Devoir 07

Variables aléatoires discrètes
Lundi 15/01/2024

Exercice I

Méthode de rejet.

On suppose que l'on dispose d'un ensemble fini A et d'une v.a. U uniforme dans cet ensemble A.

Etant donnée une partie B de A, le but de l'exercice est de construire une v.a. V uniforme sur B.

On propose l'algorithme suivant, qui consiste à tirer au sort (avec remise) des éléments de A tant qu'on n'a pas obtenu un élément de B :

1. On dispose d'une suite $(U_k)_{k\geq 1}$ de variables aléatoires indépendantes de même loi que U (un échantillon de U.
2. On pose K le plus petit indice k tel que $U_k \in B$.
3. On pose alors $V = U_K$.

1. Montrer que K suit une loi géométrique sur $\mathbb{N}^*$ dont on déterminera le paramètre de succès en fonction de $\#A$ et $\#B$.
2. En conditionnant sur les valeurs de K, montrer que V suit la loi uniforme sur B.
3. En Python/Numpy, l'expression booléenne `( np.gcd(a,b) == 1 )` vaut `True` si les entiers strictement positifs a et b sont « *premiers entre eux* » (la signification exacte de cette expression n'a aucune importance pour faire l'exercice).

Sachant cela, en utilisant la méthode décrite précédemment, construire une fonction `XY(N)` retournant un couple d'entiers « premiers entre eux » (a,b) tiré uniformément parmi les couples de ce type vérifiant de plus $1 \leq a \leq N$ et $1 \leq b \leq N$.

On pourra utiliser la fonction `np.random.randint(N)` qui retourne un entier tiré uniformément entre 0 et $N-1$.
4. Evaluer (informatiquement, en précisant comment vous procédez), pour N grand ($N = 100$ ou $N = 1000$ suffisent), l'espérance de K, c'est à dire le nombre moyen de tirages de couples d'entiers à faire avant de trouver un couple d'entiers premiers entre eux.

En lisant la page wikipedia indiquée ci-dessous (juste cette section) et avec ce que vous savez des variables géométriques que devrait être la limite théorique de $\mathbb{E}(K)$ lorsque $N \to +\infty$?

`https://fr.wikipedia.org/wiki/Nombres_premiers_entre_eux#Probabilit%C3%A9s`

Problème II

On notera $\mathbb{R}$ l'ensemble des nombres réels, $\mathbb{N}$ l'ensemble des entiers naturels et $\mathbb{N}^*$ l'ensemble des entiers naturels non nuls.

Dans tout le problème, p désignera un nombre réel de l'intervalle $]0,1[$ et $q = 1-p$.

La notation $\lfloor x \rfloor$ désigne la partie entière du nombre réel x.

Toutes les variables aléatoires (v.a.) de ce problème sont définies sur un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$. Pour $X : \Omega \to \mathbb{R}$, une v.a. à valeurs réelles, on note $\mathbb{E}(X)$, resp. $\mathbb{V}(X)$, l'espérance, resp. la variance, de X lorsque celle-ci existe.

Le but de ce problème est de donner deux modélisations, *in fine* équivalentes, du scintillement obtenu lors d'une autoradiographie β.

Une substance subissant une désintégration β, par exemple du cobalt 60 radioactif se désintégrant en du nickel 60 stable suivant

$$^{60}_{27}Co \to \ ^{60}_{27}Ni + e^- + \bar{\nu}_e,$$

émet à chaque désintégration un électron, électron qui frappe notre détecteur à un instant donné (scintillement).

On enregistre, les instants de scintillement en nanosecondes à partir de l'instant initial de la mesure et ces mesures constituent notre ensemble de données à analyser.

Partie A

Préliminaire : Instant du s-ième scintillement

Dans cette première modélisation de notre problème de mesure de l'activité de désintégration β, on observe X_k qui vaut 1 si il y a scintillement à l'instant k (en nanoseconde), 0 sinon.

Statistiquement ces observations ont l'air indépendantes (elles passent des tests d'indépendance usuels).

Soit X une variable aléatoire de BERNOULLI de paramètre de succès p et $(X_k)_{k\in\mathbb{N}^*}$ un échantillon de X, *i.e.* une suite de v.a. indépendantes, de même loi que X.

On rappelle que $\mathbb{P}(X=1)=p$ et $\mathbb{P}(X=0)=q=1-p$.

A.1. On pose $S_0=0$ et, pour un entier naturel m, $m\in\mathbb{N}^*$, on note

$$S_m = \sum_{k=1}^{m} X_k.$$

Rappeler, en justifiant brièvement, la loi de S_m, son espérance et sa variance.

A.2. Pour $m\in\mathbb{N}^*$, $s\in\mathbb{N}^*$, on définit l'événement $E_{s,m}=\{S_m<s\}$ et, pour $s\in\mathbb{N}^*$, l'événement

$$E_s = \bigcap_{m=1}^{\infty} \{S_m < s\}.$$

A.2.a. Montrer que

$$E_1 = \{\forall k\in\mathbb{N}^*, X_k=0\}$$

Plus généralement, exprimer l'événement E_s en terme des $(X_k)_{k\in\mathbb{N}^*}$ et montrer que : $\forall s\in\mathbb{N}^*$, $E_s\subset E_{s+1}$.

A.2.b. Rappeler l'inégalité de BIENAYMÉ–TCHEBYCHEFF et montrer que pour tout $m>\frac{s}{p}$,

$$\mathbb{P}(E_{s,m}) \le \frac{m.p.(1-p)}{(m.p-s)^2}$$

A.2.c. Montrer que : $\forall m\in\mathbb{N}^*$, $E_s\subset E_{s,m}$ et en déduire la valeur de $\mathbb{P}(E_s)$.

A.3. On note, pour un entier naturel s, $s\in\mathbb{N}^*$,

$$U_s = \begin{cases} \min\{m\in\mathbb{N}^*, S_m\ge s\} & \text{sur } \overline{E_s} \\ 0 & \text{sur } E_s \end{cases}$$

On admet que les U_s ainsi définies sont des variables aléatoires.

A.3.a. Que vaut $\mathbb{P}(U_1=0)$? Que vaut $\mathbb{P}(U_1\ge 1)$?

A.3.b. Que valent (sur $\overline{E_1}$) X_{U_1}, S_{U_1} et S_{U_1-1} ?

A.3.c. Montrer l'égalité d'événements

$$\{U_1>n\}\cup E_1 = \{S_n=0\}$$

A.3.d. Quelle est la loi de U_1 ? En donner espérance et variance.

A.3.e. Que vaut $\mathbb{P}(U_s=0)$? Que vaut $\mathbb{P}(U_s\ge 1)$?

A.3.f. Que valent $(\overline{E_s})$ X_{U_s}, S_{U_s} et S_{U_s-1} ?
A.3.g. Soit $s \in \mathbb{N}^*$, $n \in \mathbb{N}^*$, montrer l'égalité d'événements

$$\{U_s > n\} \cup E_s = \{S_n < s\}$$

A.3.h. En déduire que $\mathbb{P}(U_s > s-1) = 1$.
A.3.i. En utilisant les lois de S_n et S_{n-1} et la formule du triangle de PASCAL, en déduire, pour $n \geq s$, que

$$\mathbb{P}(U_s = n) = \binom{n-1}{s-1} p^s.q^{n-s}$$

Partie B

Première implémentation informatique

Dans les questions suivantes, le seul "hasard informatique" ne peut provenir que d'appels à la fonction `np.random.rand()` du module Python `numpy`. Cette fonction retourne, à chaque appel, et indépendamment des appels précédents, un nombre réel aléatoire tiré uniformément sur l'intervalle $]0,1[$.
B.1. Ecrire une fonction Python `ExperienceBernoulli(K=100,p=0.5)` retournant une liste de `K` valeurs simulant un tirage de la famille de variables $(X_k)_{1\leq k\leq K}$.
B.2. Ecrire une fonction Python `SommesPartielles(X)` qui, étant donnés une liste Python `X` comportant $K =$ `len(X)` valeurs 0 ou 1 `X=[`$X_1,\dots,X_K$`]`, retourne la liste des sommes `S=[`$S_1,\dots,S_K$`]` telles que définies dans la partie A.
B.3. Ecrire une fonction Python `U(X,s=1)` qui, étant donnée une liste Python `X` comportant $K =$ `len(X)` valeurs 0 ou 1 `X=[`$X_1,\dots,X_K$`]` et un entier `s` $= s$, retourne la valeur de U_s correspondante telle que définie dans la partie A.
B.4. Indiquer la procédure permettant de tracer, via simulations, une approximation de la loi de U_s ?

