DEVOIR MAISON # 5

à rendre le Jeudi 03/03/2022

Consignes La qualité de la rédaction, la clarté et la précision des raisonnements, entreront pour une part importante dans l'appréciation de la copie. Les abréviations, sigles ou phrases nominales sont à proscrire. La numérotation des exercices (et des questions) doit être respectée et mise en évidence. Les résultats doivent être encadrés proprement.

Vous avez la possibilité de rendre un devoir pour deux, mais les écritures doivent alors apparaître en parts égales. Il est rappelé également que recopier une correction sur internet <u>est complètement inutile pour tout le monde.</u>

Problème 1 | G2E 2015 Dans tout le problème λ désigne un réel strictement positif, n un entier naturel, p un entier naturel supérieur ou égal à 2 et $k \in [\![1, p]\!]$. Si X désigne une variable aléatoire à densité alors F_X désigne la fonction de répartition de X, f_X une densité de X et on note $\overline{F_X}$ la fonction d'antirépartition de X définie sur $\mathbf{R}$ par

$$\forall t \in \mathbf{R}, \quad \overline{F_X}(t) = 1 - F_X(t).$$

Dans la partie I, on étudie quelques propriétés classiques des lois exponentielles. La partie II est consacrée au calcul d'une limite d'une probabilité conditionnelle. Enfin, dans la partie III, on étudie la somme de certaines variables aléatoires. Ces parties peuvent être abordées indépendamment les unes des autres.

Partie I — Des lois exponentielles

On considère p techniciens qui interviennent auprès d'une machine. Pour tout $k \in [\![1, p]\!]$, on note X_k la variable aléatoire donnant la durée (en heure (s)) de l'intervention du k-ième technicien. On admet que les variables aléatoires $(X_k)_{k\in [\![1,p]\!]}$ sont mutuellement indépendantes et suivent la même loi exponentielle de paramètre λ.

1. **1.1)** Démontrer par récurrence sur n que la variable aléatoire X admet un moment à tout ordre n égal à :

$$m_n(X_k) = \frac{n!}{\lambda^n}.$$

1.2) Déterminer la nature et la somme de la série de terme général $\left(\sum \frac{1}{m_n(X_k)}\right)_{n\geqslant 0}$.

2. On note $Y_p = \min\left(X_1, \dots, X_p\right)$.

2.1) Calculer $\overline{F}_{X_1}(t)$ pour tout $t \in \mathbf{R}$.

2.2) En considérant $\overline{F}_{X_p}(t)$ pour tout $t \in \mathbf{R}$, démontrer que la variable aléatoire Y_p, suit une loi exponentielle de paramètre λp.

2.3) Calculer $\mathbf{E}\left(Y_p\right)$ et $\mathbf{Var}\left(Y_p\right)$ en fonction de λ.

3. On suppose, dans cette question uniquement, que $p = 2$ et que la durée moyenne de l'intervention du premier technicien est de deux heures. Calculer alors les probabilités ci-dessous.

3.1) probabilité que la durée de l'intervention du second technicien soit inférieure ou égale à 2 heures.

3.2) probabilité que le minimum des durées des interventions des deux techniciens soit inférieur ou égal à 1 heure sachant que la durée de l'intervention du second technicien est supérieure ou égale à 2.

3.3) probabilité que le minimum des durées des interventions des deux techniciens soit inférieur ou égal à 1 heure sachant que la durée de l'intervention du second technicien est inférieure ou égale à 2.

Partie II — Calcul d'une limite de probabilité

4. On note $Z_2 = \max(X_1, X_2)$.

4.1) Démontrer que la variable aléatoire Z_2 admet une densité que l'on exprimera à l'aide de densités de lois exponentielles.

4.2) Démontrer que la variable aléatoire Z_2 admet une espérance et obtenir une égalité entre $\mathbf{E}(Y_2) + \mathbf{E}(Z_2)$ et $\mathbf{E}(X_1 + X_2)$. Pouvait-on anticiper cette égalité ?

4.3) Justifier enfin l'existence de $\mathbf{Var}(Z_2)$ et la calculer.

5. Soient $x \in \mathbf{R}^+$ et $n \in \mathbf{N}$.

5.1) Calculer $\mathbf{P}(Z_2 \geqslant n+x)$ et démontrer que $\mathbf{P}(Z_2 \geqslant n+x)$ est équivalent (quand n tend vers $+\infty$) au terme général d'une suite géométrique dont on précisera la raison et le terme initial.

5.2) Démontrer que :

$$\lim_{n\to+\infty} \mathbf{P}\left(Z_2 \geqslant n+x \mid Z_2 \geqslant n\right) = \mathbf{P}\left(X_1 \geqslant x\right).$$

5.3) La formule précédente est-elle encore exacte si on remplace la variable aléatoire Z_2 par la variable aléatoire X_1 ?

Partie III — Sommes de variables aléatoires

6. Soit $(a,b) \in \mathbf{R}_+^* \backslash \{1\} \times \mathbf{R}$.

6.1) Démontrer que la variable aléatoire $aX_2 + b$ admet une densité définie par :

$$\forall t \in \mathbf{R}, \quad f_{aX_2+b}(t) = \begin{cases} \frac{\lambda}{a} e^{-\frac{\lambda}{a}(t-b)} & \text{si } t \geqslant b, \\ 0 & \text{sinon.} \end{cases}$$

6.2) Démontrer que les variables aléatoires X_1 et aX_2+b sont indépendantes.

