

Compléments de probabilités : couples et n-uplets de variables aléatoires réelles

Fiche d'exercices corrigés — ECG maths approfondies

Difficulté : ★ application directe · ★★ standard · ★★★ approfondissement · ★★★★★ défi.

Un conseil : cherchez d'abord, le corrigé ensuite — c'est là que ça progresse.

Exercice 1 ★ Lire une fonction de répartition

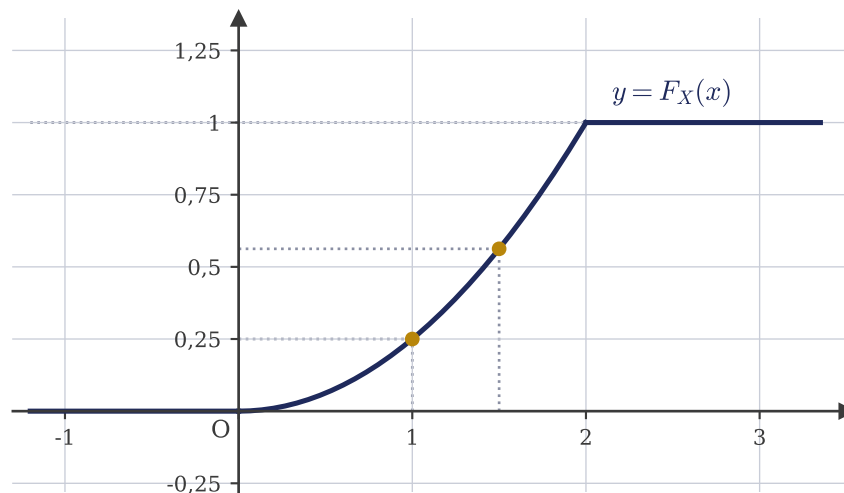
Notions : Fonction de répartition d'une variable aléatoire réelle quelconque, Variable aléatoire à densité : définition et reconnaissance d'une densité

Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, dont la fonction de répartition F_X est donnée par

$$F_X(x) = 0 \quad \text{si } x < 0,$$

$$F_X(x) = \frac{x^2}{4} \quad \text{si } 0 \leq x \leq 2,$$

$$F_X(x) = 1 \quad \text{si } x > 2.$$



1. Vérifier que F_X possède les trois propriétés caractéristiques d'une fonction de répartition.
2. Calculer $P(X \leq 1)$, $P(X > 1,5)$, $P(0,5 \leq X \leq 1,5)$ et $P(X = 1)$. Préciser à chaque fois ce que l'on lit sur la courbe.
3. Justifier que X est une variable à densité, puis en donner une densité.
4. Calculer $E(X)$ et $V(X)$.

Corrigé

1. Une fonction de répartition est caractérisée par trois propriétés : elle est croissante sur $\mathbb{R}$, elle a pour limites 0 en $-\infty$ et 1 en $+\infty$, et elle est continue à droite en tout point. On les vérifie l'une après l'autre.

Croissance. Sur $]-\infty, 0[$, F_X est constante égale à 0, donc croissante. Sur $[0, 2]$, F_X est dérivable de dérivée $F'_X(x) = \frac{x}{2}$, positive sur cet intervalle : F_X y est croissante. Sur $]2, +\infty[$, F_X est constante égale à 1, donc croissante.

Il reste à vérifier les raccords, sans quoi la croissance sur chaque morceau ne prouverait rien.

Sur $[0, 2]$, F_X croît de $F_X(0) = 0$ à $F_X(2) = \frac{4}{4} = 1$: les valeurs prises sur les trois morceaux s'enchaînent bien dans l'ordre 0, puis $[0, 1]$, puis 1. Donc F_X est croissante sur $\mathbb{R}$ tout entier.

Limites. Pour $x < 0$, $F_X(x) = 0$, donc $\lim_{x \rightarrow -\infty} F_X(x) = 0$. Pour $x > 2$, $F_X(x) = 1$, donc

$$\lim_{x \rightarrow +\infty} F_X(x) = 1.$$

Continuité à droite. F_X est continue sur chacun des trois intervalles ouverts, comme fonction

constante ou polynomiale. En 0 : $\lim_{x \rightarrow 0^-} F_X(x) = 0$ et $F_X(0) = \frac{0^2}{4} = 0$. En 2 : $F_X(2) = 1$ et $\lim_{x \rightarrow 2^+} F_X(x) = 1$. La fonction F_X est donc continue sur $\mathbb{R}$ tout entier, et en particulier continue à droite en tout point.

Les trois propriétés sont vérifiées : F_X est bien une fonction de répartition. $\square$

2. Remarque préliminaire, valable pour tout réel a . Pour $h > 0$, l'inclusion $(X = a) \subset (a - h < X \leq a)$ donne

$$0 \leq P(X = a) \leq P(a - h < X \leq a) = F_X(a) - F_X(a - h)$$

Quand h tend vers 0 par valeurs positives, la continuité de F_X en a montre que $F_X(a - h)$ tend vers $F_X(a)$, donc le majorant tend vers 0. Par encadrement, $P(X = a) = 0$. Autrement dit, aucune valeur n'est chargée : c'est ce qui permet, dans la suite, de passer librement des inégalités strictes aux inégalités larges.

Première probabilité.

$$P(X \leq 1) = F_X(1) = \frac{1^2}{4} = \frac{1}{4} = 0,25$$

Sur la courbe, c'est l'ordonnée du point d'abscisse 1.

Deuxième probabilité. L'événement $(X > 1,5)$ est le contraire de $(X \leq 1,5)$. Or $F_X(1,5) = \frac{1}{4} \times \left(\frac{3}{2}\right)^2 = \frac{1}{4} \times \frac{9}{4} = \frac{9}{16}$, donc

$$P(X > 1,5) = 1 - F_X(1,5) = 1 - \frac{9}{16} = \frac{7}{16} = 0,4375$$

Sur la courbe, c'est l'écart vertical entre le point d'abscisse 1,5 et la droite horizontale d'équation $y = 1$.

Troisième probabilité. La formule du cours est $P(a < X \leq b) = F_X(b) - F_X(a)$. Ici l'inégalité de gauche est large, mais la remarque préliminaire donne $P(X = 0,5) = 0$, donc les deux événements $(0,5 \leq X \leq 1,5)$ et $(0,5 < X \leq 1,5)$ ont la même probabilité. Avec $F_X(0,5) = \frac{1}{4} \times \left(\frac{1}{2}\right)^2 = \frac{1}{16}$:

$$P(0,5 \leq X \leq 1,5) = F_X(1,5) - F_X(0,5) = \frac{9}{16} - \frac{1}{16} = \frac{8}{16} = \frac{1}{2}$$

Sur la courbe, c'est le dénivelé de la courbe entre les abscisses 0,5 et 1,5.

Quatrième probabilité. D'après la remarque préliminaire, $P(X = 1) = 0$. Sur la courbe, cela se lit à l'absence de saut : la courbe est tracée d'un seul trait, sans marche d'escalier, contrairement à ce qui se produirait pour une variable discrète.

3. Le cours définit une variable à densité par deux conditions portant sur sa fonction de répartition : F_X est continue sur $\mathbb{R}$, et F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points.

La continuité de F_X sur $\mathbb{R}$ a été établie à la question 1. Par ailleurs, F_X est de classe $\mathcal{C}^1$ sur $]-\infty, 0[$ et sur $]2, +\infty[$ comme fonction constante, et sur $]0, 2[$ comme fonction polynomiale. Les seuls points litigieux sont 0 et 2, en nombre fini. Donc X est une variable à densité.

Une densité de X s'obtient en dérivant F_X là où c'est licite, et en attribuant aux points restants une valeur positive quelconque, par exemple 0 :

$$\begin{aligned} f(x) &= \frac{x}{2} & \text{si } 0 \leq x \leq 2, \\ f(x) &= 0 & \text{sinon.} \end{aligned}$$

Contrôle. La fonction f est positive sur $\mathbb{R}$, continue sauf en 2, donc sauf en un nombre fini de points, et elle est nulle en dehors du segment $[0, 2]$, si bien que son intégrale n'est pas impropre :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^2 \frac{t}{2} dt = \left[\frac{t^2}{4} \right]_0^2 = \frac{4}{4} = 1$$

f est bien une densité de probabilité, et c'est bien une densité de X .

4. Existence de l'espérance. Par définition, $E(X)$ existe si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t f(t) dt$ converge absolument. La fonction $t \mapsto t f(t)$ est nulle en dehors de $[0, 2]$, donc cette intégrale se réduit à $\int_0^2 \frac{t^2}{2} dt$, intégrale d'une fonction continue sur un segment : elle n'est pas impropre, la convergence est immédiate. $E(X)$ existe.

$$E(X) = \int_0^2 t \times \frac{t}{2} dt = \int_0^2 \frac{t^2}{2} dt = \left[\frac{t^3}{6} \right]_0^2 = \frac{8}{6} = \frac{4}{3} \approx 1,333$$

Moment d'ordre deux. Le théorème de transfert, admis par le programme, appliqué à la fonction $t \mapsto t^2$, ramène le calcul à une intégrale sur le même segment, donc convergente pour la même raison :

$$E(X^2) = \int_0^2 t^2 \times \frac{t}{2} dt = \int_0^2 \frac{t^3}{2} dt = \left[\frac{t^4}{8} \right]_0^2 = \frac{16}{8} = 2$$

Variance. La formule de König-Huygens donne

$$V(X) = E(X^2) - (E(X))^2 = 2 - \left(\frac{4}{3} \right)^2 = 2 - \frac{16}{9} = \frac{18 - 16}{9} = \frac{2}{9}$$

d'où l'écart-type

$$\sigma(X) = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3} \approx 0,471$$

Cohérence. La densité $t \mapsto \frac{t}{2}$ est croissante sur $[0, 2]$: elle charge davantage les grandes valeurs.

Il est donc normal de trouver $E(X) = \frac{4}{3} \approx 1,33$, strictement au-dessus du milieu 1 du segment. Et la variance est bien strictement positive, comme pour toute variable non constante.

Méthode à retenir. Tout part de F_X . Les probabilités d'intervalles se lisent en différences de valeurs de F_X , la nature de la variable se décide sur deux critères (continuité sur $\mathbb{R}$, classe $\mathcal{C}^1$ sauf en un nombre fini de points), et une densité s'obtient en dérivant morceau par morceau. Une fois la densité écrite, vérifier que son intégrale vaut 1 coûte trois lignes et rattrape la plupart des erreurs de calcul.

Erreur classique. Dérivée F_X « partout » et annoncer $f(x) = \frac{x}{2}$ sur $\mathbb{R}$: la densité est nulle en dehors de $[0, 2]$, sans quoi son intégrale divergerait. Deuxième piège : traiter les inégalités strictes et larges comme sur une variable discrète. Pour une variable à densité elles donnent la même probabilité, puisque $P(X = a) = 0$, mais c'est un résultat qui se cite ou se démontre, jamais une évidence.

Exercice 2 ★ Reconnaître une densité de probabilité

Notions : Variable aléatoire à densité : définition et reconnaissance d'une densité

On rappelle qu'une fonction f définie sur $\mathbb{R}$ est une densité de probabilité lorsqu'elle vérifie les trois conditions suivantes : f est positive sur $\mathbb{R}$; f est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points ; l'intégrale $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut 1.

- Pour chacune des fonctions suivantes, dire si elle est une densité de probabilité, en vérifiant les trois conditions.
 - $f(x) = \frac{1}{2}$ sur $[-1, 1]$, nulle ailleurs.
 - $f(x) = x$ sur $[0, 1]$, nulle ailleurs.
 - $f(x) = 3x^2$ sur $[0, 1]$, nulle ailleurs.
 - $f(x) = \frac{1}{x^2}$ sur $[1, +\infty[$, nulle ailleurs.
 - $f(x) = -e^{-x}$ sur $[0, +\infty[$, nulle ailleurs.
- Déterminer le réel c pour que la fonction définie par $f(x) = c x(1-x)$ sur $[0, 1]$ et nulle ailleurs soit une densité de probabilité.
- Déterminer le réel c pour que la fonction définie par $f(x) = c e^{-3x}$ sur $[0, +\infty[$ et nulle ailleurs soit une densité de probabilité. Reconnaître alors la loi dont f est une densité.

Corrigé

1. On examine les trois conditions dans l'ordre où elles coûtent le moins cher : la positivité d'abord, car elle élimine parfois la fonction en une ligne, puis la continuité, puis l'intégrale. Dès qu'une condition tombe, la conclusion est acquise et il est inutile de poursuivre.

a. La fonction f ne prend que les valeurs $\frac{1}{2}$ et 0 : elle est positive sur $\mathbb{R}$. Elle est continue sur $\mathbb{R}$ sauf en -1 et en 1 , soit en deux points, donc en nombre fini. Enfin, f étant nulle en dehors du segment $[-1, 1]$, son intégrale n'est pas impropre :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{-1}^1 \frac{1}{2} dt = \frac{1}{2} \times (1 - (-1)) = \frac{1}{2} \times 2 = 1$$

Les trois conditions sont vérifiées : f est une densité de probabilité. On reconnaît la densité de la loi uniforme $\mathcal{U}([-1, 1])$.

b. Sur $[0, 1]$ on a $x \geq 0$, donc f est positive sur $\mathbb{R}$. Elle est continue sauf en 1. En revanche :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^1 t dt = \left[\frac{t^2}{2} \right]_0^1 = \frac{1}{2}$$

La masse totale vaut $\frac{1}{2}$ et non 1 : f n'est pas une densité de probabilité. On notera que $2f$, elle, en serait une, puisque son intégrale vaudrait $2 \times \frac{1}{2} = 1$.

c. La fonction f est positive, car $3x^2 \geq 0$, et continue sur $\mathbb{R}$ sauf en 1 : en 0 le raccord est continu puisque $3 \times 0^2 = 0$, tandis qu'en 1 elle saute de 3 à 0. Enfin :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^1 3t^2 dt = [t^3]_0^1 = 1$$

f est une densité de probabilité.

d. Pour $x \geq 1$, $\frac{1}{x^2} > 0$, donc f est positive sur $\mathbb{R}$, et elle est continue sauf en 1. Cette fois le support n'est pas borné : l'intégrale est impropre en $+\infty$ et sa convergence doit être établie.

Pour $A > 1$:

$$\int_1^A \frac{1}{t^2} dt = \int_1^A t^{-2} dt = \left[-\frac{1}{t} \right]_1^A = 1 - \frac{1}{A}$$

Lorsque A tend vers $+\infty$, cette quantité admet la limite finie 1. L'intégrale $\int_1^{+\infty} \frac{dt}{t^2}$ converge donc et vaut 1 : f est une densité de probabilité.

e. Pour tout $x \geq 0$, $e^{-x} > 0$, donc $f(x) = -e^{-x} < 0$. La condition de positivité tombe dès le premier point : f n'est pas une densité de probabilité, et il est inutile de calculer son intégrale (elle convergerait, mais vers -1).

Bilan. Les fonctions **a.**, **c.** et **d.** sont des densités. La fonction **b.** échoue sur la valeur de l'intégrale, la fonction **e.** sur la positivité.

2. Positivité. Sur $[0, 1]$ on a $x \geq 0$ et $1 - x \geq 0$, donc le produit $x(1 - x)$ est positif. La fonction f est donc positive sur $\mathbb{R}$ si et seulement si $c \geq 0$. On garde cette contrainte en réserve.

Continuité. Sur $[0, 1]$, f est polynomiale, donc continue, et elle est nulle ailleurs. Aux points de raccord, $f(0) = f(1) = c \times 0 = 0$: les valeurs coïncident. La fonction f est donc continue sur $\mathbb{R}$ tout entier, ce qui est plus que nécessaire.

Intégrale. La fonction f étant nulle en dehors du segment $[0, 1]$, l'intégrale n'est pas impropre :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(t) dt &= c \int_0^1 (t - t^2) dt \\ &= c \left[\frac{t^2}{2} - \frac{t^3}{3} \right]_0^1 \\ &= c \left(\frac{1}{2} - \frac{1}{3} \right) \\ &= \frac{c}{6} \end{aligned}$$

Cette masse vaut 1 si et seulement si $\frac{c}{6} = 1$, c'est-à-dire $c = 6$. Cette valeur est bien positive, la contrainte mise en réserve est donc satisfaite.

Conclusion : $c = 6$.

3. Positivité. Pour tout réel x , $e^{-3x} > 0$, donc f est positive sur $\mathbb{R}$ si et seulement si $c \geq 0$.

Continuité. f est continue sur $]-\infty, 0[$ (elle y est nulle) et sur $]0, +\infty[$ (composée de fonctions continues). Le seul point de discontinuité possible est 0, où f saute de 0 à c : cela fait un point, donc un nombre fini de points.

Intégrale. Le support n'est pas borné, l'intégrale est impropre en $+\infty$. Pour $A > 0$:

$$\int_0^A c e^{-3t} dt = c \left[-\frac{1}{3} e^{-3t} \right]_0^A = \frac{c}{3} (1 - e^{-3A})$$

Comme e^{-3A} tend vers 0 quand A tend vers $+\infty$, l'intégrale converge et vaut $\frac{c}{3}$. Elle vaut 1 si et seulement si $c = 3$, valeur bien positive.

Conclusion : $c = 3$.

Reconnaissance de la loi. Avec $c = 3$, la fonction f vaut $3e^{-3x}$ sur $[0, +\infty[$ et 0 ailleurs. C'est exactement la forme $\lambda e^{-\lambda x}$ avec $\lambda = 3$: f est une densité de la loi exponentielle de paramètre

3. Si X est une variable de densité f , alors $X \hookrightarrow \mathcal{E}(3)$, avec $E(X) = \frac{1}{3}$ et $V(X) = \frac{1}{9}$.

Méthode à retenir. Trois conditions, toujours dans le même ordre : positivité, continuité sauf en un nombre fini de points, convergence de l'intégrale et valeur 1. Sur un support borné, l'intégrale d'une fonction continue par morceaux n'est pas impropre : il suffit de le signaler, puis de calculer une primitive. Sur un support non borné, on intègre entre une borne finie et A , puis on fait tendre A vers l'infini : c'est la seule rédaction acceptée. Enfin, pour une constante multiplicative

à déterminer, on calcule l'intégrale en fonction de c , on résout l'équation « masse totale égale à 1 », et on vérifie que la valeur trouvée rend bien f positive.

Erreur classique. Se contenter de « f est positive et continue, donc c'est une densité » sans calculer l'intégrale : c'est précisément la condition qui échoue le plus souvent, comme au cas **b**.

Autre erreur, plus grave : écrire $\int_1^{+\infty} \frac{dt}{t^2} = \left[-\frac{1}{t} \right]_1^{+\infty}$ en remplaçant t par $+\infty$ dans la primitive.

On ne substitue pas $+\infty$ à une variable : on passe par une borne finie A , puis par une limite, et c'est cette limite qui établit la convergence.

Exercice 3 ★ Calculer des probabilités à partir d'une densité

Notions : Variable aléatoire à densité : définition et reconnaissance d'une densité, Fonction de répartition d'une variable aléatoire réelle quelconque

Soit X une variable aléatoire réelle admettant pour densité la fonction f définie sur $\mathbb{R}$ par $f(x) = 3x^2$ si $x \in [0, 1]$, et $f(x) = 0$ sinon.

- Vérifier que f est bien une densité de probabilité.
- Déterminer la fonction de répartition F_X de X sur $\mathbb{R}$, en distinguant les trois cas.
- Calculer les probabilités suivantes.

a. $P\left(X \leq \frac{1}{2}\right)$	c. $P\left(\frac{1}{4} \leq X \leq \frac{1}{2}\right)$
b. $P\left(X \geq \frac{3}{4}\right)$	d. $P\left(X = \frac{1}{2}\right)$
- Calculer la probabilité conditionnelle $P_{(X \geq \frac{1}{2})}\left(X \geq \frac{3}{4}\right)$ et l'interpréter.

Corrigé

1. Positivité. Pour tout réel x , $3x^2 \geq 0$, et f est nulle ailleurs : f est positive sur $\mathbb{R}$.

Continuité. Sur $]-\infty, 0[$ et sur $]1, +\infty[$, f est nulle donc continue ; sur $]0, 1[$, elle est polynomiale donc continue. En 0 le raccord est continu, puisque $3 \times 0^2 = 0$. En 1, f saute de 3 à 0 : c'est le seul point de discontinuité, donc f est continue sur $\mathbb{R}$ sauf en un nombre fini de points.

Intégrale. La fonction f étant nulle en dehors du segment $[0, 1]$, l'intégrale n'est pas impropre et se réduit à

$$\int_{-\infty}^{+\infty} f(t) dt = \int_0^1 3t^2 dt = [t^3]_0^1 = 1 - 0 = 1$$

Les trois conditions sont vérifiées : f est une densité de probabilité.

2. Par définition, pour tout réel x ,

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

Cas $x < 0$. Sur $]-\infty, x]$, la fonction f est identiquement nulle, donc $F_X(x) = 0$.

Cas $0 \leq x \leq 1$. La relation de Chasles donne

$$F_X(x) = \int_{-\infty}^0 f(t) dt + \int_0^x 3t^2 dt = 0 + [t^3]_0^x = x^3$$

Cas $x > 1$. De la même manière,

$$F_X(x) = \int_{-\infty}^1 f(t) dt + \int_1^x 0 dt = 1 + 0 = 1$$

En résumé :

$$\begin{aligned} F_X(x) &= 0 & \text{si } x < 0, \\ F_X(x) &= x^3 & \text{si } 0 \leq x \leq 1, \\ F_X(x) &= 1 & \text{si } x > 1. \end{aligned}$$

Contrôles. La fonction F_X est croissante sur chaque morceau, elle se raccorde continûment en 0 (deux fois la valeur 0) et en 1 (deux fois la valeur 1), et ses limites en $-\infty$ et en $+\infty$ valent 0 et 1. Tout est cohérent.

