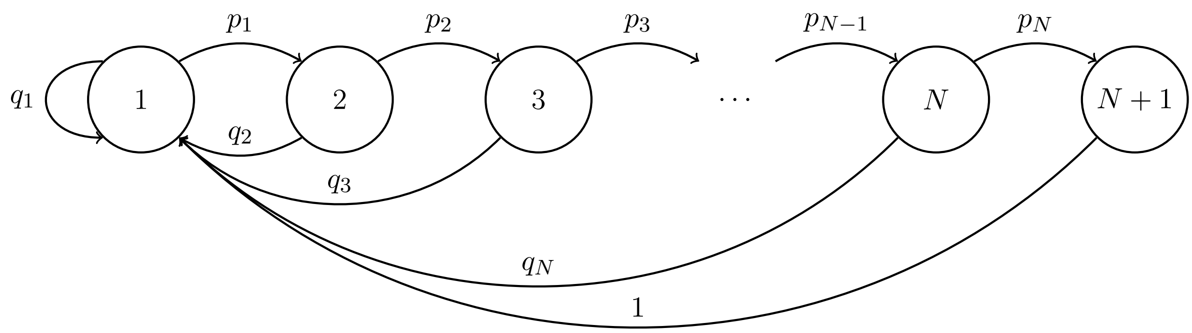
DS8 barème (version B)

Problème 1 (`sujet maison`)

Soit N un entier supérieur ou égal à 1.
On considère un joueur (ou une joueuse) qui joue à un *rogue-like*. Le rogue-like est un sous-genre de jeu vidéo dans lequel le joueur explore un donjon infesté de monstres qu'il doit combattre pour progresser vers la dernière salle, contenant le *boss final* et la promesse de récompenses épiques[1]. Une des caractéristiques du rogue-like est que toute mort est définitive, ce qui oblige le joueur à recommencer du début s'il souhaite continuer à jouer. Nous considérons dans la suite un joueur soumis aux règles suivantes :

- Le joueur commence au *niveau* 1 et a pour objectif d'atteindre le niveau $N+1$ afin d'y affronter le boss final et de le vaincre.
- A chaque niveau, le joueur effectue un unique *combat* qu'il peut soit gagner soit perdre.
  Plus précisément :
  - × pour tout $i \in [\![1, N]\!]$, on note $p_i \in]0,1[$ la probabilité que le joueur gagne le combat effectué au niveau i, et on pose $q_i = 1 - p_i$. On suppose que, pour tout $i \in [\![1, N-1]\!]$, $p_i \geqslant p_{i+1}$.
  - × si le combat mené au niveau $i \in [\![1, N]\!]$ est gagné, le joueur passe au niveau supérieur, sinon, le joueur meurt et retourne au niveau 1.
  - × quelque soit l'issue du combat effectué au niveau $N+1$, le joueur recommence au niveau 1 ensuite.
- On appelle *partie* toute séquence minimale de jeu commençant au niveau 1 et se terminant soit par la mort du joueur, soit par la victoire du joueur contre le boss final.
- Le joueur enchaîne indéfiniment les parties, même en cas de victoire contre le boss final.

On représente cette chaîne de Markov à $N+1$ états par le graphe probabiliste :



On admet que toutes les v.a.r. sont définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. On note :

- pour tout $n \in \mathbb{N}^*$, X_n la variable aléatoire égale au numéro du niveau où se trouve le joueur lors de son n^{e} **combat**. En particulier, $X_1 = 1$.
- pour tout $j \in \mathbb{N}^*$, Y_j la variable aléatoire égale au niveau maximal atteint lors de la j^{e} **partie** jouée.
- Z la variable aléatoire égale au numéro de la **partie** où le joueur combat le boss final pour la 1^{re} fois.
- T la variable aléatoire égale au numéro du premier **combat** effectué contre le boss final.

L'objectif principal du problème est de calculer $\mathbb{E}(T)$.
Les parties **I** et **II** sont indépendantes. La partie **III** utilise les résultats de la partie **II**.

(1). Voir `https://fr.wikipedia.org/wiki/Roguelike`

Question préliminaire. Donner une interprétation de l'hypothèse : pour tout $i \in [\![1, N-1]\!]$, $p_i \geqslant p_{i+1}$.

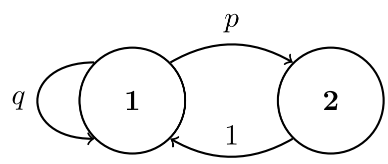
- **1 pt : Cette hypothèse traduit le fait que la difficulté des combats augmente à chaque niveau**

Partie I : le cas $N = 1$

On pose pour simplifier $p = p_1$ et $q = 1 - p$.

1. Dessiner le graphe probabiliste dans ce contexte et avec ces notations.

- **1 pt :**



2. On note, pour tout $n \in \mathbb{N}^*$, $\alpha_n = \mathbb{P}([X_n = 1])$.

a) Montrer que, pour tout $n \in \mathbb{N}^*$, $\alpha_{n+1} = 1 - p\alpha_n$.

- **2 pts : cas $n \geqslant 2$**
  - × **1 pt : On applique la formule des probabilités totales avec le système complet d'événements** $([X_n = 1], [X_n = 2])$
  - × **1 pt : calcul**
- **1 pt : cas $n = 1$**

b) En déduire que, pour tout $n \in \mathbb{N}^*$, $\alpha_n = \dfrac{1-(-p)^n}{1+p}$.