7. On note $T = X_1 + aX_2 + b$ et on rappelle que si X et Y sont deux variables aléatoires indépendantes de densités f_X et f_Y alors $X+Y$ est une variable aléatoire dont une densité est définie par :

$$\forall x \in \mathbf{R}, \quad f_{X+Y}(x) = \int_{-\infty}^{+\infty} f_X(x-t) f_Y(t)\,dt = \int_{-\infty}^{+\infty} f_X(t) f_Y(x-t)\,dt.$$

7.1) Démontrer que T admet une densité définie par :

$$\forall x \in \mathbf{R}, \quad f_T(x) = \begin{cases} \frac{\lambda}{1-a}\left(e^{-\lambda(x-b)} - e^{-\frac{\lambda}{a}(x-b)}\right) & \text{si } x \geqslant b, \\ 0 & \text{sinon.} \end{cases}$$

7.2) Démontrer qu'il existe un unique couple $(a,b) \in \mathbf{R}_+^* \backslash \{1\} \times \mathbf{R}$ tel que Z_2 et T suivent la même loi. On rappelle que Z_2 a été définie dans la partie II.

8. On prolonge l'étude précédente et on considère la variable aléatoire :

$$T_p = \sum_{k=1}^{p} \frac{1}{k} X_k.$$

8.1) Calculer $\mathbf{E}\left(T_p\right)$ et $\mathbf{Var}\left(T_p\right)$. En déduire la nature (convergence ou divergence) des suites $\left(\mathbf{E}\left(T_p\right)\right)$ et $\left(\mathbf{Var}\left(T_p\right)\right)$.

8.2) Démontrer par récurrence que T_p admet une densité définie par :

$$\forall x \in \mathbf{R}, \quad f_{T_p}(x) = \begin{cases} \lambda p e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{p-1} & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases}$$

Solution (problème 1)

1. **1.1)** ■ **Initialisation.** Pour $n = 1$, le résultat est vrai car $m_1(X_k) = \frac{1}{\lambda}$, on reconnaît bien ici l'espérance d'une loi exponentielle de paramètre λ.

■ **Hérédité.** Supposons-le vrai au rang n, *i.e.* que

$$m_n(X_k) = \frac{n!}{\lambda^n}.$$

Alors d'après le théorème de transfert, il s'agit ensuite d'étudier la convergence absolue de

$$\int_0^{\infty} t^{n+1}\left(\lambda e^{-\lambda t}\right) dt.$$

L'intégrale est généralisée en $+\infty$, soit donc $A > 0$. Alors en faisant une intégration par parties, on obtient puisque les fonctions $t \longmapsto -e^{-\lambda t}$ et $t \longmapsto t^{n+1}$ sont $\mathscr{C}^1$:

$$\begin{aligned} \int_0^{A} t^{n+1}\left(\lambda e^{-\lambda t}\right) dt &= \int_0^{A} (n+1)t^n\left(e^{-\lambda t}\right) dt - \left[t^{n+1}e^{-\lambda t}\right]_0^A, \\ &\xrightarrow{A\to\infty} (n+1)\frac{1}{\lambda}\int_0^{\infty} (n+1)t^n\left(\lambda e^{-\lambda t}\right) dt, \\ &= \frac{(n+1)n!}{\lambda^{n+1}} = \frac{(n+1)!}{\lambda^{n+1}} = . \end{aligned}$$

croissances comparées

hyp. de récurrence

D'où le résultat par principe de récurrence. Pour tout $n \geqslant 0$,

$$\boxed{m_n(X_k) = \frac{n!}{\lambda^n}}.$$

1.2) La série $\left(\sum \frac{1}{m_n(X_k)}\right)_{n\geqslant 0}$ $\boxed{\text{converge et est de somme } e^{\lambda}}$ car on reconnaît directement une série exponentielle.

2. On note $Y_p = \min\left(X_1,\dots,X_p\right)$.

2.1) Calculons $\overline{F}_{X_1}(t)$ pour tout $t \in \mathbf{R}$. C'est une conséquence directe du cours, puisque

$$F_{X_1}(t) = \begin{cases} 1-e^{-\lambda t} & \text{si } t \geqslant 0,\\ 0 & \text{sinon.}\end{cases}$$

Donc

$$\boxed{\overline{F}_{X_1}(t) = \begin{cases} e^{-\lambda t} & \text{si } t \geqslant 0,\\ 1 & \text{sinon.}\end{cases}}$$

2.2) Soit $t \in \mathbf{R}$, alors

$$\begin{aligned}\mathbf{P}\left(Y_p > t\right) &= \mathbf{P}\left(\min\left(X_1,\dots,X_p\right) > t\right)\\ &= \mathbf{P}\left(X_1 > t,\dots,X_p > t\right),\\ &= \mathbf{P}\left(X_1 > t\right)^p \quad \textit{indépendance et même loi}\\ &= \begin{cases} e^{-\lambda p t} & \text{si } t \geqslant 0,\\ 1 & \text{sinon.}\end{cases}\end{aligned}$$

Donc

$$F_{Y_p}(t) = \begin{cases} 1-e^{-\lambda p t} & \text{si } t \geqslant 0,\\ 0 & \text{sinon.}\end{cases}$$

La fonction F_{Y_p} est continue car $1-e^{-\lambda p 0} = 0$ et $\mathscr{C}^1$ sauf éventuellement en 0. Donc Y_p est à densité et une densité est donnée par

$$f_{Y_p}(t) = \begin{cases} \lambda p e^{-\lambda p t} & \text{si } t \geqslant 0,\\ 0 & \text{sinon.}\end{cases}$$

On reconnaît ici une densité exponentielle. En conclusion : $\boxed{Y_p, \text{ suit une loi exponentielle de paramètre } \lambda p.}$

2.3) Ainsi $\boxed{\mathbf{E}\left(Y_p\right) = \frac{1}{\lambda p} \text{ et } \mathbf{Var}\left(Y_p\right) = \frac{1}{\lambda^2 p^2}.}$

3. On suppose, dans cette question uniquement, que $p = 2$ et que la durée moyenne de l'intervention du premier technicien est de deux heures. Calculer alors les probabilités ci-dessous.

