TD14

Vecteurs aléatoires

Maximum, minimum

Exercice 14.1 (★)

Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes suivant toutes une loi $\mathscr{B}(p)$. Déterminer la loi de $Z = \max(X_1, \dots, X_n)$.

> Puisque $Z(\Omega) = \{0, 1\}$, Z suit une loi de Bernoulli. Pour obtenir le paramètre, remarquons que :
>
> $$[Z = 0] = [X_1 = 0] \cap \dots \cap [X_n = 0].$$
>
> D'où par indépendance des X_i et du fait qu'elles suivent la même loi :
>
> $$P(Z = 0) = P(X_1 = 0)^n = (1 - p)^n.$$
>
> Z suit donc une loi de Bernoulli de paramètre $1 - (1 - p)^n$.

Exercice 14.2 (★)

Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes suivant toutes une loi $\mathscr{G}(p)$ avec $p \in]0, 1[$, et soit $Z = \min(X_1, \dots, X_n)$. Montrer que Z suit une loi usuelle dont on précisera les paramètres. Calculer $E(Z)$ et $V(Z)$.

Exercice 14.3 (★★)

Un sac contient N boules numérotées de 1 à N. On effectue dans ce sac n tirages d'une boule avec remise et on note Z_n le plus petit numéro obtenu. Déterminer la loi de Z_n.

> Pour tout $i \in [\![1, n]\!]$, notons X_i la variable aléatoire donnant le numéro de la boule obtenue au i-ème tirage. Les variables $X_1, \dots, X_n$ sont indépendantes (tirage avec remise) et suivent toutes une loi uniforme discrète sur $[\![1, N]\!]$. Par définition, $Z_n = \min(X_1, \dots, X_n)$. En particulier, $Z_n(\Omega) = [\![1, N]\!]$ et Z est une variable discrète.
> Pour tout $k \in [\![1, N]\!]$:
>
> $$[Z_n > k] = [X_1 > k] \cap \dots \cap [X_n > k].$$
>
> Les variables X_i étant de même loi et indépendantes :
>
> $$P(Z_n > k) = P(X_1 > k)^n = \left(\frac{N - k}{N}\right)^n,$$
>
> formule valable pour $k = 0$ également. Reste à se ramener à la probabilité ponctuelle :
>
> $$P(Z_n = k) = P(Z_n > k - 1) - P(Z_n > k) = \left(\frac{N - k + 1}{N}\right)^n - \left(\frac{N - k}{N}\right)^n.$$

Exercice 14.4 (★ - Minimum et maximum de lois uniformes)
Soit $a > 0$ et soient $X_1, \dots, X_n$ n variables aléatoires indépendantes suivant toutes la loi uniforme $\mathscr{U}([0, a])$. On pose

$$Y = \max(X_1, \dots, X_n) \text{ et } Z = \min(X_1, \dots, X_n).$$

Montrer que Y et Z sont des variables aléatoires à densité et en déterminer une densité.

Exercice 14.5 (★★ - Lois uniformes et loi de Pareto)
Soit $n \geq 3$ et soient $Z_1, \dots, Z_n$ des variables aléatoires indépendantes sur $(\Omega, \mathscr{A}, P)$ suivant toutes la loi uniforme sur $]0, a]$, $a > 0$. On pose alors $X_i = \dfrac{1}{Z_i}$.

1. Montrer que X_1 est une variable aléatoire à densité, et qu'une densité de X_1 est :
$$f : x \in \mathbb{R} \mapsto \begin{cases} \dfrac{1}{ax^2} & \text{si } x \geq \frac{1}{a} \\ 0 & \text{sinon} \end{cases}$$
2. Soit $Y_n = \min(X_1, \dots, X_n)$. Montrer que Y_n est une variable à densité, et qu'une densité de Y_n est :
$$g : x \in \mathbb{R} \mapsto \begin{cases} \dfrac{n}{a^n x^{n+1}} & \text{si } x \geq \frac{1}{a} \\ 0 & \text{sinon} \end{cases}$$
3. Montrer que Y_n admet une espérance et une variance, et les calculer.
4. `Python`. Écrire une fonction `pareto` qui simule la loi de X_n pour $a > 0$ et $n \in \mathbb{N}^*$.

1. On est dans le cas d'une variable fonction d'une variable aléatoire à densité. On procède donc par étapes.

Étape 1. Support de X_1. Puisque $0 < Z_1 \leq a$, il suit $X_1 = \dfrac{1}{Z_1} \geq \dfrac{1}{a}$. Ainsi $X_1(\Omega) \subset [\frac{1}{a}, +\infty[$.
En particulier, $F_{X_1}(x) = 0$ pour tout $x < \dfrac{1}{a}$.

Étape 2. Fonction de répartition de X_1. Pour tout $x \geq \dfrac{1}{a}$:

$$\begin{aligned} F_{X_1}(x) &= P(X_1 \leq x) = P\left(\frac{1}{Z_1} \leq x\right) \\ &= P\left(Z_1 \geq \frac{1}{x}\right) = 1 - P\left(Z_1 < \frac{1}{x}\right) \\ &= 1 - P\left(Z_1 \leq \frac{1}{x}\right) \quad \text{car } Z_1 \text{ à densité} \\ &= 1 - F_{Z_1}\left(\frac{1}{x}\right) = 1 - \frac{\frac{1}{x} - 0}{a - 0} = 1 - \frac{1}{xa} \end{aligned}$$

car $\frac{1}{x} \in]0, a]$. Ainsi :

$$F_{X_1} : x \in \mathbb{R} \mapsto \begin{cases} 1 - \dfrac{1}{xa} & \text{si } x \geq \dfrac{1}{a}, \\ 0 & \text{sinon.} \end{cases}$$

Étape 3. X_1 variable à densité. La fonction F_{X_1} est continue sur $\mathbb{R}$ sauf éventuellement en $\dfrac{1}{a}$

comme composée de fonctions continues. Et en $\dfrac{1}{a}$:

$$\lim_{x\to\frac{1}{a}^-} F_{X_1}(x) = 0 = F_{X_1}(1/a) = \lim_{x\to\frac{1}{a}^+} F_{X_1}(x).$$

Ainsi, F_{X_1} est continue sur $\mathbb{R}$. Elle est aussi $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en $\dfrac{1}{a}$ comme composée de fonctions $\mathscr{C}^1$. Donc X_1 est une variable à densité.

