfrac{q^2}{(1-q^2)^2}$
– **1 pt :** $T + Z = X_1 + X_2$
– **1 pt : existence de** $\mathbb{E}(T)$ **par somme de v.a.r. admettant une espérance**
– **1 pt :** $\mathbb{E}(T) = \dfrac{1+2q}{1-q^2}$

b) Soit k un entier de $\mathbb{N}^*$. Justifier l'égalité : $[Z = k] \cup [T = k] = [X_1 = k] \cup [X_2 = k]$.

En déduire la relation suivante : $\mathbb{P}([T = k]) = 2\ \mathbb{P}([X_1 = k]) - \mathbb{P}([Z = k])$.

– **2 pts : justification de l'égalité avec les** ω **(1 pt si justification moins rigoureuse)**
– **1 pt : crible** $\mathbb{P}([Z = k] \cup [T = k]) = \mathbb{P}([Z = k]) + \mathbb{P}([T = k]) - \mathbb{P}([X_1 = k] \cap [X_2 = k])$
– **1 pt :** $[Z = k] \cap [T = k] \ = \ [X_1 = k] \cap [X_2 = k]$
– **1 pt :** $\mathbb{P}([X_1 = k] \cup [X_2 = k]) = \mathbb{P}([X_1 = k]) + \mathbb{P}([X_2 = k]) - \mathbb{P}([X_1 = k] \cap [X_2 = k])$
– **1 pt :** X_1 **et** X_2 **suivent la même loi**
– **1 pt :** $\mathbb{P}([T = k]) = 2\ \mathbb{P}([X_1 = k]) - \mathbb{P}([Z = k])$

c) Établir la formule : $\mathbb{V}(T) = \dfrac{q(2q^2 + q + 2)}{(1-q^2)^2}$.

– **1 pt :** $T(\Omega) = \big(\max(X_1, X_2)\big)(\Omega) \subset \mathbb{N}^*$
– **1 pt : convergence absolue citée**
– **1 pt : reconnaître les moments d'ordre** 2 **de** X_1 **et** Z **(**$\mathbb{E}(T^2) = 2\mathbb{E}(X_1^2) - \mathbb{E}(Z^2)$**)**
– **1 pt : formule de KH** $\mathbb{E}(X_1^2) = \mathbb{V}(X_1) + (\mathbb{E}(X_1))^2 = \dfrac{q+1}{p^2}$
– **1 pt : formule de KH** $\mathbb{E}(Z^2) = \dfrac{q^2+1}{(1-q^2)^2}$
– **1 pt : formule de KH** $\mathbb{V}(T) = \mathbb{E}(T^2) - \big(\mathbb{E}(T)\big)^2$ **(l'écrire)**
– **2 pts : valeur de** $\mathbb{V}(T) = \dfrac{q(2q^2 + q + 2)}{(1-q^2)^2}$

3. a) Préciser $(T - Z)(\Omega)$. Exprimer pour tout j de $\mathbb{N}^*$, l'évènement $[Z = j] \cap [Z = T]$ en fonction des évènements $[X_1 = j]$ et $[X_2 = j]$.
En déduire pour tout j de $\mathbb{N}^*$, l'expression de $\mathbb{P}([Z = j] \cap [Z = T])$.

- **2 pts :** $(T-Z)(\Omega)$ **(1 pt valeurs entières, 1 pt valeurs positives)**
- **1 pt :** $[Z=j] \ \cap \ [Z=T] = [X_1=j] \ \cap \ [X_2=j]$
- **1 pt :** $\mathbb{P}([Z=j] \ \cap \ [Z=T]) = p^2q^{2j-2}$ **par indépendance de** X_1 **et** X_2

b) Montrer que pour tout couple (j,l) de $(\mathbb{N}^*)^2$, on a : $\mathbb{P}([Z=j] \cap [T-Z=l]) = 2p^2q^{2j+l-2}$.

- **1 pt :** $[Z=j] \ \cap \ [T-Z=l] = \big([X_1=j] \cap [X_2=j+l]\big) \ \cup \ \big([X_1=j+l] \cap [X_2=j]\big)$
- **1 pt : incompatibilité**
- **1 pt : indépendance de** X_1 **et** X_2
- **1 pt :** $\mathbb{P}([Z=j] \cap [T-Z=l]) = 2p^2q^{2j+l-2}$

c) Montrer que pour tout k de $\mathbb{Z}$, $\mathbb{P}([X_1-X_2=k]) = \dfrac{pq^{|k|}}{1+q}$.
(on distinguera trois cas : $k=0$, $k>0$ et $k<0$)

- **1 pt : cas** $k=0$ **:** $\mathbb{P}([X_1-X_2=k]) = \mathbb{P}([X_1-X_2=0]) = \mathbb{P}([X_1=X_2]) = \dfrac{p}{1+q} = \dfrac{pq^0}{1+q}$
- **1 pt : FPT** $\mathbb{P}([X_1-X_2=k]) = \sum\limits_{i=1}^{+\infty} \mathbb{P}([X_2=i]) \ \times \ \mathbb{P}([X_1=i+k])$
- **1 pt : indépendance**
- **2 pts : cas** $k>0$
  - ▸ **1 pt : série géométrique de raison** q^2
  - ▸ **1 pt : reste du calcul et résultat** $\mathbb{P}([X_1-X_2=k]) = \dfrac{p\ q^k}{1+q}$
- **4 pts : cas** $k<0$
  - ▸ **1 pt : réduction des indices** $(i \in [\![-k+1,+\infty[\![)$
  - ▸ **1 pt : décalage d'indice** $(i \in [\![1,+\infty[\![)$
  - ▸ **1 pt : retrouver la série géométrique du cas** $k>0$
  - ▸ **1 pt : reste du calcul et résultat** $\mathbb{P}([X_1-X_2=k]) = \dfrac{p\ q^{-k}}{1+q}$

d) En déduire la loi de la variable aléatoire $|X_1-X_2|$.