Partie C

Un développement en série

On cherche à démontrer la formule suivante–dite du binôme de NEWTON négatif–

$$\forall n \in \mathbb{N}^*, \forall x \in]-1,+1[, (1-x)^{-n} = \sum_{k=0}^{+\infty} \binom{n-1+k}{k}.x^k$$

C.1.a. Vérifier que cette formule est correcte pour $n = 1$.
C.1.b. Vérifier que cette formule est correcte pour $n = 2$.
C.2. Soit $n \in \mathbb{N}^*$, $x \in]-1,+1[$, $x \neq 0$. Pour $k \in \mathbb{N}$, on pose

$$u_k = \binom{n-1+k}{k}.|x|^k.$$

C.2.a. Calculer $\frac{u_{k+1}}{u_k}$ et sa limite lorsque $k \to +\infty$. Montrer alors qu'il existe $0 < q < 1$ et $k_0 \in \mathbb{N}$ (dépendants potentiellement de n et de x, ce q n'est pas celui de l'introduction) tels que

$$\forall k \geq k_0, 0 \leq u_{k+1} \leq q.u_k$$

C.2.b. En déduire que

$$\forall k \geq k_0, 0 \leq u_k \leq q^{k-k_0}.u_{k_0}$$

C.2.c. et conclure quant à la convergence de la série $\sum_{k=0}^{+\infty} \binom{n-1+k}{k}.x^k$

C.3.a. Soit $n \in \mathbb{N}^*$, $x \in]-1,+1[$. Montrer, par récurrence sur N la formule

$$(1-x)^{-n} = \sum_{k=0}^{N} \binom{n-1+k}{k}.x^k + x^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x)$$

où

$$R_N(x) = (N+1).\int_0^1 (1-t)^N.(1-t.x)^{-N-n-1}\, dt$$

Indication:Par une i.p.p. bien choisie, montrer que

$$R_N(x) = 1 + x.\frac{N+n+1}{N+2} R_{N+1}(x)$$

C.3.b. Montrer que pour $t \in [0,1]$,

$$0 \leq (1-t)^N.(1-t.x)^{-N} \leq 1$$

et que, pour tout $x \in]-1,+1[$, la suite $\frac{R_N(x)}{N})_{N\in\mathbb{N}^*}$ est bornée.
C.3.c. En déduire que

$$x^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x) \overset{N\to+\infty}{\to} 0.$$

C.3.d. et conclure quant à la formule à démontrer.
C.4. Développements limités.
C.4.a. Soit $N \in \mathbb{N}$ et $n \in \mathbb{N}^*$. Donner le développement limité à l'ordre N de $x \mapsto (1-x)^{-n}$ lorsque $x \to 0$.
C.4.b. Soit $\gamma = (\gamma_k)_{k\in\mathbb{N}}$ une suite de nombres réels compris entre 0 et 1.
C.4.b.i. Montrer que pour tout $x \in]-1,+1[$ la série $\sum_{k\geq 0} \gamma_k.x^k$ est convergente.
On définit alors la fonction $g :]-1,+1[\to \mathbb{R}$ par la formule

$$g(x) = \sum_{k=0}^{+\infty} \gamma_k.x^k.$$

C.4.b.ii. Montrer que la fonction g est bornée sur $\left[-\frac{1}{2},+\frac{1}{2}\right]$, plus précisémment que

$$\forall x \in \left[-\frac{1}{2},+\frac{1}{2}\right], |g(x)| \leq 2.$$

C.4.b.iii. Soit $N \in \mathbb{N}$. En écrivant

$$g(x) = \sum_{k=0}^{N} \gamma_k.x^k + x^{N+1}\sum_{k=0}^{+\infty} \gamma_{k+N+1}.x^k.$$

donner, en le justifiant, le développement limité à l'ordre N de $g(x)$ lorsque $x \to 0$.

Partie D

Loi binomiale négative.

Dans cette modélisation de notre problème de mesure de l'activité de désintégration β, on mesure Δ_k, la durée en nanosecondes entre le $(k-1)$-ème et le k-ième scintillement. On constate statistiquement que ces durées sont indépendantes et ont l'air de suivre une loi géométrique sur $\mathbb{N}^*$ de paramètre $p \in]0,1[$, on pose $q = 1-p$.
D.1. Soit Δ_1 une v.a. suivant la loi géométrique sur $\mathbb{N}^*$ de paramètre p. Rappeler la définition de la loi de Δ_1, son espérance, sa variance.
D.2. Soient X et Y deux v.a. *indépendantes* prenant leurs valeurs dans $\mathbb{N}$ et $Z = X+Y$.
D.2.a. Dans quel ensemble Z prend-elle naturellement ses valeurs ?
D.2.b. Montrer que

$$\forall k \in \mathbb{N}, \mathbb{P}(Z=k) = \sum_{i=0}^{k} \mathbb{P}(X=i)\mathbb{P}(Y=k-i) = \sum_{j=0}^{k} \mathbb{P}(X=k-j)\mathbb{P}(Y=j).$$

$\star\star\star$

On se donne une suite $(\Delta_n)_{n\in\mathbb{N}^*}$ de v.a. *indépendantes*, suivant la loi géométrique sur $\mathbb{N}^*$ de paramètre p. On pose $q = 1 - p$.

D.3. On définit $U_1 = \Delta_1$ et $U_2 = \Delta_1 + \Delta_2 = U_1 + \Delta_2$.
D.3.a. Dans quel ensemble U_2 prend-elle naturellement ses valeurs ?
D.3.b. Donner espérance et variance de U_2.
D.3.c. Montrer que

$$\forall k \in \mathbb{N}, \mathbb{P}(U_2 = k+2) = (k+1)p^2.q^k.$$

D.4. On pose $U_3 = \Delta_1 + \Delta_2 + \Delta_3 = U_2 + \Delta_3$.
D.4.a. Dans quel ensemble U_3 prend-elle naturellement ses valeurs ?
D.4.b. Donner espérance et variance de U_3.
D.4.c. Rappeler la valeur de $\sum_{j=0}^{k}(j+1)$ et montrer que

$$\forall k \in \mathbb{N}, \mathbb{P}(U_3 = k+3) = p^3.q^k.\frac{1}{2}.(k+1).(k+2).$$

D.5. On pose $U_0 = 0$ et, pour $s \in \mathbb{N}^*$,

$$U_s = \Delta_1 + \cdots + \Delta_s = \sum_{k=1}^{s} \Delta_k = U_{s-1} + \Delta_n.$$

D.5.a. Dans quel ensemble U_s prend-elle naturellement ses valeurs ?
D.5.b. Donner espérance et variance de U_s.
D.6.a. Soit $x \in [-1,+1]$ et U une variable aléatoire à valeurs dans $\mathbb{N}$. Justifier que la v.a. x^U admet une espérance, *i.e.* que $\mathbb{E}(x^U)$ existe et donner une formule reliant $\mathbb{E}(x^U)$ et loi de U.

On définit, pour $s \in \mathbb{N}^*$, $x \in [-1,+1]$,

$$f_s(x) = \mathbb{E}(x^{U_s}).$$

D.6.b. Donner une formule simple pour f_1.
D.6.c. Donner une formule simple pour f_2.
D.6.d. Montrer que pour $s \in \mathbb{N}^*$,

$$\forall x \in [-1,+1], f_s(x) = \frac{(p.x)^s}{(1-q.x)^s} = f_1(x)^s$$

D.6.e. Montrer, en se servant de la partie C, que

$$\forall x \in \left]-1,+1\right[, f_s(x) = p^s.x^s.\sum_{k=0}^{+\infty}\binom{s-1+k}{k}.q^k.x^k = \sum_{k=0}^{+\infty}\binom{s-1+k}{k}.p^s.q^k.x^{k+s}$$

D.7. Démontrer que

$$\forall s \in \mathbb{N}^*, \forall k \in \mathbb{N}, \mathbb{P}(U_s = k+s) = p^s.q^k.\binom{s-1+k}{k}$$

Indication:Raisonner sur l'unicité du développement limité de f_s en 0 vu en C.4.b.
D.8. Comparer les lois des variables U_s de cette partie et celles des variables U_s de la partie A.

D.9. Soit $n \in \mathbb{N}^*$ et, pour $s \in \{0, \dots, n\}$, l'événement

$$F_{n,s} = \{U_{s+1} > n\} \cap \{U_s \leq n\}$$

On définit la variable N_n par la formule

$$N_n = \sum_{s=0}^{n} s.\mathbb{1}_{F_{n,s}}$$

D.9.a. Vérifier que $\forall \omega \in \Omega$, $U_1(\omega) < U_2(\omega) < \dots < U_s(\omega) < \dots$, que $(F_{n,s},\, s \in \{0, \dots, n\})$ est un système complet d'événements incompatibles, et que

$$N_n = \begin{cases} 0 & \text{si } U_1 > n \\ 1 & \text{si } U_2 > n \text{ et } U_1 \leq n \\ 2 & \text{si } U_3 > n \text{ et } U_2 \leq n \\ \vdots & \\ n-1 & \text{si } U_n > n \text{ et } U_{n-1} \leq n \\ n & \text{si } U_{n+1} > n \text{ et } U_n \leq n \end{cases}.$$

D.9.b. Montrer que pour $s \in \mathbb{N}^*$, $n \in \mathbb{N}^*$, on a

$$\{U_s > n\} = \{N_n < s\}$$

et déduire de A.3 que N_n suit une loi binomiale dont on précisera les paramètres.
D.10. On suppose que λ est un nombre réel > 0 fixé, n est très grand et $p = \frac{\lambda}{n}$. Justifier les affirmations suivantes en précisant leur sens :
D.10.a. N_n suit approximativement une loi de Poisson de paramètre λ.
D.10.b. Les variables $\frac{1}{n}\Delta_k$ suivent approximativement une loi exponentielle de paramètre λ.