- **1 pt : initialisation**
- **2 pts : hérédité**

3. *a)* Montrer que la suite (X_n) converge en loi vers une v.a.r. X dont on explicitera la loi.

- **1 pt :** $p \in]0,1[$ **donc** $\lim\limits_{n \to +\infty} (-p)^{n+1} = 0$
- **1 pt :** $\mathbb{P}([X_n = 1]) = \alpha_n \underset{n \to +\infty}{\longrightarrow} \dfrac{1}{1+p}$ **et** $\mathbb{P}([X_n = 2]) = 1 - \alpha_n \underset{n \to +\infty}{\longrightarrow} 1 - \dfrac{1}{1+p} = \dfrac{p}{1+p}$
- **1 pt :** (X_n) **converge en loi vers une variable aléatoire** X **de loi :**
  - × $X(\Omega) = \{1, 2\}$
  - × $\mathbb{P}([X = 1]) = \dfrac{1}{1+p}$ **et** $\mathbb{P}([X = 2]) = \dfrac{p}{1+p}$

b) En déduire un état stable de la chaîne de Markov (X_n).

- **1 pt :** $\begin{pmatrix} \dfrac{1}{1+p} & \dfrac{p}{1+p} \end{pmatrix}$ **est un état stable de la chaîne de Markov**

4. *a)* Reconnaître, pour tout $n \in \mathbb{N}^*$, la loi de $X_n - 1$.

- **1 pt :** $X_1 - 1$ **suit la loi certaine égale à** 0
- **1 pt : Pour** $n \geqslant 2$**,** $X_n - 1 \hookrightarrow \mathcal{B}(1 - \alpha_n)$

b) Démontrer : $\forall n \in \mathbb{N}^*$, $(X_n - 1)(X_{n+1} - 1) = 0$.

- **1 pt** ***(pas de point si la disjonction de cas se fait sur*** X_n ***au lieu de*** $X_n(\omega)$***)***

c) En déduire : $\forall n \in \mathbb{N}^*$, $\mathrm{Cov}(X_n, X_{n+1}) = -\dfrac{(p+(-p)^n)(p+(-p)^{n+1})}{(1+p)^2}$.

- **1 pt :** $\mathrm{Cov}(X_n - 1, X_{n+1} - 1) = -(1-\alpha_n)(1-\alpha_{n+1}) = -\dfrac{(p+(-p)^n)(p+(-p)^{n+1})}{(1+p)^2}$

- **1 pt : par bilinéarité de la covariance :**

$$\begin{aligned}\text{Cov}(X_n - 1, X_{n+1} - 1) &= \text{Cov}(X_n, X_{n+1}) - \text{Cov}(1, X_{n+1}) - \text{Cov}(X_n, 1) + \text{Cov}(1, 1) \\ &= \text{Cov}(X_n, X_{n+1})\end{aligned}$$

- **cas $n = 1$ justifié (pendant la preuve ou à part)**

d) Pour quelle(s) valeur(s) de n les variables aléatoires X_n et X_{n+1} sont indépendantes ?

- **1 pt : si $n \geqslant 2$, $\text{Cov}(X_n, X_{n+1}) \neq 0$. En effet, $p \in]0, 1[$ donc**

$$|(-p)^n| = p^n < p$$

On en déduit que les variables X_n et X_{n+1} ne sont pas indépendantes.

- **1 pt : si $n = 1$, X_1 est constante égale à 1 donc X_1 et X_2 sont indépendantes**

5. *a)* Reconnaître la loi de Z.

- **1 pt : L'expérience (du point de vue des parties) consiste en une succession infinie d'épreuves de Bernoulli indépendantes et identiques, dont le succès est « le joueur combat le boss final ». La probabilité de succès est p.**
- **1 pt : La variable aléatoire Z est égale au rang du premier succès, donc $Z \hookrightarrow \mathcal{G}(p)$**

b) Montrer que $T = Z + 1$. En déduire $\mathbb{E}(T)$ et $\mathbb{V}(T)$.

- **1 pt : Soit $\omega \in \Omega$. Le numéro de la première partie où le joueur combat le boss final est $Z(\omega)$. Chaque partie où le joueur ne combat pas le boss final est constituée d'un unique combat (perdu) au niveau 1, tandis qu'une partie où le joueur combat le boss final est constituée de deux parties.**
  On en déduit que

$$T(\omega) = 1 \times (Z(\omega) - 1) + 2 \times 1 = Z(\omega) + 1$$

- **1 pt : Z admet une espérance et une variance donc T également par transformation affine**
- **1 pt :**

$$\mathbb{E}(T) = \mathbb{E}(Z + 1) = \mathbb{E}(Z) + 1 = \frac{1}{p} + 1 = \frac{1 + p}{p} \qquad \textbf{\textit{(par linéarité)}}$$

$$\mathbb{V}(T) = \mathbb{V}(Z + 1) = \mathbb{V}(Z) = \frac{1 - p}{p^2} \qquad \textbf{\textit{(par propriété de la variance)}}$$

Partie II : espérance conditionnelle

Soient X et Y deux variables aléatoires discrètes. On suppose que, pour tout $y \in Y(\Omega)$, $\mathbb{P}([Y = y]) \neq 0$. Pour tout événement A tel que $\mathbb{P}(A) \neq 0$, on définit, sous réserve d'existence, l'*espérance conditionnelle de X sachant A* :

$$\mathbb{E}_A(X) = \sum_{x \in X(\Omega)} x \, \mathbb{P}_A([X = x])$$

6. Soit A un événement tel que $\mathbb{P}(A) \neq 0$.

a) Démontrer que $\mathbb{P}_A$ est une application probabilité.