- **1 pt :** $(|X_1-X_2|)(\Omega) \subset \mathbb{N}$
- **1 pt :** $[\ |X_1-X_2|=k] \ = \ [X_1-X_2=k] \ \cup \ [X_1-X_2=-k]$
- **1 pt : cas** $k \neq 0$ **(incompatibilité des événements) et** $\mathbb{P}([\ |X_1-X_2|=k]) = 2\ \dfrac{p\ q^k}{1+q}$
- **1 pt : cas** $k=0$ **et** $\mathbb{P}([\ |X_1-X_2|=0]) = \dfrac{p}{1+q}$

e) Établir à l'aide des questions précédentes que les variables Z et $T-Z$ sont indépendantes.

- **2 pts : cas** $l \neq 0$
- **2 pts : cas** $l=0$

4. a) À l'aide du résultat de la question ***3.e)***, calculer $\mathrm{Cov}(Z,T)$.
Les variables Z et T sont-elles indépendantes ?

- **1 pt :** $\mathrm{Cov}(Z,T-Z)=0$ **par indépendance de** Z **et** $T-Z$
- **1 pt : linéarité à droite** $\mathrm{Cov}(Z,T-Z) = \mathrm{Cov}(Z,T) - \mathrm{Cov}(Z,Z)$
- **1 pt :** $\mathrm{Cov}(Z,Z) = \mathbb{V}(Z) = \dfrac{q^2}{(1-q^2)^2}$

- **1 pt : T et Z ne sont pas indépendantes car** $\mathrm{Cov}(Z,T) = \mathbb{V}(Z) \neq 0$

b) Calculer en fonction de q, le coefficient de corrélation linéaire ρ de Z et T.

- **1 pt : existence car Z et T ont des variances non nulles !**
- **1 pt :** $\rho(Z,T) = \dfrac{\mathrm{Cov}(Z,T)}{\sqrt{\mathbb{V}(Z)}\ \sqrt{\mathbb{V}(T)}}$
- **2 pts : calcul et** $\rho(Z,T) = \sqrt{\dfrac{q}{2+q+2q^2}}$

c) Déterminer la loi de probabilité du couple (Z,T).

- **1 pt : cas $i < j$ alors** $[Z=j] \ \cap\ [T=i] = \varnothing$
- **2 pts : cas $i = j$**
  - **1 pt :** $[Z=j] \ \cap\ [T=j] = [X_1=j] \ \cap\ [X_2=j]$
  - **1 pt :** $\mathbb{P}([Z=j] \ \cap\ [T=j]) \ = \ p^2 \ q^{2j-2}$ **par indépendance**
- **2 pts : cas $i > j$**
  - **1 pt :** $[Z=j] \ \cap\ [T=i] = ([X_1=j] \cap [X_2=i]) \ \cup\ ([X_1=i] \cap [X_2=j])$
  - **1 pt :** $\mathbb{P}([Z=j] \ \cap\ [T=i]) \ = \ 2\ p^2 \ q^{i+j-2}$ **par indépendance**

d) Déterminer pour tout j de $\mathbb{N}^*$, la loi de probabilité conditionnelle de T sachant l'évènement $[Z=j]$.

- **1 pt : formule proba conditionnelle** $\mathbb{P}_{[Z=j]}([T=i]) \ = \ \dfrac{\mathbb{P}([Z=j] \cap [T=i])}{\mathbb{P}([Z=j])} \ = \ \dfrac{\mathbb{P}([Z=j] \cap [T=i])}{\left(q^2\right)^{j-1}\ (1-q^2)}$
- **1 pt : cas $i < j$ alors** $\mathbb{P}([Z=j] \ \cap\ [T=i]) = \mathbb{P}(\varnothing) = 0$
- **2 pts : cas $i = j$**
  - **1 pt :** $\mathbb{P}([Z=j] \ \cap\ [T=j]) \ = \ p^2 \ q^{2j-2}$
  - **1 pt :** $\mathbb{P}_{[Z=j]}([T=j]) \ = \ \dfrac{p}{1+q}$
- **2 pts : cas $i > j$**
  - **1 pt :** $\mathbb{P}([Z=j] \ \cap\ [T=i]) \ = \ 2\ p^2 q^{i+j-2}$
  - **1 pt :** $\mathbb{P}_{[Z=j]}([T=j]) \ = \ \dfrac{2\ p\ q^{i-j}}{1+q}$

e) Soit j un élément de $\mathbb{N}^*$. On suppose qu'il existe une variable aléatoire D_j à valeur dans $\mathbb{N}^*$, dont la loi de probabilité est la loi conditionnelle de T sachant l'évènement $[Z=j]$.
Calculer $\mathbb{E}(D_j)$.