Partie E

Deuxième implémentation informatique

E.1. Montrer que si D est une variable aléatoire suivant une loi $\mathscr{E}(\lambda)$ pour un certain nombre réel $\lambda > 0$ alors $\Delta = \lfloor D \rfloor + 1$ suit une loi géométrique sur $\mathbb{N}^*$ dont on précisera le paramètre p en fonction de λ.
E.2. Les fonctions `numpy.floor`, resp. `numpy.log` retournent la partie entière, resp. le logarithme, d'un nombre réel. En utilisant la question précédente, ces deux fonctions et la fonction `numpy.random.rand()`, écrire une fonction Python `Delta(p=0.5)` retournant une valeur aléatoire tirée suivant la loi géométrique sur $\mathbb{N}^*$ de paramètre p, valant par défaut 0.5.
E.3.a. Ecrire alors une fonction Python `U2(s=1,p=0.5)` qui, étant donné un entier `s` $= s$, retourne une valeur aléatoire tirée suivant la loi de U_s.
E.3.b. Comparer l'efficacité de cette méthode avec celle de la méthode présentée en partie B.

Correction DM 07

Correction Ex.–1

1. Posons, pour $k \in \mathbb{N}^*$, B_k la variable indicatrice de l'événement $\{U_k \in B\}$. La suite $(B_k)_{k\in\mathbb{N}^*}$ est une suite de v.a. de BERNOULLI, indépendantes, de même paramètre de succès

$$p = \mathbb{P}(U_k \in B) = \frac{\#B}{\#A}$$

(L'intervention des cardinaux dans cette formule est due au caractère uniforme sur A de la v.a. U_k)

La v.a. K est la valeur de l'instant du premie succès dans cette suite et donc K suit une loi géométrique sur $\mathbb{N}^*$ de paramètre $p = \frac{\#B}{\#A}$.

2. Soit $b \in B$. Evaluons $\mathbb{P}(V = b)$ en conditionnant sur K, c'est à dire en utilisant la formule des probabilités totales avbec le SCEI $(\{K = k\}, k \in \mathbb{N}^*)$:

$$\begin{array}{rcl}
\mathbb{P}(V=b) & = & \displaystyle\sum_{k=1}^{\infty} \mathbb{P}(V=b|K=k)\mathbb{P}(K=k) \\
 & = & \displaystyle\sum_{k=1}^{\infty} \mathbb{P}(\{V=b\}\cap\{K=k\}) \\
\textbf{[lorsque } K=k\textbf{, } V=U_k\textbf{]} & = & \displaystyle\sum_{k=1}^{\infty} \mathbb{P}(U_k=b\cap K=k) \\
\textbf{[traduction } K=k\textbf{]} & = & \displaystyle\sum_{k=1}^{\infty} \mathbb{P}(U_1\notin B,\ldots,U_{k-1}\notin B, U_k=b) \\
\textbf{[indep. } U_1,\ldots,U_k\textbf{]} & = & \displaystyle\sum_{k=1}^{\infty} (1-p)^{k-1}\frac{1}{\#A} \\
\textbf{[un peu de calcul]} & = & \displaystyle\frac{1}{1-(1-p)}\frac{1}{\#A} = \frac{1}{p}\frac{1}{\#A} = \frac{1}{\#B}
\end{array}$$

Et donc, V est bien uniforme à valeurs dans B.

3.

```
import numpy as np
import matplotlib.pyplot as plt
EntierUniforme = np.random.randint

def CoupleEntiersUniforme(N):
    return EntierUniforme(N)+1,EntierUniforme(N)+1

def XY(N):
    while True :
        x,y  = CoupleEntiersUniforme(N)
        if np.gcd(x,y) == 1:
            return x,y
```

4.

```
#Evaluation du nombre moyen de tirages
def K(N):
    cpt = 1
    while True :
        x,y  = CoupleEntiersUniforme(N)
        if np.gcd(x,y) == 1:
            return cpt
        cpt += 1
#l'espérance de K, pour N grand c'est ~ pi**2/6 : merci wikipedia !!
def EK(N, NS = 10000) :
    moy = 0
    for _ in range(NS) :
       moy += K(N)
    return moy / NS

N = 100
print("N = {}, EK(N) = {}, pi^2/6 = {}".format(N,EK(N),np.pi**2/6))
```

Correction Ex.–2 On notera $\mathbb{R}$ l'ensemble des nombres réels, $\mathbb{N}$ l'ensemble des entiers naturels et $\mathbb{N}^*$ l'ensemble des entiers naturels non nuls.

Dans tout le problème, p désignera un nombre réel de l'intervalle $]0,1[$ et $q = 1-p$.

La notation $\lfloor x \rfloor$ désigne la partie entière du nombre réel x.

Toutes les variables aléatoires (v.a.) de ce problème sont définies sur un espace probabilisé $(\Omega, \mathscr{T}, \mathbb{P})$. Pour $X : \Omega \to \mathbb{R}$, une v.a. à valeurs réelles, on note $\mathbb{E}(X)$, resp. $\mathbb{V}(X)$, l'espérance, resp. la variance, de X lorsque celle-ci existe.

Le but de ce problème est de donner deux modélisations, *in fine* équivalentes, du scintillement obtenu lors d'une autoradiographie β.

Une substance subissant une désintégration β, par exemple du cobalt 60 radioactif se désintégrant en du nickel 60 stable suivant

$$ {}^{60}_{27}Co \to \quad {}^{60}_{27}Ni + e^- + \bar{\nu}_e, $$

émet à chaque désintégration un électron, électron qui frappe notre détecteur à un instant donné (scintillement).

On enregistre, les instants de scintillement en nanosecondes à partir de l'instant initial de la mesure et ces mesures constituent notre ensemble de données à analyser.

Partie A

Préliminaire : Instant du s-ième scintillement

A.1. On pose $S_0 = 0$ et, pour un entier naturel m, $m \in \mathbb{N}^*$, on note

$$ S_m = \sum_{k=1}^{m} X_k. $$

S_m est une somme de m v.a. de Bernoulli *indépendantes* de même paramètre p, S_m suit donc une loi binomiale de paramètres m et p. On a

$$ \mathbb{E}(S_m) = m.p, \mathbb{V}(S_m) = m.p.(1-p) $$

A.2. Pour $m \in \mathbb{N}^*$, $s \in \mathbb{N}^*$, on définit l'événement $E_{s,m} = \{S_m < s\}$ et, pour $s \in \mathbb{N}^*$, l'événement $E_s = \bigcap_{m=1}^{\infty} \{S_m < s\}$.

A.2.a. L'évenement E_1 se produit si et seulement si tous les événements $E_{1,m}$ se sont produits, *i.e.* si et seulement si

$$\forall m \in \mathbb{N}^*, S_m < 1$$

Lorsque l'un des X_k vaut 1, on a $S_k \geq 1$ et donc l'évenement E_1 se produit si et seulement si tous les X_K valent 0, *i.e.*

$$E_1 = \{\forall k \in \mathbb{N}^*, X_k = 0\}$$

Si $s \in \mathbb{N}^*$, l'évenement E_s se produit si et seulement si tous les événements $E_{s,m}$ se sont produits, *i.e.* si et seulement si

$$\forall m \in \mathbb{N}^*, S_m < s$$

Lorsque au moins s des X_k valent 1, on a $S_m \geq s$ pour m un entier supérieurs aux indices k concernés (et réciproquement) et donc l'évenement E_s se produit si et seulement si au plus $s-1$ des X_K valent 1, *i.e.*

$$E_s = \{\exists K = \{k_1, \ldots, k_{s-1}\}, \forall k \in \mathbb{N}^* \setminus K, X_k = 0\}$$

Il est clair que pour $m \in \mathbb{N}^*$, $s \in \mathbb{N}^*$,

$$\{S_m < s\} \subset \{S_m < s+1\}$$

En prenant, pour $s \in \mathbb{N}^*$ fixé, l'intersection sur touttes les valeurs de $m \in \mathbb{N}^*$, on a donc $E_s \subset E_{s+1}$.