Étape 4. Densité de X_1. Pour tout $x \neq 1/a$:

$$F'_{X_1}(x) = \begin{cases} \dfrac{1}{ax^2} & \text{si } x > \frac{1}{a}, \\ 0 & \text{si } x < \frac{1}{a}. \end{cases}$$

Une densité de X_1 est donnée par (en choisissant arbitrairement la valeur en $1/a$) :

$$f_{X_1} : x \in \mathbb{R} \mapsto \begin{cases} \dfrac{1}{ax^2} & \text{si } x \geq \frac{1}{a}, \\ 0 & \text{sinon.} \end{cases},$$

2. Pour tout $x \in \mathbb{R}$:

$$P(Y_n > x) = P(X_1 > x,\ X_2 > x,\ \dots\ ,\ X_n > x) = P(X_1 > x)^n$$

car les X_i sont indépendantes et suivent la même loi. On en déduit :

$$F_{Y_n}(x) = 1 - P(Y_n > x) = 1 - P(X_1 > x)^n = 1 - (1 - F_{X_1}(x))^n.$$

On a vu que F_{X_1} est continue sur $\mathbb{R}$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en $1/a$. Donc F_{Y_n} est aussi continue sur $\mathbb{R}$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en $1/a$, en tant que composée de fonctions qui le sont. Ainsi, Y_n est une variable à densité. De plus, pour tout $x \neq 1/a$:

$$F'_{Y_n}(x) = nF'_{X_1}(x)(1 - F_{X_1}(x))^{n-1}.$$

Une densité de Y_n est donc :

$$f_{Y_n} : x \in \mathbb{R} \mapsto nf_{X_1}(x)(1-F_{X_1}(x))^{n-1} = \begin{cases} n\dfrac{1}{ax^2}\dfrac{1}{(xa)^{n-1}} & \text{si } x \geq \frac{1}{a} \\ 0 & \text{sinon} \end{cases} = \begin{cases} \dfrac{n}{a^n x^{n+1}} & \text{si } x \geq \frac{1}{a} \\ 0 & \text{sinon} \end{cases}.$$

3. Y_n admet une espérance si, et seulement si, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} xg(x)\mathrm{d}x = \frac{n}{a^n}\int_{1/a}^{+\infty} \frac{\mathrm{d}x}{x^n}$ converge absolument, donc converge car la fonction intégrée est positive. On reconnait ici une intégrale de Riemann en $+\infty$ qui converge bien car $n > 1$. Ainsi $E(Y_n)$ existe et vaut :

$$E(Y_n) = \frac{n}{a^n}\int_{1/a}^{+\infty} \frac{\mathrm{d}x}{x^n} = \frac{n}{a^n}\lim_{A\to+\infty}\left[\frac{x^{-n+1}}{-n+1}\right]_{1/a}^{A} = \frac{n}{a^n}\frac{a^{n-1}}{n-1} = \frac{n}{(n-1)a}.$$

Y_n admet un moment d'ordre 2 si, et seulement si, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^2 g(x)\mathrm{d}x = \frac{n}{a^n}\int_{1/a}^{+\infty} \frac{\mathrm{d}x}{x^{n-1}}$ converge absolument, donc converge car la fonction intégrée est positive. On reconnait là aussi une intégrale de Riemann en $+\infty$ qui converge bien car $n-1>1$. Ainsi $E(Y_n^2)$ existe bien, et :

$$E(Y_n^2) = \frac{n}{a^n}\int_{1/a}^{+\infty}\frac{\mathrm{d}x}{x^{n-1}} = \frac{n}{a^n}\lim_{A\to+\infty}\left[\frac{x^{-n+2}}{-n+2}\right]_{1/a}^{A} = \frac{n}{a^n}\frac{a^{n-2}}{n-2} = \frac{n}{(n-2)a^2}.$$

Par la formule de Huygens, $V(Y_n)$ existe et vaut :

$$V(Y_n) = E(Y_n^2) - E(Y_n)^2 = \frac{n}{(n-2)a^2} - \frac{n^2}{(n-1)^2a^2} = \frac{(n-1)^2 - n(n-2)}{(n-2)(n-1)^2}\frac{n}{a^2}$$
$$= \frac{n}{(n-2)(n-1)^2a^2}.$$

4. En suivant la construction des Y_n, on propose la fonction suivante :

```python
def pareto(a,n):
    x = a*rd.random(n) #simule n réalisations indép. de X_1
    y = np.min(x)
    return y
```

Exercice 14.6 (★★★)

Soit N une variable aléatoire suivant la loi binomiale de paramètres n et p ($n \in \mathbb{N}^*$, $p \in]0,1[$).
Soit $X_0, X_1, \ldots, X_n$ $n+1$ variables aléatoires suivant la loi uniforme sur $[0,1]$.
On suppose que $X_0, X_1, \ldots, X_n, N$ sont mutuellement indépendantes.

1. Soit $k \in \{0,\ldots,n\}$. Déterminer la fonction de répartition de la variable aléatoire $T_k = \min(X_0, X_1, \ldots, X_k)$.

2. On pose $T = \min(X_0, X_1, \ldots, X_N)$.

   (a) Montrer que T est une variable aléatoire.

   (b) Montrer que T est une variable à densité et déterminer une densité de T.

   (c) `Python`. Écrire une fonction `simulation` qui simule la loi de T pour $n \in \mathbb{N}$ et $p \in]0,1[$.

1. Pour tout $x \in \mathbb{R}$:

$$P(T_k > x) = P(X_0 > x, \ldots, X_k > x) = P(X_0 > x)^{k+1}$$

car les variables aléatoires sont indépendantes et de même loi. On obtient :

$$F_{T_k}(x) = 1 - (1 - F_{X_0}(x))^{k+1}.$$

2. (a) Ramenons nous à la définition : T est une variable aléatoire réelle si, et seulement si, pour tout $x \in \mathbb{R}$, $[T \leq x]$ appartient à $\mathscr{A}$, ce qui est équivalent à $[T > x] \in \mathscr{A}$ (puisqu'une tribu est stable par passage au complémentaire). En utilisant le système complet d'évènements $([N = k])$:

$$[T > x] = \bigcup_{k=0}^{n}[T > x] \cap [N = k] = \bigcup_{k=0}^{n}[T_k > x] \cap [N = k]$$
$$= \bigcup_{k=0}^{n}[X_1 > x] \cap \cdots \cap [X_k > x] \cap [N = k]$$

Or $X_1, \dots, X_n$ et N sont des variables aléatoires, donc $[X_j > x] \in \mathscr{A}$ pour tout $j = 0, \dots, n$ et $[N = k] \in \mathscr{A}$ aussi. Puisque $\mathscr{A}$ est stable par intersection et union finie ou dénombrable, $[T > x]$ est bien un évènement. Ainsi T est bien une variable aléatoire.