- **1 pt : absolue convergence de la série** $\sum\limits_{i \geqslant 1} i\ \mathbb{P}_{[Z=j]}([T=i])$
- **2 pts : découpage de somme**

$$\sum_{i=1}^{N} i\ \mathbb{P}_{[Z=j]}([T=i]) = \sum_{i=1}^{j-1} i\ \mathbb{P}_{[Z=j]}([T=i]) + \sum_{i=j}^{j} i\ \mathbb{P}_{[Z=j]}([T=i]) + \sum_{i=j+1}^{N} i\ \mathbb{P}_{[Z=j]}([T=i])$$

- **2 pts :** $\sum\limits_{i=j+1}^{N} i\ \dfrac{2\ p\ q^{i-j}}{1+q} = \dfrac{2\ p\ q}{1+q} \sum\limits_{i=1}^{N-j} i\ q^{i-1}$
- **2 pts : calcul** $\sum\limits_{i=1}^{+\infty} i\ \mathbb{P}_{[Z=j]}([T=i]) = \dfrac{p}{1+q} + \dfrac{2\ p\ q}{1+q} \sum\limits_{i=1}^{+\infty} i\ q^{i-1} = \dfrac{1+q^2}{1-q^2}$

Problème 2 (ESSEC II 2013, parties I et II)

L'objet du problème est l'étude de la concentration en un type de bactéries d'un bassin destiné à la baignade.

Une municipalité doit effectuer un prélèvement et l'analyser afin de décider d'autoriser ou non l'utilisation du bassin.
Dans une première partie, on étudiera des propriétés reliant la loi binomiale et la loi de Poisson. Dans une deuxième partie, on regardera la modélisation de la concentration en bactéries du bassin. Enfin, la troisième partie étudiera le principe d'un test destiné à prendre une décision d'utilisation.

Les trois parties peuvent être traitées indépendamment en admettant les résultats des parties précédentes.

Toutes les variables aléatoires sont définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. Si elles existent, on note $\mathbb{E}(T)$ et $\mathbb{V}(T)$ l'espérance et la variance d'une variable aléatoire T.

I. Lien entre loi binomiale et loi de Poisson

1. Pour tout entier n strictement positif, on se donne un réel p_n strictement positif et n variables aléatoires $(X_k)_{k\in[\![1,n]\!]}$ indépendantes et suivant une loi de Bernoulli de paramètre p_n. On suppose que np_n a une limite finie strictement positive et on pose $\lim\limits_{n\to+\infty} np_n = \lambda$.

a) Quelle est la loi de $S_n = \sum\limits_{k=1}^{n} X_k$?

- **1 pt : les v.a.r. X_k sont indépendantes et pour tout $k \in [\![1,n]\!]$, $X_k \hookrightarrow \mathcal{B}(p_n)$**
- **1 pt :** $S_n \hookrightarrow \mathcal{B}(n, p_n)$

b) Soit k un entier naturel.

(i) Donner l'expression de $\mathbb{P}\big([S_n = k]\big)$ pour n supérieur ou égal à k.

- **1 pt :** $\mathbb{P}\big([S_n = k]\big) = \binom{n}{k} p_n^k (1-p_n)^{n-k}$

(ii) Que peut-on dire de la limite de p_n quand n tend vers l'infini ?
Étudier la limite de $(1-p_n)^n$ quand n tend vers l'infini.

- **1 pt :** $np_n \underset{n\to+\infty}{\sim} \lambda$ **car** $\lambda > 0$
- **1 pt :** $p_n \underset{n\to+\infty}{\longrightarrow} 0$
- **1 pt :** $n\ln(1-p_n) \underset{n\to+\infty}{\sim} -np_n$
- **1 pt :** $(1-p_n)^n \underset{n\to+\infty}{\longrightarrow} \mathrm{e}^{-\lambda}$

(iii) Montrer : $\lim\limits_{n\to+\infty} \mathbb{P}\big([S_n = k]\big) = \mathrm{e}^{-\lambda}\, \dfrac{\lambda^k}{k!}$.

- **1 pt :** $\mathbb{P}\big([S_n = k]\big) = \frac{n!}{k!(n-k)!} p_n^k \frac{(1-p_n)^n}{(1-p_n)^k}$
- **1 pt :** $\frac{n!}{(n-k)!} = n(n-1)\dots(n-k+1) = \prod_{i=0}^{k-1}(n-i) \underset{n\to+\infty}{\sim} n^k$
- **1 pt : fin du calcul**

c) On pose $N_n = \max(X_1, \dots, X_n)$.

(i) Quelles sont les valeurs que peut prendre la variable aléatoire N_n ?

- **1 pt :** $N_n(\Omega) = \{0, 1\}$

(ii) Calculer $\lim\limits_{n\to+\infty} \mathbb{P}\big([N_n = 0]\big)$.