3.1) $\mathbf{P}\left(X_2 \leq 2\right) = 1 - e^{-2\lambda} = 1 - \frac{1}{e} \approx \boxed{0.632,2}$

3.2)
$$\begin{aligned}\mathbf{P}(Y_2 \leq 1 | X_2 \geq 2) &= \frac{\mathbf{P}\left(\min(X_1,X_2) \leq 1, X_2 \geq 2\right)}{\mathbf{P}(X_2 \geq 2)},\\ &= \frac{\mathbf{P}\left(X_1 \leq 1, X_2 \geq 2\right)}{\mathbf{P}(X_2 \geq 2)}\\ &= \frac{\mathbf{P}\left(X_1 \leq 1\right)(\mathbf{P}(X_2 \geq 2)}{\mathbf{P}(X_2 \geq 2)}\\ &= \mathbf{P}(X_1 \leq 1) = 1 - e^{-1/2}\\ &\approx \boxed{0.393}.\end{aligned}$$

3.3)
$$\begin{aligned}\mathbf{P}(Y_2 \leq 1 | X_2 \leq 2) &= 1 - \mathbf{P}(Y_2 > 1 | X_2 \leq 2)\\ &= 1 - \frac{\mathbf{P}\left(\min(X_1,X_2) > 1, X_2 \leq 2\right)}{\mathbf{P}(X_2 \leq 2)},\\ &= \frac{\mathbf{P}\left(X_1 > 1, X_2 > 1, X_2 \leq 2\right)}{\mathbf{P}(X_2 \leq 2)},\\ &= \frac{\mathbf{P}\left(X_1 > 1\right)(\mathbf{P}(1 < X_2 \leq 2)}{\mathbf{P}(X_2 \leq 2)},\\ &= 1 - \frac{\left(1-e^{-1/2}\right)\left(e^{-1/2}-e^{-1}\right)}{1-e^{-1}} = \boxed{0.85145}\end{aligned}$$

Partie I — Calcul d'une limite de probabilité

4. On note $Z_2 = \max\left(X_1, X_2\right)$.

4.1) Cette fois-ci, on calcule la fonction de répartition de Z_2. Soit donc $t \in \mathbf{R}$,

$$\begin{aligned}\mathbf{P}(Z_2 \leqslant t) &= \mathbf{P}(Y_1 \leqslant t, Y_2 \leqslant t) \quad \textit{indépendance}\\ &= \mathbf{P}(Y_1 \leqslant t)\mathbf{P}(Y_2 \leqslant t)\\ &= \mathbf{P}(Y_1 \leqslant t)^2\\ &= \begin{cases}\left(1-e^{-\lambda t}\right)^2 & \text{si } t \geqslant 0,\\ 0 & \text{si } t < 0.\end{cases}\end{aligned}$$

La fonction ainsi trouvée est continue car $\left(1-e^{-\lambda 0}\right)^2 = 1^2 = 1$. De plus, elle est $\mathscr{C}^1$ sauf éventuellement en 0, donc Z_2 est à densité et une densité est donnée par :

$$f_{Z_2}(t) = \begin{cases}2\lambda e^{-\lambda t}\left(1-e^{-\lambda t}\right) & \text{si } t \geqslant 0,\\ 0 & \text{si } t < 0.\end{cases}$$

En rentrant le terme exponentielle dans la parenthèse, on obtient

$$\boxed{f_{Z_2} = 2f_\lambda - f_{2\lambda}},$$

où f_λ désigne une densité de $\mathscr{E}(\lambda)$ pour tout $\lambda \geqslant 0$.

4.2) La variable aléatoire Z_2 admet une espérance si et seulement si l'intégrale

$$\int_0^{\infty} t\,(2f_\lambda - f_{2\lambda})(t)\,dt \text{ converge.}$$

Ceci est bien le cas puisque les intégrales ci-après convergent d'après le cours :

$$2\int_0^{\infty} t f_\lambda(t)\,dt = \frac{2}{\lambda}, \qquad \int_0^{\infty} t f_{2\lambda}(t)\,dt = \frac{1}{2\lambda}.$$

Ainsi,

$$\mathbf{E}(Z_2) = \frac{2}{\lambda} - \frac{1}{2\lambda} = \boxed{\frac{3}{2\lambda}}.$$

Pouvait-on anticiper cette égalité ? Oui, car $Y_2 + Z_2 = X_1 + X_2$. Donc $Z_2 = X_1 + X_2 - Y_2$ admet donc une espérance en tant que somme de telles variables aléatoires, et elle vaut :

$$\mathbf{E}(Z_2) = \mathbf{E}(X_1) + \mathbf{E}(X_2) - \mathbf{E}(Y_2) = \frac{1}{\lambda} + \frac{1}{\lambda} - 2\lambda = \frac{3}{2\lambda}.$$

4.3) On étudie la convergence de

$$\int_0^{\infty} t^2\,(2f_\lambda - f_{2\lambda})(t)\,dt \text{ converge.}$$

Comme précédemment, en développant l'intégrande on constate que toutes les intégrales apparaissant sont convergentes, et on a

$$\int_0^{\infty} t^2\,(2f_\lambda - f_{2\lambda})(t)\,dt = 2\left(\mathbf{Var}(X_1) + \mathbf{E}(X_1)^2\right) - \left(\frac{1}{2^2\lambda^2} + \frac{1}{4\lambda^2}\right) = \frac{7}{2\lambda^2}.$$