A.2.b. L'inégalité de BIENAYMÉ–TCHEBYCHEFF s'énonce pour X une v.a. réelle admettant une variance σ^2, une espérance μ et $\varepsilon > 0$ comme

$$\mathbb{P}(|X-\mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Soit $m > \frac{s}{p}$ et appliquons cette inégalité à S_m avec $\mu = m.p$, $\sigma^2 = m.p.(1-p)$ et $\varepsilon = mp - s > 0$. On obtient alors

$$\mathbb{P}(|S_m - mp| > mp - s) \leq \frac{m.p.(1-p)}{(m.p-s)^2}$$

Si $S_m < s$ alors $S_m - mp < s - mp < 0$ et donc $|S_m - mp| > mp - s$, ce qui signifie que $E_{m,s} \subset \{|S_m - mp| > mp - s\}$. Par croissance de $\mathbb{P}$, il vient donc

$$\mathbb{P}(E_{s,m}) \leq \frac{m.p.(1-p)}{(m.p-s)^2}.$$

A.2.c. On a, par définition de $E_s = \bigcap_{m=1}^{+\infty} E_{s,m}$ que $\forall m \in \mathbb{N}^*, E_s \subset E_{s,m}$ De la croissance de $\mathbb{P}$, on tire donc que

$$\forall m \in \mathbb{N}^*, 0 \leq \mathbb{P}(E_s) \leq \mathbb{P}(E_{s,m}) \leq \frac{m.p.(1-p)}{(m.p-s)^2}$$

Comme le terme de droite de cette inégalité tend vers 0 lorsque $m \to +\infty$, par le théorème des gendarmes, on a alors $\mathbb{P}(E_s) = 0$.

A.3. On note, pour un entier naturel s, $s \in \mathbb{N}^*$,

$$U_s = \begin{cases} \min\{m \in \mathbb{N}^*, S_m \geq s\} & \text{sur } \overline{E_s} \\ 0 & \text{sur } E_s \end{cases}$$

On admet que les U_s ainsi définies sont des variables aléatoires. **A.3.a.** U_1 est à valeurs dans $\mathbb{N}$. On a $\{U_1 = 0\} = E_1$ et donc $\mathbb{P}(U_1 = 0) = 0$. Par complémentarité, $\mathbb{P}(U_1 \geq 1) = 1$

A.3.b. Par définition du minimum, lorsque $U_1 \neq 0$, (et donc sur $\overline{E_1}$) on a,

— pour $m < U_1$, $S_m < 1$, *i.e.* $S_m = 0$ et donc, notamment pour $m = U_1 - 1$, $S_{U_1 - 1} = 0$

— pour $m \geq U_1$, $S_m \geq 1$ (remarquer que la suite (S_m), sur un tirage donné, est croissante), et donc, notamment pour $m = U_1$, $S_{U_1} \geq 1$.

Sachant que $S_{U_1} - S_{U_1-1} = X_{U_1}$ et vaut donc 0 ou 1, la seule possibilité est

$$X_{U_1} = 1, S_{U_1} = 1 \text{ et } S_{U_1-1} = 0$$

A.3.c. Lorsque l'événement $\{U_1 > n\}$ a lieu, c'est que

1. U_1 est défini par la formule de minimum,
2. n étant inférieur à ce minimum, on a $S_n < 1$, *i.e.* $S_n = 0$

Par ailleurs, si E_1 a lieu alors $S_n = 0$.

Réciproquement, si l'événement $S_n = 0$ a lieu alors soit E_1 a lieu, soit $n < U_1$ lorsque E_1 n'a pas lieu et U_1 est défini par la formule de minimum. Au final, on a la double inclusion et

$$\{U_1 > n\} \cup E_1 = \{S_n = 0\}$$

A.3.d. Comme $\mathbb{P}(E_1) = 0$ alors, pour tout $n \in \mathbb{N}$, on a donc

$$\mathbb{P}(U_1 > n) = \mathbb{P}(S_n = 0) = q^n$$

car $S_n \sim \mathscr{B}(n, p)$ et $q = 1 - p$, on a donc, pour $n \in \mathbb{N}^*$, que

$$\mathbb{P}(U_1 = n) = \mathbb{P}(U_1 > n-1) - \mathbb{P}(U_1 > n) = q^{n-1} - q^n = p.q^{n-1}$$

La loi de U_1 est donc la loi géométrique sur $\mathbb{N}^*$ et on a

$$\mathbb{E}(U_1) = \frac{1}{p}, \mathbb{V}(U_1) = \frac{q}{p^2}$$

A.3.e. Soit $s \in \mathbb{N}^*$. Sur le même modèle que précédemment en A.3.a, $\mathbb{P}(U_s = 0) = \mathbb{P}(E_s) = 0$ et $\mathbb{P}(U_s \geq 1) = 1$.

A.3.f. De même qu'en A.3.b, sur $\overline{E_s}$, $X_{U_s} = 1$, $S_{U_s} = s$ et $S_{U_s-1} = s-1$ car « au passage de l'indice U_s, S_n passe d'une valeur $< s$ à une valeur $\geq s$ en augmentant au maximum de 1 », cette remarque force les valeurs indiquées.

A.3.g. De même qu'en A.3.c, l'égalité d'événements

$$\{U_s > n\} \cup E_s = \{S_n < s\}$$

pour $s \in \mathbb{N}^*$, $n \in \mathbb{N}^*$ relève de l'alternative exclusive entre le fait que E_s ait lieu et la définition de U_s par minimalité.

A.3.h. On a donc $\mathbb{P}(U_s > s-1) = \mathbb{P}(S_{s-1} < s)$. Comme S_{s-1} suit une loi binomiale de paramètres $s-1$ et p, elle est à valeurs dans $\{0, \ldots, s-1\}$ et donc

$$\mathbb{P}(U_s > s-1) = \mathbb{P}(S_{s-1} < s) = 1$$

U_s est donc à valeurs dans $\{s, s+1, \ldots,\}$. **A.3.i.** Pour $n \in \{s, s+1, \ldots,\}$, on a donc

$$\mathbb{P}(U_s = n) = \mathbb{P}(U_s > n-1) - \mathbb{P}(U_s > n) = \mathbb{P}(S_{n-1} < s) - \mathbb{P}(S_n < s) = \sum_{k=0}^{s-1} \binom{n-1}{k} p^k.q^{n-1-k} - \sum_{k=0}^{s-1} \binom{n}{k} p^k.q^{n-k}$$

En utilisant la formule du triangle de Pascal, valable pour $1 \leq k \leq n-1$,

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

on a alors

$$
\begin{aligned}
\mathbb{P}(U_s=n) &= \sum_{k=0}^{s-1}\binom{n-1}{k}p^k.q^{n-1-k}-\sum_{k=0}^{s-1}\binom{n}{k}p^k.q^{n-k}\\
\text{(les termes } k=0 \text{ sont sortis)} &= \sum_{k=1}^{s-1}\binom{n-1}{k}p^k.q^{n-1-k}-\sum_{k=1}^{s-1}\binom{n}{k}p^k.q^{n-k}+q^{n-1}-q^n\\
\text{(formule triangle)} &= \sum_{k=1}^{s-1}\binom{n-1}{k}p^k.q^{n-1-k}-\sum_{k=1}^{s-1}\left(\binom{n-1}{k-1}+\binom{n-1}{k}\right)p^k.q^{n-k}+q^{n-1}p\\
\text{(regroupement et décalage indice)} &= \sum_{k=1}^{s-1}\binom{n-1}{k}(1-q).p^k.q^{n-1-k}-\sum_{k=0}^{s-2}\binom{n-1}{k}p.p^k.q^{n-1-k}+q^{n-1}p\\
\text{(termes centraux s'annulent)} &= \binom{n-1}{s-1}p^s.q^{n-s}-p.q^{n-1}+q^{n-1}p\\
&= \binom{n-1}{s-1}p^s.q^{n-s}
\end{aligned}
$$

Partie B

Première implémentation informatique

Listing 1 – python/simul1.py

```python
#Simulations désintégration, méthode 1
import numpy as np
def ExperienceBernoulli(K=100,p=0.5): #QB.1
    """
    retourne une liste de K valeurs, un tirage de la famille de
    variables (X_k)_{1\leq k \leq K} indépendantes, Bernoulli(p)
    """
    X=[]
    for k in range(K):
        Xk=0
        if np.random.rand()<p:
            Xk=1
        X.append(Xk)
    return X
def SommesPartielles(X):#QB.2
    """X liste de 0/1: retourne la liste des sommes partielles"""
    S=[X[0]]
    for k in range(len(X)-1):
        S.append(S[-1]+X[k+1])
    return S
def U(X,s=1):#QB.3
    """X liste de valeurs 0 ou 1 retourne la valeur de $U_s$ """
    S=SommesPartielles(X)
    for U in range(len(S)):
        if S[U] > s:
            return U
    return len(S)+1 #Cas où on dépasse, on retourne 1 de plus
```