- **1 pt : $0 \leqslant \mathbb{P}_A(B) \leqslant 1$**
- **1 pt : $\mathbb{P}_A(\Omega) = \dfrac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$**
- **1 pt : $\mathbb{P}_A\left(\bigcup\limits_{n=0}^{+\infty} B_n\right) = \sum\limits_{n=0}^{+\infty} \mathbb{P}_A(B_n)$**

b) En déduire un argument non calculatoire pour justifier le fait que l'opérateur d'espérance conditionnelle sachant A est linéaire.

- **1 pt : La définition de l'espérance conditionnelle sachant A est identique à ceci près que l'on remplace $\mathbb{P}$ par $\mathbb{P}_A$ dans la formule.**
- **1 pt : Ceci veut dire que l'espérance conditionnelle sachant A sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ coïncide avec l'espérance usuelle sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P}_A)$**

7. Démontrer, sous réserve d'existence et en admettant que l'on peut intervertir les sommes, la formule des espérances totales :

$$\mathbb{E}(X) = \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y]) \mathbb{E}_{[Y=y]}(X)$$

- **1 pt : interversion des sommes bien effectuée**
- **1 pt : formule des probabilités totales avec le système complet d'événements $([Y = y])_{y \in Y(\Omega)}$**

Partie III : cas général

On note, pour tout $m \in [\![0, N]\!]$, $u_m = \prod_{i=1}^{m} p_i$. Par convention, $u_0 = 1$.

8. a) Justifier : $Y_1(\Omega) = [\![1, N+1]\!]$.

- **1 pt : Y_1 est la variable aléatoire égale au niveau maximal atteint lors de la première partie jouée. Il y a $N + 1$ niveaux en tout et le joueur a une chance de mourir à n'importe quel niveau**

b) Démontrer que, pour tout $m \in [\![1, N]\!]$, $\mathbb{P}([Y_1 = m]) = u_{m-1} - u_m$.
(on pourra décomposer l'événement $[Y_1 = m]$ à l'aide des variables aléatoires X_n)

- **1 pt :** $[Y_1 = m] = [X_1 = 1] \cap [X_2 = 2] \cap \cdots \cap [X_m = m] \cap [X_{m+1} = 1]$
- **1 pt : d'après la formule des probabilités composées :**

$$\begin{aligned}&\mathbb{P}([Y_1 = m])\\ =&\mathbb{P}([X_1 = 1])\mathbb{P}_{[X_1=1]}([X_2 = 2]) \ldots \mathbb{P}_{[X_1=1]\cap\cdots\cap[X_{m-1}=m-1]}([X_m = m])\mathbb{P}_{[X_1=1]\cap\cdots\cap[X_m=m]}([X_{m+1} = 1])\end{aligned}$$

- **1 pt : (X_n) est une chaîne de Markov**

$$\begin{aligned}&\mathbb{P}([X_1 = 1])\mathbb{P}_{[X_1=1]}([X_2 = 2]) \ldots \mathbb{P}_{[X_1=1]\cap\cdots\cap[X_{m-1}=m-1]}([X_m = m])\mathbb{P}_{[X_1=1]\cap\cdots\cap[X_m=m]}([X_{m+1} = 1])\\ =&1 \times \mathbb{P}_{[X_1=1]}([X_2 = 2]) \ldots \mathbb{P}_{[X_{m-1}=m-1]}([X_m = m])\mathbb{P}_{[X_m=m]}([X_{m+1} = 1])\end{aligned}$$

- **1 pt :** $\mathbb{P}_{[X_m=m]}([X_{m+1} = 1]) = q_m$ **car** $m \leqslant N$
- **1 pt :** $\left(\prod_{i=1}^{m-1} p_i\right) \times q_m = u_{m-1} - u_m$

c) Calculer $\mathbb{P}([Y_1 = N+1])$.

- **1 pt :** $[Y_1 = N+1] = [X_1 = 1] \cap [X_2 = 2] \cap \cdots \cap [X_{N+1} = N+1]$
- **1 pt :** $\mathbb{P}([Y_1 = N+1]) = u_N$

9. a) Expliquer pourquoi les variables aléatoires Y_j suivent toutes la même loi et sont indépendantes.

- **1 pt : Commençons par remarquer que les p_i ne dépendent pas du temps (la chaîne de Markov est homogène). Puisque chaque partie commence au niveau 1, chaque partie suit les même « règles » de passage d'un niveau à un autre et donc les variables aléatoires Y_j suivent toutes la même loi**
- **1 pt : De plus, les résultats de deux combats différents sont toujours indépendants (il s'agit d'une chaîne de Markov, le futur dépend uniquement du présent et pas du passé). Puisque chaque partie concerne des combats différents de toutes les autres parties, il vient que les variables aléatoires Y_j sont mutuellement indépendantes**

b) En déduire que : $Z \hookrightarrow \mathcal{G}(u_N)$. On explicitera le calcul de $\mathbb{P}([Z=k])$ pour tout $k \in Z(\Omega)$.