- **1 pt :** $[N_n = 0] = \bigcap_{k=1}^{n} [X_k = 0]$
- **1 pt : argument d'indépendance**
- **1 pt :** $\mathbb{P}([N_n = 0]) \underset{n\to+\infty}{\longrightarrow} \mathbf{e}^{-\lambda}$

(iii) `(CUBES UNIQUEMENT)` En déduire la limite en loi de la suite $(N_n)_{n\in\mathbb{N}^*}$.

- **1 pt :** $([N_n = 0], [N_n = 1])$ **est un système complet d'événements**
- **1 pt :** $N_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} N$ **où** $N \hookrightarrow \mathcal{B}\left(1 - \mathbf{e}^{-\lambda}\right)$

2. Soit λ un réel strictement positif et n un entier strictement positif tels que $0 < \lambda \leqslant n$. On considère U une variable aléatoire de loi de Bernoulli de paramètre $\dfrac{\lambda}{n}$. On a donc $\mathbb{P}\big([U = 1]\big) = \dfrac{\lambda}{n}$, $\mathbb{P}\big([U = 0]\big) = 1 - \dfrac{\lambda}{n}$ et pour tout i entier naturel supérieur ou égal à 2, $\mathbb{P}\big([U = i]\big) = 0$.

a) Écrire une fonction **Python** nommée `simulU(lam, n)` qui simule la variable aléatoire U (`lam` représente ici le réel λ).

```
def simulU(lam, n):
    if rd.random() < lam / n:
        return 1
    else:
        return 0
```

- **2 pts : fonction correcte (tout ou rien)**

b) Montrer que pour tout réel positif u : $1 - u \leqslant \mathrm{e}^{-u}$.

- **1 pt : La fonction exponentielle est convexe sur $\mathbb{R}$. Ainsi, son graphe est au dessus de toutes ses tangentes et en particulier celle en 0, d'équation $y = x + 1$.**
- **1 pt :** $\mathbf{e}^{x} \geqslant 1 + x$ **donc** $\mathbf{e}^{-u} \geqslant 1 + (-u) = 1 - u$

c) Montrer :

$$\sum_{k=0}^{+\infty} \left| \mathbb{P}\big([U = k]\big) - \left(\frac{\lambda}{n}\right)^k \frac{1}{k!}\, \mathrm{e}^{-\frac{\lambda}{n}} \right| = \frac{2\lambda}{n} \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right)$$

- **1 pt :** $\displaystyle\sum_{k=0}^{+\infty} \left| \mathbb{P}\big([U = k]\big) - \left(\frac{\lambda}{n}\right)^k \frac{1}{k!}\, \mathbf{e}^{-\frac{\lambda}{n}} \right| = \left|1 - \tfrac{\lambda}{n} - \mathbf{e}^{-\frac{\lambda}{n}}\right| + \tfrac{\lambda}{n}\left|1 - \mathbf{e}^{-\frac{\lambda}{n}}\right| + \sum_{k=2}^{+\infty} \left(\frac{\lambda}{n}\right)^k \frac{1}{k!} \mathbf{e}^{-\frac{\lambda}{n}}$
- **1 pt :** $\left|1 - \frac{\lambda}{n} - \mathbf{e}^{-\frac{\lambda}{n}}\right| = \mathbf{e}^{-\frac{\lambda}{n}} + \dfrac{\lambda}{n} - 1$ **d'après la question *2.b)* avec** $u = \frac{\lambda}{n} \geqslant 0$
- **1 pt :** $\left|1 - \mathbf{e}^{-\frac{\lambda}{n}}\right| = 1 - \mathbf{e}^{-\frac{\lambda}{n}}$ **par croissance de l'exponentielle**
- **1 pt : la série** $\sum \left| \mathbb{P}\big([U = k]\big) - \left(\dfrac{\lambda}{n}\right)^k \dfrac{1}{k!}\, \mathbf{e}^{-\frac{\lambda}{n}} \right|$ **(où k décrit $\mathbb{N}$) converge**
- **1 pt :** $\displaystyle\lim_{N\to+\infty} \sum_{k=2}^{N} \left(\frac{\lambda}{n}\right)^k \frac{1}{k!}\, \mathbf{e}^{-\frac{\lambda}{n}} = \mathbf{e}^{-\frac{\lambda}{n}} \left(\mathbf{e}^{\frac{\lambda}{n}} - \frac{\lambda}{n} - 1\right) = 1 - \tfrac{\lambda}{n}\mathbf{e}^{-\frac{\lambda}{n}} - \mathbf{e}^{-\frac{\lambda}{n}}$
- **1 pt : fin du calcul correct**

d) En déduire :

$$\sum_{k=0}^{+\infty} \left| \mathbb{P}\big([U = k]\big) - \left(\frac{\lambda}{n}\right)^k \frac{1}{k!}\, \mathrm{e}^{-\frac{\lambda}{n}} \right| \leqslant 2 \left(\frac{\lambda}{n}\right)^2$$

- **1 pt : On utilise la question *2.b)* avec** $u = \frac{\lambda}{n}$ **:** $1 - \mathbf{e}^{-\frac{\lambda}{n}} \leqslant \frac{\lambda}{n}$

– **1 pt : vérification de l'hypothèse** $u \in \mathbb{R}^+$ **ou** $2\frac{\lambda}{n} \geqslant 0$