(b) Toujours avec le même SCE, on obtient pour tout $x \in \mathbb{R}$:

$$\begin{aligned}
F_T(x) &= \sum_{k=0}^{n} P([T \leq x] \cap [N = k] = \sum_{k=0}^{n} P([T_k \leq x] \cap [N = k]) \\
&= \sum_{k=0}^{n} P(T_k \leq x) P(N = k) \text{ par indépendance} \\
&= \sum_{k=0}^{n} \left(1 - (1 - F_{X_0}(x))^{k+1}\right) P(N = k) \\
&= \sum_{k=0}^{n} P(N = k) - \sum_{k=0}^{n} (1 - F_{X_0}(x))^{k+1} P(N = k) \\
&= 1 - (1 - F_{X_0}(x)) \sum_{k=0}^{n} \binom{n}{k} (1 - F_{X_0}(x))^k p^k (1 - p)^{n-k} \\
&= 1 - (1 - F_{X_0}(x)) (p - pF_{X_0}(x) + (1 - p))^n \\
&= 1 - (1 - F_{X_0}(x)) (1 - pF_{X_0}(x))^n
\end{aligned}$$

Par composée et somme de fonctions continues sur $\mathbb{R}$, F_T est continue sur $\mathbb{R}$. Elle est aussi de classe $\mathscr{C}^1$ sur $\mathbb{R}$ sauf ENFP. Donc T est bien une variable à densité.

Précisons l'expression de F_T :

$$F_T(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 - (1 - x)(1 - px) & \text{si } 0 < x < 1, \\ 1 & \text{si } x > 0. \end{cases}$$

D'où pour tout $x \neq 0, 1$:

$$F_T(x) = \begin{cases} 0 & \text{si } x < 0, \\ (1 - px) + p(1 - x) & \text{si } 0 < x < 1, \\ 0 & \text{si } x > 0. \end{cases}$$

Une densité de T est donc :

$$f_T : x \in \mathbb{R} \mapsto \begin{cases} 0 & \text{si } x \leq 0, \\ 1 + p - 2px & \text{si } 0 < x < 1, \\ 0 & \text{si } x \geq 0. \end{cases}$$

(c) On propose la fonction suivante :

```python
def simulation(n,p):
        N = rd.binomial(n,p)
        X = rd.random(N+1)
        T = np.min(X)
        return T
```

Exercice 14.7 (★★★★ - QSP HEC 2015)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes définies sur un espace probabilisé

$(\Omega, \mathscr{A}, P)$ de même loi exponentielle de paramètre $a > 0$.
Pour $x \in \mathbb{R}_+^*$, on définit la variable aléatoire N_x par :

$$\forall \omega \in \Omega, \quad N_x(\omega) = \begin{cases} \min(k \in \mathbb{N}^*,\ X_k(\omega) > x) \text{ si cet ensemble n'est pas vide} \\ 0 \text{ sinon} \end{cases}.$$

1. Déterminer la loi de N_x et préciser son espérance $E(N_x)$.
2. Déterminer $\lim\limits_{x \to +\infty} P(N_x > E(N_x))$.

1. Notons tout d'abord que $N_x(\Omega) = \mathbb{N}$ (puisqu'on fait un minimum sur des entiers !). Calculons :

$$\begin{aligned} P(N_x = 0) &= P\left(\bigcap_{k \in \mathbb{N}^*} [X_k \leq x] \right) = \lim_{N \to +\infty} P\left(\bigcap_{k=1}^{N} X_k \leq x \right) \\ &= \lim_{N \to +\infty} P(X_1 \leq x)^N \text{ car les v.a. sont indép.} \\ &= 0 \quad \text{car} \quad 0 \leq P(X_1 \leq x) < 1. \end{aligned}$$

Pour tout $k \in \mathbb{N}^*$:

$$[N_x = k] = [X_1 \leq x] \cap [X_2 \leq x] \cap \cdots \cap [X_{k-1} \leq x] \cap [X_k > x]$$

puisque $k = \min(k \in \mathbb{N}^*,\ X_k(\omega) > x)$. Comme les variables X_i sont indépendantes, on obtient :

$$\begin{aligned} P(N_x = k) &= P(X_1 \leq x) \dots P(X_{k-1} \leq x) P(X_k > x) \\ &= P(X_1 \leq x)^{k-1} (1 - P(X_1 \leq x)). \end{aligned}$$

Ainsi N_x suit une loi géométrique de paramètre $1 - F_{X_1}(x) = e^{-ax}$. En particulier, $E(N_x)$ existe et vaut $\dfrac{1}{1 - F_{X_1}(x)} = e^{ax}$.

2. Soit $x \in \mathbb{R}_+^*$. Notons $n = \lfloor \frac{1}{1 - F_{X_1}(x)} \rfloor + 1$. Alors :

$$\begin{aligned} P(N_x > E(N_x)) &= \sum_{k=n}^{+\infty} P(N_x = k) = \sum_{k=n}^{+\infty} F_{X_1}(x)^{k-1} (1 - F_{X_1}(x)) \\ &= (1 - F_{X_1}(x)) \frac{F_{X_1}(x)^{n-1}}{1 - F_{X_1}(x)} = F_{X_1}(x)^{n-1} = \exp\left((n-1) \ln(F_{X_1}(x))\right) \end{aligned}$$

Puisque $n - 1 = \lfloor \frac{1}{1 - F_{X_1}(x)} \rfloor$, on dispose des inégalités suivantes :

$$n - 1 \leq \frac{1}{1 - F_{X_1}(x)} < n.$$

En particulier :

$$\frac{1}{1 - F_{X_1}(x)} - 1 < n - 1 \leq \frac{1}{1 - F_{X_1}(x)}.$$

Puisque $\ln(F_{X_1}(x)) \leq 0$, on en déduit :

$$\frac{\ln(F_{X_1}(x))}{1 - F_{X_1}(x)} - \ln(F_{X_1}(x)) > (n-1) \ln(F_{X_1}(x)) \geq \frac{\ln(F_{X_1}(x))}{1 - F_{X_1}(x)}.$$

Comme $F_{X_1}(x) \to 1$ lorsque $x \to +\infty$, il suit $\ln(F_{X_1}(x)) \underset{x\to+\infty}{\sim} F_{X_1}(x) - 1$. Ainsi :

$$\lim_{x\to+\infty} \frac{\ln(F_{X_1}(x))}{1 - F_{X_1}(x)} = -1 = \lim_{x\to+\infty} \frac{\ln(F_{X_1}(x))}{1 - F_{X_1}(x)} - \ln(F_{X_1}(x)).$$

Par le théorème des gendarmes, on en déduit que $\lim_{x\to+\infty} (n-1)\ln(F_{X_1}(x)) = -1$. Par continuité de l'exponentielle, on peut donc conclure que :

$$\lim_{x\to+\infty} P(N_x > E(N_x)) = e^{-1}.$$

Vecteurs aléatoires, fonction d'un vecteur

Exercice 14.8 (★★)

Une urne contient n boules numérotées de 1 à n ($n \geq 3$). On tire simultanément trois boules dans l'urne, et on note X_1 le plus petit des trois numéros, X_3 le plus grand et X_2 le numéro intermédiaire.