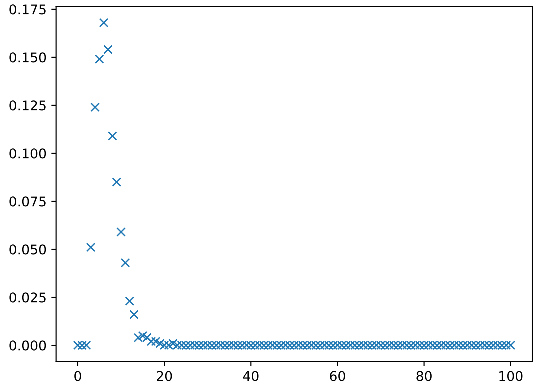

FIGURE 1 – Résultat graphique de la simulation de loi pour U_3. (Non demandé par l'énoncé.)

```
def SimulLoiU(s=1,p=0.5,NS=1000,Umax=100):#QB.4
    """ retourne une estimation de la loi de U_s avec proba. succ. p,
    basée sur NS simulations
    Umax est la valeur maximale imposée ie si retourne Umax, c'est que en fait U_s
    """
    LoiU=np.zeros((Umax+1,)) #Un vecteur de zeros
    for simul in range(NS):
        X=ExperienceBernoulli(K=Umax,p=p)
        LoiU[U(X,s=s)]+=1
    return LoiU/NS
####Tests
import matplotlib.pyplot as plt
LoiU=SimulLoiU(s=3,p=0.5,NS=1000,Umax=100)
plt.plot(LoiU,'x')
plt.savefig('LoiU3.pdf',format='pdf')
plt.show()
```

B.1. *cf.* Script.
B.2. *cf.* Script.
B.3. *cf.* Script.
B.4. Il suffit, en faisant un grand nombre de simulations de U_s utilisant la fonction construite en B.3, de calculer, sur cette liste de nombres, la fréquence d'apparition de chaque nombre. On construit donc une liste, indicée par les valeurs possibles prises par la variable U_s et contenant, pour chaque indice, sa fréquence d'apparition dans la liste.

Ce n'était pas demandé, mais pour cette correction, on a implémenté cette méthode dans le script.

Partie C
Un développement en série

On cherche à démontrer la formule suivante–dite du binôme de NEWTON négatif–

$$\forall n \in \mathbb{N}^*, \forall x \in \left]-1,+1\right[, \ (1-x)^{-n} = \sum_{k=0}^{+\infty} \binom{n-1+k}{k} . x^k$$

C.1.a. Pour $n = 1$, cette formule se lit :

$$\forall x \in \,]-1,+1[\,,\ \frac{1}{1-x} = \sum_{k=0}^{+\infty} \binom{k}{k}.x^k = \sum_{k=0}^{+\infty} x^k$$

Il s'agit de la formule usuelle concernant la somme de séries géométriques convergentes.
C.1.b. Pour $n = 2$, cette formule se lit :

$$\forall x \in \,]-1,+1[\,,\ \frac{1}{(1-x)^2} = \sum_{k=0}^{+\infty} \binom{1+k}{k}.x^k = \sum_{k=0}^{+\infty} (k+1).x^k$$

Il s'agit, au décalage d'indice près, de la formule usuelle concernant la somme de la série géométrique dérivée.
C.2. Soit $n \in \mathbb{N}^*$, $x \in \,]-1,+1[$, $x \neq 0$. Pour $k \in \mathbb{N}$, on pose

$$u_k = \binom{n-1+k}{k}.|x|^k$$

C.2.a. On a, pour $k \in \mathbb{N}$,

$$\frac{u_{k+1}}{u_k} = \frac{(n-1+k+1)!(n-1)!}{(k+1)!}.\frac{k!}{(n-1+k)!(n-1)!}|x| = \frac{n+k}{k+1}|x| \overset{k\to+\infty}{\to} |x|$$

Choissons q tel que $0 < |x| < q$ et alors, par définitiopn de la limite, il existe $k_0 \in \mathbb{N}$ tel que

$$\forall k \geq k_0,\ 0 \leq \frac{u_{k+1}}{u_k} \leq q$$

i.e.

$$\forall k \geq k_0,\ 0 \leq u_{k+1} \leq q.u_k$$

C.2.b. On peut alors démontrer (en rédigeant la récurrence sur k correctement et complètement. On notera notamment que l'initialisation s'effectue à k_0) que

$$\forall k \geq k_0,\ 0 \leq u_k \leq q^{k-k_0}.u_{k_0}$$

C.2.c. La série $\sum_{k\geq k_0} q^{k-k_0}.u_{k_0}$ est convergente car géométrique de raison q, $0 < q < 1$ et donc, par le théorème de comparisaon pour les séries à termes positifs, $\sum_{k\geq k_0} u_k$ est convergente, par CHASLES, $\sum_{k\geq 0} u_k$ est convergente et enfin par convergence absolue, $\sum_{k=0}^{+\infty} \binom{n-1+k}{k}.x^k$ est convergente.
C.3.a. Soit $n \in \mathbb{N}^*$, $x \in \,]-1,+1[$. Posons

$$H_N : (1-x)^{-n} = \sum_{k=0}^{N} \binom{n-1+k}{k}.x^k + x^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x)$$

où

$$R_N(x) = (N+1).\int_0^1 (1-t)^N.(1-t.x)^{-N-n-1}\,dt$$

— Pour $N = 0$, cette formule se lit

$$H_0 : (1-x)^{-n} = \sum_{k=0}^{0} \binom{n-1+k}{k}.x^k + x^1.\binom{n-1+1}{1}.R_0(x) = 1 + n.x.R_0(x)$$

où

$$R_0(x) = \int_0^1 \frac{1}{(1-t.x)^{-n-1}}\,dt$$

On a, par intégration directe, $n.x.R_0(x) = [(1-t.x)^{-n}]_{t=0}^{t=1} = (1-x)^{-n} - 1$, ce qui conclut positivement l'initialisation.

— Supposons H_N vraie pour un certain entier naturel N. On a, en posant, pour $t \in [0,1]$,

$$u(t) = -(1-t)^{N+1},\, v(t) = (1-t.x)^{-N-n-1}$$

que u et v sont de classe $\mathscr{C}^1$ sur $[0,1]$ avec

$$\forall t \in [0,1]\,,\, u'(t) = (N+1).(1-t)^{N+1},\, v'(t) = +x.(N+n+1).(1-t.x)^{-N-n-2}$$

EN intégrant par parties, on obtient alors

$$\begin{aligned} R_N(x) &= \int_0^1 \underbrace{(N+1).(1-t)^N}_{u'(t)}.\underbrace{(1-t.x)^{-N-n-1}}_{v(t)}\,dt \\ &= \left[-(1-t)^{N+1}.(1-t.x)^{-N-n-1}\right]_0^1 + x.(N+n+1)\int_0^1 (1-t)^{N+1}.(1-t.x)^{-N-n-2}\,dt \\ &= 1 + x.\frac{N+n+1}{N+2}R_{N+1}(x) \end{aligned}$$

Comme (H_N) est supposée vraie, on a alors, en utilisant cette formule, que

$$\begin{aligned} (1-x)^{-n} &= \sum_{k=0}^{N}\binom{n-1+k}{k}.x^k + x^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x) \\ &= \sum_{k=0}^{N}\binom{n-1+k}{k}.x^k + x^{N+1}.\binom{n-1+N+1}{N+1} + x^{N+2}.\binom{n-1+N+1}{N+1}\frac{N+n+1}{N+2}R_{N+1}(x) \\ &= \sum_{k=0}^{N}\binom{n-1+k}{k}.x^k + x^{N+1}.\binom{n-1+N+1}{N+1} + x^{N+2}.\binom{n-1+N+2}{N+2}R_{N+1}(x) \end{aligned}$$

est donc (H_{N+1}) est vérifiée.

C.3.b. Soit $t \in [0,1]$. Comme $x \in]-1,+1[$, $t.x \in]-t,+t[$ et $1-t.x \in]1-t,1+t[$. Comme $1-t \geq 0$ et $u \mapsto u^N$ est (strictement) croissante sur $\mathbb{R}^+$, on a

$$0 \leq (1-t)^N < (1-t.x)^N$$

et, donc en divisant par $(1-t.x)^N > 0$, il vient

$$0 \leq (1-t)^N.\frac{1}{(1-t.x)^N} \leq 1$$

Par croissance de l'intégration, vu que

$$R_N(x) = (N+1).\int_0^1 \underbrace{(1-t)^N.(1-t.x)^{-N}}_{0\leq . \leq 1}.(1-t.x)^{-n-1}\,dt$$

On a alors (pour $x \neq 0$, pour $x = 0$, $R_N(0) = N$)

$$0 \leq R_N(x) \leq (N+1).\int_0^1 (1-t.x)^{-n-1}\,dt = \frac{N+1}{x.n}\left(1-(1-x)^{-n}\right)$$

et donc, comme $\frac{N+1}{N} \leq 2$,

$$0 \leq \frac{R_N(x)}{N} \leq \frac{2}{x.n}\left(1-(1-x)^{-n}\right)$$

N'oublions pas que l'on travaille à x et n fixés et cela montre que la suite $\left(\frac{R_N(x)}{N}\right)_{N\in\mathbb{N}^*}$ est bornée.