On obtient alors

$$\mathbf{Var}(Z_2) = \mathbf{E}(Z_2^2) - \mathbf{E}(Z_2)^2 == \frac{7}{2\lambda^2} - \frac{9}{4\lambda^2} = \boxed{\frac{5}{4\lambda^2}.}$$

5. Soient $x \in \mathbf{R}^+$ et $n \in \mathbf{N}$.

5.1) $\mathbf{P}(Z_2 \geq n+x) = 1 - F_{Z_2}(n+x) = 1 - \left(1-e^{-\lambda(n+x)}\right)^2$. Ainsi

$$\mathbf{P}(Z_2 \geq n+x) = 2e^{-\lambda(n+x)} - e^{-2\lambda(n+x)} \underset{n\to\infty}{\sim} 2e^{-\lambda(n+x)} = \boxed{2e^{-\lambda x}\left(e^{-\lambda}\right)^n}.$$

L'équivalent trouvé est bien une suite géométrique de premier terme $2e^{-\lambda x}$ et de raison $e^{-\lambda}$.

5.2) Nous avons :

$$\mathbf{P}(Z_2 \geqslant n+x \mid Z_2 \geqslant n) = \frac{\mathbf{P}(Z_2 \geq n+x, Z_2 \geq n)}{\mathbf{P}(Z_2 \geq n)} = \frac{\mathbf{P}(Z_2 \geq n+x)}{\mathbf{P}(Z_2 \geq n)} \sim \frac{2e - \lambda(n+x)}{2e^{-\lambda n}} = e^{-\lambda x}.$$

Donc :

$$\boxed{\lim_{n\to+\infty} \mathbf{P}(Z_2 \geqslant n+x \mid Z_2 \geqslant n) = \mathbf{P}(X_1 \geqslant x)}.$$

5.3) Oui, car X_1 est sans mémoire, donc $\mathbf{P}(X_1 \geqslant x+y | X_1 \geqslant x) = \mathbf{P}(X_1 \geqslant x)$.

6. Soit $(a,b) \in \mathbf{R}_+^* \setminus \{1\} \times \mathbf{R}$.

6.1) Soit $t \in \mathbf{R}$. Alors :

$$\begin{aligned}\mathbf{P}\left(aX_2 + b \leqslant t\right) &= \mathbf{P}\left(X_2 \leqslant \frac{t-b}{a}\right)\\ &= \begin{cases} 1 - e^{-\lambda \frac{t-b}{a}} & \text{si } \frac{t-b}{a} \geqslant 0,\\ 0 & \text{sinon.}\end{cases}\\ &= \begin{cases} 1 - e^{-\lambda \frac{t-b}{a}} & \text{si } t \geqslant b,\\ 0 & \text{sinon.}\end{cases}\end{aligned}$$

Puisque $1 - e^{-\lambda \frac{t-b}{a}}\Big|_{t=b} = 1 - 1 = 0$, la fonction F_{X_2} est bien continue, elle est de plus $\mathscr{C}^1$ sauf éventuellement en b. Donc $aX_2 + b$ est à densité et une densité est donnée par la dérivée de la fonction de répartition :

$$\forall t \in \mathbf{R}, \quad f_{aX_2+b}(t) = \begin{cases} \frac{\lambda}{a} e^{-\frac{\lambda}{a}(t-b)} \text{ si } t \geqslant b,\\ 0 \text{ sinon.}\end{cases}$$

6.2) Conséquence directe du fait que X_1 et X_2 sont indépendantes, d'où $\boxed{X_1 \text{ et } aX_2 + b \text{ sont indépendantes.}}$

7. On note $T = X_1 + aX_2 + b$ et on rappelle que si X et Y sont deux variables aléatoires indépendantes de densités f_X et f_Y alors $X + Y$ est une variable aléatoire dont une densité est définie par :

$$\forall x \in \mathbf{R}, \quad f_{X+Y}(x) = \int_{-\infty}^{+\infty} f_X(x-t) f_Y(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} f_X(t) f_Y(x-t)\,\mathrm{d}t.$$

7.1) Soit $x \in \mathbf{R}$, alors :

$$\begin{aligned} f_T(x) &= f_{X_1+(aX_2+b)}(x)\\ &= \int_{-\infty}^{+\infty} f_{X_1}(x-t) f_{aX_2+b}(t)\,\mathrm{d}t\\ &= \begin{cases} \displaystyle\int_b^x \lambda e^{-\lambda(x-t)} \frac{\lambda}{a} e^{-\frac{\lambda}{a}(t-b)}\,\mathrm{d}t \text{ si } x \geqslant b,\\ 0 \text{ sinon.}\end{cases}\\ &= \begin{cases} \frac{\lambda^2}{a} e^{\lambda\left(\frac{b}{a} - x\right)} \displaystyle\int_b^x e^{\lambda t\left(1 - \frac{1}{a}\right)}\,\mathrm{d}t \text{ si } x \geqslant b,\\ 0 \text{ sinon.}\end{cases}\end{aligned}$$

$f_{X_1}(x-t) f_{aX_2+b}(t) \neq 0$ si et seulement si $x - t \geqslant 0, t \geqslant b$

Après calcul de l'intégrale, une densité de T est alors :

$$\boxed{\forall x \in \mathbf{R}, \quad f_T(x) = \begin{cases} \frac{\lambda}{1-a}\left(e^{-\lambda(x-b)} - e^{-\frac{\lambda}{a}(x-b)}\right) & \text{si } x \geqslant b,\\ 0 & \text{sinon.}\end{cases}}$$