1. Déterminer la loi de (X_1, X_2, X_3).

2. En déduire la loi de X_2, puis son espérance.

Exercice 14.9 (★★)

Soient X, Y, Z trois variables aléatoires indépendantes suivant toutes la loi de Poisson $\mathscr{P}(\lambda)$. Montrer que $\dfrac{X+Y}{1+Z}$ admet une espérance, et la déterminer.

Exercice 14.10 (★★★★ - QSP HEC 2008)

Les variables aléatoires considérées dans cet exercice sont définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$. On considère n variables aléatoires à densité, de même loi et indépendantes $X_1, \ldots, X_n$. On note F la fonction de répartition et f une densité des X_i.
Pour tout $\omega \in \Omega$, on note $(Y_1(\omega), Y_2(\omega), \ldots, Y_n(\omega))$ la suite des $X_i(\omega)$ pour $1 \leq i \leq n$ réordonnés par ordre croissant. On a donc $Y_1(\omega) \leq Y_2(\omega) \leq \cdots \leq Y_n(\omega)$ pour tout ω de Ω.

1. Si $1 \leq k \leq n$ et $x \in \mathbb{R}$, montrer que :

$$P([Y_k \leq x]) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j (1 - F(x))^{n-j}.$$

2. En déduire que Y_k admet une densité qu'on explicitera sans le signe $\sum$.

1. Fixons $x \in \mathbb{R}$. On définit la variable aléatoire N_x en posant pour tout $\omega \in \Omega$, $N_x(\omega)$ le nombre d'indices $k \in [\![1, n]\!]$ tels que $X_k(\omega) \leq x$. N_x correspond donc au nombre de succès lors d'une répétition de n épreuves de Bernoulli identiques (car les X_k suivent la même loi) avec probabilité de succès $p = P(X_1 \leq x)$, et indépendantes (car les X_k sont indépendantes). Donc N_x suit une loi binomiale de paramètres n et p.

   D'autre part puisque les Y_i sont ordonnés, l'évènement $[Y_k \leq x]$ est réalisé si, et seulement si, au moins k variables parmi $(X_1, \ldots, X_n)$ sont inférieures à x, soit si, et seulement si,

$[N_x \geq k]$ est réalisé. Ainsi :

$$\begin{aligned} P(Y_k \leq x) = P(N_x \geq k) &= \sum_{j=k}^{n} P(N_x = j) \\ &= \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \\ &= \sum_{j=k}^{n} \binom{n}{j} F(x)^j (1-F(x))^{n-j} \end{aligned}$$

2. On a obtenu l'expression de la fonction de répartition F_k de Y_k à la question précédente. En particulier F_k est combinaison linéaire, somme et produit de fonctions continues sur $\mathbb{R}$ (car F l'est). Donc F_k est continue sur $\mathbb{R}$. Elle est de plus $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points pour les mêmes raisons (car F l'est). Donc Y_k est à densité.

   Pour tout $x \in \mathbb{R}$ où F est dérivable :

$$\begin{aligned} F_k'(x) &= \sum_{j=k}^{n} \binom{n}{j} j f(x) F(x)^{j-1} (1-F(x))^{n-j} - \sum_{j=k}^{n} \binom{n}{j} (n-j) f(x) F(x)^j (1-F(x))^{n-j-1} \\ &= \sum_{j=k}^{n} \binom{n}{j} j f(x) F(x)^{j-1} (1-F(x))^{n-j} - \sum_{j=k}^{n-1} \binom{n}{j} (n-j) f(x) F(x)^j (1-F(x))^{n-j-1} \\ &= n f(x) \sum_{j=k}^{n} \binom{n-1}{j-1} F(x)^{j-1} (1-F(x))^{n-j} - n f(x) \sum_{j=k}^{n-1} \binom{n-1}{j} F(x)^j (1-F(x))^{n-j-1} \\ &= n f(x) \left(\sum_{j=k-1}^{n-1} \binom{n-1}{j} F(x)^j (1-F(x))^{n-j-1} - \sum_{j=k}^{n-1} \binom{n-1}{j} F(x)^j (1-F(x))^{n-j-1} \right) \\ &= n f(x) \binom{n-1}{k-1} F(x)^{k-1} (1-F(x))^{n-k} \end{aligned}$$

   Une densité de Y_k est donc :

$$f_k : x \in \mathbb{R} \mapsto n f(x) \binom{n-1}{k-1} F(x)^{k-1} (1-F(x))^{n-k}.$$

Somme de variables aléatoires

Exercice 14.11 (★)

Soient X_1, X_2, X_3 des variables indépendantes de loi $\mathscr{E}(\lambda)$, et soit $T = X_1 + X_2 + X_3$. Calculer $P(T > 1)$.

Exercice 14.12 (★★)

On désigne par p un réel de $]0,1[$. On considère une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$, indépendantes, dont la loi commune est donnée, pour tout $n \in \mathbb{N}^*$, par : $P(X_n = 1) = p$ et $P(X_n = -1) = 1-p$.

1. On pose $Y_n = aX_n + b$ avec $(a,b) \in \mathbb{R}^2$.
   Déterminer a et b pour que Y_n suive une loi de Bernoulli de paramètre p.

2. Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \sum_{k=1}^{n} X_k$. Déterminer la loi de S_n.

3. *Application.* Un supporter de foot sort d'un bar après une longue soirée. À chaque seconde, il avance d'un mètre en se déplaçant d'un mètre vers la gauche avec la probabilité p, et d'un mètre vers la droite avec la probabilité $1-p$. En $t=0$, il se trouve à la sortie du bar, à l'ordonnée $y=0$.
   (a) Donner la loi de son ordonnée à l'instant $t=n$ secondes.
   (b) En moyenne, où se trouve-t-il à $t=10$ secondes ?
   (c) `Python.` Proposer une fonction d'entête `def marche(n,p)` simulant la marche de ce supporter pendant n secondes. Cette fonction renverra un vecteur de longueur $n+1$, dont la k-ème coordonnée contient l'ordonnée du supporter à l'instant $t=k+1$.
   (d) `Python.` Proposer un script qui représente la trajectoire de ce supporter pendant 1000 secondes.