3. Soit λ un réel strictement positif et n un entier strictement positif tels que $0 < \lambda \leqslant n$. Soient Z, U et V trois variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. On suppose que U suit une loi de Bernoulli de paramètre $\dfrac{\lambda}{n}$ et que V suit une loi de Poisson de paramètre $\dfrac{\lambda}{n}$. On observe en particulier que pour tout entier p strictement négatif $\mathbb{P}\big([Z=p]\big) = \mathbb{P}\big([U=p]\big) = \mathbb{P}\big([V=p]\big) = 0$.

a) Soit i un entier naturel fixé. En remarquant que pour tout réels a et b, on a : $|a-b| \leqslant |a|+|b|$, montrer que la série de terme général $\big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big|$, où k décrit $\mathbb{N}$, est convergente. On note A_i sa somme, $A_i = \sum\limits_{k=0}^{+\infty} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big|$.

Montrer que pour tout i on a : $A_i \leqslant 2\left(\dfrac{\lambda}{n}\right)^2$.

– **1 pt :** $0 \leqslant \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| \leqslant \mathbb{P}\big([U=k-i]\big) + \mathbb{P}\big([V=k-i]\big)$

– **1 pt : les séries** $\sum \mathbb{P}\big([U=k-i]\big)$ **et** $\sum \mathbb{P}\big([V=k-i]\big)$ **convergent toutes les deux, car les familles** $\big([U=k-i]\big)_{k\in\mathbb{N}}$ **et** $\big([V=k-i]\big)_{k\in\mathbb{N}}$ **sont des systèmes complets d'événements**

– **1 pt : critère de comparaison pour les séries à termes positifs**

– **1 pt :** $\sum\limits_{k=0}^{+\infty} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| = \sum\limits_{k=-i}^{+\infty} \big|\mathbb{P}\big([U=k]\big) - \mathbb{P}\big([V=k]\big)\big|$

– **1 pt :** $\sum\limits_{k=-i}^{+\infty} \big|\mathbb{P}\big([U=k]\big) - \mathbb{P}\big([V=k]\big)\big| = \sum\limits_{k=0}^{+\infty} \big|\mathbb{P}\big([U=k]\big) - \mathbb{P}\big([V=k]\big)\big|$

b) Montrer que la série de terme général $A_i\, \mathbb{P}\big([Z=i]\big)$ est convergente.

– **1 pt :** $0 \leqslant A_i\mathbb{P}([Z=i]) \leqslant 2\left(\dfrac{\lambda}{n}\right)^2 \mathbb{P}([Z=i])$

– **1 pt : la série** $\sum 2\left(\dfrac{\lambda}{n}\right)^2 \mathbb{P}([Z=i])$ **est convergente**

– **1 pt : critère de comparaison pour les séries à termes positifs**

c) Soit k un entier naturel fixé.
Montrer que la série de terme général $\big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| \times \mathbb{P}\big([Z=i]\big)$, où i décrit $\mathbb{N}$, est convergente.

– **1 pt :** $0 \leqslant \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| \times \mathbb{P}\big([Z=i]\big) \leqslant A_i\mathbb{P}([Z=i])$

– **1 pt : la série** $\sum A_i\mathbb{P}([Z=i])$ **est convergente d'après la question précédente**

– **1 pt : critère de comparaison pour les séries à termes positifs**

d) On pose $B_k = \sum\limits_{i=0}^{+\infty} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| \times \mathbb{P}\big([Z=i]\big)$.

(i) Montrer que, pour $k \geqslant 2$, on a :

$$B_k = \left|1 - \frac{\lambda}{n} - \mathrm{e}^{-\frac{\lambda}{n}}\right| \mathbb{P}\big([Z=k]\big) + \left|\frac{\lambda}{n} - \frac{\lambda}{n}\,\mathrm{e}^{-\frac{\lambda}{n}}\right| \mathbb{P}\big([Z=k-1]\big) + \sum_{i=0}^{k-2} \mathbb{P}\big([V=k-i]\big)\,\mathbb{P}\big([Z=i]\big)$$

– **1 pt :**

$$\begin{aligned}
&\sum_{i=0}^{+\infty} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| \times \mathbb{P}\big([Z=i]\big) \\
&\qquad\qquad = \sum_{i=0}^{k} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big| \times \mathbb{P}\big([Z=i]\big)
\end{aligned}$$

– **2 pts :** $U \hookrightarrow \mathcal{B}\left(\frac{\lambda}{n}\right)$ **et** $V \hookrightarrow \mathcal{P}\left(\frac{\lambda}{n}\right)$ **bien utilisées et fin du calcul**

(ii) Justifier que pour $k \geqslant 2$, on a : $\sum\limits_{i=0}^{k-2} \mathbb{P}\big([V = k-i]\big)\ \mathbb{P}\big([Z = i]\big) \leqslant \mathbb{P}\big([V + Z = k]\big)$.