7.2) Les variables aléatoires Z_2 et T suivent la même loi si et seulement si elles ont même densité sauf en un nombre fini de points. On veut donc que, sauf en un nombre fini de points, l'on ait :

$$\begin{cases} \frac{\lambda}{1-a}\left(e^{-\lambda(x-b)} - e^{-\frac{\lambda}{a}(x-b)}\right) & \text{si } x \geqslant b,\\ 0 & \text{sinon.}\end{cases} = \begin{cases} 2\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right) & \text{si } x \geqslant 0,\\ 0 & \text{sinon.}\end{cases}$$

Nécessairement, il faut $\boxed{b = 0}$. On doit donc avoir

$$\begin{cases} \frac{\lambda}{1-a}\left(e^{-\lambda x} - e^{-\frac{\lambda}{a}x}\right) & \text{si } x \geqslant 0,\\ 0 & \text{sinon.}\end{cases} = \begin{cases} 2\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right) & \text{si } x \geqslant 0,\\ 0 & \text{sinon.}\end{cases}$$

En particulier, si $x \geqslant 0$,

$$\frac{\lambda}{1-a}\left(e^{-\lambda x} - e^{-\frac{\lambda}{a}x}\right) = 2\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right).$$

Cette relation est équivalente à

$$1 - e^{-\frac{\lambda}{a}x + \lambda x} = 2(1-a)\left(1 - e^{-\lambda x}\right).$$

Donc en faisant $x \longrightarrow \infty$, on obtient nécessairement que $1 = 2(1-a)$ et $\frac{\lambda}{a} - \lambda > 0$ (pour que le terme de gauche tende bien vers 1, comme celui de droite). Ce qui fournit $a = \frac{1}{2}$. La seule solution pouvant convenir est donc $\boxed{a = \frac{1}{2}, b = 0}$.

Inversement, on vérifie sans difficulté que $a = \frac{1}{2}, b = 0$, alors $f_{Z_2} = f_T$.

8. On prolonge l'étude précédente et on considère la variable aléatoire :

$$T_p = \sum_{k=1}^{p} \frac{1}{k} X_k.$$

8.1) Par linéarité de l'espérance, $\mathbf{E}\left(T_p\right) = \sum_{k=1}^{p} \frac{1}{k} \mathbf{E}\left(X_k\right) = \boxed{\frac{1}{\lambda} \sum_{k=1}^{p} \frac{1}{k}}$. et par indépendance et propriété de la variance,

$$\begin{aligned}
\mathbf{Var}\left(T_p\right) &= \mathbf{Var}\left(\sum_{k=1}^{p} \frac{1}{k} X_k\right) \quad \textit{indépendance} \\
&= \sum_{k=1}^{p} \frac{1}{k^2} \mathbf{Var}\left(X_k\right) \\
&= \boxed{\frac{1}{\lambda^2} \sum_{k=1}^{p} \frac{1}{k^2}.}
\end{aligned}$$

D'après le cours sur les séries, puisque $\left(\sum \frac{1}{k}\right)_{k \geqslant 1}$ diverge, alors $\left(\mathbf{E}\left(T_p\right)\right)$ diverge vers $+\infty$. De-même, d'après le cours, $\left(\sum \frac{1}{k^2}\right)_{k \geqslant 1}$ donc $\left(\mathbf{Var}\left(T_p\right)\right)$ converge. *On peut montrer que la limite est* $\frac{\pi^2}{6\lambda^2}$

8.2) Afin de moins souffrir dans cette question, nous allons maximiser l'utilisation des questions précédentes. D'après la question 6.1), on sait que

$$f_{\frac{1}{k}X_k}(x) = \begin{cases} k\lambda e^{-k\lambda x} & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases}$$

La formule pour T_p s'obtient en itérant les produits de convolutions. Montrons cela par récurrence sur p.

■ **Initialisation.** Pour $p = 1$, on retrouve la densité calculée précédemment avec $k = 1$.

■ **Hérédité.** Supposons l'expression de f_{T_p} vraie au rang p. Alors constatons que $T_{p+1} = T_p + \frac{X_{p+1}}{p+1}$ et que T_p et $\frac{X_{p+1}}{p+1}$ sont indépendantes. Une densité de $\frac{X_{p+1}}{p+1}$ est, encore une fois, donnée par la question 6.1), on la note f_{p+1} :

$$f_{\frac{1}{p+1}X_{p+1}}(x) = \begin{cases} (p+1)\lambda e^{-(p+1)\lambda x} & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases}$$

Donc

$$\begin{aligned}
f_{T_{p+1}}(x) &= \int_{-\infty}^{\infty} f_{T_p(t)} f_{\frac{1}{p+1}X_{p+1}}(x-t)\, dt, \\
&= \begin{cases} \displaystyle\int_0^x \lambda p e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{p-1} (p+1)\lambda e^{-(p+1)\lambda(x-t)}\, dt & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases} \\
&= \begin{cases} \displaystyle\lambda^2 p(p+1) e^{-\lambda(p+1)x} \int_0^x e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{p-1} e^{(p+1)\lambda t}\, dt & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases} \\
&= \begin{cases} \displaystyle\lambda^2 p(p+1) e^{-\lambda(p+1)x} \int_0^x e^{\lambda t}\left(e^{\lambda t} - 1\right)^{p-1} dt & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases}
\end{aligned}$$

$f_{T_p(t)} f_{\frac{1}{p+1}X_p}$ ... $t) \neq 0$ si et seulement si $t \geqslant 0, x - t \geqslant 0$

Enfin, l'intégrale apparaissant se calcule :

$$\int_0^x e^{\lambda t}\left(e^{\lambda t} - 1\right)^{p-1} dt = \frac{\left(e^{\lambda x} - 1\right)^p}{\lambda p}.$$

Donc si $x \geqslant 0$:

$$f_{T_{p+1}}(x) = \lambda(p+1) e^{-\lambda(p+1)x}\left(e^{\lambda x} - 1\right)^p = \lambda(p+1) e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^p.$$

Donc une densité de T_p est :

$$\boxed{\forall x \in \mathbf{R}, \quad f_{T_p}(x) = \begin{cases} \lambda p e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{p-1} & \text{si } x \geqslant 0, \\ 0 \text{ sinon.} \end{cases}}$$

Problème 2 | On s'intéresse dans cet exercice à l'étude d'un jeu dans une fête foraine. On admet qu'il existe un espace probabilisé $(\Omega, \mathscr{T}, \mathbf{P})$ modélisant les expériences aléatoires du problème et sur lequel sont définies les variables aléatoires.