1. On cherche a et b tels que $P(Y_n=1)=P(X_n=1)$ et $P(Y_n=0)=P(X_n=-1)$, ce qui revient à résoudre :
$$\begin{cases} a+b=1 \\ -a+b=0 \end{cases} \Leftrightarrow a=b=\frac{1}{2}.$$
La variable $Y_n=\frac{X_n+1}{2}$
2. Pour tout $k\in[\![1,n]\!]$, $X_k=2Y_k-1$ et donc :
$$S_n=2\left(\sum_{k=1}^{n} Y_k\right)-n.$$
Par stabilité de la loi binomiale, $Z_n=\sum_{k=1}^{n} Y_k$ suit une loi $\mathscr{B}(n,p)$. Il suit que $S_n(\Omega)=\{2\ell-n,\ \ell\in[\![0,n]\!]\}$ et pour tout $\ell\in[\![0,n]\!]$:
$$P(S_n=2\ell-1)=P(Z_n=\ell)=\binom{n}{\ell}p^{\ell}(1-p)^{n-\ell}.$$
3. (a) La loi de son ordonnée est précisément celle de S_n.
   (b) Il s'agit de $E(S_{10})$. Par linéarité de l'espérance :
$$E(S_{10})=2E(Z_{10})-10=2\times 10\times p-10=10(2p-1).$$
   (c) Rappelons que la commande `rp.random()<=p` renvoie le booléen `True` avec une probabilité p, et `False` avec une probabilité $1-p$. Avec cette commande, nous pouvons, à l'aide d'une boucle conditionnelle, simuler la variable X_k. Reste à sommer pour simuler la variable S_k et à conserver les résultats dans un vecteur `u`. On peut procéder ainsi :

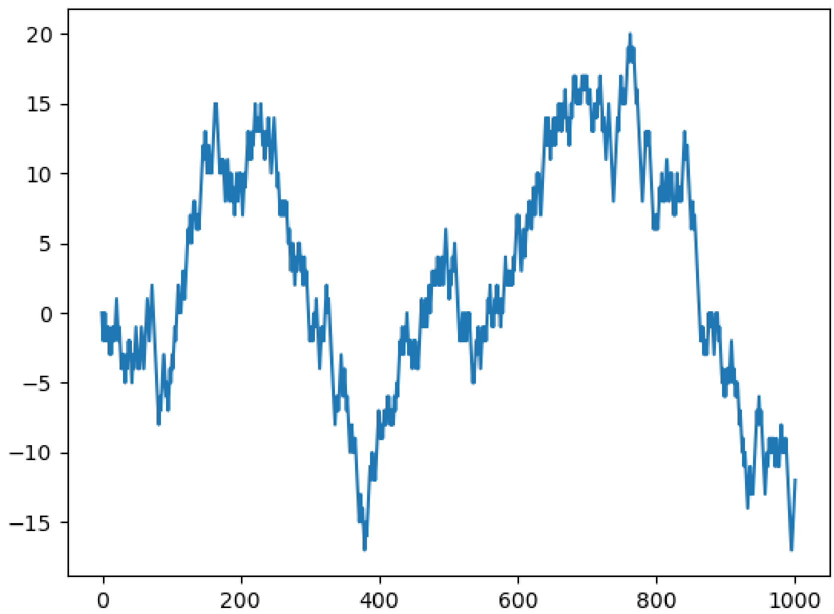
```python
def marche(n,p):
    S = 0
    v = np.zeros(n+1)
    for k in range(1,n+1):
        if rd.random()<=p :
            S = S+1
        else :
```

```
                S = S-1
            v[k] = S
        return v
```

(d) On peut procéder ainsi :

```
p = float(input('Donner p : '))
u = np.arange(1001)
v = marche(1000,p)
plt.plot(u,v)
plt.show()
```

Voici ce qu'on peut obtenir par exemple lorsque $p = 1/2$.



Exercice 14.13 (★★ - Loi du χ^2)

Soit $n \in \mathbb{N}^*$, et $X_1, \ldots, X_n$ n variables aléatoires mutuellement indépendantes, suivant toutes la loi normale centrée réduite. On note $Z = \dfrac{1}{2}\displaystyle\sum_{i=1}^{n} X_i^2$.

1. Montrer que $\dfrac{1}{2}X_1^2$ suit une loi usuelle que l'on déterminera.
2. En déduire la loi de Z.

1. Posons $Y = \frac{1}{2}X_1^2$. On procède par étapes.
   - **Ensemble image de Y.**
     Puisque $X_1(\Omega) = \mathbb{R}$, alors $Y(\Omega) = \mathbb{R}_+$ et pour tout $x < 0$, $F_Y(x) = 0$.
   - **Fonction de répartition de Y.**

Soit $x \geq 0$. Calculons :

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = P(X_1^2 \leq 2x) = P(-\sqrt{2x} \leq X_1 \leq \sqrt{2x}) \\ &= P(X_1 \leq \sqrt{2x}) - P(X_1 < -\sqrt{2x}) = P(X_1 \leq \sqrt{2x}) - P(X_1 \leq -\sqrt{2x}) \quad \text{car } X \text{ continue} \\ &= \Phi(\sqrt{2x}) - \Phi(-\sqrt{2x}) = \Phi(\sqrt{2x}) - (1 - \Phi(\sqrt{2x})) = 2\Phi(\sqrt{2x}) - 1. \end{aligned}$$

Ainsi,

$$F_Y : x \mapsto \begin{cases} 2\Phi(\sqrt{2x}) - 1 & \text{si } x \geq 0, \\ 0 & \text{si } x < 0. \end{cases}$$

- **Y est-elle à densité ?**
  F_Y est de classe $\mathscr{C}^1$ sur $]-\infty, 0[$ et sur $]0, +\infty[$ en tant que composée de fonctions qui le sont. En 0 :
  $$\lim_{x \to 0^-} F_Y(x) = 0$$
  et
  $$\lim_{x \to 0^+} F_Y(x) = F_Y(0) = 2\Phi(0) - 1 = 2 \times \frac{1}{2} - 1 = 0.$$
  Donc F_Y est continue en 0, et donc sur $\mathbb{R}$. Y est bien une variable à densité.
- **Densité de Y.**
  Si $x < 0$, $F_Y'(x) = 0$. Pour $x > 0$, calculons :
  $$\begin{aligned} F_Y'(x) &= 2 \times \frac{2}{2\sqrt{2x}} \times \Phi'(\sqrt{2x}) \\ &= \sqrt{\frac{2}{x}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{2x})^2}{2}} \\ &= \frac{1}{\sqrt{\pi}} x^{-1/2} e^{-x} = \frac{1}{\Gamma(1/2)} x^{1/2-1} e^{-x} \end{aligned}$$
  Une densité de Y est donc par exemple :
  $$f : t \mapsto \begin{cases} \dfrac{1}{\Gamma(1/2)} t^{1/2-1} e^{-t} & \text{si } t > 0, \\ 0 & \text{si } t \leq 0. \end{cases}$$
  On reconnait une densité de la loi $\gamma(1/2)$. Ainsi, Y suit une loi $\gamma(1/2)$.