3. Comme X est une variable à densité, sa fonction de répartition est continue, donc $P(X = a) = 0$ pour tout réel a . Les inégalités strictes et larges donnent par conséquent les mêmes probabilités, ce qui sera utilisé librement ci-dessous.

a. Lecture directe de la fonction de répartition :

$$P\left(X \leq \frac{1}{2}\right) = F_X\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^3 = \frac{1}{8} = 0,125$$

b. L'événement $\left(X \geq \frac{3}{4}\right)$ est le contraire de $\left(X < \frac{3}{4}\right)$, dont la probabilité est $F_X\left(\frac{3}{4}\right)$ puisque $P\left(X = \frac{3}{4}\right) = 0$. Donc

$$P\left(X \geq \frac{3}{4}\right) = 1 - F_X\left(\frac{3}{4}\right) = 1 - \left(\frac{3}{4}\right)^3 = 1 - \frac{27}{64} = \frac{37}{64} \approx 0,578$$

c. La formule $P(a < X \leq b) = F_X(b) - F_X(a)$, jointe à $P\left(X = \frac{1}{4}\right) = 0$, donne

$$P\left(\frac{1}{4} \leq X \leq \frac{1}{2}\right) = F_X\left(\frac{1}{2}\right) - F_X\left(\frac{1}{4}\right) = \frac{1}{8} - \frac{1}{64} = \frac{8}{64} - \frac{1}{64} = \frac{7}{64} \approx 0,109$$

d. La variable X est à densité, donc aucun singleton n'est chargé :

$$P\left(X = \frac{1}{2}\right) = 0$$

4. Il faut d'abord s'assurer que l'événement conditionnant n'est pas négligeable :

$$P\left(X \geq \frac{1}{2}\right) = 1 - F_X\left(\frac{1}{2}\right) = 1 - \frac{1}{8} = \frac{7}{8} \neq 0$$

La probabilité conditionnelle est donc bien définie. Par ailleurs, $\left(X \geq \frac{3}{4}\right) \subset \left(X \geq \frac{1}{2}\right)$, donc l'intersection des deux événements est le plus petit des deux, à savoir $\left(X \geq \frac{3}{4}\right)$. Ainsi

$$P_{(X \geq \frac{1}{2})}\left(X \geq \frac{3}{4}\right) = \frac{P\left(X \geq \frac{3}{4}\right)}{P\left(X \geq \frac{1}{2}\right)} = \frac{\frac{37}{64}}{\frac{7}{8}} = \frac{37}{64} \times \frac{8}{7} = \frac{37}{56} \approx 0,661$$

Interprétation. Sans information, la probabilité de dépasser $\frac{3}{4}$ vaut environ 0,578. Sachant que X dépasse déjà $\frac{1}{2}$, elle monte à environ 0,661 : l'information reçue élimine les petites valeurs et rend l'événement plus probable. C'est logique, la densité $t \mapsto 3t^2$ étant croissante.

Méthode à retenir. Sur une variable à densité, on construit d'abord F_X par $F_X(x) = \int_{-\infty}^x f(t) dt$, en séparant les trois régions imposées par le support. Une fois F_X écrite, toutes les probabilités demandées s'en déduisent sans plus aucune intégrale : $P(X \leq a) = F_X(a)$, $P(X \geq a) = 1 - F_X(a)$, $P(a \leq X \leq b) = F_X(b) - F_X(a)$. Pour une probabilité conditionnelle entre deux événements de la forme $(X \geq \cdot)$, le réflexe est de repérer l'inclusion avant tout calcul.

Erreur classique. Oublier le morceau constant à droite et écrire $F_X(x) = x^3$ sur tout $\mathbb{R}$. Pour $x = 2$ on obtiendrait $F_X(2) = 8$, alors qu'une fonction de répartition est à valeurs dans $[0, 1]$. Deuxième erreur, en question **4.** : traiter les deux événements comme s'ils étaient indépendants et multiplier leurs probabilités. Ils sont au contraire emboîtés, ce qui est le cas le plus éloigné possible de l'indépendance.

Exercice 4 ★ Attente du tramway

Notions : Loi uniforme sur un segment

Un tramway dessert un arrêt avec une régularité parfaite : il y passe exactement toutes les 12 minutes. Un voyageur se présente à cet arrêt à un instant quelconque, sans connaître les horaires. On note X son temps d'attente, exprimé en minutes, et on admet que $X \hookrightarrow \mathcal{U}([0, 12])$.

1. Donner une densité de X , puis sa fonction de répartition F_X .
2. Calculer $P(X \leq 5)$, $P(X \geq 8)$ et $P(3 \leq X \leq 7)$.
3. Calculer $E(X)$, $V(X)$ et $\sigma(X)$, puis interpréter ces trois valeurs.
4. Le voyageur attend déjà depuis 4 minutes. Calculer $P_{(X \geq 4)}(X \geq 8)$, puis comparer cette valeur à $P(X \geq 4)$. Que peut-on en conclure sur la « mémoire » de la loi uniforme ?

Corrigé

1. La loi uniforme $\mathcal{U}([a, b])$ admet pour densité la fonction constante égale à $\frac{1}{b-a}$ sur $[a, b]$ et nulle ailleurs. Ici $a = 0$ et $b = 12$, donc une densité de X est la fonction f définie par

$$f(x) = \frac{1}{12} \text{ si } 0 \leq x \leq 12, \quad f(x) = 0 \text{ sinon}$$

Contrôle. f est positive, continue sauf en 0 et en 12, et $\int_0^{12} \frac{1}{12} dt = \frac{12}{12} = 1$.

La fonction de répartition se calcule ensuite par $F_X(x) = \int_{-\infty}^x f(t) dt$, en distinguant trois cas. Pour $x < 0$, la fonction à intégrer est nulle, donc $F_X(x) = 0$. Pour $0 \leq x \leq 12$,

$$F_X(x) = \int_0^x \frac{1}{12} dt = \frac{x}{12}$$

Pour $x > 12$, on retrouve la masse totale, donc $F_X(x) = 1$. En résumé :

$$\begin{aligned} F_X(x) &= 0 & \text{si } x < 0, \\ F_X(x) &= \frac{x}{12} & \text{si } 0 \leq x \leq 12, \\ F_X(x) &= 1 & \text{si } x > 12. \end{aligned}$$

2. La variable X étant à densité, on a $P(X = a) = 0$ pour tout réel a : les inégalités strictes et larges donnent les mêmes probabilités.

$$P(X \leq 5) = F_X(5) = \frac{5}{12} \approx 0,417$$

$$P(X \geq 8) = 1 - F_X(8) = 1 - \frac{8}{12} = \frac{4}{12} = \frac{1}{3} \approx 0,333$$

$$P(3 \leq X \leq 7) = F_X(7) - F_X(3) = \frac{7}{12} - \frac{3}{12} = \frac{4}{12} = \frac{1}{3}$$

Remarque. Les deux dernières probabilités sont égales, et ce n'est pas un hasard : sur une loi uniforme, la probabilité d'un intervalle inclus dans $[0, 12]$ ne dépend que de sa longueur. Les intervalles $[3, 7]$ et $[8, 12]$ ont tous deux une longueur de 4 minutes, donc la même probabilité $\frac{4}{12}$.

3. Les formules du cours donnent $E(X) = \frac{a+b}{2}$ et $V(X) = \frac{(b-a)^2}{12}$, mais il faut savoir les retrouver. La variable étant à support borné, toutes les intégrales portent sur un segment : elles ne sont pas impropres, l'existence des moments est acquise.

$$E(X) = \int_0^{12} t \times \frac{1}{12} dt = \frac{1}{12} \left[\frac{t^2}{2} \right]_0^{12} = \frac{1}{12} \times \frac{144}{2} = \frac{72}{12} = 6$$

Par le théorème de transfert appliqué à $t \mapsto t^2$:

$$E(X^2) = \int_0^{12} t^2 \times \frac{1}{12} dt = \frac{1}{12} \left[\frac{t^3}{3} \right]_0^{12} = \frac{1}{12} \times \frac{1728}{3} = \frac{576}{12} = 48$$

La formule de König-Huygens donne alors

$$V(X) = E(X^2) - (E(X))^2 = 48 - 6^2 = 48 - 36 = 12$$

$$\sigma(X) = \sqrt{12} = 2\sqrt{3} \approx 3,46$$

On vérifie au passage la formule du cours : $\frac{(12-0)^2}{12} = \frac{144}{12} = 12$.

Interprétation. Un voyageur qui arrive au hasard attend en moyenne 6 minutes, soit la moitié de l'intervalle entre deux passages : c'est exactement ce que dit le bon sens. L'écart-type, de l'ordre de 3,5 minutes, indique que cette attente est très dispersée autour de sa moyenne. Annoncer « vous attendrez six minutes » n'a donc guère de sens pour un voyageur particulier.

4. L'événement conditionnant a pour probabilité

$$P(X \geq 4) = 1 - F_X(4) = 1 - \frac{4}{12} = \frac{8}{12} = \frac{2}{3} \neq 0$$

la probabilité conditionnelle est donc définie. Comme $(X \geq 8) \subset (X \geq 4)$, l'intersection des deux événements est $(X \geq 8)$, d'où

$$P_{(X \geq 4)}(X \geq 8) = \frac{P(X \geq 8)}{P(X \geq 4)} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

Contrôle. Sachant $X \geq 4$, tout se passe comme si le temps d'attente était uniforme sur $[4, 12]$, et l'on retrouve $\frac{12-8}{12-4} = \frac{4}{8} = \frac{1}{2}$.

Comparaison. L'absence de mémoire s'écrirait $P_{(X \geq s)}(X \geq s+t) = P(X \geq t)$ pour tous réels positifs s et t . Testons-la avec $s = t = 4$: le membre de gauche vaut $\frac{1}{2}$, tandis que le membre de droite vaut $P(X \geq 4) = \frac{2}{3}$. Ces deux nombres diffèrent, donc la propriété est en défaut.

Conclusion : contrairement à la loi exponentielle, la loi uniforme a de la mémoire. Avoir déjà attendu 4 minutes est une information qui modifie les probabilités, car elle exclut les instants d'arrivée les plus favorables et réduit l'attente restante possible à 8 minutes au maximum. Avec une loi exponentielle, au contraire, attendre encore 4 minutes aurait la même probabilité qu'au premier instant : c'est la seule loi à densité à valeurs positives qui possède cette propriété.

Méthode à retenir. Sur une loi uniforme, tout se lit en longueurs : $P(c \leq X \leq d) = \frac{d-c}{b-a}$ dès que $[c, d] \subset [a, b]$. Les formules $E(X) = \frac{a+b}{2}$ et $V(X) = \frac{(b-a)^2}{12}$ sont à connaître par cœur, mais savoir les retrouver par le calcul intégral en cinq lignes est le vrai objectif de l'exercice. Enfin, pour une probabilité conditionnelle du type $P_{(X \geq s)}(X \geq s+t)$, on commence toujours

par remarquer l'inclusion des deux événements : elle transforme l'intersection en un seul des deux.

Erreur classique. Appliquer la formule $\frac{d-c}{b-a}$ sans vérifier que $[c, d]$ est inclus dans le support.

Pour $P(10 \leq X \leq 20)$, la réponse n'est pas $\frac{10}{12}$ mais $\frac{2}{12}$, puisque X ne dépasse jamais 12. Deuxième erreur : croire que toute loi à densité est sans mémoire. Cette propriété est très particulière, et le calcul de la question 4. montre qu'elle est fausse ici.

Exercice 5 ★ Durée de vie d'une ampoule

Notions : Loi exponentielle et absence de mémoire

La durée de vie, exprimée en heures, d'une ampoule d'un certain modèle est modélisée par une variable aléatoire X telle que $X \hookrightarrow \mathcal{E}(\lambda)$, avec $\lambda = \frac{1}{5000}$.

On donne les valeurs approchées suivantes : $e^{-0,4} \approx 0,6703$, $e^{-0,6} \approx 0,5488$, $e^{-1} \approx 0,3679$, $e^{-1,6} \approx 0,2019$, $\ln(0,9) \approx -0,1054$ et $\ln(2) \approx 0,6931$.

1. Donner une densité de X et sa fonction de répartition, puis $E(X)$ et $\sigma(X)$.
2. Calculer $P(X > 5000)$, $P(X \leq 2000)$ et $P(3000 \leq X \leq 8000)$, en donnant une valeur approchée à 10^{-3} près.
3. Déterminer la durée t , en heures, telle que $P(X > t) = 0,9$.
4. La durée de vie moyenne vaut 5000 heures, et pourtant $P(X > 5000) \approx 0,37$ seulement. Expliquer ce phénomène.

Corrigé

1. La loi exponentielle de paramètre $\lambda > 0$ admet pour densité la fonction f définie par $f(x) = \lambda e^{-\lambda x}$ si $x \geq 0$ et $f(x) = 0$ si $x < 0$. Ici $\lambda = \frac{1}{5000}$, donc

$$f(x) = \frac{1}{5000} e^{-\frac{x}{5000}} \text{ si } x \geq 0, \quad f(x) = 0 \text{ si } x < 0$$

Fonction de répartition. Pour $x < 0$, la fonction à intégrer est nulle, donc $F_X(x) = 0$. Pour $x \geq 0$,

$$F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_0^x = 1 - e^{-\lambda x}$$

On retiendra surtout la forme complémentaire, qui sert dans presque tous les calculs :

$$P(X > x) = 1 - F_X(x) = e^{-\lambda x} = e^{-\frac{x}{5000}} \quad \text{pour } x \geq 0$$

Moments. Le cours donne $E(X) = \frac{1}{\lambda}$ et $V(X) = \frac{1}{\lambda^2}$, donc

$$E(X) = 5000 \text{ heures}, \quad V(X) = 5000^2 = 25\,000\,000, \quad \sigma(X) = 5000 \text{ heures}$$

La loi exponentielle a cette particularité remarquable : son écart-type est égal à son espérance. Une dispersion aussi grande que la moyenne elle-même annonce déjà que la durée de vie sera très irrégulière d'une ampoule à l'autre.

2. *Première probabilité.* Avec $\frac{5000}{5000} = 1$:

$$P(X > 5000) = e^{-1} \approx 0,368$$

Deuxième probabilité. Avec $\frac{2000}{5000} = 0,4$:

$$P(X \leq 2000) = F_X(2000) = 1 - e^{-0,4} \approx 1 - 0,6703 = 0,330$$

Troisième probabilité. Avec $\frac{3000}{5000} = 0,6$ et $\frac{8000}{5000} = 1,6$:

$$\begin{aligned} P(3000 \leq X \leq 8000) &= F_X(8000) - F_X(3000) \\ &= (1 - e^{-1,6}) - (1 - e^{-0,6}) \\ &= e^{-0,6} - e^{-1,6} \\ &\approx 0,5488 - 0,2019 = 0,3469 \end{aligned}$$

soit environ 0,347. On remarquera que les deux « 1 » se sont simplifiés : il est plus rapide, et moins risqué, d'écrire directement la différence des deux exponentielles.

3. Pour $t \geq 0$, la formule établie à la question **1.** donne $P(X > t) = e^{-\frac{t}{5000}}$. L'équation à résoudre est donc

$$e^{-\frac{t}{5000}} = 0,9$$

La fonction exponentielle étant strictement croissante sur $\mathbb{R}$, donc injective, on peut composer par le logarithme népérien :

$$-\frac{t}{5000} = \ln(0,9) \quad \text{soit} \quad t = -5000 \ln(0,9)$$

Numériquement, $t \approx -5000 \times (-0,1054) = 527,0$. Cette valeur est positive, elle est donc acceptable, et

$$t \approx 527 \text{ heures}$$

Interprétation. Neuf ampoules sur dix dépassent environ 527 heures de fonctionnement, c'est-à-dire un peu plus de trois semaines en marche continue. C'est très peu au regard des 5000 heures annoncées comme durée de vie moyenne.

4. L'explication tient à la forme de la densité. La fonction $x \mapsto \lambda e^{-\lambda x}$ est strictement décroissante sur $[0, +\infty[$: les petites durées sont les plus probables, et la distribution est fortement dissymétrique. À droite, en revanche, la loi n'est pas bornée : quelques ampoules durent très longtemps et tirent la moyenne vers le haut. La moyenne n'est donc pas la valeur « typique » de la durée de vie.

Le calcul de la médiane m le montre clairement. Elle est définie par $F_X(m) = 0,5$, soit

$$1 - e^{-\frac{m}{5000}} = 0,5 \iff e^{-\frac{m}{5000}} = 0,5 \iff m = 5000 \ln(2) \approx 5000 \times 0,6931 = 3465,5$$

Une ampoule sur deux grille donc avant environ 3466 heures, alors que la moyenne annoncée est de 5000 heures. Il n'y a aucune contradiction : sur une loi dissymétrique, moyenne et médiane sont deux indicateurs différents, et c'est la médiane qui décrit le sort de l'ampoule « du milieu ».

Remarque. La loi exponentielle est sans mémoire, ce qui s'illustre ici de façon frappante :

$$P_{(X>5000)}(X > 10\,000) = \frac{P(X > 10\,000)}{P(X > 5000)} = \frac{e^{-2}}{e^{-1}} = e^{-1} = P(X > 5000)$$

Une ampoule qui a déjà tenu 5000 heures a exactement les mêmes chances de tenir 5000 heures de plus qu'une ampoule neuve. Ce modèle convient donc à une panne accidentelle, beaucoup moins à un phénomène d'usure progressive.

Méthode à retenir. Pour la loi exponentielle, mémoriser $P(X > x) = e^{-\lambda x}$ pour $x \geq 0$: c'est la formule qui évite les cascades de $1 - (1 - \dots)$. Retenir aussi $E(X) = \frac{1}{\lambda}$, $V(X) = \frac{1}{\lambda^2}$ et $\sigma(X) = \frac{1}{\lambda}$. Une équation du type $P(X > t) = p$ se résout en composant par le logarithme, et l'on vérifie toujours que la solution obtenue est positive, faute de quoi elle serait à rejeter.

Erreur classique. Confondre le paramètre et l'espérance. Ici $\lambda = \frac{1}{5000}$ et $E(X) = 5000$: prendre $\lambda = 5000$ conduit à des probabilités absurdes, de l'ordre de e^{-10^7} . Deuxième erreur, moins visible mais tout aussi coûteuse : assimiler « durée de vie moyenne » et « durée que la moitié des ampoules dépasse ». La question **4.** est là pour installer la distinction.

Exercice 6 ★ Lecture de la table de la loi normale

Notions : Loi normale, centrage-réduction et lecture de la table

On note Φ la fonction de répartition de la loi normale centrée réduite. On dispose de l'extrait de table suivant.

t	0,5	1	1,5	2	2,5
$\Phi(t)$	0,6915	0,8413	0,9332	0,9772	0,9938

- Soit $Z \hookrightarrow \mathcal{N}(0, 1)$. Calculer les probabilités suivantes.
 - $P(Z \leq 1)$
 - $P(Z \geq 1)$
 - $P(-1 \leq Z \leq 1)$
 - $P(Z \leq -1,5)$
 - $P(|Z| \geq 2)$
- Soit $Y \hookrightarrow \mathcal{N}(20, 16)$. Calculer $P(Y \leq 24)$, $P(16 \leq Y \leq 28)$ et $P(Y \geq 30)$ en passant par la variable centrée réduite associée.
- Déterminer le réel positif a tel que $P(20 - a \leq Y \leq 20 + a) = 0,9544$.

Corrigé

Préliminaires. Trois résultats du cours servent en permanence dans cet exercice.

Symétrie. La densité φ de la loi normale centrée réduite est paire, donc pour tout réel t :

$$\Phi(-t) = 1 - \Phi(t)$$

C'est ce qui permet de se contenter d'une table donnant les seules valeurs positives.

Centrage-réduction. Si $Y \hookrightarrow \mathcal{N}(m, \sigma^2)$ avec $\sigma > 0$, alors $Z = \frac{Y - m}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$.

Continuité. La loi normale est une loi à densité, donc $P(Z = t) = 0$ pour tout réel t : les inégalités strictes et larges se valent, ce qui sera utilisé sans le redire.

1.

a. Lecture directe de la table :

$$P(Z \leq 1) = \Phi(1) = 0,8413$$

b. Par passage au complémentaire :

$$P(Z \geq 1) = 1 - \Phi(1) = 1 - 0,8413 = 0,1587$$

c. L'intervalle est symétrique par rapport à 0 :

$$P(-1 \leq Z \leq 1) = \Phi(1) - \Phi(-1) = \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1$$

soit $2 \times 0,8413 - 1 = 1,6826 - 1 = 0,6826$.

d. L'argument est négatif, on applique la formule de symétrie :

$$P(Z \leq -1,5) = \Phi(-1,5) = 1 - \Phi(1,5) = 1 - 0,9332 = 0,0668$$

e. L'événement $(|Z| \geq 2)$ est la réunion des deux événements incompatibles $(Z \leq -2)$ et $(Z \geq 2)$, qui ont la même probabilité par symétrie. Donc

$$P(|Z| \geq 2) = 2 \times (1 - \Phi(2)) = 2 \times (1 - 0,9772) = 2 \times 0,0228 = 0,0456$$

Contrôle. On peut aussi passer par le complémentaire : $1 - P(-2 \leq Z \leq 2) = 1 - (2\Phi(2) - 1) = 2 - 2 \times 0,9772 = 0,0456$. Les deux chemins concordent.

2. Dans la notation $\mathcal{N}(m, \sigma^2)$, le second paramètre est la variance. Ici $m = 20$ et $\sigma^2 = 16$, donc $\sigma = \sqrt{16} = 4$. On pose

$$Z = \frac{Y - 20}{4}, \quad \text{qui suit la loi } \mathcal{N}(0, 1)$$

Chaque calcul consiste à appliquer la même transformation aux deux membres de l'inégalité.

Première probabilité. Comme $\frac{24 - 20}{4} = 1$:

$$P(Y \leq 24) = P\left(\frac{Y - 20}{4} \leq 1\right) = P(Z \leq 1) = \Phi(1) = 0,8413$$

Deuxième probabilité. Comme $\frac{16 - 20}{4} = -1$ et $\frac{28 - 20}{4} = 2$:

$$\begin{aligned} P(16 \leq Y \leq 28) &= P(-1 \leq Z \leq 2) \\ &= \Phi(2) - \Phi(-1) \\ &= \Phi(2) - (1 - \Phi(1)) \\ &= 0,9772 - (1 - 0,8413) \\ &= 0,9772 - 0,1587 = 0,8185 \end{aligned}$$

Troisième probabilité. Comme $\frac{30 - 20}{4} = 2,5$:

$$P(Y \geq 30) = P(Z \geq 2,5) = 1 - \Phi(2,5) = 1 - 0,9938 = 0,0062$$

3. L'intervalle $[20 - a, 20 + a]$ est centré sur la moyenne $m = 20$. En retranchant 20 puis en divisant par 4, l'encadrement $20 - a \leq Y \leq 20 + a$ équivaut à

$$-\frac{a}{4} \leq Z \leq \frac{a}{4}$$

Posons $u = \frac{a}{4}$, qui est positif puisque a l'est. D'après la question **1. c.**, appliquée cette fois à u :

$$P(20 - a \leq Y \leq 20 + a) = P(-u \leq Z \leq u) = 2\Phi(u) - 1$$

L'équation devient donc

$$2\Phi(u) - 1 = 0,9544 \iff \Phi(u) = \frac{1,9544}{2} = 0,9772$$

La table donne $\Phi(2) = 0,9772$. La fonction Φ étant strictement croissante sur $\mathbb{R}$, donc injective, cette valeur est atteinte en un seul point : $u = 2$, puis

$$a = 4u = 8$$

Vérification. L'intervalle obtenu est $[12, 28]$, c'est-à-dire $[m - 2\sigma, m + 2\sigma]$. On retrouve un fait à connaître : une variable normale s'écarte de sa moyenne de moins de deux écarts-types avec une probabilité voisine de 0,95.