C.3.c. Lorsque $N \to +\infty$,

$$\binom{n-1+N+1}{N+1} = \frac{(N+n).(N+n-1).....(N+1)}{(N+1)!} = \frac{O(N^n)}{(N+1)!} \to 0$$

et $x^{N+1}.O(N) \to 0$. Au bout du compte,

$$x^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x) \overset{N\to+\infty}{\to} 0$$

(à réécrire sans les notations $O()$, ,juste avec des suites bornées.) **C.3.d.** On a donc, pour $n \in \mathbb{N}^*$ et $x \in]-1,+1[$ que

$$(1-x)^{-n} - \sum_{k=0}^{N} \binom{n-1+k}{k}.x^k \overset{N\to+\infty}{\to} 0$$

ce qui signifie que

$$(1-x)^{-n} = \sum_{k=0}^{+\infty} \binom{n-1+k}{k}.x^k$$

C.4. Développements limités. **C.4.a.** Soit $N \in \mathbb{N}$ et $n \in \mathbb{N}^*$. En reprenant la formule de C.3.a et la majoration obtenue en C.3.b, on a

$$\forall x \in]-1,+1[, x \neq 0, (1-x)^{-n} - \sum_{k=0}^{N} \binom{n-1+k}{k}.x^k = x^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x)$$

Or , $\forall x \in]-1,+1[, x \neq 0$,

$$0 \leq |x|^{N+1}.\binom{n-1+N+1}{N+1}.R_N(x) \leq |x|^{N+1}.\binom{n-1+N+1}{N+1}.\frac{N+1}{n}\frac{(1-(1-x)^{-n})}{x} \leq C_N.|x|^N\left|1-(1-x)^{-n}\right|$$

Cette dernière quantité est $o\left(x^N\right)$ lorsque $x \to 0$ et donc

$$\forall x \in]-1,+1[, (1-x)^{-n} = \sum_{k=0}^{N} \binom{n-1+k}{k}.x^k + o\left(x^N\right)$$

Ce qui est le développement limité à l'ordre N recherché.

Remarquons qu'alternativement, on aurait pu appliquer directement la formule de TAYLOR–YOUNG en calculant les dérivées successives de $x \mapsto (1-x)^{-n}$, de classe $\mathscr{C}^\infty$ sur $]-1,+1[$.

C.4.b. Soit $\gamma = (\gamma_k)_{k\in\mathbb{N}}$ une suite de nombres réels compris entre 0 et 1.

C.4.b.i. Soit $x \in]-1,+1[$. La série $\sum_{k\geq 0} \gamma_k.x^k$ est absolument convergente (et donc convergente) car

$$\forall k \in \mathbb{N}, |\gamma_k.x^k| \leq |x|^k$$

et $\sum_{k\geq 0} |x|^k$ est une série géométrique convergente, on peut donc appliquer le théorème de comparaison des séries à termes positifs.

On définit alors la fonction $g :]-1,+1[\to \mathbb{R}$ par la formule

$$g(x) = \sum_{k=0}^{+\infty} \gamma_k.x^k$$

C.4.b.ii. Lorsque $|x| \leq \frac{1}{2}$, par inégalité triangulaire (pour les séries),

$$|g(x)| \leq \sum_{k=0}^{+\infty} |x^k| \leq \sum_{k=0}^{+\infty} 2^{-k} = 2$$

La fonction g est donc bornée, en v. abs. par 2 sur $\left[-\frac{1}{2},+\frac{1}{2}\right]$
C.4.b.iii. Soit $N\in\mathbb{N}$. Par CHASLES et un décalage d'indice, on a

$$g(x)=\sum_{k=0}^{N}\gamma_k.x^k+x^{N+1}\sum_{k=0}^{+\infty}\gamma_{k+N+1}.x^k$$

En appliquant la question précédente à $\sum_{k=0}^{+\infty}\gamma_{k+N+1}.x^k$, on obtient que

$$\forall x\in\left[-\frac{1}{2},\frac{1}{2}\right],\left|\sum_{k=0}^{+\infty}\gamma_{k+N+1}.x^k\right|\leq 2$$

et donc $x^{N+1}\sum_{k=0}^{+\infty}\gamma_{k+N+1}.x^k$ est $O\left(x^{N+1}\right)$ et donc $o\left(x^N\right)$ lorsque $x\to 0$.

Le développement limité à l'ordre N de $g(x)$ lorsque $x\to 0$ est donc

$$g(x)=\sum_{k=0}^{N}\gamma_k.x^k+o\left(x^N\right)$$

Partie D
Loi binomiale négative.

Dans cette modélisation de notre problème de mesure de l'activité de désintégration β, on mesure Δ_k, la durée en nanosecondes entre le $(k-1)$-ème et le k-ième scintillement. On constate statistiquement que ces durées sont indépendantes et ont l'air de suivre une loi géométrique sur $\mathbb{N}^*$ de paramètre $p\in]0,1[$, on pose $q=1-p$.
D.1. Soit Δ_1 une v.a. suivant la loi géométrique sur $\mathbb{N}^*$ de paramètre p. Δ_1 est à valeurs dans $\mathbb{N}^*$ et on a

$$\forall n\in\mathbb{N}^*,\mathbb{P}(\Delta_1=n)=p.q^{n-1},\mathbb{E}(\Delta_1)=\frac{1}{p},\mathbb{V}(\Delta_1)=\frac{q}{p^2}$$

D.2. Soient X et Y deux v.a. *indépendantes* prenant leurs valeurs dans $\mathbb{N}$ et $Z=X+Y$. **D.2.a.** La v.a. Z est à valerus dans $\mathbb{N}$ car une somme de deux entiers naturels est un entier naturel. **D.2.b.** Montrons uniquement la première formule, la deuxième s'en déduit par symétrie entre X et Y ou par changement d'indice $j=k-i$.

Soit $k\in\mathbb{N}$, la famille d'événements $((X=i),i\in\mathbb{N})$ est un s.c.e.i car X est à valeurs dans $\mathbb{N}$. On a donc, en démarrant par la formule des probabilités totales

$$\begin{aligned}\mathbb{P}(Z=k)&=\sum_{i=0}^{+\infty}\mathbb{P}(Z=k\text{ et }X=i)\\&=\sum_{i=0}^{+\infty}\mathbb{P}(Y=k-i\text{ et }X=i)\\&=\sum_{i=0}^{+\infty}\mathbb{P}(Y=k-i)\mathbb{P}(X=i)\end{aligned}$$

La deuxième ligne est obtenue en constatant l'égalité d'événements

$$(Z=k\text{ et }X=i)=(X+Y=k\text{ et }X=i)=(Y=k-i\text{ et }X=i)$$

La troisième ligne est obtenue en utilisant l'indépendance de X et Y.

On constate pour finir que comme Y est à valeurs dans $\mathbb{N}^*$, lorsque $i>k$, on a $\mathbb{P}(Y=k-i)=0$. Les termes de la série pour $i>k$ sont donc tous nuls et il reste

$$\mathbb{P}(Z=k)=\sum_{i=0}^{k}\mathbb{P}(X=i)\mathbb{P}(Y=k-i)$$

On se donne une suite $(\Delta_n)_{n\in\mathbb{N}^*}$ de v.a. *indépendantes*, suivant la loi géométrique sur $\mathbb{N}^*$ de paramètre p.
On pose $q=1-p$.

D.3. On définit $U_1 = \Delta_1$ et $U_2 = \Delta_1 + \Delta_2 = U_1 + \Delta_2$. **D.3.a.** La v.a. U_2 est naturellement à valeurs dans $\{2,3,\ldots,\ldots,\}$ car une somme de deux entiers naturels ≥ 1 est un entier naturel ≥ 2. **D.3.b.** Par linéarité de l'espérance

$$\mathbb{E}(U_2) = \mathbb{E}(\Delta_1) + \mathbb{E}(\Delta_2) = \frac{2}{p}$$

Par la règle donnant la variance d'une somme de v.a. indépendantes (non corrélées),

$$\mathbb{V}(U_2) = \mathbb{V}(\Delta_1) + \mathbb{V}(\Delta_2) = \frac{2q}{p^2}$$

D.3.c. SOit $k \in \mathbb{N}$, appliquant la règle sur la loi de la somme de v.a. *indépendantes* démontrée en D.2, on a

$$\begin{array}{rcl}
\mathbb{P}(U_2 = k+2) & = & \displaystyle\sum_{i=0}^{k+2} \mathbb{P}(\Delta_1 = i)\mathbb{P}(\Delta_2 = k+2-i) \\
(\mathbb{P}(\Delta_1 = 0) = \mathbb{P}(\Delta_2 = 0) = 0, \text{ lois}) & = & \displaystyle\sum_{i=1}^{k+1} q^{i-1} p.p.q^{k+2-i-1} \\
 & = & (k+1)p^2.q^k
\end{array}$$