2. Par lemme de coalition, les variables $\frac{1}{2}X_i^2$ sont indépendantes. Par stabilité de la loi gamma, Z suit donc une loi $\gamma\left(\frac{n}{2}\right)$.

Exercice 14.14 (★★ - Loi binomiale négative)

On considère une expérience ayant probabilité p de réussite, et $1 - p$ d'échouer.
On répète l'expérience jusqu'à obtenir m succès, et on note X le nombre d'essais nécessaires.
On note également X_i la variable aléatoire qui vaut 1 si l'expérience i a été un succès, et 0 sinon.

1. Reconnaitre la loi de X lorsque $m = 1$.
2. Déterminer la loi de X_i.
3. Écrire l'évènement $[X = k]$ en fonction d'évènements formés à partir des variables aléatoires $X_1 + \cdots + X_{k-1}$ et X_k.
4. En déduire la loi de X dans le cas général.
5. En écrivant X comme somme de m variables aléatoires suivant une loi usuelle, déterminer l'espérance de X.

Exercice 14.15 (★★ - Loi Γ à deux paramètres - ✍)
Soient b et ν deux réels strictement positifs. On définit la fonction

$$f : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \dfrac{t^{\nu-1}e^{-\frac{t}{b}}}{b^\nu \Gamma(\nu)} & \text{si } t > 0 \end{cases}$$

1. Montrer que f est une densité de probabilité.

   Dans la suite, on considère une variable aléatoire X admettant pour densité f, et on dit que X suit la loi Γ de paramètres b et ν, et on écrit $X \hookrightarrow \Gamma(b, \nu)$.

2. Quelle est la loi de $\dfrac{1}{b}X$? Quelle est la loi de X si $\nu = 1$?
3. Reconnaitre la loi $\Gamma(1, \nu)$.
4. Montrer que X admet une espérance et une variance, et qu'on a $E(X) = b\nu$ et $V(X) = b^2\nu$.
5. Soient X_1, X_2 deux variables aléatoires réelles indépendantes sur $(\Omega, \mathscr{A}, P)$ telles que $X_1 \hookrightarrow \Gamma(b, \nu_1)$ et $X_2 \hookrightarrow \Gamma(b, \nu_2)$. Montrer que $X_1 + X_2 \hookrightarrow \Gamma(b, \nu_1 + \nu_2)$.
   On pourra commencer par considérer la loi de $\frac{X_1}{b} + \frac{X_2}{b}$.
6. Plus généralement, si $X_1, \ldots, X_n$ sont n variables mutuellement indépendantes telles que $X_i \hookrightarrow \Gamma(b, \nu_i)$, montrer que $X_1 + \cdots + X_n \hookrightarrow \Gamma(b, \nu_1 + \cdots + \nu_n)$.
7. Soient $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes sur (Ω, A, P) suivant la loi exponentielle $\mathscr{E}(\lambda)$. Montrer que $X_1 + \cdots + X_n$ suit la loi $\Gamma(\frac{1}{\lambda}, n)$.
8. Soit Y une variable aléatoire suivant la loi $\mathscr{N}(0, 1)$. Montrer que Y^2 suit une loi Γ dont on précisera les paramètres.

1. f est bien continue sur $\mathbb{R}$ sauf éventuellement en 0, et positive sur $\mathbb{R}$. On cherche à présent à montrer que :
   $$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \frac{1}{b^\nu \Gamma(\nu)} \int_0^{+\infty} t^{\nu-1}e^{-t/b}\,\mathrm{d}t \qquad (*)$$
   converge et vaut 1. C'est l'intégrale d'une fonction continue sur $]0, +\infty[$, elle est donc généralisée en 0 et en $+\infty$.

   Posons pour cela $x = \frac{t}{b}$. C'est un changement de variables affine donc licite. On a $\mathrm{d}t = b\mathrm{d}x$ et $x : 0 \to +\infty$ lorsque $t : 0 \to +\infty$. Par le théorème de changement de variables, l'intégrale $(*)$ est de même nature que l'intégrale :
   $$\int_0^{+\infty} (bx)^{\nu-1}e^{-x}b\mathrm{d}x = b^\nu \int_0^{+\infty} x^{\nu-1}e^{-x}\mathrm{d}x.$$
   On reconnait ici une intégrale $\Gamma(\nu)$ qui converge car $\nu > 0$. Ainsi $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge, et :
   $$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \frac{1}{b^\nu \Gamma(\nu)} b^\nu \Gamma(\nu) = 1.$$
   f est bien une densité de probabilité.
2. $Y = \dfrac{1}{b}X$ est une variable à densité par transformation affine d'une variable à densité, et

une densité de Y est donnée par (formule du cours) :

$$f_Y : t \mapsto bf(bt) = \begin{cases} 0 & \text{si } t \leq 0 \\ \dfrac{1}{\Gamma(\nu)} t^{\nu-1} e^{-t} & \text{si } t > 0 \end{cases}.$$

Ainsi, Y suit une loi $\gamma(\nu)$.

Si $\nu = 1$:

$$f : t \mapsto \begin{cases} 0 & \text{si } t \leq 0 \\ \dfrac{1}{b} e^{-t/b} & \text{si } t > 0 \end{cases}$$

On reconnait une densité d'une loi $\mathscr{E}(\frac{1}{b})$. Donc X suit dans ce cas une loi exponentielle de paramètre $\dfrac{1}{b}$.

3. La loi $\Gamma(1, \nu)$ n'est autre que la loi $\gamma(\nu)$.

4. Utilisons que $Y = \dfrac{1}{b} X$ suit une loi $\gamma(\nu)$.

   Par linéarité de l'espérance, $E(X)$ existe et vaut :

$$E(X) = bE(Y) = b\nu.$$

   Par propriété de la variance, $V(X)$ existe et vaut :

$$V(X) = b^2 V(Y) = b^2 \nu.$$

5. Puisque X_1 et X_2 sont indépendantes, il en est de même de $\dfrac{X_1}{b}$ et de $\dfrac{X_2}{b}$ par le lemme de coalition. De plus, ces variables suivent une loi $\gamma(\nu_1)$ et $\gamma(\nu_2)$. Par stabilité de la loi Gamma, $\dfrac{1}{b}(X_1 + X_2)$ suit une loi $\gamma(\nu_1 + \nu_2) = \Gamma(1, \nu_1 + \nu_2)$. D'après la question 2., $X_1 + X_2$ suit la loi $\Gamma(b, \nu_1 + \nu_2)$.

6. On raisonne par récurrence sur $n \in \mathbb{N}^*$.

   Init. Si $n = 1$, c'est immédiat. Donc la propriété est vraie au rang $n = 1$.