- **1 pt : $Z(\Omega) = \mathbb{N}^*$**
- **1 pt : $[Z=k] = [Y_1 \leqslant N] \cap [Y_2 \leqslant N] \cap \cdots \cap [Y_{k-1} \leqslant N] \cap [Y_k = N+1]$**
- **1 pt : par indépendance et car les Y_j suivent la même loi**
- **1 pt : $\mathbb{P}([Y_1 \leqslant N]) = \mathbb{P}\left(\overline{[Y_1 = N+1]}\right)$ car $Y_1(\Omega) = [\![1, N+1]\!]$**

10. *a)* Déterminer la loi conditionnelle de Y_1 sachant $[Y_1 \leqslant N]$.

- **1 pt : si l'événement $[Y_1 \leqslant N]$ est réalisé, alors la variable aléatoire Y_1 peut prendre les valeurs $1, \dots, N$ mais pas la valeur $N+1$**
- **1 pt : pour tout $m \in [\![1, N]\!]$, $\mathbb{P}_{[Y_1 \leqslant N]}([Y_1 = m]) = \dfrac{u_{m-1} - u_m}{1 - u_N}$**
- **1 pt : $[Y_1 = m] \subset [Y_1 \leqslant N]$**

b) Soient $1 \leqslant j < k$ et soit $m \in Y_j(\Omega)$. Montrer que

$$\mathbb{P}_{[Z=k]}([Y_j = m]) = \mathbb{P}_{[Y_j \leqslant N]}([Y_j = m]) = \mathbb{P}_{[Y_1 \leqslant N]}([Y_1 = m])$$

- **1 pt : $\mathbb{P}_{[Y_j \leqslant N]}([Y_j = m]) = \mathbb{P}_{[Y_1 \leqslant N]}([Y_1 = m])$**
- **2 pts : $\mathbb{P}_{[Z=k]}([Y_j = m]) = \mathbb{P}_{[Y_j \leqslant N]}([Y_j = m])$**
- **1 pt : par indépendance**

c) Soient $1 \leqslant j < k$. Montrer que

$$\mathbb{E}_{[Z=k]}(Y_j) = \frac{1}{1 - u_N} \left(\sum_{m=0}^{N} u_m - (N+1)u_N \right)$$

- **1 pt : Y_j étant une variable aléatoire finie (ses valeurs possibles sont $1, \dots, N+1$), elle admet une espérance conditionnelle sachant $[Z=k]$**
- **1 pt : $\mathbb{E}_{[Z=k]}(Y_j) = \displaystyle\sum_{m=1}^{N} m \frac{u_{m-1} - u_m}{1 - u_N}$**
- **1 pt : $\mathbb{E}_{[Z=k]}(Y_j) = \dfrac{1}{1 - u_N} \left(\displaystyle\sum_{m=0}^{N} u_m - (N+1)u_N \right)$**
- **1 pt : justifications du calcul**

11. On note, pour tout $k \in \mathbb{N}$, $S_k = \displaystyle\sum_{j=1}^{k} Y_j$ et on pose $W = S_{Z-1}$. Autrement dit, W est la variable aléatoire définie par :

$$\forall \omega \in \Omega,\ W(\omega) = S_{Z(\omega)-1}(\omega) = \sum_{j=1}^{Z(\omega)-1} Y_j(\omega)$$

a) Démontrer : $W(\Omega) = \mathbb{N}$.

- **1 pt : $W(\Omega) \subset \mathbb{N}$ (une somme d'entiers naturels est un entier naturel)**
- **1 pt : $\mathbb{N} \subset W(\Omega)$ en donnant une issue concrète**

b) Soit $k \in \mathbb{N}^*$. Montrer que, pour tout $n \geqslant (k-1)N$,

$$\sum_{i=0}^{n} i\ \mathbb{P}_{[Z=k]}([S_{k-1} = i]) = \mathbb{E}_{[Z=k]}(S_{k-1})$$

- **1 pt : $S_{k-1}(\Omega) \subset [\![0, (k-1)(N+1)]\!]$**

- **1 pt : La variable aléatoire S_{k-1} étant finie, elle admet une espérance conditionnelle sachant $[Z=k]$**
- **1 pt :** $\mathbb{E}_{[Z=k]}\left(S_{k-1}\right) = \displaystyle\sum_{i=0}^{(k-1)N} i\ \mathbb{P}_{[Z=k]}([S_{k-1}=i])$
- **1 pt :** $\displaystyle\sum_{i=0}^{n} i\ \mathbb{P}_{[Z=k]}([S_{k-1}=i]) = \sum_{i=0}^{(k-1)N} i\ \mathbb{P}_{[Z=k]}([S_{k-1}=i])$

c) En déduire que, pour tout $k \in \mathbb{N}^*$, $\mathbb{E}_{[Z=k]}(W)$ existe et

$$\mathbb{E}_{[Z=k]}(W) = \mathbb{E}_{[Z=k]}\left(S_{k-1}\right)$$

- **1 pt : La variable aléatoire W est à valeurs dans $\mathbb{N}$, donc W admet une espérance conditionnelle sachant $[Z=k]$ si et seulement si la série $\sum\ i\ \mathbb{P}_{[Z=k]}([W=i])$ converge absolument, ce qui revient à montrer la convergence car il s'agit d'une série à termes positifs**
- **1 pt :** $\displaystyle\sum_{i=0}^{n} i\ \mathbb{P}_{[Z=k]}([W=i]) = \mathbb{E}_{[Z=k]}\left(S_{k-1}\right)$
- **1 pt : la suite des sommes partielles est constante à partir d'un certain rang et donc converge**

12. a) Démontrer : $T = W + N + 1$.