Méthode à retenir. Trois réflexes suffisent. Premièrement, toute probabilité relative à $\mathcal{N}(m, \sigma^2)$ se ramène à Φ en appliquant $z \mapsto \frac{z - m}{\sigma}$ à toutes les bornes de l'inégalité, sans oublier aucune. Deuxièmement, la table ne donne que des arguments positifs : pour un argument négatif, on utilise $\Phi(-t) = 1 - \Phi(t)$. Troisièmement, un intervalle symétrique autour de m se traite avec $P(-u \leq Z \leq u) = 2\Phi(u) - 1$, formule qui sert aussi bien à calculer qu'à résoudre une équation, comme en question **3**.

Erreur classique. Lire $\sigma = 16$ dans $\mathcal{N}(20, 16)$ alors que 16 est la variance et que l'écart-type vaut 4. Toutes les probabilités deviennent alors fausses, sans que rien ne le signale. Autre erreur, plus grossière mais fréquente sous la pression : écrire $\Phi(-1,5) = -\Phi(1,5)$. Une fonction de répartition est à valeurs dans $[0, 1]$ et ne peut pas être négative ; c'est la densité φ qui est paire, et cette parité se traduit sur Φ par $\Phi(-t) = 1 - \Phi(t)$.

Exercice 7 ★ Transformation affine d'une variable à densité

Notions : Loi d'une transformée : transformation affine et méthode de la fonction de répartition, Loi uniforme sur un segment, Loi exponentielle et absence de mémoire

Dans tout l'exercice, la loi d'une transformée se cherche par la méthode de la fonction de répartition : on écrit $F_Y(y) = P(Y \leq y)$, on traduit l'événement en un événement portant sur X , on l'exprime à l'aide de F_X , puis on dérive.

1. Soit $X \hookrightarrow \mathcal{U}([0, 1])$ et $Y = 3X + 2$. Montrer que $Y \hookrightarrow \mathcal{U}([2, 5])$.
2. Soit $\lambda > 0$, soit $X \hookrightarrow \mathcal{E}(\lambda)$ et $Y = 2X$. Montrer que $Y \hookrightarrow \mathcal{E}\left(\frac{\lambda}{2}\right)$.
3. Soit $X \hookrightarrow \mathcal{E}(1)$ et $Y = -X$. Déterminer une densité de Y . On prendra garde au sens des inégalités.
4. Cas général. Soit X une variable à densité, de densité f et de fonction de répartition F_X , soient a et b deux réels avec $a \neq 0$, et soit $Y = aX + b$. Établir que Y est une variable à densité et que

$$f_Y(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right)$$

en traitant séparément les cas $a > 0$ et $a < 0$. Retrouver enfin $E(Y) = aE(X) + b$ et $V(Y) = a^2 V(X)$.

Corrigé

1. La fonction de répartition de la loi $\mathcal{U}([0, 1])$ vaut $F_X(x) = 0$ pour $x < 0$, $F_X(x) = x$ pour $0 \leq x \leq 1$ et $F_X(x) = 1$ pour $x > 1$.

Soit y un réel. On part de la définition, et l'on isole X :

$$F_Y(y) = P(Y \leq y) = P(3X + 2 \leq y) = P\left(X \leq \frac{y-2}{3}\right) = F_X\left(\frac{y-2}{3}\right)$$

Le sens de l'inégalité est conservé, car on divise par 3, qui est strictement positif. Il reste à discuter selon la position de $\frac{y-2}{3}$ par rapport à 0 et à 1.

Si $\frac{y-2}{3} < 0$, c'est-à-dire $y < 2$, alors $F_Y(y) = 0$. Si $0 \leq \frac{y-2}{3} \leq 1$, c'est-à-dire $2 \leq y \leq 5$, alors $F_Y(y) = \frac{y-2}{3}$. Si $\frac{y-2}{3} > 1$, c'est-à-dire $y > 5$, alors $F_Y(y) = 1$.

Or la fonction de répartition de la loi $\mathcal{U}([\alpha, \beta])$ vaut $\frac{y-\alpha}{\beta-\alpha}$ sur $[\alpha, \beta]$. Avec $\alpha = 2$ et $\beta = 5$, cela donne exactement $\frac{y-2}{5-2} = \frac{y-2}{3}$, et les deux morceaux constants coïncident également. Deux variables ayant la même fonction de répartition ont la même loi, donc

$$Y \hookrightarrow \mathcal{U}([2, 5]) \quad \square$$

Contrôle. $E(Y) = 3E(X) + 2 = 3 \times \frac{1}{2} + 2 = \frac{7}{2}$, qui est bien le milieu $\frac{2+5}{2}$ du segment $[2, 5]$.

2. La fonction de répartition de la loi $\mathcal{E}(\lambda)$ vaut $F_X(x) = 1 - e^{-\lambda x}$ pour $x \geq 0$ et $F_X(x) = 0$ pour $x < 0$.

Soit y un réel. Comme $2 > 0$, l'inégalité est conservée :

$$F_Y(y) = P(2X \leq y) = P\left(X \leq \frac{y}{2}\right) = F_X\left(\frac{y}{2}\right)$$

Si $y < 0$, alors $\frac{y}{2} < 0$ et $F_Y(y) = 0$. Si $y \geq 0$, alors $\frac{y}{2} \geq 0$ et

$$F_Y(y) = 1 - e^{-\lambda \frac{y}{2}} = 1 - e^{-\frac{\lambda}{2}y}$$

On reconnaît la fonction de répartition de la loi exponentielle de paramètre $\frac{\lambda}{2}$, qui est bien un réel strictement positif. Donc

$$Y \hookrightarrow \mathcal{E}\left(\frac{\lambda}{2}\right) \quad \square$$

Contrôle. $E(Y) = 2E(X) = \frac{2}{\lambda}$, ce qui est bien l'inverse du paramètre $\frac{\lambda}{2}$. Autrement dit, doubler une durée exponentielle revient à diviser son paramètre par deux : c'est cohérent avec l'interprétation de λ comme un taux.

3. Ici $X \hookrightarrow \mathcal{E}(1)$, donc $F_X(x) = 1 - e^{-x}$ pour $x \geq 0$ et $F_X(x) = 0$ pour $x < 0$. La transformation $x \mapsto -x$ est strictement décroissante : c'est là que le sens de l'inégalité change.

Soit y un réel.

$$F_Y(y) = P(-X \leq y) = P(X \geq -y) = 1 - P(X < -y) = 1 - F_X(-y)$$

La dernière égalité utilise $P(X < -y) = P(X \leq -y)$, valable parce que X est une variable à densité, donc $P(X = -y) = 0$.

Si $y > 0$, alors $-y < 0$, donc $F_X(-y) = 0$ et $F_Y(y) = 1$. Si $y \leq 0$, alors $-y \geq 0$, donc $F_X(-y) = 1 - e^{-(-y)} = 1 - e^y$, d'où $F_Y(y) = e^y$.

Nature de Y . La fonction F_Y est croissante (elle croît de 0 à 1 sur $]-\infty, 0]$, puis reste égale à 1), de limites 0 en $-\infty$ et 1 en $+\infty$, et elle se raccorde continûment en 0 puisque $e^0 = 1$. Elle est donc continue sur $\mathbb{R}$, et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement en 0, soit en un nombre fini de points : Y est une variable à densité.

En dérivant là où c'est licite, une densité de Y est donnée par

$$f_Y(y) = e^y \text{ si } y \leq 0, \quad f_Y(y) = 0 \text{ si } y > 0$$

Contrôle. La fonction f_Y est positive et continue sauf en 0. Son intégrale est impropre en $-\infty$: pour $B < 0$,

$$\int_B^0 e^t dt = [e^t]_B^0 = 1 - e^B$$

Quand B tend vers $-\infty$, cette quantité tend vers 1 : l'intégrale converge et vaut 1. Il s'agit bien d'une densité.

Interprétation. Y ne prend que des valeurs négatives, et sa densité est la symétrique de celle de X par rapport à l'axe des ordonnées. C'est exactement ce qu'on attend de la variable $-X$.

4. Cas $a > 0$. Pour tout réel y , diviser par $a > 0$ conserve le sens de l'inégalité :

$$aX + b \leq y \iff X \leq \frac{y-b}{a}$$

donc

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right)$$

La fonction $y \mapsto \frac{y-b}{a}$ est continue sur $\mathbb{R}$ et F_X l'est aussi, donc F_Y est continue sur $\mathbb{R}$ par composition. De plus, F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points $x_1, \dots, x_n$; par composition avec une fonction affine, F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement aux points $ax_1 + b, \dots, ax_n + b$, qui sont encore en nombre fini. Donc Y est une variable à densité, et en dérivant là où c'est licite :

$$f_Y(y) = \frac{1}{a} F'_X\left(\frac{y-b}{a}\right) = \frac{1}{a} f\left(\frac{y-b}{a}\right)$$

Cas $a < 0$. Cette fois, diviser par $a < 0$ renverse l'inégalité :

$$aX + b \leq y \iff X \geq \frac{y-b}{a}$$

En passant au complémentaire, et en utilisant $P\left(X = \frac{y-b}{a}\right) = 0$ puisque X est à densité :

$$F_Y(y) = P\left(X \geq \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right)$$

Le même argument de composition montre que F_Y est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf en un nombre fini de points, donc Y est encore une variable à densité, et

$$f_Y(y) = -\frac{1}{a} f\left(\frac{y-b}{a}\right)$$

Synthèse. Si $a > 0$, alors $\frac{1}{a} = \frac{1}{|a|}$; si $a < 0$, alors $-\frac{1}{a} = \frac{1}{|a|}$. Les deux cas se résument donc en une seule formule :

$$f_Y(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right) \quad \square$$

Le facteur $\frac{1}{|a|}$ n'est pas décoratif : il compense la dilatation de l'axe des abscisses par le facteur a , de sorte que l'aire totale sous la courbe reste égale à 1.

Espérance et variance. Dès que $E(X)$ existe, la linéarité de l'espérance, admise par le programme dans le cas général, donne

$$E(Y) = E(aX + b) = aE(X) + b$$

et dès que $V(X)$ existe, la propriété $V(\alpha X + \beta) = \alpha^2 V(X)$ donne

$$V(Y) = V(aX + b) = a^2 V(X)$$

Vérification sur les trois premières questions.

En question 1., $a = 3$ et $b = 2$: la formule générale donne $f_Y(y) = \frac{1}{3} f\left(\frac{y-2}{3}\right)$, soit $\frac{1}{3}$ lorsque $\frac{y-2}{3} \in [0, 1]$, c'est-à-dire sur $[2, 5]$. C'est bien la densité de $\mathcal{U}([2, 5])$. On a de plus $E(Y) = \frac{7}{2}$ et $V(Y) = 9 \times \frac{1}{12} = \frac{3}{4}$, alors que la loi $\mathcal{U}([2, 5])$ donne $\frac{2+5}{2} = \frac{7}{2}$ et $\frac{(5-2)^2}{12} = \frac{9}{12} = \frac{3}{4}$.

En question 2., $a = 2$ et $b = 0$: la formule donne $f_Y(y) = \frac{1}{2} \lambda e^{-\lambda \frac{y}{2}} = \frac{\lambda}{2} e^{-\frac{\lambda}{2} y}$ pour $y \geq 0$, qui est bien la densité de $\mathcal{E}\left(\frac{\lambda}{2}\right)$. Et $V(Y) = 4 \times \frac{1}{\lambda^2} = \frac{4}{\lambda^2} = \frac{1}{\left(\frac{\lambda}{2}\right)^2}$.

En question 3., $a = -1$ et $b = 0$: la formule donne $f_Y(y) = \frac{1}{|-1|} f(-y) = f(-y)$. Pour $y \leq 0$, on a $-y \geq 0$ et $f(-y) = e^{-(-y)} = e^y$. On retrouve la densité obtenue directement.

Méthode à retenir. La loi d'une transformée se cherche par la fonction de répartition, dans cet ordre invariable : écrire $F_Y(y) = P(Y \leq y)$; traduire l'événement en un événement portant sur X , en surveillant le sens de l'inégalité ; exprimer le résultat avec F_X ; discuter selon les valeurs de y ; enfin reconnaître une loi usuelle, ou bien justifier que F_Y est continue et de classe $\mathcal{C}^1$ sauf en un nombre fini de points, puis dériver. On ne dérive jamais la densité de X directement.

Erreur classique. Écrire $f_Y(y) = f\left(\frac{y-b}{a}\right)$ en oubliant le facteur $\frac{1}{|a|}$: la fonction obtenue n'a plus pour intégrale 1. Et lorsque $a < 0$, conserver $\frac{1}{a}$ au lieu de $\frac{1}{|a|}$ produit une densité négative, ce qui est impossible. Deux contrôles immédiats permettent de s'en apercevoir : une densité est positive, et son intégrale vaut 1.

Exercice 8 ★ Espérance et variance à partir d'une densité

Notions : Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Variance, écart-type et moments d'une variable à densité

Les deux questions sont indépendantes l'une de l'autre.

1. Soit X une variable aléatoire réelle de densité f définie sur $\mathbb{R}$ par $f(x) = 2x$ si $x \in [0, 1]$, et $f(x) = 0$ sinon. Justifier l'existence de $E(X)$, puis calculer $E(X)$, $E(X^2)$, $V(X)$ et $\sigma(X)$.
2. Soit Y une variable aléatoire réelle de densité g définie sur $\mathbb{R}$ par $g(y) = \frac{2}{y^3}$ si $y \geq 1$, et $g(y) = 0$ sinon.
 - a. Vérifier que g est bien une densité de probabilité.
 - b. Montrer que $E(Y)$ existe, puis calculer sa valeur.
 - c. Montrer que $E(Y^2)$ n'existe pas. Qu'en conclut-on quant à la variance de Y ?

Corrigé

1. Existence de l'espérance. Par définition, $E(X)$ existe si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t f(t) dt$ converge absolument. La fonction $t \mapsto t f(t)$ est nulle en dehors de $[0, 1]$, donc

$$\int_{-\infty}^{+\infty} |t| f(t) dt = \int_0^1 t \times 2t dt = \int_0^1 2t^2 dt$$

Cette dernière intégrale est celle d'une fonction continue sur un segment : elle n'est pas impropre, donc elle converge. L'espérance $E(X)$ existe. Sur un support borné, cette justification tient en une phrase, mais elle doit être écrite.

$$E(X) = \int_0^1 2t^2 dt = \left[\frac{2t^3}{3} \right]_0^1 = \frac{2}{3} \approx 0,667$$

Moment d'ordre deux. Le théorème de transfert, admis par le programme, appliqué à la fonction $t \mapsto t^2$, ramène le calcul à une intégrale portant sur le même segment, donc convergente pour la même raison :

$$E(X^2) = \int_0^1 t^2 \times 2t dt = \int_0^1 2t^3 dt = \left[\frac{t^4}{2} \right]_0^1 = \frac{1}{2}$$

Variance et écart-type. Le moment d'ordre deux existant, la variance existe et la formule de König-Huygens s'applique :

$$V(X) = E(X^2) - (E(X))^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}$$

$$\sigma(X) = \sqrt{\frac{1}{18}} = \frac{1}{3\sqrt{2}} = \frac{\sqrt{2}}{6} \approx 0,236$$

Cohérence. La densité $t \mapsto 2t$ est croissante sur $[0, 1]$: elle charge davantage les valeurs proches de 1. Il est donc normal que $E(X) = \frac{2}{3}$ dépasse le milieu $\frac{1}{2}$ du segment. La variance, elle, est strictement positive et petite, ce qui traduit une variable resserrée sur un intervalle de longueur 1.

2. a. Positivité. Pour $y \geq 1$, $\frac{2}{y^3} > 0$, et g est nulle ailleurs : g est positive sur $\mathbb{R}$.

Continuité. Sur $] -\infty, 1[$, g est nulle donc continue ; sur $] 1, +\infty[$, elle est continue comme quotient de fonctions continues dont le dénominateur ne s'annule pas. En 1, elle saute de 0 à 2 : c'est l'unique point de discontinuité, donc un nombre fini de points.

Intégrale. Le support n'est pas borné, l'intégrale est impropre en $+\infty$. Pour $A > 1$:

$$\int_1^A \frac{2}{t^3} dt = \int_1^A 2t^{-3} dt = [-t^{-2}]_1^A = -\frac{1}{A^2} + 1 = 1 - \frac{1}{A^2}$$

Lorsque A tend vers $+\infty$, cette quantité admet la limite finie 1. L'intégrale $\int_1^{+\infty} g(t) dt$ converge donc et vaut 1 : g est bien une densité de probabilité.

2. b. $E(Y)$ existe si et seulement si $\int_{-\infty}^{+\infty} |t| g(t) dt$ converge. La fonction $t \mapsto t g(t)$ est nulle sur $] -\infty, 1[$ et positive sur $] 1, +\infty[$, donc $|t| g(t) = t g(t)$ sur tout $\mathbb{R}$: la convergence absolue équivaut ici à la convergence simple. Il suffit donc d'étudier $\int_1^{+\infty} t g(t) dt$. Pour $A > 1$:

$$\int_1^A t \times \frac{2}{t^3} dt = \int_1^A \frac{2}{t^2} dt = \left[-\frac{2}{t} \right]_1^A = -\frac{2}{A} + 2 = 2 - \frac{2}{A}$$

Lorsque A tend vers $+\infty$, cette quantité admet la limite finie 2. L'intégrale converge, donc $E(Y)$ existe et

$$E(Y) = 2$$

2. c. Par le théorème de transfert appliqué à $t \mapsto t^2$, le moment $E(Y^2)$ existe si et seulement si $\int_{-\infty}^{+\infty} t^2 g(t) dt$ converge absolument. Là encore la fonction à intégrer est positive, la convergence absolue équivaut donc à la convergence. Pour $A > 1$:

$$\int_1^A t^2 \times \frac{2}{t^3} dt = \int_1^A \frac{2}{t} dt = [2 \ln t]_1^A = 2 \ln A - 2 \ln 1 = 2 \ln A$$

Lorsque A tend vers $+\infty$, $2 \ln A$ tend vers $+\infty$: la quantité n'admet pas de limite finie, donc l'intégrale $\int_1^{+\infty} t^2 g(t) dt$ diverge. Par conséquent, $E(Y^2)$ n'existe pas.

Conclusion sur la variance. La variance de Y est par définition $V(Y) = E((Y - E(Y))^2)$, et son existence équivaut à celle du moment d'ordre deux. Comme $E(Y^2)$ n'existe pas, Y n'admet pas de variance. La formule de König-Huygens ne peut évidemment pas être invoquée : l'un de ses deux termes n'a pas de sens.

Bilan de l'exercice. La variable Y possède une espérance mais pas de variance. La densité g décroît en $\frac{1}{t^3}$, ce qui suffit tout juste à faire converger $\int t g(t) dt$, dont l'intégrande se comporte en $\frac{1}{t^2}$, mais pas $\int t^2 g(t) dt$, dont l'intégrande se comporte en $\frac{1}{t}$. Chaque puissance supplémentaire de t rend la convergence plus difficile : l'existence d'un moment se démontre, elle ne se suppose jamais.

Méthode à retenir. Deux situations, deux rédactions. Sur un support borné, la fonction à intégrer est continue sur un segment, l'intégrale n'est pas impropre, et il suffit de le dire avant de calculer. Sur un support non borné, on intègre entre la borne finie et A , on calcule une primitive, on fait tendre A vers $+\infty$, et l'on conclut sur la convergence avant d'annoncer la moindre valeur. Quand la fonction à intégrer est positive, on signale en une ligne que convergence absolue et convergence simple coïncident : cela évite de traiter séparément la valeur absolue.

Erreur classique. Écrire « $E(Y^2) = \int_1^{+\infty} \frac{2}{t} dt = +\infty$, donc la variance est infinie ». En ECG, une intégrale divergente ne fournit aucune valeur, même infinie : elle interdit simplement de parler du moment. La formulation attendue est « $E(Y^2)$ n'existe pas », donc « Y n'admet pas de variance ». L'erreur symétrique consiste à appliquer König-Huygens sans avoir vérifié au préalable l'existence du moment d'ordre deux : le calcul produit alors un nombre, et ce nombre ne veut rien dire.

Exercice 9 ★ Formule de l'espérance totale

Notions : Espérance conditionnelle et formule de l'espérance totale

Une usine fabrique des pièces dans trois ateliers, qui assurent respectivement 50 %, 30 % et 20 % de la production totale. On prélève une pièce au hasard dans cette production, et on note A_1 , A_2 et A_3 les événements « la pièce provient du premier atelier », « du deuxième atelier », « du troisième atelier ».

On note X le nombre de défauts de la pièce prélevée, qui vaut 0 ou 1. Sachant l'atelier d'origine, X suit une loi de Bernoulli de paramètre 0,2 pour A_1 , 0,5 pour A_2 et 0,8 pour A_3 .

1. Donner $E(X | A_1)$, $E(X | A_2)$ et $E(X | A_3)$.
2. En déduire $E(X)$ par la formule de l'espérance totale.
3. Retrouver ce résultat en déterminant d'abord la loi de X .
4. Chaque défaut coûte 3 euros de reprise, de sorte que le coût de reprise d'une pièce est $Y = 3X$ euros. Calculer $E(Y)$.
5. On prélève une pièce et l'on constate qu'elle présente un défaut. Quelle est la probabilité qu'elle provienne du troisième atelier ?

Corrigé

Préliminaire. Une pièce provient d'un atelier et d'un seul, donc (A_1, A_2, A_3) est un système complet d'événements, avec

$$P(A_1) = 0,5, \quad P(A_2) = 0,3, \quad P(A_3) = 0,2$$

et $0,5 + 0,3 + 0,2 = 1$. Ces trois probabilités sont non nulles, ce qui autorise à conditionner par chacun de ces événements : c'est la condition d'emploi de l'espérance conditionnelle comme de la formule des probabilités totales.