   Hér. Soit $n \geq 1$ et supposons la propriété vraie au rang n. Montrons la au rang $n + 1$.

   Par hypothèse de récurrence, $X_1 + \cdots + X_n$ suit une loi $\Gamma(b, \nu_1 + \cdots + \nu_n)$. Par le lemme de coalition, $X_1 + \cdots + X_n$ est indépendante de X_{n+1}. Enfin en utilisant la question précédente, on en déduit que :

$$X_1 + \cdots + X_n + X_{n+1} = (X_1 + \cdots + X_n) + X_{n+1} \hookrightarrow \Gamma(b, \nu_1 + \cdots + \nu_n + \nu_{n+1}).$$

   On obtient donc le résultat voulu par principe de récurrence.

7. Soient $X_1, \ldots, X_n$ mutuellement indépendantes suivant une loi $\mathscr{E}(\lambda) = \Gamma(\frac{1}{\lambda}, 1)$. Par la question précédente, on en déduit que $X_1 + \cdots + X_n$ suit une loi $\Gamma(\frac{1}{\lambda}, n)$.

8. On a vu à l'Exercice 14.13 que $\dfrac{Y^2}{2}$ suit une loi $\gamma(1/2) = \Gamma(1, \frac{1}{2})$. Par la question 2, Y^2 suit une loi $\Gamma(2, \frac{1}{2})$.

Exercice 14.16 (★★★★ - Oral ESCP 2014)

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, identiquement distribuées, définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$ et suivant la loi exponentielle de paramètre $\lambda > 0$.

Pour tout entier $n \geq 1$, on pose $S_n = \sum_{k=1}^{n} X_k$ et $U_n = \min_{1 \leq k \leq n} X_k$

1. Soit n un entier naturel non nul.
   (a) Déterminer la loi de U_n.
   (b) Donner une densité de S_n. Montrer que pour tout $x > 0$,
$$P(S_n > x) = e^{-\lambda x} \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!}.$$
2. Soit N une variable aléatoire définie sur $(\Omega, \mathscr{A}, P)$, indépendante des (X_i) et qui suit la loi géométrique de paramètre p, avec $0 < p < 1$. On définit S et U par :

   pour tout $\omega \in \Omega$, $S(\omega) = S_{N(\omega)}(\omega)$ et $U(\omega) = U_{N(\omega)}(\omega)$.

   On admet que S et U sont des variables aléatoires.
   (a) Montrer que U est une variable aléatoire à densité et déterminer une densité de U.
   (b) Déterminer la loi de S.

1. (a) On a déjà fait ce calcul (voir cours par exemple) : U_n suit une loi $\mathscr{E}(n\lambda)$.
   (b) Là aussi, le calcul a déjà été fait (voir cours). On montre que $S_n \hookrightarrow \Gamma(\frac{1}{\lambda}, n)$. On en déduit que pour tout $x > 0$:
$$P(S_n > x) = \frac{\lambda^n}{\Gamma(n)} \int_x^{+\infty} t^{n-1} e^{-\lambda t}\,\mathrm{d}t.$$
   Soit $A > x$. On fait une intégration par partie :
$$\begin{array}{c|ccc}
+ & t^{n-1} & & e^{-\lambda t} \\
 & & \searrow & \\
- & (n-1)t^{n-2} & & -\dfrac{e^{-\lambda t}}{\lambda} \\
 & & \searrow & \\
\vdots & \vdots & & \dfrac{e^{-\lambda t}}{\lambda^2} \\
 & & & \\
-1)^{n-1} & (n-1)! & & \vdots \\
 & & \searrow & \\
(-1)^n & 0 & \longleftarrow & (-1)^n \dfrac{e^{-\lambda t}}{\lambda^n}
\end{array}$$
   Les fonctions $t \mapsto t^{n-1}$ et $t \mapsto e^{-\lambda t}$ sont de classe $\mathscr{C}^n$. On obtient donc :
$$\int_x^A t^{n-1} e^{-\lambda t}\,\mathrm{d}t = \left[-t^{n-1}\frac{e^{-\lambda t}}{\lambda} - (n-1)t^{n-2}\frac{e^{-\lambda t}}{\lambda^2} - \cdots - (n-1)!\frac{e^{-\lambda t}}{\lambda^n}\right]_x^A$$
$$\underset{A \to +\infty}{\longrightarrow} \sum_{k=0}^{n-1} \frac{(n-1)!}{(n-k-1)!} x^{n-1-k} \frac{e^{-\lambda x}}{\lambda^{k+1}}$$
   par croissances comparées. On en déduit que :
$$P(S_n > x) = \frac{\lambda^n}{\Gamma(n)} \sum_{k=0}^{n-1} \frac{(n-1)!}{(n-k-1)!} x^{n-1-k} \frac{e^{-\lambda x}}{\lambda^{k+1}} = \sum_{k=0}^{n-1} \frac{\lambda^{n-1-k}}{(n-k-1)!} x^{n-1-k} e^{-\lambda x} = \sum_{i=0}^{n-1} \frac{\lambda^i}{i!} x^i e^{-\lambda x}$$

2. (a) On cherche la fonction de répartition de U. Notons tout d'abord que $U \in]0,+\infty[$ presque surement, et donc que $F_U(x) = 0$ si $x \leq 0$. Soit $x > 0$. On applique la FPT avec le SCE $([N = n])_{n\in\mathbb{N}^*}$ (on pourra toujours procéder de cette manière dans cette situation).

$$\begin{aligned}
F_U(x) &= \sum_{n=1}^{+\infty} P([U \leq x] \cap [N = n]) \\
&= \sum_{n=1}^{+\infty} P([U_n \leq x] \cap [N = n]) \\
&= \sum_{n=1}^{+\infty} P(U_n \leq x)P(N = n) \quad \text{par indép.} \\
&= \sum_{n=1}^{+\infty} (1 - e^{-\lambda nx} P(N = n) \\
&= \sum_{n=1}^{+\infty} P(N = n) - \sum_{n=1}^{+\infty} e^{-\lambda nx}(1 - p)^{n-1}p \\
&= 1 - e^{-\lambda x}p \sum_{n=1}^{+\infty} (e^{-\lambda x}(1 - p))^{n-1} \\
&= 1 - e^{-\lambda x}p \frac{1}{1 - e^{-\lambda x}(1 - p)} \\
&= \frac{1 - e^{-\lambda x}}{1 - e^{-\lambda x}(1 - p)}
\end{aligned}$$

Ainsi la fonction de répartition de U est donnée par :

$$F_U : x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ \dfrac{1 - e^{-\lambda x}}{1 - e^{-\lambda x}(1 - p)} & \text{si } x > 0 \end{cases}$$

On vérifie que cette fonction de répartition est continue sur $\mathbb{R}$ (il faut étudier notamment ce qui se passe en 0), et de classe $\mathscr{C}^1$ sur $\mathbb{R}$ privé de 0. Donc U est à densité.