- **1 pt : $W(\omega)$ compte le nombre total de combats effectués lors des $Z(\omega)-1$ premières parties**
- **1 pt : $W(\omega)+N+1$ compte le nombre total de combats effectués lors des $Z(\omega)$ premières parties**
- **1 pt : $W(\omega)+N+1$ est précisément le numéro du premier combat effectué contre le boss final**

b) En déduire que : $\mathbb{E}(T) = \dfrac{1}{u_N}\displaystyle\sum_{m=0}^{N} u_m$. Ce résultat est-il cohérent avec celui de la partie **I** ?

- **1 pt : par transformation affine, T admet une espérance si et seulement si W admet une espérance**
- **1 pt : la variable aléatoire W admet une espérance si et seulement si la série $\sum\ \mathbb{P}([Z=k])\mathbb{E}_{[Z=k]}(W)$ converge absolument, ce qui revient à montrer la convergence car il s'agit d'une série à termes positifs. De plus, en cas de convergence**
- **1 pt :** $\displaystyle\sum_{k=1}^{n} \mathbb{P}([Z=k])\mathbb{E}_{[Z=k]}(W) = \sum_{k=1}^{n} \mathbb{P}([Z=k]) \sum_{j=1}^{k-1} \mathbb{E}_{[Z=k]}\left(Y_j\right)$
- **1 pt : par linéarité de l'espérance conditionnelle cf question 6b**
- **1 pt :** $\displaystyle\sum_{k=1}^{n} \mathbb{P}([Z=k])\mathbb{E}_{[Z=k]}(W) = \frac{1}{1-u_N}\left(\sum_{m=0}^{N} u_m - (N+1)u_N\right) \sum_{k=1}^{n} \mathbb{P}([Z=k])(k-1)$
- **1 pt :** $\displaystyle\sum_{k=1}^{n} \mathbb{P}([Z=k])(k-1) = \sum_{k=1}^{n} k\mathbb{P}([Z=k]) - \sum_{k=1}^{n} \mathbb{P}([Z=k]) \underset{n\to+\infty}{\longrightarrow} \mathbb{E}(Z)-1$
- **1 pt :** $\mathbb{E}(W) = \dfrac{1}{u_N}\displaystyle\sum_{m=0}^{N} u_m - (N+1)$
- **1 pt : lorsque $N=1$, on trouve $\mathbb{E}(T) = \dfrac{1}{u_1}\displaystyle\sum_{m=0}^{1} u_m = \dfrac{1}{p_1}(1+p_1) = \dfrac{1+p}{p}$ en posant $p=p_1$. On retrouve bien le résultat de la partie I.**

c) On suppose, **dans cette question uniquement**, que : pour tout $i \in [\![1, N]\!]$, $p_i = p \in]0, 1[$.

Démontrer : $\mathbb{E}(T) \underset{N\to+\infty}{\sim} \dfrac{1}{(1-p)p^N}$.

- **1 pt :** $\mathbb{E}(T) = \dfrac{1}{u_N} \displaystyle\sum_{m=0}^{N} u_m = \dfrac{1}{p^N} \displaystyle\sum_{m=0}^{N} p^m$
- **1 pt : On reconnait une série géométrique de raison p, vérifiant** $|p| < 1$
- **1 pt :** $\dfrac{1}{1-p} \neq 0$ **donc** $\displaystyle\sum_{m=0}^{N} p^m \underset{N\to+\infty}{\sim} \dfrac{1}{1-p}$

13. Simulations informatiques. On rappelle qu'en **Python**, si `L` est une liste, la commande `len(L)` renvoie la taille de `L` et la commande `L[-1]` permet d'accéder au dernier élément de `L`.

a) Ecrire une fonction **Python**, nommée `CalcU`, qui prend en argument une liste de probabilités `P = [`p_1`,...,`p_N`]` et qui renvoie la liste `U = [`u_0`,...,`u_N`]`.

```
def CalcU(P):
    U = [1]
    for k in range(len(P)):
        U.append(U[-1]*P[k])
    return U
```

- **1 pt : initialisation de** `U`
- **2 pts : boucle for**

b) Ecrire une fonction **Python**, nommée `CalcEsp`, qui prend en argument une liste de probabilités `P = [`p_1`,...,`p_N`]` et qui renvoie l'espérance de T (on utilisera la question 12b).

```
def CalcEsp(P):
    U = CalcU(P)
    return sum(U)/U[-1]
```

- **1 pt : création de** `U`
- **2 pts :** `return sum(U)/U[-1]`

c) Compléter la fonction **Python** suivante, afin qu'elle prenne en argument une liste de probabilités `P = [`p_1`,...,`p_N`]` et un entier `a` spécifiant l'état de la chaîne de Markov à un instant donné et qu'elle renvoie l'état à l'instant suivant.

```
def EtapeMarkov(P,a):
    N = len(P)
    if a == N+1 :
        return 1
    else:
        r = rd.random()
        if r < P[a-1] :
            return a+1
        return 1
```

- **1 pt :** `if a == N+1:`
- **1 pt :** `if r < P[a-1]:`
- **1 pt :** `return 1`

Problème (ESSEC-I 2007)

Dans tout l'exercice, les variables aléatoires sont définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.