D.4. On pose $U_3 = \Delta_1 + \Delta_2 + \Delta_3 = U_2 + \Delta_3$. **D.4.a.** La v.a. U_3 est naturellement à valeurs dans $\{3,4,\ldots,\ldots,\}$ car une somme de trois entiers naturels ≥ 1 est un entier naturel ≥ 3. **D.4.b.** Par linéarité de l'espérance

$$\mathbb{E}(U_3) = \mathbb{E}(\Delta_1) + \mathbb{E}(\Delta_2) + \mathbb{E}(\Delta_3) = \frac{3}{p}$$

Par la règle donnant la variance d'une somme de v.a. indépendantes (non corrélées),

$$\mathbb{V}(U_3) = \mathbb{V}(\Delta_1) + \mathbb{V}(\Delta_2) + \mathbb{V}(\Delta_2) = \frac{3q}{p^2}$$

D.4.c. Soit $k \in \mathbb{N}$, appliquant la règle sur la loi de la somme de v.a. *indépendantes* démontrée en D.2, sahcant que Δ_2 et U_3 sont indépendantes (coalitions), on a

$$\begin{array}{rcl}
\mathbb{P}(U_3 = k+3) & = & \displaystyle\sum_{i=0}^{k+3} \mathbb{P}(\Delta_2 = i)\mathbb{P}(U_3 = k+3-i) \\
(\mathbb{P}(\Delta_2 < 1) = \mathbb{P}(U_3 = 0) = 0) & = & \displaystyle\sum_{i=2}^{k+2} \mathbb{P}(\Delta_2 = i)\mathbb{P}(U_3 = k+3-i) \\
(\text{Chgt indice } i' = i-2) & = & \displaystyle\sum_{i=0}^{k} \mathbb{P}(\Delta_2 = i+2)\mathbb{P}(U_3 = k+1-i) \\
(\text{Lois}, \displaystyle\sum_{i=0}^{k}(i+1) = \ldots) & = & \displaystyle\sum_{j=0}^{k} p^2.q^i(i+1)p.q^{k-i} = p^3.q^k.\frac{1}{2}.(k+1).(k+2)
\end{array}$$

et donc

$$\forall k \in \mathbb{N}, \mathbb{P}(U_3 = k+3) = p^3.q^k.\frac{1}{2}.(k+1).(k+2)$$

D.5. On pose $U_0 = 0$ et, pour $s \in \mathbb{N}^*$,

$$U_s = \Delta_1 + \cdots + \Delta_s = \sum_{k=1}^{s} \Delta_k = U_{s-1} + \Delta_n$$

D.5.a. La v.a. U_s prend naturellement ses valeurs dans l'ensemble des entiers $\geq s$, car une somme de s entiers naturels ≥ 1 est un entier naturel $\geq s$. **D.5.b.** Comme en D.3.b et D.4.b,

$$\mathbb{E}(U_s) = \frac{s}{p}, \mathbb{V}(U_s) = \frac{s.q}{p^2}$$

D.6.a. Soit $x \in [-1,+1]$ et U une variable aléatoire à valeurs dans $\mathbb{N}$. La v.a. x^U admet une espérance si et seulement si la série $\sum_{u=0}^{+\infty} x^u.\mathbb{P}(U=u)$ est ACV. Sachant que $\forall u \in \mathbb{N}$, $|x|^u \leq 1$, le terme général de cette série est majoré en valeur absolue par $\mathbb{P}(U=u)$ et, comme $\sum_{u=0}^{+infty} \mathbb{P}(U=u) = 1$, par le théorème de comparaison pour les séries à termes positifs, $\sum_{u=0}^{+\infty} x^u.\mathbb{P}(U=u)$ est ACV et

$$\mathbb{E}(x^U) = \sum_{u=0}^{+\infty} x^u.\mathbb{P}(U=u)$$

On définit, pour $s \in \mathbb{N}^*$, $x \in [-1,+1]$,

$$f_s(x) = \mathbb{E}(x^{U_s})$$

D.6.b. On a, pour $x \in [-1,+1]$, que

$$f_1(x) = \sum_{u=0}^{+\infty} x^u.\mathbb{P}(U_1 = u) \sum_{u=1}^{+\infty} x^u.p.q^{u-1} = p.x.\frac{1}{1-q.x}$$

(On a reconnu une somme de termes géométriques de raison $q.x$, $|q.x| < 1$.)
D.6.c. On a, pour $x \in [-1,+1]$, que

$$f_2(x) = \sum_{u=0}^{+\infty} x^u.\mathbb{P}(U_2 = u) \sum_{u=2}^{+\infty} x^u.p^2.(u-1).q^{u-2} = (p.x)^2.\frac{1}{(1-q.x)^2}$$

(On a reconnu, après chgt d'indice $u' = u-1$ une somme géométrique dérivée de raison $q.x$, $|q.x| < 1$.)
D.6.d. Soit $s \in \mathbb{N}^*$, $x \in [-1,+1]$. On a

$$U_s = \Delta_1 + \cdots + \Delta_s, x^{U_s} = x^{\Delta_1}.\ldots.x^{\Delta_s}$$

et donc, après passage à l'espérance et utilisation de l'indépendance des Δ_k,

$$\mathbb{E}(x^{U_s}) = \mathbb{E}(x^{\Delta_1}).\ldots.\mathbb{E}(x^{\Delta_s})$$

Or

$$\mathbb{E}(x^{\Delta_1}) = \cdots = \mathbb{E}(x^{\Delta_s}) = \mathbb{E}(U_1) = f_1(x)$$

et donc

$$\forall x \in [-1,+1], f_s(x) = f_1(x)^s = \frac{(p.x)^s}{(1-q.x)^s}$$

D.6.e. En utilisant le développement en série vu en A.3.d pour $n = s$, on a

$$\forall x \in]-1,+1[, f_s(x) = p^s.x^s.\sum_{k=0}^{+\infty} \binom{s-1+k}{k}.q^k.x^k = \sum_{k=0}^{+\infty} \binom{s-1+k}{k}.p^s.q^k.x^{k+s}$$

D.7. On sait que d'une part, on a, pour tout $x \in]-1,+1[$,

$$f_s(x) = \sum_{\ell=0}^{+\infty} \mathbb{P}(U_s = \ell) x^\ell$$

et d'autre part, que pour tout $x \in]-1,+1[$,

$$f_s(x) = \sum_{k=0}^{+\infty} \binom{s-1+k}{k}.p^s.q^k.x^{k+s}$$

Le DL en 0 à un ordre N ($\geq s$) est donc, à la fois, par les questions C.4.b.iii et C.4.a

$$f_s(x) = \sum_{\ell=0}^{N} \mathbb{P}(U_s = \ell)x^\ell + o\left(x^N\right)$$

et

$$f_s(x) = \sum_{k=0}^{N-s} \binom{s-1+k}{k}.p^s.q^k.x^{k+s} + o\left(x^N\right)$$

Par unicité d'un DL en 0, on en déduit l'égalité des coefficients de même degré et donc, pour $\ell < s$, $P(U_s = \ell) = 0$ et, pour $\ell \geq s$, $\ell = k+s$,

$$\mathbb{P}(U_s = k+s) = p^s.q^k.\binom{s-1+k}{k}$$

Au final, on a bien

$$\forall s \in \mathbb{N}^*, \forall k \in \mathbb{N}, \mathbb{P}(U_s = k+s) = p^s.q^k.\binom{s-1+k}{k}$$

D.8. Chaque variable aléatoire U_s (pour $s \in \mathbb{N}^*$) de cette partie a même loi que la variable aléatoire U_s de la partie A.

D.9. Soit $n \in \mathbb{N}^*$ et, pour $s \in \{0, \dots, n\}$, l'événement

$$F_{n,s} = \{U_{s+1} > n\} \cap \{U_s \leq n\}$$

On définit la variable N_n par la formule

$$N_n = \sum_{s=0}^{n} s.\mathbb{1}_{F_{n,s}}$$

D.9.a.

- Comme, pour $\omega \in \Omega$, $U_{s+1}(\omega) = U_s(\omega) + \Delta_{s+1}(\omega)$ et que $\Delta_{s+1}(\omega) \geq 1$, on a alors

$$\forall \omega \in \Omega, U_1(\omega) < U_2(\omega) < \cdots < U_s(\omega) < \dots$$

- En explicitant la famille ($F_{n,s}$, $s \in \{0, \dots, n\}$), on a que

$$\begin{array}{rcl}
F_{n,0} & = & \text{si } U_1 > n \text{ et } U_0 \leq n, \ \textit{i.e.} \text{ si } U_1 > n \\
F_{n,1} & = & \text{si } U_2 > n \text{ et } U_1 \leq n \\
F_{n,2} & = & \text{si } U_3 > n \text{ et } U_2 \leq n \\
\vdots & & \\
F_{n,n-1} & = & \text{si } U_n > n \text{ et } U_{n-1} \leq n \\
F_{n,n} & = & \text{si } U_{n+1} > n \text{ et } U_n \leq n
\end{array}$$

Il est à noter que l'on a $U_1 < U_2 < \cdots < U_n < U_{n+1}$ et donc on a

$$U_{\ell+1} \leq n \Rightarrow U_\ell \leq n$$

De plus, $U_{n+1} \geq n+1 > n$ et $F_{n,n} = \{U_n \leq n\}$.
Pour $\omega \in \Omega$, $\omega \in F_{n,s}$ si et seulement si les valeurs de $U_1(\omega) < \cdots < U_s(\omega)$ sont toutes inférieures à n alors que les valeurs de $U_{s+1}(\omega) < U_{s+2}(\omega) < \dots$ sont toutes strictement supérieures à n. $(\star)$

Il apparait alors beaucoup clairement que ces événements sont deux à deux incompatibles et « tous les cas sont couverts », *i.e.* il s'agit d'un système complet d'événements.