– **1 pt : formule des probabilités totales avec le système complet d'événements associé à Z, c'est-à-dire la famille** $([Z = i])_{i \in \mathbb{N}}$

– **1 pt :** $\mathbb{P}\big([V + Z = k]\big) = \sum\limits_{i=0}^{k} \mathbb{P}\big([V = k-i]\big)\mathbb{P}\big([Z = i]\big)$

– **1 pt : argument d'indépendance**

– **1 pt :** $\sum\limits_{i=0}^{k-2} \mathbb{P}\big([V = k-i]\big)\ \mathbb{P}\big([Z = i]\big) \leqslant \sum\limits_{i=0}^{k} \mathbb{P}\big([V = k-i] \cap [Z = i]\big)$

(iii) Montrer finalement que la série de terme général B_k, où k décrit $\mathbb{N}$, est convergente.

On admettra alors qu'on a : $\sum\limits_{k=0}^{+\infty} B_k = \sum\limits_{i=0}^{+\infty} A_i\, \mathbb{P}\big([Z = i]\big)$, c'est-à-dire :

$$\begin{aligned} & \sum_{k=0}^{+\infty} \left(\sum_{i=0}^{+\infty} \left|\mathbb{P}\big([U = k-i]\big) - \mathbb{P}\big([V = k-i]\big)\right| \times \mathbb{P}\big([Z = i]\big) \right) \\ = \ & \sum_{i=0}^{+\infty} \left(\sum_{k=0}^{+\infty} \left|\mathbb{P}\big([U = k-i]\big) - \mathbb{P}\big([V = k-i]\big)\right| \right) \times \mathbb{P}\big([Z = i]\big) \end{aligned}$$

– **1 pt :** $0 \leqslant B_k \leqslant \left|1 - \dfrac{\lambda}{n} - \mathrm{e}^{-\frac{\lambda}{n}}\right|\ \mathbb{P}\big([Z = k]\big) + \left|\dfrac{\lambda}{n} - \dfrac{\lambda}{n}\ \mathrm{e}^{-\frac{\lambda}{n}}\right|\ \mathbb{P}\big([Z = k-1]\big) + \mathbb{P}\big([V + Z = k]\big)$

– **1 pt : les trois familles** $([Z = k])_{k \in \mathbb{N}}$**,** $([Z = k-1])_{k \in \mathbb{N}}$ **et** $([V + Z = k])_{k \in \mathbb{N}}$ **sont des systèmes complets d'événements donc les trois séries** $\sum \mathbb{P}([Z = k])$**,** $\sum \mathbb{P}([Z = k-1])$ **et** $\sum \mathbb{P}([V + Z = k])$ **convergent**

– **1 pt : critère de comparaison pour les séries à termes positifs**

4. On conserve dans cette question les notations et les hypothèses de la question 3 concernant les variables aléatoires Z, U et V.

a) Montrer que la série de terme général : $\left|\mathbb{P}\big([Z + U = k]\big) - \mathbb{P}\big([Z + V = k]\big)\right|$ est convergente.

– **1 pt :** $0 \leqslant \left|\mathbb{P}\big([Z + U = k]\big) - \mathbb{P}\big([Z + V = k]\big)\right| \leqslant \mathbb{P}\big([Z + U = k]\big) + \mathbb{P}\big([Z + V = k]\big)$

– **1 pt : les familles** $([Z + U = k])_{k \in \mathbb{N}}$ **et** $([Z + V = k])_{k \in \mathbb{N}}$ **sont des systèmes complets d'événements et donc les séries** $\sum \mathbb{P}\big([Z + U = k]\big)$ **et** $\sum \mathbb{P}\big([Z + V = k]\big)$ **convergent**

– **1 pt : critère de comparaison pour les séries à termes positifs**

b) Déduire de la question 3 :

$$\sum_{k=0}^{+\infty} \left|\mathbb{P}\big([Z + U = k]\big) - \mathbb{P}\big([Z + V = k]\big)\right| \leqslant \sum_{i=0}^{+\infty} \left(\sum_{k=0}^{+\infty} \left|\mathbb{P}\big([U = k-i]\big) - \mathbb{P}\big([V = k-i]\big)\right| \right) \mathbb{P}\big([Z = i]\big)$$

puis :

$$\sum_{k=0}^{+\infty} \left|\mathbb{P}\big([Z + U = k]\big) - \mathbb{P}\big([Z + V = k]\big)\right| \leqslant 2 \left(\frac{\lambda}{n}\right)^2$$

– **1 pt : formule des probabilités totales avec le système complet d'événements** $([Z = i])_{i \in \mathbb{N}}$

– **1 pt : argument d'indépendance**

– **1 pt :** $\left|\mathbb{P}\big([Z + U = k]\big) - \mathbb{P}\big([Z + V = k]\big)\right| = \left|\sum\limits_{i=0}^{+\infty} \Big(\mathbb{P}\big([U = k-i]\big) - \mathbb{P}\big([V = k-i]\big)\Big)\, \mathbb{P}\big([Z = i]\big)\right|$

– **1 pt : inégalité triangulaire :**

$$\left|\sum_{i=0}^{+\infty} \big(\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big)\,\mathbb{P}\big([Z=i]\big)\right| \leqslant \sum_{i=0}^{+\infty} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big|\,\mathbb{P}\big([Z=i]\big)$$