1. Sachant A_i , la variable X suit une loi de Bernoulli de paramètre p_i , dont l'espérance vaut p_i . La variable ne prenant que les valeurs 0 et 1, elle est bornée et toutes ces espérances existent, sans hypothèse à vérifier. Donc

$$E(X | A_1) = 0,2, \quad E(X | A_2) = 0,5, \quad E(X | A_3) = 0,8$$

Le détail, pour le premier atelier, tient en une ligne :

$$E(X | A_1) = 0 \times P_{A_1}(X = 0) + 1 \times P_{A_1}(X = 1) = P_{A_1}(X = 1) = 0,2$$

2. La formule de l'espérance totale, appliquée au système complet (A_1, A_2, A_3) , s'écrit

$$E(X) = P(A_1) E(X | A_1) + P(A_2) E(X | A_2) + P(A_3) E(X | A_3)$$

Chaque terme est un produit « poids de l'atelier multiplié par la moyenne dans cet atelier ». Numériquement :

$$\begin{aligned} E(X) &= 0,5 \times 0,2 + 0,3 \times 0,5 + 0,2 \times 0,8 \\ &= 0,10 + 0,15 + 0,16 \\ &= 0,41 \end{aligned}$$

En moyenne, une pièce prélevée au hasard présente 0,41 défaut. Comme X ne vaut que 0 ou 1, ce nombre s'interprète aussi comme la proportion de pièces défectueuses dans la production.

3. La variable X prend ses valeurs dans $\{0, 1\}$. La formule des probabilités totales, appliquée au même système complet, donne

$$\begin{aligned} P(X = 1) &= P(A_1) P_{A_1}(X = 1) + P(A_2) P_{A_2}(X = 1) + P(A_3) P_{A_3}(X = 1) \\ &= 0,5 \times 0,2 + 0,3 \times 0,5 + 0,2 \times 0,8 \\ &= 0,41 \end{aligned}$$

et par passage au complémentaire, $P(X = 0) = 1 - 0,41 = 0,59$. La variable X suit donc la loi de Bernoulli de paramètre 0,41, d'où

$$E(X) = 0 \times 0,59 + 1 \times 0,41 = 0,41$$

On retrouve exactement la valeur de la question **2.**, ce qui valide les deux calculs. Les deux formules, celle des probabilités totales et celle de l'espérance totale, font d'ailleurs apparaître les mêmes produits : sur une variable de Bernoulli, elles coïncident terme à terme, puisque l'espérance conditionnelle y est égale à la probabilité conditionnelle de l'événement ($X = 1$).

4. La variable $Y = 3X$ ne prend que les valeurs 0 et 3, elle est donc bornée et son espérance existe. Par linéarité de l'espérance :

$$E(Y) = E(3X) = 3 E(X) = 3 \times 0,41 = 1,23$$

Le coût moyen de reprise s'élève donc à 1,23 euro par pièce produite, soit 12 300 euros pour un lot de 10 000 pièces. Il n'était pas nécessaire de déterminer la loi de Y : la linéarité suffit, et c'est tout l'intérêt de la propriété.

5. On cherche $P_{(X=1)}(A_3)$. L'événement conditionnant a pour probabilité $P(X = 1) = 0,41$, qui est non nulle : la probabilité conditionnelle est bien définie. La formule de Bayes donne

$$P_{(X=1)}(A_3) = \frac{P(A_3) P_{A_3}(X = 1)}{P(X = 1)} = \frac{0,2 \times 0,8}{0,41} = \frac{0,16}{0,41} \approx 0,390$$

Interprétation. Le troisième atelier ne fabrique que 20 % des pièces, mais il produit environ 39 % des pièces défectueuses. Constaté un défaut rend donc cet atelier presque deux fois plus « suspect » qu'il ne l'était avant l'observation.

Contrôle. Les deux autres probabilités a posteriori se calculent de la même façon :

$$P_{(X=1)}(A_1) = \frac{0,10}{0,41} \approx 0,244, \quad P_{(X=1)}(A_2) = \frac{0,15}{0,41} \approx 0,366$$

Leur somme avec la précédente vaut $\frac{0,10 + 0,15 + 0,16}{0,41} = \frac{0,41}{0,41} = 1$, comme il se doit puisque (A_1, A_2, A_3) reste un système complet une fois conditionné.

Méthode à retenir. La formule de l'espérance totale est à l'espérance ce que la formule des probabilités totales est à la probabilité. Le schéma est toujours le même : repérer un système complet d'événements (A_i) de probabilités non nulles, calculer chaque espérance conditionnelle $E(X | A_i)$, puis sommer les produits $P(A_i) E(X | A_i)$. Quand la variable est simple, il est prudent de recalculer l'espérance à partir de la loi complète : les deux résultats doivent coïncider, et cette double lecture repère immédiatement une erreur de pondération.

Erreur classique. Moyenner les trois espérances conditionnelles sans les pondérer, et annoncer $\frac{0,2 + 0,5 + 0,8}{3} = 0,5$. Cette valeur ne serait correcte que si chaque atelier produisait un tiers des pièces, ce qui n'est pas le cas ici. Autre confusion, très fréquente en question **5.** : écrire $P_{(X=1)}(A_3) = P_{A_3}(X = 1) = 0,8$. Ces deux probabilités conditionnelles sont inverses l'une de l'autre, et c'est précisément la formule de Bayes qui permet de passer de la seconde à la première.

Exercice 10 ★★ Constante de normalisation et étude complète

Notions : Variable aléatoire à densité : définition et reconnaissance d'une densité, Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Variance, écart-type et moments d'une variable à densité

Soit c un réel. On considère la fonction f définie sur $\mathbb{R}$ par

$$f(x) = c(1 - x^2) \quad \text{si } x \in [-1, 1], \quad f(x) = 0 \quad \text{sinon.}$$

1. Déterminer la valeur de c pour laquelle f est une densité de probabilité.
2. Dans toute la suite, X désigne une variable aléatoire à densité admettant f pour densité. Déterminer la fonction de répartition F_X de X sur chacun des trois intervalles $]-\infty, -1[$, $[-1, 1]$ et $]1, +\infty[$.
3. Montrer que X admet une espérance et la calculer.
4. Montrer que X admet une variance et la calculer.
5. Calculer $P(|X| \leq \frac{1}{2})$.
6. Calculer la probabilité conditionnelle $P_{(X \geq 0)}(X \geq \frac{1}{2})$.

Corrigé

1. Une fonction f est une densité de probabilité lorsqu'elle vérifie les trois conditions suivantes : elle est positive sur $\mathbb{R}$, elle est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points, et l'intégrale $\int_{-\infty}^{+\infty} f(t) dt$ converge et vaut 1. On vérifie ces trois points dans cet ordre, systématiquement.

Positivité. Pour $x \in [-1, 1]$, on a $x^2 \leq 1$, donc $1 - x^2 \geq 0$. La positivité de f sur $[-1, 1]$ équivaut donc à $c \geq 0$. En dehors de $[-1, 1]$, f est nulle, donc positive.

Continuité. Sur $]-\infty, -1[$ et sur $]1, +\infty[$, f est nulle donc continue. Sur $]-1, 1[$, f est polynomiale donc continue. Aux points -1 et 1 , la valeur du polynôme est $c(1 - 1) = 0$, qui coïncide avec la valeur nulle prise à l'extérieur : f est en fait continue sur $\mathbb{R}$ tout entier. C'est confortable, mais le programme n'en demande pas tant.

Intégrale. La fonction f est nulle en dehors du segment $[-1, 1]$, et continue sur ce segment : l'intégrale $\int_{-\infty}^{+\infty} f(t) dt$ n'est impropre en aucune borne, elle converge donc sans discussion et se réduit à une intégrale sur $[-1, 1]$.

$$\int_{-\infty}^{+\infty} f(t) dt = c \int_{-1}^1 (1 - t^2) dt = c \left[t - \frac{t^3}{3} \right]_{-1}^1 = c \left(\left(1 - \frac{1}{3}\right) - \left(-1 + \frac{1}{3}\right) \right) = c \times \frac{4}{3}$$

Cette quantité vaut 1 si et seulement si $c = \frac{3}{4}$, valeur qui est bien positive. Ainsi

$$c = \frac{3}{4}$$

2. Par définition, $F_X(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$. On traite les trois zones séparément, car l'expression de f change.

Cas $x < -1$. Sur $]-\infty, x]$, la fonction f est identiquement nulle, donc

$$F_X(x) = 0$$

Cas $-1 \leq x \leq 1$. La relation de Chasles coupe l'intégrale en -1 , et la première partie est nulle :

$$F_X(x) = \int_{-\infty}^{-1} 0 dt + \frac{3}{4} \int_{-1}^x (1 - t^2) dt = \frac{3}{4} \left[t - \frac{t^3}{3} \right]_{-1}^x$$

En évaluant :

$$\begin{aligned} F_X(x) &= \frac{3}{4} \left(x - \frac{x^3}{3} - \left(-1 + \frac{1}{3} \right) \right) \\ &= \frac{3}{4} \left(x - \frac{x^3}{3} + \frac{2}{3} \right) \\ &= \frac{3}{4} \times \frac{3x - x^3 + 2}{3} \\ &= \frac{1}{4} (-x^3 + 3x + 2) \end{aligned}$$

Cas $x > 1$. Toute la masse est déjà accumulée, donc

$$F_X(x) = 1$$

Contrôle. Un raccordement faux est l'erreur la plus fréquente sur ce type de question, il faut donc toujours vérifier les valeurs aux points de jonction. En $x = -1$: $\frac{1}{4}(-1 - 3 + 2) = 0$. En $x = 1$: $\frac{1}{4}(-1 + 3 + 2) = \frac{4}{4} = 1$. La fonction F_X est bien continue sur $\mathbb{R}$, croissante, de limite 0 en $-\infty$ et 1 en $+\infty$: elle a toutes les propriétés attendues d'une fonction de répartition.

3. L'espérance de X existe par définition lorsque l'intégrale $\int_{-\infty}^{+\infty} t f(t) dt$ converge absolument. Ici la fonction $t \mapsto t f(t)$ est continue sur $\mathbb{R}$ et nulle en dehors du segment $[-1, 1]$: l'intégrale se réduit à $\int_{-1}^1 t f(t) dt$, intégrale d'une fonction continue sur un segment, qui n'est impropre nulle part. Elle converge donc, et il en va de même de l'intégrale des valeurs absolues. **L'espérance $E(X)$ existe.**

Une fois cette existence acquise, on peut invoquer la parité. La fonction $g : t \mapsto t \times \frac{3}{4}(1 - t^2)$ est impaire, car $g(-t) = -t \times \frac{3}{4}(1 - t^2) = -g(t)$. L'intégrale d'une fonction impaire, continue, sur un intervalle symétrique par rapport à 0, est nulle. Donc

$$E(X) = 0$$

Le calcul direct confirme, si l'on préfère ne rien invoquer :

$$E(X) = \frac{3}{4} \int_{-1}^1 (t - t^3) dt = \frac{3}{4} \left[\frac{t^2}{2} - \frac{t^4}{4} \right]_{-1}^1 = \frac{3}{4} \left(\left(\frac{1}{2} - \frac{1}{4} \right) - \left(\frac{1}{2} - \frac{1}{4} \right) \right) = 0$$

Ce résultat était prévisible : la densité est paire, la loi est symétrique par rapport à 0, son centre de gravité est donc en 0.

4. La variance existe lorsque X admet un moment d'ordre 2, c'est-à-dire lorsque l'intégrale $\int_{-\infty}^{+\infty} t^2 f(t) dt$ converge. Le même argument qu'à la question précédente s'applique : la fonction $t \mapsto t^2 f(t)$ est continue et nulle hors de $[-1, 1]$, l'intégrale porte donc sur un segment et converge.

La variance $V(X)$ existe.

Le théorème de transfert, appliqué à la fonction $t \mapsto t^2$, donne

$$\begin{aligned} E(X^2) &= \frac{3}{4} \int_{-1}^1 t^2 (1 - t^2) dt \\ &= \frac{3}{4} \int_{-1}^1 (t^2 - t^4) dt \\ &= \frac{3}{4} \left[\frac{t^3}{3} - \frac{t^5}{5} \right]_{-1}^1 \\ &= \frac{3}{4} \times 2 \times \left(\frac{1}{3} - \frac{1}{5} \right) \\ &= \frac{3}{2} \times \frac{2}{15} = \frac{1}{5} \end{aligned}$$

Le facteur 2 vient de la parité de la fonction intégrée. La formule de König-Huygens fournit alors

$$V(X) = E(X^2) - (E(X))^2 = \frac{1}{5} - 0^2 = \frac{1}{5}$$

L'écart-type vaut $\sigma(X) = \frac{1}{\sqrt{5}} \approx 0,447$. Cette valeur est cohérente : la variable vit dans $[-1, 1]$, un écart-type de l'ordre de 0,45 est plausible pour une loi centrée sur un tel support.

5. L'événement $(|X| \leq \frac{1}{2})$ s'écrit $(-\frac{1}{2} \leq X \leq \frac{1}{2})$. Comme X est une variable à densité, on a $P(X = -\frac{1}{2}) = 0$, si bien que les inégalités larges ou strictes donnent la même probabilité. On peut donc utiliser directement la fonction de répartition :

$$P\left(|X| \leq \frac{1}{2}\right) = F_X\left(\frac{1}{2}\right) - F_X\left(-\frac{1}{2}\right)$$

Or, avec l'expression obtenue à la question **2.** :

$$F_X\left(\frac{1}{2}\right) = \frac{1}{4} \left(-\frac{1}{8} + \frac{3}{2} + 2\right) = \frac{1}{4} \times \frac{-1 + 12 + 16}{8} = \frac{1}{4} \times \frac{27}{8} = \frac{27}{32}$$

$$F_X\left(-\frac{1}{2}\right) = \frac{1}{4} \left(\frac{1}{8} - \frac{3}{2} + 2\right) = \frac{1}{4} \times \frac{1 - 12 + 16}{8} = \frac{1}{4} \times \frac{5}{8} = \frac{5}{32}$$

D'où

$$P\left(|X| \leq \frac{1}{2}\right) = \frac{27}{32} - \frac{5}{32} = \frac{22}{32} = \frac{11}{16} = 0,6875$$

Un peu plus des deux tiers de la masse se concentre sur la moitié centrale du support, ce qui traduit bien la forme en cloche de la parabole $1 - x^2$.

6. Par définition de la probabilité conditionnelle, et puisque $P(X \geq 0) \neq 0$ comme on le vérifie ci-dessous,

$$P_{(X \geq 0)}\left(X \geq \frac{1}{2}\right) = \frac{P((X \geq 0) \cap (X \geq \frac{1}{2}))}{P(X \geq 0)}$$

L'inclusion $(X \geq \frac{1}{2}) \subset (X \geq 0)$ simplifie l'intersection : elle vaut $(X \geq \frac{1}{2})$. C'est le point à ne pas manquer, il évite tout calcul supplémentaire.

Dénominateur. La densité étant paire, la symétrie donne immédiatement $P(X \geq 0) = \frac{1}{2}$. Le calcul le confirme :

$$P(X \geq 0) = 1 - F_X(0) = 1 - \frac{1}{4}(0 + 0 + 2) = 1 - \frac{1}{2} = \frac{1}{2}$$

Numérateur.

$$P\left(X \geq \frac{1}{2}\right) = 1 - F_X\left(\frac{1}{2}\right) = 1 - \frac{27}{32} = \frac{5}{32}$$

Conclusion.

$$P_{(X \geq 0)}\left(X \geq \frac{1}{2}\right) = \frac{\frac{5}{32}}{\frac{1}{2}} = \frac{5}{32} \times 2 = \frac{5}{16} = 0,3125$$

Le conditionnement a bien augmenté la probabilité : sans information, $P(X \geq \frac{1}{2}) = \frac{5}{32} \approx 0,156$, alors qu'en sachant la variable positive on passe à $\frac{5}{16} \approx 0,313$, exactement le double, puisque l'on a restreint l'univers à un événement de probabilité $\frac{1}{2}$.

Méthode à retenir. L'étude complète d'une variable à densité suit toujours le même canevas. On détermine la constante de normalisation en écrivant $\int_{-\infty}^{+\infty} f = 1$, sans oublier de contrôler la positivité qui contraint le signe de la constante. On construit ensuite F_X morceau par morceau, en vérifiant les raccordements aux points de jonction. Enfin, chaque espérance et chaque moment commencent par une phrase d'existence : sur un support borné avec une densité continue, cette phrase est brève (intégrale d'une fonction continue sur un segment), mais elle doit figurer. Quand

la densité est paire, la parité donne $E(X) = 0$ en une ligne, ce qui fait gagner un temps précieux le jour du concours.

Erreur classique. Écrire $E(X) = 0$ « parce que la densité est paire » sans avoir d'abord justifié l'existence de l'espérance. La parité est un argument de calcul, jamais un argument d'existence : il existe des densités paires pour lesquelles l'espérance n'existe pas, l'intégrale $\int t f(t) dt$ pouvant diverger aux deux bornes. La seconde erreur, à la question **6.**, consiste à écrire l'intersection $(X \geq 0) \cap (X \geq \frac{1}{2})$ comme $(X \geq 0)$: c'est la plus petite des deux conditions qui l'emporte, donc $(X \geq \frac{1}{2})$.

Exercice 11 ★★ Loi du carré d'une variable uniforme

Notions : Loi d'une transformée : transformation affine et méthode de la fonction de répartition, Fonction de répartition d'une variable aléatoire réelle quelconque, Loi uniforme sur un segment

Soit X une variable aléatoire telle que $X \hookrightarrow \mathcal{U}([0, 1])$.

1. On pose $Y = X^2$. Déterminer $Y(\Omega)$, puis la fonction de répartition F_Y de Y sur $\mathbb{R}$. En déduire que Y est une variable à densité et en donner une densité f_Y . Vérifier que f_Y est bien une densité de probabilité, bien qu'elle ne soit pas bornée au voisinage de 0.
2. Calculer $E(Y)$ de deux façons : par le théorème de transfert appliqué à X , puis à partir de la densité f_Y obtenue à la question précédente. Les deux résultats doivent coïncider.
3. Calculer $V(Y)$.
4. On pose maintenant $Z = \sqrt{X}$. Déterminer F_Z , en déduire une densité f_Z , puis calculer $E(Z)$.
5. Expliquer pourquoi on ne peut pas obtenir la loi de $Y = X^2$ en « dérivant la densité de X », et énoncer la méthode correcte.

Corrigé

1. Rappelons d'abord la loi de X . Comme $X \hookrightarrow \mathcal{U}([0, 1])$, une densité de X est la fonction f_X valant 1 sur $[0, 1]$ et 0 ailleurs, et sa fonction de répartition est

$$F_X(x) = 0 \quad \text{si } x < 0, \quad F_X(x) = x \quad \text{si } 0 \leq x \leq 1, \quad F_X(x) = 1 \quad \text{si } x > 1.$$

Support de Y . On a $X(\Omega) = [0, 1]$, donc $Y = X^2$ prend ses valeurs dans $[0, 1]$ également, puisque la fonction carré envoie $[0, 1]$ sur $[0, 1]$. Ainsi $Y(\Omega) = [0, 1]$.

Fonction de répartition. On applique la méthode de la fonction de répartition, qui est la seule méthode au programme pour obtenir la loi d'une transformée. Soit $y \in \mathbb{R}$. Par définition,

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$$

Si $y < 0$. L'événement $(X^2 \leq y)$ est impossible, car un carré est positif. Donc $F_Y(y) = 0$.

Si $0 \leq y \leq 1$. L'inégalité $X^2 \leq y$ équivaut à $-\sqrt{y} \leq X \leq \sqrt{y}$. Il faut ici prendre garde à ne pas écrire simplement $X \leq \sqrt{y}$: la fonction carré n'est pas croissante sur $\mathbb{R}$, l'équivalence fait apparaître les deux racines. Cela dit, X est à valeurs positives, donc l'événement $(X \geq -\sqrt{y})$ est certain, et

$$F_Y(y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = P(X \leq \sqrt{y}) = F_X(\sqrt{y}) = \sqrt{y}$$

la dernière égalité venant de $\sqrt{y} \in [0, 1]$.

Si $y > 1$. Alors $\sqrt{y} > 1$ et $F_Y(y) = F_X(\sqrt{y}) = 1$.

En résumé :

$$F_Y(y) = 0 \quad \text{si } y < 0, \quad F_Y(y) = \sqrt{y} \quad \text{si } 0 \leq y \leq 1, \quad F_Y(y) = 1 \quad \text{si } y > 1.$$

Caractère à densité. La fonction F_Y est continue sur $\mathbb{R}$: elle est continue sur chaque morceau, et les raccordements sont bons, puisque $\sqrt{0} = 0$ et $\sqrt{1} = 1$. Elle est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé des deux points 0 et 1. Ces deux points sont en nombre fini, donc Y est bien une variable aléatoire à densité.

Une densité s'obtient en dérivant F_Y là où c'est possible, et en attribuant une valeur arbitraire aux points de non-dérivabilité :

$$f_Y(y) = \frac{1}{2\sqrt{y}} \quad \text{si } y \in]0, 1[, \quad f_Y(y) = 0 \quad \text{sinon.}$$

Vérification. La fonction f_Y est positive, continue sur $\mathbb{R}$ sauf en 0 et en 1. Reste à vérifier que son intégrale vaut 1. Attention : l'intégrale $\int_0^1 \frac{dy}{2\sqrt{y}}$ est **impropre en 0**, car $f_Y(y)$ tend vers $+\infty$ quand y tend vers 0 par valeurs positives. Une densité n'a aucune obligation d'être bornée, mais son intégrale doit converger, et cela se justifie. Pour $\varepsilon \in]0, 1[$,

$$\int_{\varepsilon}^1 \frac{dy}{2\sqrt{y}} = [\sqrt{y}]_{\varepsilon}^1 = 1 - \sqrt{\varepsilon} \xrightarrow{\varepsilon \rightarrow 0^+} 1$$

L'intégrale converge donc et vaut 1 : f_Y est bien une densité de probabilité. On pouvait aussi invoquer le critère de Riemann : au voisinage de 0, $f_Y(y) = \frac{1}{2}y^{-1/2}$, avec $\frac{1}{2} < 1$, donc l'intégrale converge.

2. Première méthode : le théorème de transfert. Ce théorème, admis, permet de calculer $E(g(X))$ sans connaître la loi de $g(X)$. Avec $g : t \mapsto t^2$,

$$E(Y) = E(X^2) = \int_{-\infty}^{+\infty} t^2 f_X(t) dt = \int_0^1 t^2 dt$$

Cette intégrale porte sur une fonction continue sur un segment : elle converge, donc $E(Y)$ existe, et

$$E(Y) = \left[\frac{t^3}{3} \right]_0^1 = \frac{1}{3}$$

Deuxième méthode : par la densité de Y. Par définition de l'espérance d'une variable à densité,

$$E(Y) = \int_{-\infty}^{+\infty} y f_Y(y) dy = \int_0^1 \frac{y}{2\sqrt{y}} dy = \frac{1}{2} \int_0^1 \sqrt{y} dy$$

Cette dernière intégrale n'est plus impropre, car $y \mapsto \sqrt{y}$ se prolonge par continuité en 0. Elle converge donc, et

$$E(Y) = \frac{1}{2} \left[\frac{2}{3} y^{3/2} \right]_0^1 = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$$

Les deux méthodes donnent bien $E(Y) = \frac{1}{3}$, ce qui valide la densité trouvée à la question 1. C'est un contrôle qu'il faut prendre l'habitude de faire quand le temps le permet.