— En explicitant et simplifiant la formule donnant N_n, dans chacun des cas énumérés, on a aussi que

$$N_n = \begin{cases} 0 & \text{si } U_1 > n \\ 1 & \text{si } U_2 > n \text{ et } U_1 \leq n \\ 2 & \text{si } U_3 > n \text{ et } U_2 \leq n \\ \vdots & \\ n-1 & \text{si } U_n > n \text{ et } U_{n-1} \leq n \\ n & \text{si } U_{n+1} > n \text{ et } U_n \leq n \end{cases}$$

La v.a. N_n compte le nombre de valeurs $U_1(\omega) < \cdots < U_?(\omega)$ qui sont inférieures ou égales à n.

D.9.b. Soit $s \in \mathbb{N}^*$, $n \in \mathbb{N}^*$, par la remarque $(\star)$, $\omega \in \{U_s > n\}$ si et seulement si les valeurs de $U_s(\omega) < U_{s+1}(\omega) < \ldots$ sont toutes strictement supérieures à n, *i.e.* s'il y moins (au sens large) de $s-1$ valeurs de $U_1(\omega) < \cdots < U_?(\omega)$ qui sont inférieures ou égales à n. Comme N_n compte le nombre de telles valeurs, cela signifie exactement que $N_n(\omega) \leq s-1$, *i.e.* $N_n(\omega) < s$.

On a donc, pour tout $s \in \mathbb{N}^*$,

$$\mathbb{P}(N_n < s) = \mathbb{P}(U_s > n)$$

Par A.3.g, comme les variables U_s de cette partie et celles de la partie A ont même loi, il apparait (remarquer que $\mathbb{P}(E_s) = 0$) que N_n à la même a la même fonction de répartition que la variable aléatoire S_n de la partie A et donc la même loi.

Par A.1, N_n est donc une variable binomiale de paramètres n et p.

D.10. On suppose que λ est un nombre réel > 0 fixé, n est très grand et $p = \frac{\lambda}{n}$? Justifier les affirmations suivantes en précisant leur sens : **D.10.a.** L'énoncé correct traduisant « N_n suit approximativement une loi de POISSON de paramètre λ » est

$$\forall k \in \mathbb{N}, \mathbb{P}(N_n = k) \overset{n \to +\infty}{\to} e^{-\lambda} \frac{\lambda^n}{n!}$$

Pour la démonstration, *cf.* cours.

D.10.b. Un énoncé correct traduisant « les variables $\frac{1}{n}\Delta_k$ suivent approximativement une loi exponentielle de paramètre λ » pourrait être (convergence des fonctions de répartition) :

$$\forall x > 0, \mathbb{P}(\frac{1}{n}\Delta_k \leq x) \overset{n \to +\infty}{\to} 1 - e^{-\lambda . x}$$

Soit donc $x > 0$, on a, vu que $p = \frac{\lambda}{n}$,

$$\begin{aligned} \mathbb{P}(\Delta_k \leq n.x) &= \frac{\lambda}{n} \sum_{k=1}^{\lfloor n.x \rfloor} (1 - \frac{\lambda}{n})^{k-1} \\ &= \frac{\lambda}{n} . \frac{1 - (1 - \frac{\lambda}{n})^{\lfloor n.x \rfloor}}{\frac{\lambda}{n}} \\ &= 1 - (1 - \frac{\lambda}{n})^{\lfloor n.x \rfloor} \end{aligned}$$

Il s'agit maintenant de passer à la limite lorsque $n \to +\infty$. Prenons le log de $(1 - \frac{\lambda}{n})^{\lfloor n.x \rfloor}$ et cherchons sa limite.

On a

$$\ln(1 - \frac{\lambda}{n})^{\lfloor n.x \rfloor} = 0 \lfloor n.x \rfloor . \ln(1 - \frac{\lambda}{n})$$

et on a donc l'encadrement

$$n.x.\ln(1-\frac{\lambda}{n}) \leq \lfloor n.x \rfloor.\ln(1-\frac{\lambda}{n}) \leq (n.x+1).\ln(1-\frac{\lambda}{n})$$

Lorsque $n \to +\infty$, $\ln(1-\frac{\lambda}{n}) = -\frac{\lambda}{n} + o\left(\frac{1}{n}\right)$ et $n.x.\ln(1-\frac{\lambda}{n}) \overset{n\to+\infty}{\to} -\lambda.x$ et $\ln(1-\frac{\lambda}{n}) \overset{n\to+\infty}{\to} 0$ Par le théorème des gendarmes, il vient que

$$\lfloor n.x \rfloor.\ln(1-\frac{\lambda}{n}) \overset{n\to+\infty}{\to} -\lambda.x$$

et enfin,

$$\mathbb{P}(\frac{1}{n}\Delta_k \leq x)1-(1-\frac{\lambda}{n})^{\lfloor n.x \rfloor} \overset{n\to+\infty}{\to} 1-e^{-\lambda.x}$$

Partie E

Deuxième implémentation informatique

E.1. Soit $\lambda > 0$ et $D \sim \mathscr{E}(\lambda)$, $\Delta = \lfloor D \rfloor + 1$. D est à valeurs dans $\mathbb{R}^+$, donc $\lfloor D \rfloor$ est à valeurs dans $\mathbb{N}$ et Δ à valeurs dans $\mathbb{N}^*$.

Pour $k \in \mathbb{N}^*$, on a

$$\begin{array}{rcl} \mathbb{P}(\Delta = k) & = & \mathbb{P}(\lfloor D \rfloor + 1 = k) \\ \text{(def. partie entière)} & = & \mathbb{P}(k-1 \leq D < k) \\ \text{(loi de } D\text{)} & = & \displaystyle\int_{k-1}^{k} \lambda.e^{-\lambda.x}\,dx \\ & = & e^{-\lambda.(k-1)} - e^{-\lambda.k} = (1-e^{-\lambda}).\left(e^{-\lambda}\right)^{k-1} \end{array}$$

et donc Δ suit une loi géométrique sur $\mathbb{N}^*$ de paramètre $p = 1-e^{-\lambda}$.

Listing 2 – python/simul2.py

```
#2e set de simulations
import numpy as np
def Delta(p=0.5): #QE.2
    """
    Tirage de loi géométrique sur N* de paramètre p
    """
    lambada = -np.log(1-p)
    return np.floor(-np.log(np.random.rand())/lambada)+1
def U2(s=1,p=0.5): #QE.3.a
    """
    Simulation de U_s par somme de géométriques
    """
    u=0
    for k in range(s):
        u+=Delta(p=p)
    return u
####Tests
import matplotlib.pyplot as plt
NS=10000
U=[U2(s=3,p=0.5) for n in range(NS)]
plt.hist(U,normed=True,bins=np.arange(0,25,1)+0.5)
plt.savefig('LoiU3-2.pdf',format='pdf')
plt.show()
```

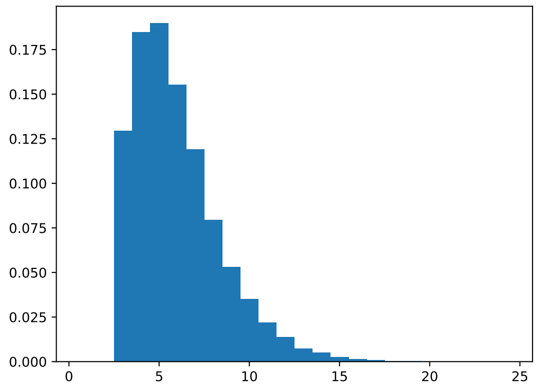

FIGURE 2 – Résultat graphique de la simulation de loi pour U_3 par la deuxième méthode. (Non demandé par l'énoncé.)

E.2. On a dans la question précédente que $p = 1 - e^{-\lambda}$ et donc $\lambda = -\ln(1-p)$. Il suffit donc de simuler (par la méthode de la fonction des quantiles une v.a. exponentielle D de paramètre $\lambda = -\ln(1-p)$ et d'en prendre la partie entière $+1$. *cf.* Script.
E.3.a. *cf.* Script.
E.3.b. Cette méthode est plus rapide que celle de la méthode présentée en partie B car on n'a pas à traiter la grande liste X et plus précise, car *a priori* , la variable simulée n'est pas limitée à une valeur maximale.