– **1 pt :**

$$\sum_{k=0}^{+\infty} \big|\mathbb{P}\big([Z+U=k]\big) - \mathbb{P}\big([Z+V=k]\big)\big| \leqslant \sum_{i=0}^{+\infty} \left(\sum_{k=0}^{+\infty} \big|\mathbb{P}\big([U=k-i]\big) - \mathbb{P}\big([V=k-i]\big)\big|\right) \times \mathbb{P}\big([Z=i]\big)$$

– **1 pt :** $\sum\limits_{i=0}^{+\infty} A_i \mathbb{P}\big([Z=i]\big) \leqslant 2\left(\frac{\lambda}{n}\right)^2$

5. Soient $U_1, \dots, U_n, V_1, \dots, V_n$, $2n$ variables aléatoires indépendantes, telles que pour tout entier i tel que $1 \leqslant i \leqslant n$, U_i suit une loi de Bernoulli de paramètre $\dfrac{\lambda}{n}$ et V_i suit une loi de Poisson de paramètre $\dfrac{\lambda}{n}$.

a) Montrer que pour tout entier naturel k, on a :

$$\begin{array}{rl} & \big|\mathbb{P}\big([U_1+\cdots+U_n=k]\big) - \mathbb{P}\big([V_1+\cdots+V_n=k]\big)\big| \\[2mm] \leqslant & \big|\mathbb{P}\big([U_1+\cdots+U_n=k]\big) - \mathbb{P}\big([U_1+\cdots+U_{n-1}+V_n=k]\big)\big| \\[2mm] + & \big|\mathbb{P}\big([U_1+\cdots+U_{n-1}+V_n=k]\big) - \mathbb{P}\big([U_1+\cdots+U_{n-2}+V_{n-1}+V_n=k]\big)\big| \\[2mm] + & \cdots + \big|\mathbb{P}\big([U_1+V_2+\cdots+V_n=k]\big) - \mathbb{P}\big([V_1+\cdots+V_n=k]\big)\big| \end{array}$$

– **1 pt : argument de télescopage**
– **1 pt : inégalité triangulaire**
– **1 pt : formalisation / explications**

b) En déduire :

$$\sum_{k=0}^{+\infty} \big|\mathbb{P}\big([U_1+\cdots+U_n=k]\big) - \mathbb{P}\big([V_1+\cdots+V_n=k]\big)\big| \leqslant 2\frac{\lambda^2}{n}$$

– **1 pt : formalisation $U = U_{j+1}$, $V = V_{j+1}$ et $Z = U_1 + \cdots + U_j + V_{j+2} + \cdots + V_n$**
– **1 pt : vérification que l'on peut appliquer la question *4.b)***
– **1 pt : lemme des coalitions**
– **1 pt :** $\sum\limits_{k=0}^{K} \big|\mathbb{P}\big([U_1+\cdots+U_n=k]\big) - \mathbb{P}\big([V_1+\cdots+V_n=k]\big)\big| \leqslant 2\frac{\lambda^2}{n}$
– **1 pt : la suite des sommes partielles associée à la série de terme général $\big|\mathbb{P}\big([U_1+\cdots+U_n=k]\big) - \mathbb{P}\big([V_1+\cdots+V_n=k]\big)\big|$ est majorée. S'agissant d'une série à termes positifs, on peut conclure qu'elle converge**

c) Soit X une variable aléatoire de loi binomiale de paramètres n et $\dfrac{\lambda}{n}$. Conclure de ce qui précède :

$$\sum_{k=0}^{+\infty} \left|\mathbb{P}\big([X=k]\big) - \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\right| \leqslant 2\frac{\lambda^2}{n}$$

- **1 pt : argument d'indépendance**
- **1 pt :** $U_1 + \cdots + U_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ **et donc** $U_1 + \cdots + U_n$ **et** X **suivent la même loi**
- **1 pt :** $V_1 + \cdots + V_n \hookrightarrow \mathcal{P}(\lambda)$
- **1 pt :** $\sum\limits_{k=0}^{+\infty} \left| \mathbb{P}\big([X = k]\big) - \dfrac{\lambda^k}{k!}\mathbf{e}^{-\lambda} \right| = \sum\limits_{k=0}^{+\infty} \left| \mathbb{P}\big([U_1 + \cdots + U_n = k]\big) - \mathbb{P}\big([V_1 + \cdots + V_n = k]\big) \right|$

II. Modélisation de la concentration en bactéries.

Le bassin qu'on étudie est supposé de volume V, en m^3, et on effectue un prélèvement de volume ΔV, en m^3. Dans cette situation, la probabilité pour qu'une bactérie spécifique du bassin se trouve dans le prélèvement est égale à $\dfrac{\Delta V}{V}$.

Supposons que le bassin contienne n bactéries numérotées de 1 à n, $n \in \mathbb{N}^*$. On considère alors n variables aléatoires $X_1, \ldots, X_n$ à valeurs 0 ou 1 telles que $X_i = 1$ si la bactérie i se trouve dans le prélèvement et 0 sinon. Les variables en question sont supposées indépendantes.

On pose $c = \dfrac{n}{V}$ qui représente la concentration en bactéries du bassin par m^3.

6. ***a)*** Quelle est la loi de X_i pour $i \in [\![1, n]\!]$?