Pour tout $x \neq 0$:

$$F_U'(x) = \begin{cases} 0 & \text{si } x < 0, \\ \dfrac{\lambda e^{-\lambda x}(1 - e^{-\lambda x}(1 - p)) + (1 - e^{-\lambda x})(1 - p)\lambda e^{-\lambda x}}{(1 - e^{-\lambda x}(1 - p))^2} & \text{si } x > 0. \end{cases}$$

Une densité de U est donc :

$$f : x \mapsto \begin{cases} 0 & \text{si } x \leq 0, \\ \dfrac{\lambda e^{-\lambda x}((2 - p) - (2 - 2p)e^{-\lambda x})}{(1 - e^{-\lambda x}(1 - p))^2} & \text{si } x > 0. \end{cases}$$

(b) Puisque $S > 0$ presque surement, $F_S(x) = 0$ pour tout $x \leq 0$. Supposons $x > 0$. Par la FPT :

$$\begin{aligned}
P(S > x) &= \sum_{n=1}^{+\infty} P(S > x,\ N = n) = \sum_{n=1}^{+\infty} P(S_n > x,\ N = n) \\
&= \sum_{n=1}^{+\infty} P(S_n > x)P(N = n) \quad \text{par indépendance} \\
&= e^{-\lambda x} p \sum_{n=1}^{+\infty} \left(\sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!} \right) (1-p)^{n-1} \\
&= e^{-\lambda x} p \sum_{n=1}^{+\infty} \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!} (1-p)^{n-1}
\end{aligned}$$

La série double considérée est à termes positifs. De plus, elle converge. Elle converge donc absolument, et on peut appliquer Fubini.

$$\begin{aligned}
P(S > x) &= e^{-\lambda x} p \sum_{k=0}^{+\infty} \sum_{n=k+1}^{+\infty} \frac{\lambda^k x^k}{k!} (1-p)^{n-1} = e^{-\lambda x} p \sum_{k=0}^{+\infty} \frac{\lambda^k x^k}{k!} \sum_{n=k+1}^{+\infty} (1-p)^{n-1} \\
&= e^{-\lambda x} p \sum_{k=0}^{+\infty} \frac{\lambda^k x^k}{k!} (1-p)^k \frac{1}{1-(1-p)} = e^{-\lambda x} \sum_{k=0}^{+\infty} \frac{\lambda^k x^k}{k!} (1-p)^k \\
&= e^{-\lambda x} e^{\lambda x(1-p)} = e^{-\lambda p x}.
\end{aligned}$$

Finalement, on obtient que $P(S \leq x) = 1 - e^{-\lambda p x}$. S suit donc une loi $\mathscr{E}(\lambda p)$.

Exercice 14.17 (★★★★ - QSP ESCP 2014)

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, mutuellement indépendantes et suivant toutes la loi géométrique de paramètre p.

Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \displaystyle\sum_{k=1}^{n} X_k$.

1. Soit $n \in \mathbb{N}^*$. Montrer que la variable aléatoire $\dfrac{1}{S_n}$ admet une espérance, que l'on notera par la suite m_n.
2. Soient $(k, n) \in \mathbb{N} \times \mathbb{N}^*$. Déterminer l'espérance de $\dfrac{S_k}{S_n}$ en fonction de m_n.

1. Puisque les X_i suivent des lois géométriques, il suit que $S_n \geq 1$ presque surement. Ainsi $0 \leq \dfrac{1}{S_n} \leq 1$ presque surement, et $E(\frac{1}{S_n})$ existe par théorème d'existence de l'espérance par domination.
2. On étudie plusieurs cas :
   - Cas $k = n$. Dans ce cas $\dfrac{S_k}{S_n} = 1$ et son espérance existe et vaut 1.
   - Cas $k \geq n$. Dans ce cas :

$$\frac{S_k}{S_n} = 1 + \frac{X_{n+1} + \cdots + X_k}{S_n}.$$

Par lemme de coalition, les variables $X_{n+1} + \cdots + X_k$ et $\dfrac{1}{S_n}$ sont indépendantes, et admettent toutes les deux une espérance. Par produit de variables indépendantes,

$\dfrac{X_{n+1}+\cdots+X_k}{S_n}$ admet une espérance, et :

$$E\left(\frac{X_{n+1}+\cdots+X_k}{S_n}\right)=E(X_{n+1}+\cdots+X_k)E(\frac{1}{S_n})=\frac{k-n}{p}m_n$$

L'espérance de $\dfrac{S_k}{S_n}$ existe donc, et vaut $1+\dfrac{k-n}{p}m_n$.

- Cas $k<n$. Pour tout $i\in[\![1,k]\!]$, commençons par remarquer que les variables $\dfrac{X_i}{S_n}$ suivent la même loi. Pour le voir, notons d'abord que les vecteurs $(X_1,\ldots,X_i,\ldots X_n)$ et $(X_i,\ldots,X_1,\ldots X_n)$ ont même loi (quitte à se ramener à la définition de la loi d'un vecteur du cours). Comme la fonction $g\,:\,(x_1,\ldots,x_n)\mapsto\dfrac{x_1}{x_1+\cdots+x_n}$ est continue sur $(\mathbb{R}_+^*)^n$ (cette notion sera précisée dans un prochain chapitre), on en déduit (propriété du cours) que :

$$g(X_1,\ldots,X_i,\ldots,X_n)=\frac{X_1}{X_1+\cdots+X_n}\quad\text{et}\quad g(X_i,\ldots,X_1,\ldots,X_n)=\frac{X_i}{X_1+\cdots+X_n}$$

ont même loi pour tout i.

De plus, pour tout $i\in[\![1,k]\!]$, les variables $\dfrac{X_i}{S_n}$ admettent une espérance puisque $0\leq\dfrac{X_i}{S_n}\leq 1$ (existence de l'espérance par domination). Par linéarité de l'espérance :

$$\begin{aligned}1=E\left(\frac{S_n}{S_n}\right)&=\sum_{i=1}^{n}E\left(\frac{X_i}{S_n}\right)\\&=\sum_{i=1}^{n}E\left(\frac{X_1}{S_n}\right)\qquad\text{car elles sont de même loi}\\&=nE\left(\frac{X_1}{S_n}\right)\end{aligned}$$

Donc pour tout $i\in[\![1,n]\!]$, $E\left(\frac{X_i}{S_n}\right)=\dfrac{1}{n}$, et par linéarité de l'espérance, on obtient :

$$E\left(\frac{S_k}{S_n}\right)=\sum_{i=1}^{k}E\left(\frac{X_i}{S_n}\right)=\frac{k}{n}$$