Pour ce jeu, le participant lance successivement n boules au hasard dans N cases numérotées de 1 à N avec $N \geqslant 2$. On suppose que les différents lancers sont indépendants et que la probabilité pour qu'une boule tombe dans une case donnée est $\frac{1}{N}$ (toutes les boules finissent donc dans une case de manière équiprobable). Une case peut contenir plusieurs boules.
Le gain étant fonction du nombre de cases atteintes, on étudie la variable aléatoire T_n égale au nombre de cases non vides à l'issue de ces n lancers.

1. **(Modélisation Informatique)**
   - **1.1)** Écrire une fonction Python `Liste` qui prend en argument les nombres entiers n et N et renvoie une liste qui contient le nombre de boules dans chaque case à l'issue du jeu. <u>*Indication :*</u> *On s'autorisera à utiliser la fonction* `randint` *du module* `random`.
   - **1.2)** À l'aide de cette fonction, écrire une fonction `T` qui prend en argument n, N et renvoie la valeur prise par la variable aléatoire T_n lors de l'expérience.
   - **1.3)** Deux personnes jouent à ce jeu. Elles souhaitent comparer les résultats obtenues car elles pensent que le jeu est truqué. Écrire une fonction `compar` qui prend en argument les nombres entiers n et N et renvoie le pourcentage de cases qui contiennent le même nombre de boules pour deux expériences.
2. Déterminer les valeurs prises par T_n en fonction de n et N. <u>*Indication :*</u> *On pourra distinguer deux cas.*
3. Donner les lois de T_1 et T_2. <u>*Indication :*</u> *On pourra considérer les évènements B_k « les deux boules sont dans la case k » avec $1 \leqslant k \leqslant N$, et exprimer $\{T_2 = 1\}$ à l'aide des B_k.*
4. - **4.1)** Déterminer, lorsque $n \geqslant 2$: $\mathbf{P}(T_n = 1)$.
   - **4.2)** Montrer que : $\mathbf{P}(T_n = n) = \frac{N!}{(N-n)!N^n}$ si $n \leqslant N$, et $\mathbf{P}(T_n = n) = 0$ si $n > N$.
   On supposera dans la suite que $n \leqslant N$.
5. Justifier l'égalité $(\star)$ pour tout entier $k \in [\![1\,,\,n]\!]$:
$$\mathbf{P}(T_{n+1} = k) = \frac{k}{N}\mathbf{P}(T_n = k) + \frac{N-k+1}{N}\mathbf{P}(T_n = k-1) \qquad (\star).$$
Vérifier qu'elle est encore valable pour $k = n+1$.
6. **(Calcul de l'espérance par série génératrice)** Afin de calculer l'espérance $\mathbf{E}(T_n)$, on considère la fonction polynomiale G_n suivante appelée *fonction génératrice de* T_n :
$$\forall x \in \mathbf{R}, \quad G_n(x) = \sum_{k=1}^{n} \mathbf{P}(T_n = k)\, x^k.$$
   - **6.1)** Que vaut $G_n(1)$?
   - **6.2)** Calculer G_n' puis exprimer $\mathbf{E}(T_n)$ en fonction de $G_n'(1)$.
   - **6.3)** En utilisant la relation $(\star)$, monter que :
$$\forall x \in \mathbf{R}, \quad G_{n+1}(x) = \frac{1}{N}x(1-x)G_n'(x) + xG_n(x).$$
   - **6.4)** En dérivant l'égalité précédente, en déduire que :
$$\mathbf{E}(T_{n+1}) = \left(1 - \frac{1}{N}\right)\mathbf{E}(T_n) + 1.$$
   Quelle est la nature de $(\mathbf{E}(T_n))_{n \geqslant 0}$?
   - **6.5)** Conclure en donnant l'expression de $\mathbf{E}(T_n)$ en fonction de n et N.

Solution (problème 2)

1. **1.1)**

```
import random as rd
def Liste(n, N):
    """
    (n,N)-> la liste des nombres de boules dans chacune des
    N cases en n lancers
    """
```

```python
    L = [0 for _ in range(N)]
    for _ in range(n):
        L[rd.randint(0, N-1)] += 1
    return L
```

Par exemple, `Liste(10, 5)` renvoie [2, 2, 1, 1, 4].

1.2)
```python
import random as rd
def T_valeur(n, N):
    """
    (n,N)-> une simulation de Tn
    """
    S = 0
    L = Liste(n, N)
    for x in L:
        if x == 0:
            S += 1
    return S
```

Par exemple, `T_valeur(10, 5)` renvoie 1.

1.3)
```python
import random as rd
def Compar(n, N):
    """
    (n,N)-> compare les résultats de deux joueurs pour le
    ↪ jeu de paramètres n,N.
    Résultat sous forme de pourcentage
    """
    S = 0
    L_Joueur1 = Liste(n, N)
    L_Joueur2 = Liste(n, N)
    for k in range(N):
        if L_Joueur1[k] == L_Joueur2[k]:
            S += 1
    return (S/N)*100
```

Par exemple, `Compar(100, 60)` renvoie 20.0.