- **1 pt :** $X_i(\Omega) = \{0, 1\}$ **donc** X_i **suit une loi de Bernoulli**
- **1 pt :** $X_i \hookrightarrow \mathcal{B}\left(\dfrac{\Delta V}{V}\right)$

b) Soit N le nombre de bactéries présentes dans le prélèvement.
Montrer que N suit une loi binomiale de paramètres n et $\dfrac{\Delta V}{V}$.

- **1 pt :** $N = \sum\limits_{i=1}^{n} X_i$
- **1 pt :** $X_1, \ldots, X_n$ **sont indépendantes**

7. Soit U une variable aléatoire qui suit la loi de Poisson de paramètre $a > 0$.
On appelle F sa fonction de répartition.

a) Montrer que, pour tout $x \in \mathbb{R}$, $F(x) = \mathrm{e}^{-a} \sum\limits_{k=0}^{\lfloor x \rfloor} \dfrac{a^k}{k!}$.

- **1 pt :** $F(x) = \mathbb{P}([U \leqslant x]) = \mathbb{P}\left(\bigcup\limits_{k=0}^{\lfloor x \rfloor} [U = k] \right)$
- **1 pt : argument d'incompatibilité**
- **1 pt :** $F(x) = \mathbf{e}^{-a} \sum\limits_{k=0}^{\lfloor x \rfloor} \dfrac{a^k}{k!}$

b) Recopier et compléter la fonction **Python** suivante pour qu'elle prenne en argument d'entrées les réels x et a et qu'elle renvoie la valeur de $F(x)$. On remarquera qu'on s'attache dans cette fonction à minimiser le nombre d'opérations.

```
def repartition_Poisson(x, a):
    S = 0
    terme = 1
    k = 0
    while  k <= x:
        S = S + terme
        k =  k + 1
        terme =  terme * (a / k)
    return  np.exp(-a) * S
```

- **1 pt par ligne**
- **1 pt bonus si tout est correct**

8. Soit U une variable aléatoire de loi de Poisson de paramètre $c\Delta V$. Soit $K \in \mathbb{N}^*$ fixé.

a) Montrer : $\left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right| \leqslant \dfrac{2c\,(\Delta V)^2}{V}$.

- **1 pt : argument d'incompatibilité**
- **1 pt : inégalité triangulaire**
- **1 pt :** $\left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right| \leqslant \sum\limits_{k=0}^{K} \left|\mathbb{P}\big([N = k]\big) - \mathbf{e}^{-c\Delta V}\frac{(c\Delta V)^k}{k!}\right|$
- **1 pt :** $N \hookrightarrow \mathcal{B}\left(n, \dfrac{c\Delta V}{n}\right)$
- **1 pt :** $\left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right| \leqslant \dfrac{2c\,(\Delta V)^2}{V}$

b) On suppose que $V = 1000$, V volume en m^3, et que le prélèvement est de volume égal à 1 litre soit $\Delta V = 10^{-3}$.
Trouver un majorant de l'erreur commise en approximant $\mathbb{P}\big([N \leqslant K]\big)$ par $\mathbb{P}\big([U \leqslant K]\big)$.

- **1 pt : l'erreur commise en approximant $\mathbb{P}\big([N \leqslant K]\big)$ par $\mathbb{P}\big([U \leqslant K]\big)$ est :**

$$\varepsilon = \left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right|$$

- **1 pt :** $\varepsilon \leqslant 2c\dfrac{(10^{-3})^2}{10^3} = 2c10^{-9}$

c) On suppose que la concentration dans le bassin reste inférieure à 10^6 bactéries par mètre cube, et que le prélèvement réalisé est encore de 1 litre.
Quelle est la valeur minimale du volume V garantissant que l'erreur commise dans l'approximation précédente soit inférieure à 10^{-6} ?

- **1 pt :** $\varepsilon \leqslant 2 \times 10^6 \dfrac{(10^{-3})^2}{V}$
- **1 pt : la valeur minimale est $V = 2 \times 10^6$ en m^3**

On <u>admettra</u> dans la suite que le résultat de la question ***5.c)*** peut être amélioré de la façon suivante :

$$\sum_{k=0}^{+\infty} \left|\mathbb{P}\big([X = k]\big) - \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda}\right| \leqslant 2\lambda\,\frac{\min(2, \lambda)}{n}$$

d) Montrer que si U suit une loi de Poisson de paramètre $c\Delta V$:

$$\left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right| \leqslant 4\frac{\Delta V}{V}$$

- **1 pt :** $\left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right| \leqslant 2\lambda\,\dfrac{\min(2, \lambda)}{n}$ **où** $\lambda = c\Delta V$
- **1 pt :** $\min(2, \lambda) \leqslant 2$
- **1 pt :** $\left|\mathbb{P}\big([N \leqslant K]\big) - \mathbb{P}\big([U \leqslant K]\big)\right| \leqslant 4\dfrac{\Delta V}{V}$

e) On suppose toujours que le prélèvement réalisé est de un litre. À l'aide de l'inégalité précédente, trouver la valeur minimale du volume V garantissant que l'erreur commise dans l'approximation précédente soit inférieure a 10^{-6}.

- **1 pt : valeur minimale $V = 4 \times 10^3$ en m^3**

III. Construction d'une procédure de test.

Non proposé ici.