I. Préliminaires

Dans cette partie **I.**, λ désigne un réel strictement positif.

1. Soit X une variable aléatoire suivant la loi exponentielle de paramètre λ.

a) Déterminer la fonction : $x \mapsto \mathbb{P}([X > x])$ (appelée *fonction de survie de* X).

- 1 pt : survie $x \mapsto \begin{cases} e^{-\lambda x} & \textbf{si } x \geqslant 0 \\ 1 & \textbf{si } x < 0 \end{cases}$

b) Pour tous nombres réels strictement positifs x et y, calculer la probabilité conditionnelle $\mathbb{P}_{[X>x]}([X > x + y])$; justifier alors que, si X modélise la durée de vie d'un phénomène, on dise de ce dernier qu'il est « sans vieillissement ».

- 1 pt : $\mathbb{P}([X > x]) > 0$

- 1 pt : $[X > x + y] \subset [X > x]$ **car** $y > 0$

- 1 pt : $x + y \geqslant 0$ **et** $x \geqslant 0$

- 1 pt : $\mathbb{P}_{[X>x]}([X > x + y]) = \mathbb{P}([X > y])$

- 1 pt : interprétation

2. Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes, de même loi exponentielle de paramètre λ. Pour tout entier naturel n non nul, on pose :

$$S_n = \sum_{k=1}^{n} X_k.$$

a) Déterminer l'espérance et la variance de S_n.

- 1 pt : existence espérance / variance

- 1 pt : $\mathbb{E}(S_n) = \dfrac{n}{\lambda}$ **(1 pt pour linéarité)**

- 2 pts : $\mathbb{V}(S_n) = \dfrac{n}{\lambda^2}$ **(1 pt pour indépendance, 1 pt pour** (X_i) **suivent** $\mathcal{E}(\lambda)$

b) Démontrer par récurrence que, pour tout n de $\mathbb{N}^*$, la variable aléatoire S_n admet pour densité la fonction :

$$f_n : t \mapsto \begin{cases} 0 & \text{si } t < 0\,; \\ \dfrac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} & \text{si } t \geqslant 0. \end{cases}$$

Pour cela, on admettra que, si U et V sont des variables aléatoires indépendantes admettant respectivement pour densité les fonctions f_U et f_V, alors la variable aléatoire $U + V$ admet pour densité la fonction f_{U+V} définie sur $\mathbb{R}$ par :

$$f_{U+V}(t) = \int_{-\infty}^{+\infty} f_U(x) f_V(t - x)\, dx$$

- 1 pt : initialisation

- 5 pts : hérédité (1 pt pour hypothèse v.a.r. densité, 1 pt pour lemme des coalitions, 1 pt pour $f_{S_{n+1}}$ **sur** $]-\infty, 0[$**, 2 pts pour cas** $[0, +\infty[$**)**

II. Loi de Pareto (Vilfredo Pareto (1848-1923), sociologue et économiste italien)

Soient a et b des réels strictement positifs. Par définition, on dit d'une variable aléatoire qu'elle suit la loi de Pareto de paramètres a et b si elle admet pour densité la fonction f définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} 0 & \text{si } t < b\,; \\ a\dfrac{b^a}{t^{a+1}} & \text{si } t \geqslant b. \end{cases}$$

Soit alors X une variable aléatoire de loi de Pareto de paramètres a et b.

3. Vérifier que l'égalité : $\displaystyle\int_{-\infty}^{+\infty} f(t)\ dt = 1$ est bien satisfaite ; calculer l'espérance et la variance de X, en précisant à quelles conditions chacune de ces quantités existe.

- **1 pt : f nulle en dehors de $[b, +\infty[$**
- **1 pt : $t \mapsto a\dfrac{b^a}{t^{a+1}}$ continue sur $[b, +\infty[$**
- **1 pt : $\displaystyle\int_{-\infty}^{+\infty} f(t)\ dt = 1$**
- **1 pt : convergence absolue**
- **1 pt : X admet une espérance ssi $a > 1$**
- **1 pt : $\mathbb{E}(X) = \dfrac{a\,b}{a-1}$**
- **2 pts : $\mathbb{E}(X^2) = \dfrac{a\,b^2}{a-2}$ (dont 1 pt pour existence ssi $a > 2$)**
- **1 pt : $\mathbb{V}(X) = \dfrac{a\,b^2}{(a-2)(a-1)^2}$**

4. Déterminer la fonction de répartition de X. Préciser la fonction de survie : $x \mapsto \mathbb{P}([X > x])$.

- **1 pt : $F_X(x) = 0$ si $x < b$**
- **2 pts : $F_X(x) = 1 - \left(\dfrac{b}{x}\right)^a$ si $x \geqslant b$**
- **1 pt : fonction de survie**

5. Démontrer que, pour tout réel y positif ou nul, la probabilité conditionnelle $\mathbb{P}_{[X>x]}([X > x + y])$ tend vers 1 quand x tend vers $+\infty$. De façon analogue à la question ***I.1.b)***, que peut-on dire d'un phénomène dont la durée de vie est modélisée par X ?