2. Il y a forcément une case ayant une boule.

2.1) Si $\mathrm{N} \geqslant n$: alors on lance moins de boules que de cases disponibles, donc on a $\boxed{\mathrm{T}_n(\Omega) = [\![1\,,\,n]\!]}$.

2.2) Si $\mathrm{N} < n$: alors le nombre de boules dépasse le nombre cases. On déduit alors que $\boxed{\mathrm{T}_n(\Omega) = [\![1\,,\,\mathrm{N}]\!]}$.

Finalement : $\boxed{\mathrm{T}_n(\Omega) = [\![1\,,\,\min(n,\mathrm{N})]\!]}$.

3. $\mathrm{T}_1 = 1$ p.s puisqu'une seule case est remplie. Pour T_2 : suivons l'indication. On a $\{\mathrm{T}_2 = 1\} = \mathrm{B}_1 \cap \mathrm{B}_2 \cup \mathrm{B}_3 \cap \cdots \cup \mathrm{B}_\mathrm{N}$. La réunion est même disjointe donc :

$$\mathbf{P}(\mathrm{T}_2 = 1) = \sum_{k=1}^{\mathrm{N}} \mathbf{P}(\mathrm{B}_k) = \sum_{k=1}^{\mathrm{N}} \frac{1}{\mathrm{N}^2} = \boxed{\frac{1}{\mathrm{N}}}.$$

Donc $\boxed{\mathrm{T}_2(\Omega) = \{0, 1\} \text{ et } \mathbf{P}(\mathrm{T}_2 = 1) = \frac{1}{\mathrm{N}}, \quad \mathbf{P}(\mathrm{T}_2 = 0) = 1 - \frac{1}{\mathrm{N}}}.$

4. 4.1) Le même raisonnement que précédemment en considérant B_k « les n boules sont dans la case k », on déduit que $\boxed{\mathbf{P}(\mathrm{T}_n = 1) = \frac{1}{\mathrm{N}}}$.

4.2) Dans le second, l'évènement $\{\mathrm{T}_n = n\}$ signifie que les boules sont arrivées dans n cases distinctes.

- Si $n > \mathrm{N}$, la probabilité est nulle car forcément une case contiendra au moins deux boules.
- Si $n < \mathrm{N}$, cela revient à sélectionner, de manière simultanée, les n cases accueillant une boules parmi les N disponible (le caractère simultané garantit l'unicité de la boule dans chaque case). Ainsi

$$\begin{aligned}\mathbf{P}(\mathrm{T}_n = n) &= \frac{\text{Nb répartition 1 boule par case}}{\text{Nb total de répartition des boules}} \\ &= \frac{\mathrm{N}(\mathrm{N}-1)(\mathrm{N}-2)\ldots(\mathrm{N}-n+1)}{\mathrm{N}^n} = \boxed{\frac{\mathrm{N}!}{(\mathrm{N}-n)!\mathrm{N}^n}}.\end{aligned}$$

5. Remarquons que puisque $n \leqslant \mathrm{N}$, $\{\mathrm{T}_n = k\}_{k=1,\ldots,n}$ forme un système complet d'évènements. Soit alors $k \in [\![0, \mathrm{N}]\!]$. Pour $k = 0$ la formule est évidemment vraie puisque toutes les probabilités sont alors nulles dans la formule. Supposons que $k \in [\![1, \mathrm{N}]\!]$, d'après la formule des probabilités totales :

$$\begin{aligned}
&\mathbf{P}\left(\mathrm{T}_{n+1}=k\right)\\
&=\sum_{\ell=1}^{\mathrm{N}} \mathbf{P}\left(\mathrm{T}_{n+1}=k \middle| \mathrm{T}_n=\ell\right) \mathbf{P}\left(\mathrm{T}_n=\ell\right)\\
&=\mathbf{P}\left(\mathrm{T}_{n+1}=k \middle| \mathrm{T}_n=k-1\right) \mathbf{P}\left(\mathrm{T}_n=k-1\right)+\mathbf{P}\left(\mathrm{T}_{n+1}=k \middle| \mathrm{T}_n=k\right) \mathbf{P}\left(\mathrm{T}_n=k\right)
\end{aligned}$$

car si en $n+1$ lancers on a k cases différentes occupées, alors sur les n premiers lancers on avait :

- soit une de moins (et le dernier lancer fait augmenter le nombre de cases occupées), la probabilité que cela arrive est de lancer la boule dans l'une des cases restantes en nombre $\mathrm{N}-k+1$ sur le nombre total de cases N.
- Dans le deux$i$-ème cas, la boule arrive dans une case déjà occupée, qui sont en nombre k, il y a donc une probabilité $\frac{k}{\mathrm{N}}$ que cela arrive.

Finalement, on obtient la formule voulue :

$$\mathbf{P}\left(\mathrm{T}_{n+1}=k\right)=\frac{k}{\mathrm{N}} \mathbf{P}\left(\mathrm{T}_n=k\right)+\frac{\mathrm{N}-k+1}{\mathrm{N}} \mathbf{P}\left(\mathrm{T}_n=k-1\right) \quad (\star).$$

De plus, pour $k=n+1$, on constate en utilisant une question précédente, qu'elle est équivalente à

$$\frac{\mathrm{N}!}{(\mathrm{N}-n-1) \mathrm{N}^{n+1}}=\frac{n+1}{\mathrm{N}} \times 0+\frac{\mathrm{N}-n}{\mathrm{N}} \frac{\mathrm{N}!}{(\mathrm{N}-n)!\mathrm{N}^n}.$$

Cette dernière formule est vraie.