3. Pour la variance, il faut le moment d'ordre 2 de Y , c'est-à-dire $E(Y^2) = E(X^4)$. Le transfert appliqué à $t \mapsto t^4$ donne

$$E(Y^2) = \int_0^1 t^4 dt = \left[\frac{t^5}{5} \right]_0^1 = \frac{1}{5}$$

l'intégrale d'une fonction continue sur un segment convergeant sans discussion. La formule de König-Huygens donne alors

$$V(Y) = E(Y^2) - (E(Y))^2 = \frac{1}{5} - \left(\frac{1}{3}\right)^2 = \frac{1}{5} - \frac{1}{9} = \frac{9-5}{45} = \frac{4}{45}$$

Numériquement, $V(Y) = \frac{4}{45} \approx 0,0889$ et $\sigma(Y) \approx 0,298$. Pour comparaison, $V(X) = \frac{1}{12} \approx 0,0833$: élever au carré une variable uniforme concentre la masse près de 0 tout en étirant légèrement la dispersion relative.

4. On procède exactement de la même manière, avec $Z = \sqrt{X}$. Comme X est à valeurs dans $[0, 1]$, la racine carrée est bien définie et $Z(\Omega) = [0, 1]$.

Soit $z \in \mathbb{R}$.

Si $z < 0$. L'événement $(\sqrt{X} \leq z)$ est impossible, donc $F_Z(z) = 0$.

Si $0 \leq z \leq 1$. La fonction racine carrée est **strictement croissante** sur $\mathbb{R}_+$, et X y prend ses valeurs, donc l'inégalité $\sqrt{X} \leq z$ équivaut à $X \leq z^2$. Ici, contrairement à la question 1., il n'y a pas de seconde branche à considérer. Donc

$$F_Z(z) = P(X \leq z^2) = F_X(z^2) = z^2$$

puisque $z^2 \in [0, 1]$.

Si $z > 1$. Alors $z^2 > 1$ et $F_Z(z) = 1$.

La fonction F_Z est continue sur $\mathbb{R}$ (les raccordements se font en 0 et en 1 sans saut) et de classe $\mathcal{C}^1$ sauf éventuellement en 1 : Z est bien une variable à densité, et

$$f_Z(z) = 2z \text{ si } z \in]0, 1[, \quad f_Z(z) = 0 \text{ sinon.}$$

Cette fonction est positive et son intégrale vaut $[z^2]_0^1 = 1$, comme attendu.

L'espérance existe, l'intégrale portant sur une fonction continue sur un segment, et vaut

$$E(Z) = \int_0^1 z \times 2z \, dz = 2 \int_0^1 z^2 \, dz = 2 \times \frac{1}{3} = \frac{2}{3}$$

On peut évidemment retrouver ce résultat par transfert, sans passer par f_Z :

$$E(\sqrt{X}) = \int_0^1 \sqrt{t} \, dt = \left[\frac{2}{3} t^{3/2} \right]_0^1 = \frac{2}{3}$$

Remarque. On a d'une part $E(X^2) = \frac{1}{3} > \frac{1}{4} = (E(X))^2$, et d'autre part $E(\sqrt{X}) = \frac{2}{3} < \sqrt{E(X)} = \frac{1}{\sqrt{2}} \approx 0,707$. Ces deux inégalités vont dans le sens de la convexité : la fonction carré est convexe, la racine carrée est concave.

5. L'erreur consiste à croire qu'une densité se transporte comme une fonction, c'est-à-dire à écrire quelque chose comme $f_Y(y) = f_X(\sqrt{y})$, ou à « dériver f_X ». C'est faux, et pour une raison de fond : une densité n'est pas une probabilité, c'est une **densité de probabilité par unité de longueur**. Quand on transforme la variable par g , on dilate ou contracte l'axe des abscisses, et la masse doit se redistribuer en conséquence. Ce facteur de dilatation est précisément le facteur $\frac{1}{2\sqrt{y}}$ qui est apparu à la question 1., et rien dans f_X ne le contient.

Sur cet exemple, la fausse formule donnerait $f_Y(y) = 1$ sur $[0, 1]$, c'est-à-dire $Y \hookrightarrow \mathcal{U}([0, 1])$, ce qui est manifestement absurde : $E(Y) = \frac{1}{3}$ et non $\frac{1}{2}$, et Y est bien plus concentrée près de 0 que X , puisque le carré d'un nombre de $[0, 1]$ est plus petit que ce nombre.

La méthode correcte, la seule au programme, tient en trois temps.

Premier temps. Déterminer $Y(\Omega)$, afin de savoir sur quel intervalle la fonction de répartition sera non triviale.

Deuxième temps. Écrire $F_Y(y) = P(g(X) \leq y)$ et **transformer l'événement** $(g(X) \leq y)$ en un événement portant sur X seul, en surveillant le sens de variation de g (une fonction décroissante retourne l'inégalité, une fonction non monotone comme le carré fait apparaître deux branches). On exprime alors F_Y à l'aide de F_X .

Troisième temps. Vérifier que F_Y est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf en un nombre fini de points, puis **dériver** F_Y , et seulement elle, pour obtenir f_Y .

Méthode à retenir. La loi d'une transformée $g(X)$ se cherche toujours par la fonction de répartition, jamais par manipulation directe des densités. Le point technique se joue à l'étape de transformation de l'événement : avec g strictement croissante, $(g(X) \leq y)$ devient $(X \leq g^{-1}(y))$; avec g strictement décroissante, l'inégalité change de sens ; avec la fonction carré, il faut écrire l'encadrement $-\sqrt{y} \leq X \leq \sqrt{y}$ avant de simplifier grâce au support de X . Une fois F_Y obtenue, on justifie la continuité et le caractère $\mathcal{C}^1$ par morceaux avant de dériver, sans quoi la dérivation n'est pas légitime.

Erreur classique. Oublier la branche négative dans $(X^2 \leq y)$. Elle disparaît ici parce que X est positive, mais si l'on avait pris $X \hookrightarrow \mathcal{U}([-1, 1])$, on aurait obtenu $F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) = \sqrt{y}$ sur $[0, 1]$, la simplification étant alors une coïncidence numérique et non une évidence. Seconde erreur, très répandue : conclure que f_Y n'est pas une densité sous prétexte qu'elle tend vers $+\infty$ en 0. Une densité peut parfaitement être non bornée, seule la convergence de son intégrale est requise.

Exercice 12 ★★ Absence de mémoire de la loi exponentielle

Notions : Loi exponentielle et absence de mémoire, Fonction de répartition d'une variable aléatoire réelle quelconque

1. Soit $\lambda > 0$ et soit X une variable aléatoire telle que $X \hookrightarrow \mathcal{E}(\lambda)$. Démontrer que, pour tous réels $s \geq 0$ et $t \geq 0$,

$$P_{(X>s)}(X > s+t) = P(X > t)$$

2. Interpréter concrètement cette propriété lorsque X modélise la durée de vie d'une machine.
3. La durée de vie, en années, d'un composant électronique est modélisée par une variable X telle que $X \hookrightarrow \mathcal{E}(\frac{1}{8})$. Calculer $P(X > 4)$, puis $P_{(X>4)}(X > 10)$. On donnera les valeurs décimales arrondies à 10^{-4} près.
4. On remplace le modèle précédent par une variable Y telle que $Y \hookrightarrow \mathcal{U}([0, 20])$. Calculer $P_{(Y>4)}(Y > 10)$ et comparer à $P(Y > 6)$. Que peut-on en conclure sur la propriété de la question 1. ?
5. On revient au modèle exponentiel de la question 3. Montrer que, conditionnellement à l'événement $(X > 4)$, la durée de vie résiduelle $X - 4$ a la même fonction de répartition qu'une variable de loi $\mathcal{E}(\frac{1}{8})$. En déduire l'espérance de vie résiduelle d'un composant qui a déjà fonctionné 4 ans.

On donne les valeurs approchées suivantes : $e^{-0,5} \approx 0,6065$, $e^{-0,75} \approx 0,4724$ et $e^{-1,25} \approx 0,2865$.

Corrigé

1. Commençons par le seul outil dont on aura besoin : la **fonction de survie** de la loi exponentielle. Si $X \hookrightarrow \mathcal{E}(\lambda)$, une densité de X est $f_X(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$ et $f_X(x) = 0$ pour $x < 0$, et sa fonction de répartition est

$$F_X(x) = 0 \text{ si } x < 0, \quad F_X(x) = 1 - e^{-\lambda x} \text{ si } x \geq 0.$$

Par conséquent, pour tout $x \geq 0$,

$$P(X > x) = 1 - F_X(x) = e^{-\lambda x}$$

C'est cette formule, et elle seule, qui fait toute la démonstration. On la retiendra : pour une loi exponentielle, c'est la probabilité de **survie** qui est simple, pas la fonction de répartition.

Soient maintenant $s \geq 0$ et $t \geq 0$. La probabilité $P(X > s) = e^{-\lambda s}$ est strictement positive, donc le conditionnement par l'événement $(X > s)$ est licite. Cette vérification n'est pas une formalité : sans elle, la probabilité conditionnelle n'est pas définie.

Comme $t \geq 0$, on a $s+t \geq s$, donc l'événement $(X > s+t)$ est **inclus** dans l'événement $(X > s)$. Leur intersection se réduit donc au plus petit des deux :

$$(X > s) \cap (X > s+t) = (X > s+t)$$

Il vient alors, par définition de la probabilité conditionnelle,

$$\begin{aligned} P_{(X>s)}(X > s+t) &= \frac{P((X > s) \cap (X > s+t))}{P(X > s)} \\ &= \frac{P(X > s+t)}{P(X > s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda(s+t)+\lambda s} \\ &= e^{-\lambda t} \\ &= P(X > t) \end{aligned}$$

□

La propriété repose entièrement sur l'équation fonctionnelle de l'exponentielle, $e^{a+b} = e^a e^b$. C'est cette identité, et rien d'autre, qui permet à s de disparaître.

2. Supposons que X modélise la durée de vie d'une machine. L'événement $(X > s)$ signifie « la machine fonctionne encore à l'instant s », et l'égalité démontrée dit que la probabilité qu'elle tienne encore t unités de temps de plus **ne dépend pas de s** , c'est-à-dire pas de son âge.

Autrement dit, une machine qui a déjà fonctionné dix ans a exactement la même probabilité de tenir encore trois ans qu'une machine neuve d'en tenir trois. La machine **ne vieillit pas** : elle n'accumule ni usure ni fatigue, sa panne survient par accident pur, sans mémoire du passé. On dit que la loi exponentielle est **sans mémoire**, ou que le taux de panne est constant.

Ce modèle est réaliste pour un composant électronique qui tombe en panne à cause d'une surtension aléatoire, ou pour la désintégration d'un atome radioactif. Il ne l'est pas du tout pour une pièce mécanique soumise à l'usure, ni pour une durée de vie humaine : dans ces derniers cas, la probabilité de survivre un an de plus décroît avec l'âge, et il faut un autre modèle.

3. Ici $\lambda = \frac{1}{8}$, donc $P(X > x) = e^{-x/8}$ pour $x \geq 0$.

Première probabilité.

$$P(X > 4) = e^{-4/8} = e^{-0,5} \approx 0,6065$$

Deux composants sur cinq environ tombent en panne avant quatre ans.

Seconde probabilité. On applique la propriété de la question **1.** avec $s = 4$ et $t = 6$, puisque $10 = 4 + 6$:

$$P_{(X>4)}(X > 10) = P(X > 6) = e^{-6/8} = e^{-0,75} \approx 0,4724$$

On peut aussi refaire le calcul à la main, ce qui est plus prudent tant que la propriété n'est pas parfaitement maîtrisée :

$$P_{(X>4)}(X > 10) = \frac{P(X > 10)}{P(X > 4)} = \frac{e^{-10/8}}{e^{-4/8}} = e^{-1,25+0,5} = e^{-0,75} \approx 0,4724$$

Sachant que le composant a tenu quatre ans, il a donc environ 47,2 % de chances de dépasser dix ans, soit exactement la probabilité qu'un composant neuf dépasse six ans.

4. Si $Y \hookrightarrow \mathcal{U}([0, 20])$, une densité de Y vaut $\frac{1}{20}$ sur $[0, 20]$ et 0 ailleurs, et pour $y \in [0, 20]$,

$$F_Y(y) = \frac{y-0}{20-0} = \frac{y}{20}, \quad \text{donc} \quad P(Y > y) = 1 - \frac{y}{20} = \frac{20-y}{20}$$

On calcule les trois quantités utiles.

$$P(Y > 4) = \frac{16}{20} = 0,8, \quad P(Y > 10) = \frac{10}{20} = 0,5, \quad P(Y > 6) = \frac{14}{20} = 0,7$$

Comme $(Y > 10) \subset (Y > 4)$ et $P(Y > 4) = 0,8 \neq 0$,

$$P_{(Y>4)}(Y > 10) = \frac{P(Y > 10)}{P(Y > 4)} = \frac{0,5}{0,8} = \frac{5}{8} = 0,625$$

Or $P(Y > 6) = 0,7$. Ces deux nombres sont différents :

$$P_{(Y>4)}(Y > 10) = 0,625 \neq 0,7 = P(Y > 6)$$

La propriété d'absence de mémoire est donc **fausse** pour la loi uniforme. Et le sens de l'écart est instructif : $0,625 < 0,7$, ce qui signifie qu'un composant uniforme qui a déjà tenu quatre ans a **moins** de chances de tenir six ans de plus qu'un composant neuf. C'est logique, puisque la durée de vie est bornée par 20 ans : le temps qui passe rapproche mécaniquement de la date butoir, le composant vieillit.

La conclusion est importante : l'absence de mémoire n'est pas une propriété générale des durées de vie, c'est une propriété très particulière. On démontre d'ailleurs qu'elle **caractérise** la loi exponentielle parmi les lois à densité sur $\mathbb{R}_+$: une durée de vie sans mémoire est nécessairement exponentielle. Vérifier l'absence de mémoire sur des données, c'est donc valider ou invalider d'un coup tout le modèle exponentiel.

5. Notons $R = X - 4$ la durée de vie résiduelle du composant, sachant qu'il a déjà fonctionné quatre ans. On veut la loi de R **sous la probabilité conditionnelle** $P_{(X>4)}$, c'est-à-dire dans l'univers restreint aux composants ayant survécu quatre ans.

Soit $t \geq 0$. En appliquant directement la propriété de la question 1. avec $s = 4$:

$$P_{(X>4)}(R > t) = P_{(X>4)}(X > 4 + t) = P(X > t) = e^{-t/8}$$

Par passage au complémentaire, pour tout $t \geq 0$,

$$P_{(X>4)}(R \leq t) = 1 - e^{-t/8}$$

et pour $t < 0$ cette probabilité est nulle, puisque $R > 0$ dans l'univers considéré. On reconnaît exactement la fonction de répartition d'une variable de loi $\mathcal{E}(\frac{1}{8})$.

Sous la probabilité conditionnelle $P_{(X>4)}$, la durée de vie résiduelle R suit donc la même loi que la durée de vie totale d'un composant neuf. Comme l'espérance d'une variable de loi $\mathcal{E}(\lambda)$ vaut $\frac{1}{\lambda}$, l'espérance de vie résiduelle vaut

$$\frac{1}{\lambda} = 8 \text{ ans}$$

Un composant qui a déjà tenu quatre ans a donc encore huit années de vie en moyenne devant lui, exactement comme un composant sorti d'usine. Sa durée de vie totale moyenne, sachant qu'il a passé le cap des quatre ans, est de $4 + 8 = 12$ ans, alors que celle d'un composant neuf est de 8 ans : ce n'est pas une contradiction, c'est simplement l'effet de la sélection opérée par le conditionnement, qui élimine tous les composants morts avant quatre ans.

Pour la loi uniforme de la question 4., la conclusion serait tout autre. Un composant neuf a une espérance de vie de 10 ans, tandis que sous $P_{(Y>4)}$ la variable Y se comporte comme une uniforme sur $[4, 20]$, d'espérance 12 ans, soit une vie résiduelle moyenne de 8 ans seulement au lieu de 10. La perte est bien réelle : le composant a vieilli.

Méthode à retenir. Sur une loi exponentielle, on ne travaille jamais avec F_X mais avec la survie $P(X > x) = e^{-\lambda x}$, qui transforme tout quotient de probabilités en soustraction d'exposants. Toute probabilité conditionnelle du type $P_{(X>s)}(X > s+t)$ se traite en trois gestes : vérifier que $P(X > s) \neq 0$, remarquer l'inclusion $(X > s+t) \subset (X > s)$ pour simplifier l'intersection, puis simplifier le quotient d'exponentielles. Pour une durée de vie résiduelle, on pose $R = X - s$ et l'on calcule sa fonction de répartition conditionnelle : c'est la seule façon rigoureuse de parler de « ce qu'il reste à vivre » sans sortir du programme.

Erreur classique. Écrire $(X > 4) \cap (X > 10) = (X > 4)$. C'est le contraire : quand on impose deux minoration, c'est la plus exigeante qui commande, donc l'intersection vaut $(X > 10)$. Autre erreur, plus subtile : croire que l'absence de mémoire signifie $P_{(X>4)}(X > 10) = P(X > 10)$. Ce serait dire que le conditionnement n'apporte rien du tout, ce qui est faux ($0,4724 \neq 0,2865$). La propriété porte sur la durée **supplémentaire**, comptée à partir de l'instant s , jamais sur la date absolue.

Exercice 13 ★★ La fonction Gamma d'Euler et les lois gamma

Notions : Lois gamma à un paramètre et fonction Gamma d'Euler, Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Variance, écart-type et moments d'une variable à densité

Pour $\nu > 0$, on pose

$$\Gamma(\nu) = \int_0^{+\infty} t^{\nu-1} e^{-t} dt$$

1. Montrer que cette intégrale converge pour tout $\nu > 0$. On étudiera séparément le comportement au voisinage de 0 et au voisinage de $+\infty$, en utilisant en $+\infty$ la relation $t^{\nu-1}e^{-t} = o\left(\frac{1}{t^2}\right)$.
2. Calculer $\Gamma(1)$. Établir ensuite, pour tout $\nu > 0$, la relation $\Gamma(\nu+1) = \nu \Gamma(\nu)$, par une intégration par parties sur un segment $[a, A]$ suivie d'un passage à la limite. En déduire par récurrence que $\Gamma(n) = (n-1)!$ pour tout entier $n \geq 1$.
3. Soit $\nu > 0$. Vérifier que la fonction f définie par

$$f(x) = \frac{x^{\nu-1} e^{-x}}{\Gamma(\nu)} \quad \text{si } x > 0, \quad f(x) = 0 \quad \text{si } x \leq 0$$

est une densité de probabilité. La loi correspondante est appelée **loi gamma de paramètre ν** , et notée $\gamma(\nu)$.

4. Soit X une variable aléatoire telle que $X \hookrightarrow \gamma(\nu)$. Montrer que X admet une espérance et une variance, et établir $E(X) = \nu$ puis $V(X) = \nu$.
5. Vérifier que la loi $\gamma(1)$ n'est autre que la loi $\mathcal{E}(1)$. En déduire que si $Y \hookrightarrow \mathcal{E}(\lambda)$ avec $\lambda > 0$, alors $\lambda Y \hookrightarrow \gamma(1)$.

Corrigé

1. La fonction $g : t \mapsto t^{\nu-1} e^{-t}$ est continue et positive sur $]0, +\infty[$. L'intégrale proposée est donc doublement impropre : en 0, car g n'est pas définie en ce point et peut y tendre vers $+\infty$ lorsque $\nu < 1$, et en $+\infty$, car l'intervalle d'intégration est non borné. On étudie séparément les deux intégrales

$$\int_0^1 t^{\nu-1} e^{-t} dt \quad \text{et} \quad \int_1^{+\infty} t^{\nu-1} e^{-t} dt$$

et l'intégrale de départ converge si et seulement si les deux convergent. Le point de coupure 1 est arbitraire, n'importe quel réel strictement positif conviendrait.

Au voisinage de 0. Quand t tend vers 0 par valeurs positives, e^{-t} tend vers 1, donc

$$t^{\nu-1} e^{-t} \underset{t \rightarrow 0^+}{\sim} t^{\nu-1} = \frac{1}{t^{1-\nu}}$$

Les deux fonctions étant continues et **positives** sur $]0, 1]$, le critère d'équivalence s'applique : les intégrales $\int_0^1 t^{\nu-1} e^{-t} dt$ et $\int_0^1 \frac{dt}{t^{1-\nu}}$ sont de même nature. Or l'intégrale de Riemann $\int_0^1 \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha < 1$. Ici $\alpha = 1 - \nu$, et l'hypothèse $\nu > 0$ donne exactement $1 - \nu < 1$. La première intégrale converge donc, et c'est précisément là que sert la condition $\nu > 0$.

Au voisinage de $+\infty$. On utilise la croissance comparée : pour tout réel β , $t^\beta e^{-t} \rightarrow 0$ quand $t \rightarrow +\infty$, l'exponentielle l'emportant sur toute puissance. En appliquant ceci avec $\beta = \nu + 1$,

$$t^2 \times t^{\nu-1} e^{-t} = t^{\nu+1} e^{-t} \xrightarrow{t \rightarrow +\infty} 0$$

ce qui s'écrit exactement $t^{\nu-1}e^{-t} = o\left(\frac{1}{t^2}\right)$ au voisinage de $+\infty$. Comme l'intégrale de Riemann $\int_1^{+\infty} \frac{dt}{t^2}$ converge (l'exposant 2 étant strictement supérieur à 1) et que les fonctions en jeu sont positives, le critère de négligeabilité assure la convergence de $\int_1^{+\infty} t^{\nu-1}e^{-t} dt$.

Les deux intégrales convergeant, $\Gamma(\nu)$ est bien définie pour tout $\nu > 0$. $\square$

Remarquons au passage que $\Gamma(\nu) > 0$, comme intégrale convergente d'une fonction continue, positive et non identiquement nulle. Ce point servira à la question 3.

2. Valeur de $\Gamma(1)$. Pour $\nu = 1$, l'exposant $\nu - 1$ est nul et l'intégrale n'est plus impropre en 0 :

$$\Gamma(1) = \int_0^{+\infty} e^{-t} dt = \lim_{A \rightarrow +\infty} [-e^{-t}]_0^A = \lim_{A \rightarrow +\infty} (1 - e^{-A}) = 1$$

Donc $\Gamma(1) = 1$.