- **1 pt : $[X > x + y] \subset [X > x]$**
- **1 pt : $\mathbb{P}_{[X>x]}([X > x + y]) = \left(\dfrac{x}{x+y}\right)^a$**
- **1 pt : $\displaystyle\lim_{x\to+\infty} \left(\dfrac{x}{x+y}\right)^a = 1$**
- **1 pt : interprétation**

6. On pose dans cette question : $Y = \ln \dfrac{X}{b}$.

a) Démontrer que Y suit une loi exponentielle dont on précisera le paramètre.

- **1 pt : $Y(\Omega) = [0, +\infty[$**

- **1 pt : cas** $]-\infty, 0[$
- **2 pts : cas** $[0,+\infty[$ **(1 pt pour stricte croissance** exp**, 1 pt pour** $b\,\mathrm{e}^x \geqslant b$**)**
- **1 pt :** $X \hookrightarrow \mathcal{E}(a)$

b) Déduire de la question précédente une fonction **Python**, nommée `SimulX(a,b)`, qui permette de simuler la variable aléatoire X.

```
def SimulX(a,b):
    Y = rd.exponential(1/a)
    return b * np.exp(Y)
```

- **1 pt :** `Y = rd.exponential(1/a)`
- **1 pt :** `return b * np.exp(Y)`

III. Estimation des paramètres d'une loi de Pareto

Les instants aléatoires des arrivées de paquets (symboles binaires représentant de l'information de type audio, vidéo, données, ...) dans un canal de communication sont modélisés par une variable aléatoire X suivant une loi de Pareto de paramètres α et β ($\alpha > 0$, $\beta > 0$).
Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes, toutes de même loi que X.
On dit qu'un estimateur T_n de θ est *sans biais* si $\mathbb{E}(T_n) = \theta$.

7. On suppose tout d'abord que le paramètre β fait partie des caractéristiques connues du canal de communication ; on se propose de déterminer un estimateur de α par une méthode dite du « maximum de vraisemblance ». Pour cela, n désignant un entier naturel non nul et x_1, ..., x_n des réels supérieurs ou égaux à β, on introduit la fonction $\mathcal{L}$, à valeurs dans $\mathbb{R}$ et définie sur $]0,+\infty[$ par :

$$\mathcal{L}(a) = f_a(x_1) \times \cdots \times f_a(x_n) = \prod_{k=1}^{n} f_a(x_k),$$

où f_a est la fonction définie sur $\mathbb{R}$ par :

$$f_a(t) = \begin{cases} 0 & \text{si } t < \beta\,; \\ a\dfrac{\beta^a}{t^{a+1}} & \text{si } t \geqslant \beta. \end{cases}$$

a) Exprimer $\mathcal{L}(a)$, puis $\ln(\mathcal{L}(a))$.

- **1 pt :** $\mathcal{L}(a) = \dfrac{(a\,\beta^a)^n}{(x_1 \cdots x_n)^{a+1}}$
- **1 pt :** $\ln(\mathcal{L}(a)) = n\,\ln(a) + n\,a\,\ln(\beta) - (a+1)\sum\limits_{k=1}^{n} \ln(x_k)$

b) On considère la fonction φ de $]0,+\infty[$ dans $\mathbb{R}$:

$$a \mapsto n\ln(a) + na\ln(\beta) - (a+1)\sum_{k=1}^{n} \ln(x_k)\,.$$

(i) Démontrer que la fonction φ admet un maximum, atteint en un unique réel w que l'on exprimera en fonction de x_1, ..., x_n et de β.

- **1 pt :** φ **est dérivable sur** $]0,+\infty[$
- **1 pt :** $\varphi'(a) = \dfrac{n}{a} - \sum\limits_{k=1}^{n} \ln\left(\dfrac{x_k}{\beta}\right)$

- **2 pts : $\varphi'(a) \geqslant 0$ ssi $a \leqslant w$, où $w = \dfrac{n}{\sum\limits_{k=1}^{n} \ln\left(\dfrac{x_k}{\beta}\right)}$**
- **1 pt : w est l'unique maximum de φ**

(ii) Que peut-on dire de w pour la fonction $\mathcal{L}$?

- **1 pt**

c) On pose dorénavant, pour tout n de $\mathbb{N}^*$:

$$W_n = \frac{n}{\sum\limits_{k=1}^{n} \ln\left(\dfrac{X_k}{\beta}\right)}.$$

(La suite $(W_n)_{n\geqslant 1}$ est appelée *estimateur du maximum de vraisemblance.*)

(i) Justifier que la variable aléatoire $\sum\limits_{k=1}^{n} \ln\left(\dfrac{X_k}{\beta}\right)$ admet pour densité la fonction f_n définie dans ***I.2.b)*** en prenant $\lambda = \alpha$.

- **1 pt : $\ln\left(\dfrac{X_k}{\beta}\right) \hookrightarrow \mathcal{E}(\alpha)$**
- **1 pt : $(\ln\left(\dfrac{X_i}{\beta}\right))$ indépendantes par lemme des coalitions**
- **1 pt : question *I.2.b)***

(ii) À l'aide du théorème de transfert, en déduire que W_n admet pour espérance $\dfrac{n\alpha}{n-1}$ lorsque $n \geqslant 2$, puis proposer un estimateur sans biais de α construit sur W_n.