6. **6.1)** Pour $x=1$, on obtient $\boxed{\mathrm{G}_n(1)=\sum_{k=1}^n \mathbf{P}\left(\mathrm{T}_n=k\right)=1}$ par définition d'une probabilité.

6.2) En dérivant une fois, on trouve $\mathrm{G}_n'(x)=\sum\limits_{k=1} \mathbf{P}\left(\mathrm{T}_n=k\right) k x^{k-1}$. En faisant $x=1$, on récupère : $\boxed{\mathbf{E}\left(\mathrm{T}_n\right)=\mathrm{G}_n'(1).}$

6.3) Sommons la relation $(\star)$ multipliée par x^k avec $x \in \mathbf{R}$:

$$\begin{aligned}
&\sum_{k=1}^{n+1} \mathbf{P}\left(\mathrm{T}_{n+1}=k\right) x^k\\
&=\sum_{k=1}^{n+1} \frac{k}{\mathrm{N}} \mathbf{P}\left(\mathrm{T}_n=k\right) x^k+\sum_{k=1}^{n+1} \frac{\mathrm{N}-k+1}{\mathrm{N}} \mathbf{P}\left(\mathrm{T}_n=k-1\right) x^k\\
&=\frac{x}{\mathrm{N}} \sum_{k=1}^{n} k \mathbf{P}\left(\mathrm{T}_n=k\right) x^{k-1}+\frac{\mathrm{N}+1}{\mathrm{N}} \sum_{k=1}^{n+1} \mathbf{P}\left(\mathrm{T}_n=k-1\right) x^k\\
&-\frac{1}{\mathrm{N}} \sum_{k=1}^{n+1} k \mathbf{P}\left(\mathrm{T}_n=k-1\right) x^k\\
&=\frac{x}{\mathrm{N}} \mathrm{G}_n'(x)+\frac{\mathrm{N}+1}{\mathrm{N}} \sum_{k=0}^{n} \mathbf{P}\left(\mathrm{T}_n=k\right) x^{k+1}\\
&-\frac{1}{\mathrm{N}} \sum_{k=0}^{n}(k+1) \mathbf{P}\left(\mathrm{T}_n=k\right) x^{k+1}\\
&=\frac{x}{\mathrm{N}} \mathrm{G}_n'(x)+\frac{x(\mathrm{N}+1)}{\mathrm{N}} \sum_{k=1}^{n} \mathbf{P}\left(\mathrm{T}_n=k\right) x^k-\frac{x}{\mathrm{N}} \sum_{k=1}^{n} k \mathbf{P}\left(\mathrm{T}_n=k\right) x^k\\
&-\frac{x}{\mathrm{N}} \sum_{k=1}^{n} \mathbf{P}\left(\mathrm{T}_n=k\right) x^k\\
&=\frac{x}{\mathrm{N}} \mathrm{G}_n'(x)+\frac{x(\mathrm{N}+1)}{\mathrm{N}} \mathrm{G}_n(x)-\frac{x^2}{\mathrm{N}} \mathrm{G}_n'(x)-\frac{x}{\mathrm{N}} \mathrm{G}_n(x)
\end{aligned}$$

où :

- la première somme a été arrêtée à n puisque $\mathbf{P}\left(\mathrm{T}_n=n+1\right)=0$,
- dans la seconde nous avons fait le changement $\ell=k-1$ afin de reconnaître la fonction génératrice de T_n,
- et enfin, le même changement de variable dans la dernière somme.

En factorisant par $\mathrm{G}_n'(x)$ et $\mathrm{G}_n(x)$, on trouve la relation cherchée :

$$\boxed{\forall x \in \mathbf{R}, \quad \mathrm{G}_{n+1}(x)=\frac{1}{\mathrm{N}} x(1-x) \mathrm{G}_n'(x)+x \mathrm{G}_n(x).}$$

6.4) Dérivons-là par rapport à x et faisons $x=1$. Alors :

$$\begin{aligned}
&\mathrm{G}_{n+1}'(x)=\frac{1}{\mathrm{N}}(1-2 x) \mathrm{G}_n'(x)+\mathrm{G}_n(x)+x \mathrm{G}_n'(x),\\
&\mathrm{G}_{n+1}'(1)=-\frac{1}{\mathrm{N}} \mathrm{G}_n'(1)+\mathrm{G}_n(1)+\mathrm{G}_n'(1).
\end{aligned}$$

D'où : $\boxed{\mathbf{E}\left(\mathrm{T}_{n+1}\right)=\left(1-\frac{1}{\mathrm{N}}\right)\mathbf{E}\left(\mathrm{T}_n\right)+1}$. C'est la formule demandée. La suite $\left(\mathbf{E}\left(\mathrm{T}_n\right)\right)$ est donc arithmético-géométrique.

6.5) On cherche α tel que : $\alpha=\left(1-\frac{1}{\mathrm{N}}\right)\alpha+1$, on trouve $\alpha=\mathrm{N}$. Ainsi la suite $(\mathbf{E}\left(\mathrm{T}_n\right)-\mathrm{N})$ est géométrique de raison $1-\frac{1}{\mathrm{N}}$, donc pour tout entier $n\in\mathbf{N}$:

$$\mathbf{E}\left(\mathrm{T}_n\right)-\mathrm{N}=\left(1-\frac{1}{\mathrm{N}}\right)\left(\mathbf{E}\left(\mathrm{T}_1\right)-\mathrm{N}\right)=\left(1-\frac{1}{\mathrm{N}}\right)(1-\mathrm{N})$$

donc

$$\boxed{\mathbf{E}\left(\mathrm{T}_n\right)=\mathrm{N}+\left(1-\frac{1}{\mathrm{N}}\right)(1-\mathrm{N})=2-\frac{1}{\mathrm{N}}}.$$