Relation fonctionnelle. Soit $\nu > 0$. Les deux intégrales $\Gamma(\nu + 1)$ et $\Gamma(\nu)$ convergent d'après la question 1., puisque $\nu + 1 > 0$ et $\nu > 0$. On ne peut pas intégrer par parties directement sur $]0, +\infty[$: l'intégration par parties est un théorème valable sur un **segment**. On travaille donc sur $[a, A]$ avec $0 < a < A$, puis on passe à la limite.

Posons, sur $[a, A]$, $u(t) = t^\nu$ et $v'(t) = e^{-t}$, donc $u'(t) = \nu t^{\nu-1}$ et $v(t) = -e^{-t}$. Ces deux fonctions sont de classe $\mathcal{C}^1$ sur $[a, A]$, l'intégration par parties est licite :

$$\begin{aligned} \int_a^A t^\nu e^{-t} dt &= [-t^\nu e^{-t}]_a^A + \nu \int_a^A t^{\nu-1} e^{-t} dt \\ &= -A^\nu e^{-A} + a^\nu e^{-a} + \nu \int_a^A t^{\nu-1} e^{-t} dt \end{aligned}$$

Examinons les deux termes tout intégrés.

Quand A tend vers $+\infty$, la croissance comparée donne $A^\nu e^{-A} \rightarrow 0$.

Quand a tend vers 0 par valeurs positives, $a^\nu \rightarrow 0$ car $\nu > 0$, et $e^{-a} \rightarrow 1$, donc $a^\nu e^{-a} \rightarrow 0$. C'est encore l'hypothèse $\nu > 0$ qui fait le travail.

Les deux intégrales convergeant par la question 1., on peut faire tendre simultanément a vers 0^+ et A vers $+\infty$ dans l'égalité ci-dessus, ce qui donne

$$\Gamma(\nu + 1) = 0 + 0 + \nu \Gamma(\nu) = \nu \Gamma(\nu)$$

$\square$

Valeur de $\Gamma(n)$. Montrons par récurrence sur $n \geq 1$ la propriété $\mathcal{P}(n) : \Gamma(n) = (n-1)!$.

Initialisation. Pour $n = 1$: $\Gamma(1) = 1 = 0!$, donc $\mathcal{P}(1)$ est vraie.

Hérédité. Supposons $\mathcal{P}(n)$ vraie pour un entier $n \geq 1$. La relation fonctionnelle, appliquée avec $\nu = n > 0$, donne

$$\Gamma(n + 1) = n \Gamma(n) = n \times (n-1)! = n!$$

ce qui est exactement $\mathcal{P}(n+1)$.

Conclusion. Par récurrence, $\Gamma(n) = (n-1)!$ pour tout entier $n \geq 1$. $\square$

La fonction Γ prolonge donc la factorielle aux réels strictement positifs, avec ce décalage d'une unité qu'il faut mémoriser une fois pour toutes : $\Gamma(n)$ vaut $(n-1)!$ et non $n!$. On admet par ailleurs la valeur $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, qui se déduit de l'intégrale de Gauss.

3. On vérifie les trois conditions caractérisant une densité.

Positivité. Pour $x > 0$, on a $x^{\nu-1} > 0$ et $e^{-x} > 0$, et l'on a noté à la question 1. que $\Gamma(\nu) > 0$. Donc $f(x) > 0$. Pour $x \leq 0$, $f(x) = 0$. La fonction f est bien positive sur $\mathbb{R}$.

Continuité. La fonction f est continue sur $]-\infty, 0[$ (elle y est nulle) et sur $]0, +\infty[$ (produit et quotient de fonctions continues, le dénominateur étant une constante non nulle). Elle présente au plus une discontinuité, en 0, ce qui est un nombre fini de points : la condition du programme

est remplie. Notons que pour $\nu < 1$, f tend vers $+\infty$ en 0^+ , ce qui n'empêche nullement d'être une densité.

Intégrale. Par définition de f et par linéarité,

$$\int_{-\infty}^{+\infty} f(x) dx = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^{\nu-1} e^{-x} dx = \frac{\Gamma(\nu)}{\Gamma(\nu)} = 1$$

la convergence étant précisément celle établie à la question 1. Ainsi f est une densité de probabilité. $\square$

On voit ici la raison d'être de la constante $\Gamma(\nu)$ au dénominateur : c'est exactement la constante de normalisation qui ramène l'intégrale à 1. La fonction Γ n'a pas été inventée pour la probabilité, mais elle s'y trouve parfaitement adaptée.

4. Espérance. Par définition, $E(X)$ existe si l'intégrale $\int_0^{+\infty} x f(x) dx$ converge absolument. La fonction intégrée est positive, la convergence absolue équivaut donc à la convergence simple. Or

$$x f(x) = \frac{x^\nu e^{-x}}{\Gamma(\nu)} \quad \text{pour } x > 0$$

et l'intégrale $\int_0^{+\infty} x^\nu e^{-x} dx$ n'est autre que $\Gamma(\nu+1)$, qui converge d'après la question 1. puisque $\nu+1 > 0$. **L'espérance existe donc**, et

$$E(X) = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^\nu e^{-x} dx = \frac{\Gamma(\nu+1)}{\Gamma(\nu)} = \frac{\nu \Gamma(\nu)}{\Gamma(\nu)} = \nu$$

la relation fonctionnelle de la question 2. faisant tout le travail. Donc

$$E(X) = \nu$$

Moment d'ordre 2. Le théorème de transfert, appliqué à la fonction $t \mapsto t^2$, donne

$$E(X^2) = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^2 \times x^{\nu-1} e^{-x} dx = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^{\nu+1} e^{-x} dx$$

Cette dernière intégrale vaut $\Gamma(\nu+2)$, qui converge puisque $\nu+2 > 0$. **La variance existe donc.** En appliquant deux fois la relation fonctionnelle,

$$\Gamma(\nu+2) = (\nu+1) \Gamma(\nu+1) = (\nu+1) \nu \Gamma(\nu)$$

d'où

$$E(X^2) = \frac{(\nu+1) \nu \Gamma(\nu)}{\Gamma(\nu)} = \nu(\nu+1) = \nu^2 + \nu$$

Variance. La formule de König-Huygens conclut :

$$V(X) = E(X^2) - (E(X))^2 = \nu^2 + \nu - \nu^2 = \nu$$

Donc $E(X) = V(X) = \nu$. Cette coïncidence entre espérance et variance est une signature de la loi gamma, comme elle l'est de la loi de Poisson dans le monde discret. L'écart-type vaut $\sqrt{\nu}$, il croît donc moins vite que la moyenne : plus ν est grand, plus la loi est resserrée **en valeur relative** autour de son espérance.

5. Identification de $\gamma(1)$. On prend $\nu = 1$ dans la densité de la question 3. Comme $\Gamma(1) = 1$ et $x^{1-1} = x^0 = 1$ pour $x > 0$, on obtient

$$f(x) = \frac{x^0 e^{-x}}{1} = e^{-x} \quad \text{si } x > 0, \quad f(x) = 0 \quad \text{si } x \leq 0$$

C'est exactement la densité de la loi exponentielle de paramètre 1, à savoir $x \mapsto \lambda e^{-\lambda x}$ sur $\mathbb{R}_+^*$ avec $\lambda = 1$. Deux variables ayant la même densité ont la même loi, donc

$$\gamma(1) = \mathcal{E}(1)$$

On contrôle la cohérence avec la question 4. : pour $\nu = 1$, elle prédit $E(X) = 1$ et $V(X) = 1$, ce qui correspond bien à $\frac{1}{\lambda}$ et $\frac{1}{\lambda^2}$ pour $\lambda = 1$.

Cas d'une exponentielle quelconque. Soit $Y \hookrightarrow \mathcal{E}(\lambda)$ avec $\lambda > 0$, et posons $T = \lambda Y$. Il s'agit d'une transformation affine, que l'on traite comme toujours par la fonction de répartition. Soit $x \in \mathbb{R}$. Comme $\lambda > 0$, la division par λ ne change pas le sens de l'inégalité :

$$F_T(x) = P(\lambda Y \leq x) = P\left(Y \leq \frac{x}{\lambda}\right) = F_Y\left(\frac{x}{\lambda}\right)$$

Si $x < 0$, alors $\frac{x}{\lambda} < 0$ et $F_T(x) = 0$. Si $x \geq 0$, alors

$$F_T(x) = 1 - e^{-\lambda \times \frac{x}{\lambda}} = 1 - e^{-x}$$

On reconnaît la fonction de répartition de la loi $\mathcal{E}(1)$, c'est-à-dire de $\gamma(1)$ d'après ce qui précède. Donc

$$\lambda Y \hookrightarrow \gamma(1)$$

□

Ce résultat est le pont indispensable entre les deux mondes. Le programme ne retient que les lois gamma **à un seul paramètre**, si bien qu'une variable $\mathcal{E}(\lambda)$ avec $\lambda \neq 1$ n'est pas une loi gamma au sens du cours. La multiplication par λ la ramène dans le cadre autorisé, et c'est ce changement d'échelle que l'on fera systématiquement avant d'invoquer la stabilité des lois gamma par somme de variables indépendantes.

Méthode à retenir. Toute question sur la fonction Γ se ramène à trois outils. La convergence s'établit en coupant l'intégrale en deux, avec le critère de Riemann en 0 (c'est là qu'intervient $\nu > 0$) et la croissance comparée en $+\infty$. La relation $\Gamma(\nu + 1) = \nu \Gamma(\nu)$ s'obtient par intégration par parties **sur un segment**, jamais directement sur un intervalle ouvert non borné, puis par passage à la limite aux deux bornes. Enfin, tous les moments d'une loi $\gamma(\nu)$ se calculent sans aucune primitive nouvelle : $E(X^k) = \frac{\Gamma(\nu+k)}{\Gamma(\nu)}$, qu'il suffit de dérouler avec la relation fonctionnelle.

Erreur classique. Écrire $\Gamma(n) = n!$ au lieu de $(n-1)!$. Le décalage se contrôle en une seconde sur $n = 1$: $\Gamma(1) = 1$, et $(1-1)! = 0! = 1$ convient, tandis que $1! = 1$ convient aussi, ce qui ne tranche pas. On teste donc plutôt sur $n = 3$: $\Gamma(3) = 2\Gamma(2) = 2 \times 1 \times \Gamma(1) = 2 = 2!$, et non $3! = 6$. Seconde erreur : intégrer par parties directement entre 0 et $+\infty$ sans passer par un segment, ce qui manipule des termes tout intégrés potentiellement infinis et fait perdre tout le bénéfice de la rigueur attendue.

Exercice 14 ★★ Le théorème de transfert en pratique

Notions : Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Loi exponentielle et absence de mémoire, Loi uniforme sur un segment

- Soit X une variable aléatoire telle que $X \hookrightarrow \mathcal{U}([0, 1])$. Montrer que les espérances suivantes existent et les calculer.
 - $E\left(\frac{1}{X+1}\right)$
 - $E(e^X)$
 - $E(X^n)$ pour $n \in \mathbb{N}$
- Soit $\lambda > 0$ et soit X une variable aléatoire telle que $X \hookrightarrow \mathcal{E}(\lambda)$.
 - Calculer $E(X^2)$ à l'aide d'une intégration par parties.
 - Soit t un réel. Discuter, selon la valeur de t , l'existence de $E(e^{tX})$, et calculer cette espérance lorsqu'elle existe.
- Soit X une variable aléatoire telle que $X \hookrightarrow \mathcal{U}([0, 1])$. Calculer $E(-\ln X)$, puis déterminer la loi de la variable $-\ln X$.
- Expliquer en quoi le théorème de transfert dispense de déterminer la loi de $g(X)$, et préciser dans quelles situations on n'a pas le choix et doit malgré tout déterminer cette loi.

Corrigé

Rappelons l'énoncé du **théorème de transfert**, admis par le programme. Si X est une variable à densité f_X et si g est une fonction continue sur $X(\Omega)$ sauf éventuellement en un nombre fini de points, alors $g(X)$ admet une espérance si et seulement si l'intégrale

$$\int_{-\infty}^{+\infty} g(t) f_X(t) dt$$

converge absolument, et dans ce cas $E(g(X))$ est égale à cette intégrale. Le point capital est qu'on n'a **jamais besoin de la loi de $g(X)$** : tout se calcule avec la densité de X .

1. Ici f_X vaut 1 sur $[0, 1]$ et 0 ailleurs. Toutes les intégrales de la question se réduisent donc à des intégrales sur le segment $[0, 1]$.

a. La fonction $t \mapsto \frac{1}{t+1}$ est continue sur $[0, 1]$, puisque $t+1$ ne s'annule pas sur cet intervalle. L'intégrale $\int_0^1 \frac{dt}{t+1}$ porte sur une fonction continue sur un segment : elle n'est impropre en aucune borne et converge donc, tout comme l'intégrale des valeurs absolues, la fonction étant positive. L'espérance existe, et

$$E\left(\frac{1}{X+1}\right) = \int_0^1 \frac{dt}{t+1} = [\ln(t+1)]_0^1 = \ln 2 - \ln 1 = \ln 2 \approx 0,693$$

On notera au passage que $E\left(\frac{1}{X+1}\right) = \ln 2 \approx 0,693$, alors que $\frac{1}{E(X)+1} = \frac{1}{1,5} \approx 0,667$. Ces deux nombres diffèrent, comme il se doit : l'espérance ne commute pas avec les fonctions non affines.

b. La fonction exponentielle est continue sur $[0, 1]$, l'intégrale porte encore sur un segment et converge. L'espérance existe, et

$$E(e^X) = \int_0^1 e^t dt = [e^t]_0^1 = e - 1 \approx 1,718$$

c. Pour tout entier naturel n , la fonction $t \mapsto t^n$ est continue sur $[0, 1]$, donc l'intégrale converge et le moment d'ordre n existe :

$$E(X^n) = \int_0^1 t^n dt = \left[\frac{t^{n+1}}{n+1} \right]_0^1 = \frac{1}{n+1}$$

Ce résultat contient les cas usuels : $E(X) = \frac{1}{2}$ pour $n = 1$, et $E(X^2) = \frac{1}{3}$ pour $n = 2$, d'où l'on retrouve $V(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$, valeur bien connue de la loi $\mathcal{U}([0, 1])$.

2. Ici une densité de X est $f_X(t) = \lambda e^{-\lambda t}$ pour $t \geq 0$ et $f_X(t) = 0$ pour $t < 0$. Toutes les intégrales portent donc sur $[0, +\infty[$ et sont impropres en $+\infty$: la convergence doit être discutée à chaque fois, ce qui contraste avec la question 1.

a. La fonction intégrée $t \mapsto t^2 \lambda e^{-\lambda t}$ est continue et positive sur $[0, +\infty[$. Sa convergence en $+\infty$ résulte de la croissance comparée : $t^2 \times t^2 \lambda e^{-\lambda t} = \lambda t^4 e^{-\lambda t} \rightarrow 0$, donc $t^2 \lambda e^{-\lambda t} = o\left(\frac{1}{t^2}\right)$, et l'intégrale de Riemann $\int_1^{+\infty} \frac{dt}{t^2}$ converge. **Le moment d'ordre 2 existe.**

Pour le calculer, on intègre par parties sur un segment $[0, A]$, avec $u(t) = t^2$ et $v'(t) = \lambda e^{-\lambda t}$, donc $u'(t) = 2t$ et $v(t) = -e^{-\lambda t}$. Ces fonctions sont de classe $\mathcal{C}^1$ sur $[0, A]$:

$$\begin{aligned} \int_0^A t^2 \lambda e^{-\lambda t} dt &= [-t^2 e^{-\lambda t}]_0^A + \int_0^A 2t e^{-\lambda t} dt \\ &= -A^2 e^{-\lambda A} + \frac{2}{\lambda} \int_0^A t \lambda e^{-\lambda t} dt \end{aligned}$$

Quand A tend vers $+\infty$, la croissance comparée donne $A^2 e^{-\lambda A} \rightarrow 0$, tandis que l'intégrale restante converge vers $E(X) = \frac{1}{\lambda}$, résultat du cours sur la loi exponentielle. On obtient donc

$$E(X^2) = \frac{2}{\lambda} \times \frac{1}{\lambda} = \frac{2}{\lambda^2}$$

Contrôle immédiat : $V(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$, ce qui est bien la variance de la loi $\mathcal{E}(\lambda)$ donnée par le cours.

b. On étudie l'intégrale

$$\int_0^{+\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{+\infty} e^{-(\lambda-t)x} dx$$

La fonction intégrée est continue et positive sur $[0, +\infty[$, la convergence absolue équivaut donc à la convergence, et tout se joue sur le signe de $\lambda - t$.

Cas $t < \lambda$. Alors $\lambda - t > 0$. Pour $A > 0$,

$$\lambda \int_0^A e^{-(\lambda-t)x} dx = \lambda \left[\frac{-e^{-(\lambda-t)x}}{\lambda-t} \right]_0^A = \frac{\lambda}{\lambda-t} (1 - e^{-(\lambda-t)A})$$

Comme $\lambda - t > 0$, on a $e^{-(\lambda-t)A} \rightarrow 0$ quand $A \rightarrow +\infty$. L'intégrale converge donc et

$$E(e^{tX}) = \frac{\lambda}{\lambda-t}$$

Cas $t = \lambda$. L'intégrande vaut $\lambda e^0 = \lambda$, constante non nulle. Alors $\int_0^A \lambda dx = \lambda A \rightarrow +\infty$: l'intégrale **diverge**, et $E(e^{\lambda X})$ n'existe pas.

Cas $t > \lambda$. Alors $\lambda - t < 0$, donc $-(\lambda - t) > 0$ et

$$\lambda \int_0^A e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t} (1 - e^{(t-\lambda)A})$$

Le terme $e^{(t-\lambda)A}$ tend vers $+\infty$, et comme $\lambda - t < 0$ le quotient tend vers $+\infty$. L'intégrale **diverge** là encore, et l'espérance n'existe pas.

Conclusion. L'espérance $E(e^{tX})$ existe si et seulement si $t < \lambda$, et vaut alors $\frac{\lambda}{\lambda-t}$.

Cet exemple montre qu'une espérance peut parfaitement ne pas exister, et qu'une réponse comme « $\frac{\lambda}{\lambda-t}$ » écrite sans condition serait fautive : pour $t = 2\lambda$ elle donnerait la valeur négative -1 , alors que e^{tX} est une variable positive. L'absurdité du résultat signale l'oubli de la discussion.

3. Calcul de l'espérance. On applique le transfert avec $g : t \mapsto -\ln t$, continue sur $]0, 1]$:

$$E(-\ln X) = \int_0^1 (-\ln t) \times 1 dt = - \int_0^1 \ln t dt$$

Attention, cette intégrale est **impropre en 0**, car $\ln t$ tend vers $-\infty$ quand $t \rightarrow 0^+$. La fonction $t \mapsto -\ln t$ étant positive sur $]0, 1]$, convergence et convergence absolue coïncident. Une primitive de $\ln$ est $t \mapsto t \ln t - t$, donc pour $\varepsilon \in]0, 1[$,

$$- \int_{\varepsilon}^1 \ln t dt = - [t \ln t - t]_{\varepsilon}^1 = - ((0 - 1) - (\varepsilon \ln \varepsilon - \varepsilon)) = 1 + \varepsilon \ln \varepsilon - \varepsilon$$

Par croissance comparée, $\varepsilon \ln \varepsilon \rightarrow 0$ quand $\varepsilon \rightarrow 0^+$, et bien sûr $\varepsilon \rightarrow 0$. L'intégrale converge donc, **l'espérance existe** et

$$E(-\ln X) = 1$$

Loi de $-\ln X$. Posons $T = -\ln X$. Comme X est à valeurs dans $]0, 1]$ (l'événement $(X = 0)$ est de probabilité nulle, on peut l'écarter), $\ln X$ est négatif ou nul, donc T est positive : $T(\Omega) = [0, +\infty[$. On applique la méthode de la fonction de répartition. Soit $x \in \mathbb{R}$.

Si $x < 0$, l'événement $(T \leq x)$ est impossible et $F_T(x) = 0$.

Si $x \geq 0$, on transforme l'événement. L'inégalité $-\ln X \leq x$ équivaut à $\ln X \geq -x$, puis, en composant par la fonction exponentielle qui est **strictement croissante** donc conserve le sens de l'inégalité, à $X \geq e^{-x}$. Donc

$$F_T(x) = P(X \geq e^{-x}) = 1 - P(X < e^{-x}) = 1 - F_X(e^{-x})$$

la variable X étant à densité, on a $P(X = e^{-x}) = 0$ et les inégalités strictes ou larges donnent le même résultat. Comme $x \geq 0$, on a $e^{-x} \in]0, 1]$, donc $F_X(e^{-x}) = e^{-x}$, d'où

$$F_T(x) = 1 - e^{-x}$$

On reconnaît la fonction de répartition de la loi exponentielle de paramètre 1 :

$$-\ln X \hookrightarrow \mathcal{E}(1)$$

Ce résultat fournit un second contrôle du calcul d'espérance : l'espérance d'une loi $\mathcal{E}(1)$ vaut $\frac{1}{1} = 1$, ce qui confirme $E(-\ln X) = 1$. Il est aussi le point de départ de la méthode d'inversion pour simuler une loi exponentielle.

4. Le théorème de transfert affirme que

$$E(g(X)) = \int_{-\infty}^{+\infty} g(t) f_X(t) dt$$

Le membre de droite ne fait intervenir que g et la densité de X . On peut donc calculer l'espérance de $g(X)$ **sans jamais déterminer la loi de $g(X)$** , ce qui économise à chaque fois toute une étude : support, fonction de répartition, transformation d'événement, vérification de continuité, dérivation. La question **3.** l'illustre parfaitement : le calcul de $E(-\ln X)$ n'a demandé qu'une primitive, alors que la détermination de la loi de $-\ln X$ a demandé une page.

C'est aussi le théorème qui rend possible l'ensemble du cours sur les moments : $E(X^2)$, $E(X^n)$, $E(e^{tX})$ se calculent tous avec f_X seule, sans connaître les lois de X^2 , X^n ou e^{tX} , dont certaines seraient d'ailleurs pénibles à écrire.

En revanche, le transfert ne suffit pas dans les situations suivantes.

L'énoncé demande la loi. Si la question est « déterminer la loi de $g(X)$ », ou « montrer que $g(X)$ suit une loi exponentielle », aucune espérance ne répondra : il faut la fonction de répartition. C'est le cas de la seconde partie de la question **3**.

On veut une probabilité, pas une espérance. Une quantité comme $P(g(X) \leq a)$ ne s'obtient pas par transfert. Elle demande soit la loi de $g(X)$, soit, plus simplement, la transformation directe de l'événement en un événement portant sur X .

On a besoin de la densité de $g(X)$ pour la suite. Par exemple pour étudier une somme de variables indépendantes par produit de convolution, ou pour reconnaître une loi usuelle et invoquer ensuite ses propriétés.