- **1 pt : théorème de transfert**
- **1 pt : convergence absolue**
- **1 pt : $g(t)\, f_n(t) = \dfrac{n}{n-1}\, \alpha\, f_{n-1}(t)$**
- **1 pt : $\mathbb{E}(W_n) = \dfrac{n}{n-1}\, \alpha$**
- **1 pt : $\dfrac{n-1}{n}\, W_n$ sans biais**

d) On pose, pour tout n de $\mathbb{N}^*$:

$$W'_n = \frac{n-1}{n} W_n.$$

(i) Soit n un entier supérieur ou égal à 3.

En admettant que le moment d'ordre 2 de W'_n est égal à $\dfrac{(n-1)\alpha^2}{n-2}$, calculer la variance de W'_n puis établir, à l'aide de l'inégalité de Bienaymé-Tchebychev, que pour tout réel ε strictement positif, on a l'inégalité :

$$\mathbb{P}\left(\left[W'_n - \varepsilon < \alpha < W'_n + \varepsilon\right]\right) \geqslant 1 - \frac{\alpha^2}{\varepsilon^2 (n-2)}.$$

- **1 pt : $\mathbb{V}(W'_n) = \dfrac{\alpha^2}{n-2}$**
- **1 pt : inégalité de Bienaymé-Tchebychev (0 sans argument existence de la variance)**
- **1 pt : événement contraire**

- **1 pt :** $|W'_n - \alpha| < \varepsilon$ **ssi** $W'_n - \varepsilon < \alpha < W'_n - \varepsilon$

(ii) On suppose dans cette question (et elle seule) que α est strictement compris entre 1 et 2. Déterminer un entier naturel N tel que, pour tout entier n supérieur ou égal à N, $\left[W'_n - \dfrac{1}{10}\ ,\ W'_n + \dfrac{1}{10}\right]$ soit un intervalle de confiance du paramètre réel α au niveau de confiance 0,95.

- **1 pt : question précédente avec** $\varepsilon = \dfrac{1}{10}$
- **3 pts :** $n \geqslant 8002$

8. On suppose maintenant que seul le paramètre α est déjà identifié et qu'il vérifie : $\alpha > 2$.

a) Pour tout entier strictement positif n, on pose :

$$Y_n = c_n \sum_{k=1}^{n} X_k,$$

où le réel c_n est choisi de sorte que $(Y_n)_{n\geqslant 1}$ soit un estimateur sans biais de β.

(i) Calculer c_n.

- **1 pt :** $c_n = \dfrac{\alpha - 1}{\alpha\, n}$

(ii) Quelle est la limite de la variance de Y_n quand n tend vers $+\infty$?
(On dit que l'estimateur est convergent.)

- **1 pt : existence** $\mathbb{V}(Y_n)$
- **2 pts :** $\mathbb{V}(Y_n) = \dfrac{\beta^2}{\alpha(\alpha-2)}\dfrac{1}{n}$
- **1 pt :** $\mathbb{V}(Y_n) \underset{n\to+\infty}{\longrightarrow} 0$

b) Pour tout entier strictement positif n, on pose : $Z_n = \min(X_1, \dots, X_n)$.

(i) Déterminer la fonction de répartition de Z_n, puis reconnaître sa loi et préciser son espérance. Quelle est la limite de cette dernière quand n tend vers $+\infty$?

- **1 pt :** $Z_n(\Omega) \subset [\beta, +\infty[$
- **1 pt :** $F_{Z_n}(x) = 0$ **si** $x < \beta$
- **2 pts :** $F_{Z_n}(x) = 1 - \left(\dfrac{\beta}{x}\right)^{\alpha\, n}$ **si** $x \geqslant \beta$ **(1 pt pour** X_i **indépendantes, 1 pt pour** (X_i) **ont même loi)**
- **1 pt :** $\mathbb{E}(Z_n) = \dfrac{\alpha\, n\, \beta}{\alpha\, n - 1}$
- **1 pt :** $\mathbb{E}(Z_n) \underset{n\to+\infty}{\longrightarrow} \beta$

(ii) Pour tout entier strictement positif n, on pose : $Z'_n = d_n Z_n$, où le réel d_n est choisi de telle sorte que $(Z'_n)_{n\geqslant 1}$ soit un estimateur sans biais de β.
Quelle est la limite de la variance de Z'_n quand n tend vers $+\infty$?

- **1 pt :** $d_n = \dfrac{n\,\alpha - 1}{n\,\alpha}$
- **1 pt :** $\mathbb{V}(Z'_n) = \dfrac{\beta^2}{n\,\alpha(n\,\alpha - 2)}$
- **1 pt :** $\mathbb{V}(Z'_n) \underset{n\to+\infty}{\longrightarrow} 0$

(iii) Démontrer que l'estimateur $(Z'_n)_{n\geqslant 1}$ est plus efficace que l'estimateur $(Y_n)_{n\geqslant 1}$, c'est-à-dire, qu'à partir d'un certain rang, la variance de Z'_n est inférieure à celle de Y_n.

- **3 pts :** $\mathbb{V}(Z'_n) \leqslant \mathbb{V}(Y_n)$ **ssi** $n \geqslant 1$