Méthode à retenir. Avant tout calcul par transfert, on écrit une phrase d'existence, et cette phrase dépend du support. Sur un **segment**, avec une fonction continue, l'intégrale n'est impropre nulle part et la justification tient en une ligne. Sur un intervalle **non borné**, ou lorsque la fonction n'est pas bornée au voisinage d'une borne ouverte, il faut un vrai argument : croissance comparée contre l'exponentielle, critère de Riemann, comparaison ou négligeabilité. La discussion selon un paramètre, comme au **2. b.**, n'est jamais un ornement : elle décide de l'existence même de l'objet que l'on calcule.

Erreur classique. Confondre $E(g(X))$ et $g(E(X))$. La question **1. a.** le montre : $E\left(\frac{1}{X+1}\right) = \ln 2 \approx 0,693$ tandis que $\frac{1}{E(X)+1} \approx 0,667$. Seules les fonctions affines commutent avec l'espérance. Seconde erreur, plus coûteuse en points : appliquer la formule du transfert avec la densité de $g(X)$ au lieu de celle de X , en écrivant $\int g(t)f_{g(X)}(t) dt$. La bonne écriture est $\int g(t)f_X(t) dt$, ou bien $\int y f_{g(X)}(y) dy$ si l'on tient à passer par la loi de $g(X)$, mais jamais un mélange des deux.

Exercice 15 ★★ Contrôle de fabrication et loi normale

Notions : Loi normale, centrage-réduction et lecture de la table

Une machine produit des pièces cylindriques. Le diamètre, exprimé en millimètres, d'une pièce prise au hasard est modélisé par une variable aléatoire X telle que

$$X \hookrightarrow \mathcal{N}(20; 0,04)$$

Le second paramètre est la variance, l'écart-type vaut donc $\sigma = \sqrt{0,04} = 0,2$ mm. Une pièce est dite **conforme** lorsque son diamètre appartient à l'intervalle de tolérance $[19,6; 20,4]$.

On note Φ la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0, 1)$. On donne les valeurs suivantes, à lire dans la table.

u	1	1,5	2	2,5	2,58	3,09
$\Phi(u)$	0,8413	0,9332	0,9772	0,9938	0,9951	0,9990

1. Calculer la probabilité qu'une pièce soit conforme, puis le pourcentage de rebut.
2. On souhaite fixer un intervalle de tolérance symétrique autour de 20, de la forme $[20 - a; 20 + a]$, qui accepte 99 % des pièces. Déterminer a .
3. La machine se dérègle et produit désormais des pièces de diamètre $X' \hookrightarrow \mathcal{N}(20,1; 0,04)$, l'intervalle de tolérance restant $[19,6; 20,4]$. Calculer la nouvelle proportion de rebut.
4. On règle la machine sur une moyenne m , l'écart-type restant égal à 0,2. Déterminer les valeurs de m pour lesquelles $P(X < 19,6) \leq 0,001$.
5. La machine est de nouveau réglée sur $\mathcal{N}(20; 0,04)$. On prélève une pièce au hasard parmi les pièces conformes. Quelle est la probabilité que son diamètre dépasse 20,2 mm ?

Corrigé

Tout le corrigé repose sur un seul geste, le **centrage-réduction**. Si $X \hookrightarrow \mathcal{N}(m, \sigma^2)$ avec $\sigma > 0$, alors la variable

$$T = \frac{X - m}{\sigma}$$

suit la loi normale centrée réduite $\mathcal{N}(0, 1)$, dont la fonction de répartition est Φ . On en tire, pour tout réel x ,

$$P(X \leq x) = P\left(T \leq \frac{x - m}{\sigma}\right) = \Phi\left(\frac{x - m}{\sigma}\right)$$

Deux propriétés de Φ seront utilisées en permanence, et découlent de la parité de la densité φ :

$$\Phi(-u) = 1 - \Phi(u) \quad \text{et} \quad P(-u \leq T \leq u) = 2\Phi(u) - 1$$

1. Ici $m = 20$ et $\sigma = 0,2$. La pièce est conforme si $19,6 \leq X \leq 20,4$. Les deux bornes sont symétriques par rapport à 20, à une distance de 0,4, soit exactement deux écarts-types :

$$\frac{20,4 - 20}{0,2} = 2 \quad \text{et} \quad \frac{19,6 - 20}{0,2} = -2$$

Donc, en posant $T = \frac{X - 20}{0,2}$, qui suit $\mathcal{N}(0, 1)$,

$$P(19,6 \leq X \leq 20,4) = P(-2 \leq T \leq 2) = 2\Phi(2) - 1 = 2 \times 0,9772 - 1 = 0,9544$$

La probabilité qu'une pièce soit conforme vaut donc 0,9544, et la proportion de rebut est

$$1 - 0,9544 = 0,0456 \quad \text{soit} \quad 4,56 \%$$

Environ quarante-six pièces sur mille sont rejetées. On retrouve ici la règle des deux écarts-types, à connaître de tête : une loi normale place environ 95 % de sa masse dans l'intervalle $[m - 2\sigma; m + 2\sigma]$.

2. On cherche $a > 0$ tel que $P(20 - a \leq X \leq 20 + a) = 0,99$. En centrant et réduisant, l'intervalle devient symétrique autour de 0 :

$$P(20 - a \leq X \leq 20 + a) = P\left(-\frac{a}{0,2} \leq T \leq \frac{a}{0,2}\right) = 2\Phi\left(\frac{a}{0,2}\right) - 1$$

L'équation à résoudre s'écrit donc

$$2\Phi\left(\frac{a}{0,2}\right) - 1 = 0,99, \quad \text{c'est-à-dire} \quad \Phi\left(\frac{a}{0,2}\right) = \frac{1 + 0,99}{2} = 0,995$$

La table donne $\Phi(2,58) = 0,9951$, valeur la plus proche de 0,995 parmi celles fournies. On retient donc $\frac{a}{0,2} = 2,58$, d'où

$$a = 2,58 \times 0,2 = 0,516 \text{ mm}$$

L'intervalle de tolérance cherché est $[19,484; 20,516]$. Il faut donc élargir la tolérance actuelle, qui n'était que de 0,4 mm de demi-largeur, pour passer de 95,44 % à 99 % de pièces acceptées. Le passage de 0,99 à 0,995 est le point délicat de la question. Une probabilité 0,99 pour un intervalle **bilatéral** laisse 0,01 de masse répartie en deux queues égales, donc 0,005 de chaque côté, et la table donne la masse cumulée à gauche : il faut donc lire Φ en $1 - 0,005 = 0,995$, jamais en 0,99.

3. La machine produit maintenant $X' \hookrightarrow \mathcal{N}(20,1; 0,04)$. L'écart-type est inchangé, seule la moyenne s'est déplacée de 0,1 mm. L'intervalle de tolérance, lui, n'a pas bougé : il n'est donc plus symétrique autour de la moyenne, et il faudra deux lectures de table au lieu d'une.

On pose $T' = \frac{X' - 20,1}{0,2}$, qui suit $\mathcal{N}(0, 1)$. Les deux bornes réduites valent

$$\frac{20,4 - 20,1}{0,2} = \frac{0,3}{0,2} = 1,5 \quad \text{et} \quad \frac{19,6 - 20,1}{0,2} = \frac{-0,5}{0,2} = -2,5$$

D'où

$$\begin{aligned} P(19,6 \leq X' \leq 20,4) &= P(-2,5 \leq T' \leq 1,5) \\ &= \Phi(1,5) - \Phi(-2,5) \\ &= \Phi(1,5) - (1 - \Phi(2,5)) \\ &= 0,9332 - (1 - 0,9938) \\ &= 0,9332 - 0,0062 \\ &= 0,9270 \end{aligned}$$

La proportion de rebut devient

$$1 - 0,9270 = 0,0730 \quad \text{soit} \quad 7,30 \%$$

Un dérèglement d'un demi-écart-type seulement, soit un dixième de millimètre, fait passer le rebut de 4,56 % à 7,30 % : il augmente de plus de moitié. C'est exactement ce que quantifie un contrôle de fabrication. On observe aussi le déséquilibre des rejets : la queue supérieure représente $1 - \Phi(1,5) = 0,0668$, contre 0,0062 pour la queue inférieure. Les rebuts sont désormais presque tous des pièces trop grandes, ce qui est l'indice concret du sens du dérèglement.

4. L'écart-type reste 0,2 et la moyenne vaut m . On pose $T = \frac{X - m}{0,2}$, qui suit $\mathcal{N}(0, 1)$. Comme X est une variable à densité, $P(X < 19,6) = P(X \leq 19,6)$, et

$$P(X < 19,6) = \Phi\left(\frac{19,6 - m}{0,2}\right)$$

La condition demandée s'écrit donc

$$\Phi\left(\frac{19,6 - m}{0,2}\right) \leq 0,001$$

La table fournit $\Phi(3,09) = 0,9990$, d'où $\Phi(-3,09) = 1 - 0,9990 = 0,0010$. La fonction Φ étant **strictement croissante** sur $\mathbb{R}$, l'inégalité $\Phi(u) \leq \Phi(-3,09)$ équivaut à $u \leq -3,09$. Ainsi

$$\frac{19,6 - m}{0,2} \leq -3,09$$

En multipliant par 0,2, qui est positif et ne change donc pas le sens de l'inégalité :

$$19,6 - m \leq -3,09 \times 0,2 = -0,618$$

puis, en isolant m :

$$m \geq 19,6 + 0,618 = 20,218$$

Il faut donc régler la machine sur une moyenne d'au moins 20,218 mm pour que moins d'une pièce sur mille soit trop petite. Ce résultat est à prendre avec le recul de l'ingénieur : plus on éloigne la moyenne de la borne basse, plus on se rapproche de la borne haute 20,4, et plus on produit de pièces trop grandes. Le réglage $m = 20,218$ n'est optimal que si l'on ne se soucie que de la borne inférieure.

5. La machine est réglée sur $\mathcal{N}(20; 0,04)$. On cherche une probabilité **conditionnelle** : la condition « la pièce est conforme » restreint l'univers. Notons C l'événement « la pièce est conforme », c'est-à-dire $(19,6 \leq X \leq 20,4)$. La question 1. a établi $P(C) = 0,9544 \neq 0$, le conditionnement est donc licite.

On cherche $P_C(X > 20,2)$. Par définition,

$$P_C(X > 20,2) = \frac{P(C \cap (X > 20,2))}{P(C)}$$

L'intersection ne se simplifie pas ici en un seul des deux événements : elle vaut $(20,2 < X \leq 20,4)$, car il faut à la fois dépasser 20,2 et rester conforme, donc sous 20,4. C'est le point à ne pas bâcler.

En centrant et réduisant, avec $T = \frac{X - 20}{0,2}$:

$$\frac{20,2 - 20}{0,2} = 1 \quad \text{et} \quad \frac{20,4 - 20}{0,2} = 2$$

d'où

$$P(20,2 < X \leq 20,4) = P(1 < T \leq 2) = \Phi(2) - \Phi(1) = 0,9772 - 0,8413 = 0,1359$$

Finalement

$$P_C(X > 20,2) = \frac{0,1359}{0,9544} \approx 0,1424$$

Environ 14,2 % des pièces conformes ont un diamètre supérieur à 20,2 mm. Sans conditionnement, on aurait trouvé $P(X > 20,2) = 1 - \Phi(1) = 0,1587$, soit 15,87 %. Le conditionnement

fait baisser la probabilité, ce qui est logique : en écartant les pièces non conformes, on élimine des pièces très grandes, donc des pièces qui dépassaient 20,2 mm, et l'on ne retire aucune pièce petite de la catégorie recherchée.

Méthode à retenir. Sur un exercice de loi normale, on n'écrit jamais une ligne sans avoir centré et réduit. Le réflexe est de calculer immédiatement le nombre d'écart-types séparant chaque borne de la moyenne, puis de traduire l'événement en une inégalité sur la variable centrée réduite. Quand l'intervalle est symétrique autour de la moyenne, on utilise $2\Phi(u) - 1$; sinon, on écrit une différence $\Phi(u_2) - \Phi(u_1)$ et l'on transforme les valeurs négatives par $\Phi(-u) = 1 - \Phi(u)$, la table ne donnant que les valeurs positives. Pour un problème inverse (trouver un seuil, une tolérance, un réglage), on isole $\Phi(\cdot)$, on lit la table à l'envers, puis on invoque la stricte croissance de Φ pour transformer l'inégalité sur les probabilités en inégalité sur les quantiles.

Erreur classique. Confondre variance et écart-type dans la notation $\mathcal{N}(20; 0,04)$. Le second paramètre est la **variance**, l'écart-type vaut donc 0,2 et non 0,04. Prendre $\sigma = 0,04$ multiplierait par cinq tous les quotients et donnerait une probabilité de conformité pratiquement égale à 1, résultat absurde dans un problème de contrôle qualité. Seconde erreur, à la question **2.** : lire directement $\Phi^{-1}(0,99) = 2,33$ et conclure $a = 0,466$. Cette valeur répond à la question unilatérale $P(X \leq 20 + a) = 0,99$, pas à la question bilatérale posée ici.

Exercice 16 ★★ Mélange de deux fournisseurs

Notions : Espérance conditionnelle et formule de l'espérance totale, Loi exponentielle et absence de mémoire, Fonction de répartition d'une variable aléatoire réelle quelconque

Une entreprise s'approvisionne en composants auprès de deux fournisseurs. Une proportion de 60 % des composants provient du fournisseur A , dont la durée de vie, exprimée en années, suit la loi $\mathcal{E}(\frac{1}{4})$. Les 40 % restants proviennent du fournisseur B , dont la durée de vie suit la loi $\mathcal{E}(\frac{1}{6})$.

On prélève un composant au hasard dans le stock et l'on note X sa durée de vie, en années. On note A l'événement « le composant provient du fournisseur A » et B l'événement « le composant provient du fournisseur B ».

1. À l'aide de la formule des probabilités totales, montrer que pour tout réel $t \geq 0$,

$$F_X(t) = 1 - 0,6 e^{-t/4} - 0,4 e^{-t/6}$$

Préciser également $F_X(t)$ pour $t < 0$.

2. En déduire que X est une variable à densité et en donner une densité. Montrer que X ne suit pas une loi exponentielle.
3. Montrer que X admet une espérance et la calculer directement à partir de la densité obtenue à la question 2.
4. On admet, pour une variable à densité, la formule de l'espérance totale

$$E(X) = P(A) E(X | A) + P(B) E(X | B)$$

Vérifier qu'elle redonne bien le résultat de la question 3.

5. Un composant prélevé au hasard a fonctionné plus de 10 ans. Quelle est la probabilité qu'il provienne du fournisseur A ? On donnera une valeur arrondie à 10^{-4} près.
6. Commenter le résultat de la question 5. au regard de la proportion initiale de 60 %.

On donne les valeurs approchées suivantes : $e^{-2,5} \approx 0,08208$ et $e^{-5/3} \approx 0,18888$.

Corrigé

1. Le couple d'événements (A, B) forme un **système complet d'événements** : le composant provient de l'un ou de l'autre fournisseur, et pas des deux. On a $P(A) = 0,6$ et $P(B) = 0,4$, deux probabilités non nulles, ce qui autorise à conditionner par chacun d'eux.

Soit $t \geq 0$. La formule des probabilités totales appliquée à ce système complet donne

$$F_X(t) = P(X \leq t) = P(A) P_A(X \leq t) + P(B) P_B(X \leq t)$$

Sachant A , la durée de vie suit la loi $\mathcal{E}(\frac{1}{4})$, dont la fonction de répartition vaut $1 - e^{-t/4}$ pour $t \geq 0$. Sachant B , elle suit $\mathcal{E}(\frac{1}{6})$, de fonction de répartition $1 - e^{-t/6}$ pour $t \geq 0$. Donc

$$\begin{aligned} F_X(t) &= 0,6 (1 - e^{-t/4}) + 0,4 (1 - e^{-t/6}) \\ &= 0,6 + 0,4 - 0,6 e^{-t/4} - 0,4 e^{-t/6} \\ &= 1 - 0,6 e^{-t/4} - 0,4 e^{-t/6} \end{aligned}$$

Pour $t < 0$, les deux lois conditionnelles sont portées par $\mathbb{R}_+$, donc $P_A(X \leq t) = P_B(X \leq t) = 0$ et la même formule donne

$$F_X(t) = 0$$

Contrôles. En $t = 0$: $F_X(0) = 1 - 0,6 - 0,4 = 0$, ce qui raccorde bien avec la valeur nulle à gauche. En $+\infty$: les deux exponentielles tendent vers 0, donc $F_X(t) \rightarrow 1$. La fonction obtenue a bien le comportement d'une fonction de répartition.

Il vaut la peine de retenir la forme équivalente de ce résultat, écrite en termes de survie :

$$P(X > t) = 0,6 e^{-t/4} + 0,4 e^{-t/6} \quad \text{pour } t \geq 0$$

C'est elle qui servira à la question 5.

2. Caractère à densité. La fonction F_X est continue sur $]-\infty, 0[$ et sur $[0, +\infty[$, et le raccordement en 0 est bon puisque $F_X(0) = 0$: elle est donc continue sur $\mathbb{R}$. Elle est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ privé du seul point 0, comme somme de fonctions exponentielles sur $]0, +\infty[$ et comme fonction nulle sur $]-\infty, 0[$. Un point isolé étant un nombre fini de points, X est bien une variable à densité.

Densité. On dérive F_X là où elle est dérivable. Pour $t > 0$,

$$f_X(t) = 0,6 \times \frac{1}{4} e^{-t/4} + 0,4 \times \frac{1}{6} e^{-t/6} = 0,15 e^{-t/4} + \frac{1}{15} e^{-t/6}$$

et $f_X(t) = 0$ pour $t \leq 0$, la valeur en 0 étant choisie arbitrairement.

On reconnaît la structure d'un mélange : la densité de X est la combinaison $0,6 f_A + 0,4 f_B$ des deux densités exponentielles, avec pour coefficients les proportions des fournisseurs. Vérifions d'ailleurs que l'intégrale vaut bien 1, en utilisant $\int_0^{+\infty} e^{-t/a} dt = a$ pour $a > 0$:

$$\int_0^{+\infty} f_X(t) dt = 0,15 \times 4 + \frac{1}{15} \times 6 = 0,6 + 0,4 = 1$$

La loi de X n'est pas exponentielle. Raisonnons par l'absurde et supposons qu'il existe $\mu > 0$ tel que $X \hookrightarrow \mathcal{E}(\mu)$. La densité de X serait alors $t \mapsto \mu e^{-\mu t}$ sur $]0, +\infty[$. Deux densités d'une même loi coïncident en tout point où elles sont toutes deux continues, donc pour tout $t > 0$,

$$\mu e^{-\mu t} = 0,15 e^{-t/4} + \frac{1}{15} e^{-t/6}$$

En faisant tendre t vers 0 par valeurs positives, les trois exponentielles tendent vers 1 et l'on obtient

$$\mu = 0,15 + \frac{1}{15} = \frac{9}{60} + \frac{4}{60} = \frac{13}{60}$$

L'espérance de X vaudrait alors $\frac{1}{\mu} = \frac{60}{13} \approx 4,615$ années. Or la question 3. établit $E(X) = 4,8$ années, et $4,615 \neq 4,8$. Contradiction. **La loi de X n'est donc pas une loi exponentielle.**

Ce fait est important et va à l'encontre de l'intuition : mélanger deux populations sans mémoire produit une population qui, elle, a de la mémoire. On peut le voir directement sur le modèle. Sachant qu'un composant a survécu longtemps, il est probable qu'il vienne du fournisseur le plus fiable, et cette information change ses perspectives d'avenir. Le mélange crée donc une dépendance au passé qui n'existait dans aucune des deux populations d'origine.

3. L'espérance existe si l'intégrale $\int_{-\infty}^{+\infty} t f_X(t) dt$ converge absolument. La fonction intégrée est continue et positive sur $[0, +\infty[$, et nulle ailleurs, donc convergence et convergence absolue coïncident. La convergence en $+\infty$ résulte de la croissance comparée : chacun des deux termes $t e^{-t/4}$ et $t e^{-t/6}$ est négligeable devant $\frac{1}{t^2}$, dont l'intégrale converge en $+\infty$. **L'espérance existe donc.**

Pour le calcul, on utilise le résultat classique, obtenu par intégration par parties ou directement à partir de l'espérance de la loi $\mathcal{E}(\frac{1}{a})$:

$$\int_0^{+\infty} t e^{-t/a} dt = a^2 \quad \text{pour } a > 0$$

En effet, $\frac{1}{a} e^{-t/a}$ est la densité de la loi $\mathcal{E}(\frac{1}{a})$, dont l'espérance vaut a , ce qui donne $\frac{1}{a} \int_0^{+\infty} t e^{-t/a} dt = a$. Par linéarité de l'intégrale, les deux intégrales convergeant séparément,

$$\begin{aligned} E(X) &= \int_0^{+\infty} t \left(0,15 e^{-t/4} + \frac{1}{15} e^{-t/6} \right) dt \\ &= 0,15 \int_0^{+\infty} t e^{-t/4} dt + \frac{1}{15} \int_0^{+\infty} t e^{-t/6} dt \\ &= 0,15 \times 4^2 + \frac{1}{15} \times 6^2 \\ &= 0,15 \times 16 + \frac{36}{15} \\ &= 2,4 + 2,4 = 4,8 \end{aligned}$$

La durée de vie moyenne d'un composant prélevé au hasard est donc de 4,8 années.

4. Sachant A , la durée de vie suit la loi $\mathcal{E}(\frac{1}{4})$, dont l'espérance vaut $\frac{1}{1/4} = 4$ années. Sachant B , elle suit $\mathcal{E}(\frac{1}{6})$, d'espérance 6 années. Autrement dit

$$E(X \mid A) = 4 \quad \text{et} \quad E(X \mid B) = 6$$

La formule de l'espérance totale donne alors

$$E(X) = P(A) E(X \mid A) + P(B) E(X \mid B) = 0,6 \times 4 + 0,4 \times 6 = 2,4 + 2,4 = 4,8$$

On retrouve exactement la valeur de la question 3., et l'on comprend rétrospectivement d'où venaient les deux termes égaux à 2,4 dans le calcul intégral : ils sont les contributions respectives des deux fournisseurs. La formule de l'espérance totale évite ici tout calcul d'intégrale, ce qui en fait la méthode à privilégier dès qu'un mélange est décrit par un système complet d'événements. Notons au passage que 4,8 est bien compris entre 4 et 6, et plus proche de 4 que de 6, comme le veut une moyenne pondérée qui accorde le poids 0,6 à la valeur 4. Ce contrôle d'ordre de grandeur prend une seconde et détecte la plupart des erreurs de pondération.

5. On demande $P_{(X>10)}(A)$, c'est-à-dire une probabilité de cause sachant une conséquence : c'est exactement le cadre de la **formule de Bayes**.

Numérateur. Par la formule des probabilités composées,

$$P(A \cap (X > 10)) = P(A) P_A(X > 10) = 0,6 \times e^{-10/4} = 0,6 e^{-2,5}$$

Avec $e^{-2,5} \approx 0,08208$, ce produit vaut environ 0,04925.

Dénominateur. On reprend la fonction de survie établie à la question 1. :

$$P(X > 10) = 0,6 e^{-2,5} + 0,4 e^{-10/6} = 0,6 e^{-2,5} + 0,4 e^{-5/3}$$

Avec $e^{-5/3} \approx 0,18888$, le second terme vaut environ 0,07555, d'où

$$P(X > 10) \approx 0,04925 + 0,07555 = 0,12480$$

Cette probabilité est non nulle, le conditionnement est licite.

Conclusion.

$$P_{(X>10)}(A) = \frac{0,6 e^{-2,5}}{0,6 e^{-2,5} + 0,4 e^{-5/3}} \approx \frac{0,04925}{0,12480} \approx 0,3946$$

Par complémentaire, $P_{(X>10)}(B) \approx 1 - 0,3946 = 0,6054$.

6. Avant toute observation, un composant a 60 % de chances de venir du fournisseur A . Après avoir constaté qu'il a tenu plus de dix ans, cette probabilité tombe à 39,5 % environ, tandis que

celle de venir du fournisseur B monte de 40 % à 60,5 %. Le rapport des deux fournisseurs s'est donc **inversé**.

L'explication tient en une phrase : le fournisseur B produit des composants de durée de vie moyenne 6 ans, contre 4 ans pour le fournisseur A , et l'écart se creuse dans les queues de distribution. Un composant de B a $e^{-5/3} \approx 0,189$ de chances de dépasser dix ans, contre $e^{-2,5} \approx 0,082$ pour un composant de A , soit plus du double. La condition « avoir survécu dix ans » sélectionne donc préférentiellement les composants du fournisseur le plus fiable, au point de renverser un avantage initial pourtant net en nombre.

C'est un mécanisme général et très utile à reconnaître : **survivre longtemps est un indice de provenance**. Le raisonnement est le même que celui d'un assureur qui, observant un conducteur sans sinistre depuis dix ans, révisé à la hausse sa probabilité d'appartenir au groupe des conducteurs prudents. Concrètement, si l'entreprise cherchait à identifier a posteriori l'origine de ses composants les plus durables, elle aurait tout intérêt à parier sur B .

Méthode à retenir. Un mélange de populations se traite toujours en deux temps. On identifie d'abord le système complet d'événements qui décrit l'origine (ici les deux fournisseurs), puis on applique la formule des probabilités totales à la **fonction de répartition**, jamais directement aux densités si l'on veut rester dans le cadre du programme. La densité s'obtient ensuite par dérivation, et l'on constate qu'elle est le mélange des densités avec les mêmes poids. Pour l'espérance, la formule de l'espérance totale est presque toujours plus rapide que l'intégrale, et fournit en tout cas un contrôle immédiat. Enfin, dès qu'une question demande la probabilité d'une **origine** sachant une **observation**, c'est Bayes, et le dénominateur est toujours la probabilité totale de l'observation.

Erreur classique. Croire qu'un mélange de lois exponentielles est encore une loi exponentielle, et écrire directement $X \hookrightarrow \mathcal{E}(\mu)$ avec $\mu = 0,6 \times \frac{1}{4} + 0,4 \times \frac{1}{6}$. Ce qui se mélange, ce sont les **lois**, c'est-à-dire les fonctions de répartition et les densités, jamais les paramètres. La preuve la plus parlante est numérique : le paramètre ainsi fabriqué vaut $\frac{13}{60}$ et donnerait une espérance de 4,615 ans, alors que la vraie espérance est 4,8 ans. Seconde erreur : confondre $P_A(X > 10)$, qui vaut $e^{-2,5}$, avec $P_{(X>10)}(A)$, qui vaut 0,3946. Ces deux quantités conditionnent en sens inverse, et c'est précisément ce renversement que la formule de Bayes organise.

Exercice 17 ★★ Médiane d'une loi à densité

Notions : Fonction de répartition d'une variable aléatoire réelle quelconque, Loi uniforme sur un segment, Loi exponentielle et absence de mémoire, Variable aléatoire à densité : définition et reconnaissance d'une densité

Soit X une variable aléatoire à densité. On appelle **médiane** de X tout réel m vérifiant

$$F_X(m) = \frac{1}{2}$$

Dans tout l'exercice, on se place dans le cas favorable où F_X est continue sur $\mathbb{R}$ et strictement croissante sur un intervalle I contenant le support de la loi, avec F_X nulle à gauche de I et égale à 1 à droite de I .

Question préliminaire. Justifier que, sous ces hypothèses, la médiane existe et est unique.

1. Déterminer la médiane de la loi $\mathcal{U}([a, b])$, où $a < b$.
2. Déterminer la médiane de la loi $\mathcal{E}(\lambda)$, où $\lambda > 0$. Comparer cette médiane à l'espérance $\frac{1}{\lambda}$ et expliquer le sens de l'inégalité obtenue.
3. Soit X une variable de densité $f(x) = 3x^2$ sur $[0, 1]$ et $f(x) = 0$ ailleurs. Déterminer la médiane de X et en donner une valeur approchée à 10^{-3} près.
4. Le premier quartile q_1 et le troisième quartile q_3 d'une variable à densité sont définis par $F_X(q_1) = \frac{1}{4}$ et $F_X(q_3) = \frac{3}{4}$. Déterminer q_1 et q_3 pour la loi $\mathcal{E}(\lambda)$.
5. La durée de vie d'un appareil, en années, est modélisée par une variable de loi exponentielle. Une étude établit que la moitié des appareils tombe en panne avant 5 ans. Déterminer le paramètre λ , puis la durée de vie moyenne. Commenter l'écart entre ces deux indicateurs.

On donne les valeurs approchées suivantes : $2^{-1/3} \approx 0,794$, $\ln(2) \approx 0,6931$, $\ln(3) \approx 1,0986$ et $e^{-1} \approx 0,3679$.

Corrigé

Question préliminaire. Existence. La fonction F_X est continue sur $\mathbb{R}$, de limite 0 en $-\infty$ et de limite 1 en $+\infty$. Comme $\frac{1}{2}$ appartient à l'intervalle $]0, 1[$, le théorème des valeurs intermédiaires assure l'existence d'un réel m tel que $F_X(m) = \frac{1}{2}$. C'est ici que sert la continuité de F_X , propriété caractéristique des variables à densité : une variable **discrète** a une fonction de répartition en escalier, qui peut sauter par-dessus la valeur $\frac{1}{2}$ sans jamais l'atteindre, et pour laquelle la médiane doit être définie autrement.

Unicité. Supposons que m et m' soient deux médianes, avec $m < m'$. Comme $F_X(m) = \frac{1}{2} \in]0, 1[$, le réel m appartient nécessairement à l'intervalle I sur lequel F_X est strictement croissante, et de même pour m' . La stricte croissance sur I donnerait alors $F_X(m) < F_X(m')$, ce qui contredit $F_X(m) = F_X(m') = \frac{1}{2}$. Donc $m = m'$: **la médiane est unique.** $\square$

La médiane partage la loi en deux moitiés de même masse : $P(X \leq m) = \frac{1}{2}$ et, par passage au complémentaire, $P(X > m) = \frac{1}{2}$. C'est un indicateur de **position**, comme l'espérance, mais construit sur un principe entièrement différent : il coupe la population en deux, là où l'espérance calcule un centre de gravité.

1. Si $X \hookrightarrow \mathcal{U}([a, b])$, la fonction de répartition vaut

$$F_X(x) = 0 \quad \text{si } x < a, \quad F_X(x) = \frac{x-a}{b-a} \quad \text{si } a \leq x \leq b, \quad F_X(x) = 1 \quad \text{si } x > b.$$

Cette fonction est bien continue sur $\mathbb{R}$ et strictement croissante sur $[a, b]$, puisque $b - a > 0$: les hypothèses de la question préliminaire sont vérifiées. La médiane appartient donc à $[a, b]$ et

s'obtient en résolvant

$$\frac{m-a}{b-a} = \frac{1}{2}$$

soit $m-a = \frac{b-a}{2}$, puis

$$m = a + \frac{b-a}{2} = \frac{2a+b-a}{2} = \frac{a+b}{2}$$

La médiane de la loi uniforme est donc le **milieu du segment**, ce qui coïncide exactement avec son espérance $\frac{a+b}{2}$. Rien d'étonnant : la loi uniforme est symétrique par rapport à ce point, et pour toute loi symétrique admettant une espérance, médiane et espérance sont confondues au centre de symétrie.

2. Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors $F_X(x) = 1 - e^{-\lambda x}$ pour $x \geq 0$ et $F_X(x) = 0$ pour $x < 0$. La fonction est continue sur $\mathbb{R}$, et strictement croissante sur $[0, +\infty[$ puisque sa dérivée $\lambda e^{-\lambda x}$ y est strictement positive. Les hypothèses sont donc vérifiées, et la médiane est positive. On résout

$$1 - e^{-\lambda m} = \frac{1}{2}$$

ce qui équivaut successivement à $e^{-\lambda m} = \frac{1}{2}$, puis, en composant par le logarithme népérien qui est strictement croissant donc injectif, à $-\lambda m = \ln \frac{1}{2} = -\ln 2$. Comme $\lambda > 0$, on peut diviser et

$$m = \frac{\ln 2}{\lambda}$$

Comparaison avec l'espérance. L'espérance de la loi $\mathcal{E}(\lambda)$ vaut $\frac{1}{\lambda}$. Le quotient des deux indicateurs ne dépend pas de λ :

$$\frac{m}{E(X)} = \frac{\ln 2 / \lambda}{1/\lambda} = \ln 2 \approx 0,693$$

Comme $\ln 2 < 1$, on a toujours

$$m = \frac{\ln 2}{\lambda} < \frac{1}{\lambda} = E(X)$$

La médiane vaut environ 69,3 % de l'espérance, quelle que soit la valeur de λ .

Interprétation. La loi exponentielle est **étalée à droite** : son support est non borné vers le haut, tandis qu'elle est bloquée par 0 vers le bas. Les quelques réalisations très grandes, rares mais parfois considérables, tirent la moyenne vers le haut sans déplacer la médiane, laquelle ne dépend que du **rang** des observations et non de leur valeur. Concrètement, plus de la moitié des réalisations d'une loi exponentielle sont inférieures à sa moyenne : on calcule d'ailleurs

$$P\left(X \leq \frac{1}{\lambda}\right) = 1 - e^{-1} \approx 0,632$$

soit 63,2 % des réalisations sous la moyenne. Ce phénomène se retrouve dans toutes les lois asymétriques à droite : les salaires, les temps d'attente, les tailles d'entreprises. C'est la raison pour laquelle la médiane est souvent l'indicateur le plus honnête pour décrire de telles populations.

3. Vérifions d'abord que f est une densité : elle est positive sur $[0, 1]$, continue sauf éventuellement en 0 et en 1, et son intégrale vaut $\int_0^1 3x^2 dx = [x^3]_0^1 = 1$.

La fonction de répartition s'obtient par intégration. Pour $x \in [0, 1]$,

$$F_X(x) = \int_0^x 3t^2 dt = [t^3]_0^x = x^3$$

avec $F_X(x) = 0$ pour $x < 0$ et $F_X(x) = 1$ pour $x > 1$. Cette fonction est continue sur $\mathbb{R}$ et strictement croissante sur $[0, 1]$, la fonction cube étant strictement croissante. La médiane existe, est unique, et appartient à $[0, 1]$. On résout

$$m^3 = \frac{1}{2}$$

La fonction cube étant une bijection strictement croissante de $\mathbb{R}$ dans $\mathbb{R}$, il vient

$$m = \left(\frac{1}{2}\right)^{1/3} = 2^{-1/3} = \frac{1}{\sqrt[3]{2}} \approx 0,794$$

Comparaison. L'espérance existe, car la fonction $x \mapsto xf(x)$ est continue sur le segment $[0, 1]$ et nulle en dehors : l'intégrale n'est même pas impropre. Elle vaut

$$E(X) = \int_0^1 x \times 3x^2 dx = 3 \int_0^1 x^3 dx = \frac{3}{4} = 0,75$$

Ici la médiane 0,794 est **supérieure** à l'espérance 0,75, contrairement au cas exponentiel. C'est cohérent avec la forme de la densité : $f(x) = 3x^2$ est croissante, la masse s'accumule près de 1, la loi est étalée vers la gauche. Il n'y a donc aucune règle générale sur le sens de l'inégalité entre médiane et espérance : tout dépend du sens de l'asymétrie.

4. On reprend $F_X(x) = 1 - e^{-\lambda x}$ pour $x \geq 0$.

Premier quartile. On résout $1 - e^{-\lambda q_1} = \frac{1}{4}$, soit $e^{-\lambda q_1} = \frac{3}{4}$, puis $-\lambda q_1 = \ln \frac{3}{4}$, d'où

$$q_1 = -\frac{1}{\lambda} \ln \frac{3}{4} = \frac{1}{\lambda} \ln \frac{4}{3} \approx \frac{0,2877}{\lambda}$$

Troisième quartile. On résout $1 - e^{-\lambda q_3} = \frac{3}{4}$, soit $e^{-\lambda q_3} = \frac{1}{4}$, puis $-\lambda q_3 = -\ln 4$, d'où

$$q_3 = \frac{\ln 4}{\lambda} = \frac{2 \ln 2}{\lambda} \approx \frac{1,3863}{\lambda}$$

On remarque au passage la relation $q_3 = 2m$, valable pour toute loi exponentielle, puisque $m = \frac{\ln 2}{\lambda}$.

Écart interquartile. Il vaut

$$q_3 - q_1 = \frac{\ln 4 - \ln \frac{4}{3}}{\lambda} = \frac{\ln 3}{\lambda} \approx \frac{1,0986}{\lambda}$$

Cet écart, qui mesure la dispersion de la moitié centrale de la population, vaut donc environ 1,0986 fois l'espérance. Il est intéressant de constater que les trois quarts de la masse se trouvent avant $\frac{1,386}{\lambda}$, alors que la loi s'étend jusqu'à l'infini : la queue de distribution est longue mais légère.

5. L'énoncé indique que la moitié des appareils tombe en panne avant cinq ans, ce qui signifie exactement $P(X \leq 5) = \frac{1}{2}$, c'est-à-dire que 5 est la **médiane** de la loi. D'après la question 2., la médiane d'une loi $\mathcal{E}(\lambda)$ vaut $\frac{\ln 2}{\lambda}$, d'où l'équation

$$\frac{\ln 2}{\lambda} = 5$$

qui donne

$$\lambda = \frac{\ln 2}{5} \approx \frac{0,6931}{5} \approx 0,1386$$

Le paramètre s'exprime en pannes par an. La durée de vie moyenne est alors

$$E(X) = \frac{1}{\lambda} = \frac{5}{\ln 2} \approx 7,21 \text{ ans}$$

Commentaire. L'écart est considérable : la moitié des appareils rend l'âme avant cinq ans, et pourtant la durée de vie **moyenne** annoncée serait de plus de sept ans. Ces deux chiffres décrivent la même population, ils ne se contredisent pas, mais ils ne racontent pas la même histoire.

L'explication est celle de la question 2. : la moyenne est tirée vers le haut par une minorité d'appareils très durables. On peut le chiffrer. La probabilité qu'un appareil dépasse quinze ans vaut $e^{-15\lambda} = e^{-3\ln 2} = \frac{1}{8}$, soit un appareil sur huit, et celle de dépasser trente ans vaut $\frac{1}{64}$. Ce sont ces survivants qui font monter la moyenne, sans jamais déplacer la médiane.

Sur le plan commercial, la conséquence est directe. Un fabricant qui annonce « durée de vie moyenne : sept ans » communique un chiffre exact et pourtant trompeur, puisque la majorité de ses clients constatera une panne bien avant. Un consommateur avisé demandera la médiane, et un organisme de contrôle exigera les deux.

Méthode à retenir. Chercher une médiane, un quartile ou plus généralement un quantile, c'est toujours résoudre une équation de la forme $F_X(x) = p$. La marche à suivre est invariable : écrire F_X sur le bon morceau, poser l'équation, la résoudre en composant par la fonction réciproque appropriée (logarithme pour une exponentielle, racine cubique pour un cube), et vérifier que la solution appartient bien à l'intervalle sur lequel l'expression de F_X était valable. La justification d'existence et d'unicité repose sur deux propriétés distinctes qu'il ne faut pas confondre : la **continuité** de F_X donne l'existence par valeurs intermédiaires, la **stricte croissance** donne l'unicité.

Erreur classique. Confondre médiane et espérance, ou pire, croire que l'égalité $m = E(X)$ observée sur la loi uniforme vaut en général. Elle ne vaut que pour les lois symétriques. Sur une loi exponentielle, prendre la moyenne pour la médiane conduit à une erreur relative de plus de 30 %, dans le sens qui surestime systématiquement la durée typique. Seconde erreur, fréquente à la question 4. : écrire $q_1 = \frac{1}{\lambda} \ln \frac{1}{4}$. Le signe est faux, et le résultat serait négatif, ce qui est impossible pour une variable positive. Un contrôle de signe en fin de calcul suffit à repérer la faute.

Exercice 18 ★★ Simulation par la méthode d'inversion

Notions : Loi d'une transformée : transformation affine et méthode de la fonction de répartition, Loi uniforme sur un segment, Loi exponentielle et absence de mémoire

Soit U une variable aléatoire telle que $U \hookrightarrow \mathcal{U}(]0, 1[)$. On cherche à fabriquer, à partir de U seule, des variables aléatoires de lois imposées. Toutes les réponses sont des formules mathématiques, aucune ne fait appel à un programme informatique.

1. Montrer que la variable $1 - U$ suit également la loi $\mathcal{U}(]0, 1[)$.
2. Soit $\lambda > 0$. Montrer que la variable $X = -\frac{\ln U}{\lambda}$ suit la loi $\mathcal{E}(\lambda)$. On soignera particulièrement le sens des inégalités.
3. Soit F une fonction continue et strictement croissante de $\mathbb{R}$ dans $]0, 1[$, admettant donc une bijection réciproque F^{-1} de $]0, 1[$ dans $\mathbb{R}$. Montrer que la variable $F^{-1}(U)$ admet F pour fonction de répartition.
4. Soit X une variable de fonction de répartition définie par $F(x) = 1 - \frac{1}{x^2}$ pour $x \geq 1$, et $F(x) = 0$ pour $x < 1$. Construire, à partir de U , une variable de même loi que X .
5. Soient $a < b$ deux réels. Construire, à partir de U , une variable de loi $\mathcal{U}([a, b])$.

Corrigé

Rappelons d'abord la loi de départ. Si $U \hookrightarrow \mathcal{U}(]0, 1[)$, une densité de U vaut 1 sur $]0, 1[$ et 0 ailleurs, et sa fonction de répartition est

$$F_U(u) = 0 \text{ si } u \leq 0, \quad F_U(u) = u \text{ si } 0 < u < 1, \quad F_U(u) = 1 \text{ si } u \geq 1.$$

Le fait que l'intervalle soit ouvert ne change rien aux probabilités, puisque $P(U = 0) = P(U = 1) = 0$ pour une variable à densité. Il garantit en revanche que U ne prend jamais les valeurs 0 ni 1, ce qui permettra d'écrire $\ln U$ et $\frac{1}{\sqrt{U}}$ sans réserve.

Toutes les questions se traitent par la **méthode de la fonction de répartition**, seule méthode au programme pour déterminer la loi d'une transformée.

1. Posons $V = 1 - U$. Comme U prend ses valeurs dans $]0, 1[$, il en va de même de V : si $0 < U < 1$, alors $0 < 1 - U < 1$. Donc $V(\Omega) =]0, 1[$.

Soit $x \in \mathbb{R}$. On transforme l'événement $(V \leq x)$ en un événement portant sur U :

$$F_V(x) = P(1 - U \leq x) = P(-U \leq x - 1) = P(U \geq 1 - x)$$

La multiplication par -1 a **retourné l'inégalité**, c'est le point de vigilance de la question. Ensuite, U étant une variable à densité, $P(U = 1 - x) = 0$, donc

$$F_V(x) = 1 - P(U < 1 - x) = 1 - F_U(1 - x)$$

Il reste à discuter selon la position de x .

Si $x \leq 0$. Alors $1 - x \geq 1$, donc $F_U(1 - x) = 1$ et $F_V(x) = 0$.

Si $0 < x < 1$. Alors $0 < 1 - x < 1$, donc $F_U(1 - x) = 1 - x$ et

$$F_V(x) = 1 - (1 - x) = x$$

Si $x \geq 1$. Alors $1 - x \leq 0$, donc $F_U(1 - x) = 0$ et $F_V(x) = 1$.

La fonction F_V coïncide donc exactement avec F_U sur $\mathbb{R}$. Deux variables ayant la même fonction de répartition ont la même loi, donc

$$1 - U \hookrightarrow \mathcal{U}(]0, 1[)$$

□

Ce résultat, presque évident sur un dessin (la symétrie $u \mapsto 1 - u$ échange les deux extrémités de l'intervalle sans déformer la densité constante), est un outil pratique : il permet de remplacer $1 - U$ par U dans toute formule de simulation, ce qui allège les expressions. La question 4. en tirera parti.

2. Posons $X = -\frac{\ln U}{\lambda}$.

Support. Comme U prend ses valeurs dans $]0, 1[$, le logarithme $\ln U$ est strictement négatif, donc $-\ln U$ est strictement positif, et $\lambda > 0$ ne change pas le signe. Ainsi X est à valeurs dans $]0, +\infty[$, ce qui est déjà cohérent avec la loi visée.

Fonction de répartition. Soit $x \in \mathbb{R}$.

Si $x \leq 0$. L'événement $(X \leq x)$ est de probabilité nulle, car $X > 0$. Donc $F_X(x) = 0$.

Si $x > 0$. On transforme l'événement pas à pas, en surveillant chaque opération.

$$\begin{aligned}(X \leq x) &= \left(-\frac{\ln U}{\lambda} \leq x \right) \\ &= (\ln U \geq -\lambda x) \\ &= (U \geq e^{-\lambda x})\end{aligned}$$

Détaillons les deux étapes. Le passage de la première à la deuxième ligne consiste à multiplier par $-\lambda$, qui est **strictement négatif** : l'inégalité change donc de sens. Le passage de la deuxième à la troisième consiste à composer par la fonction exponentielle, qui est **strictement croissante** : l'inégalité est cette fois conservée. Ces deux surveillances sont l'essentiel du travail, et une erreur sur l'une ou l'autre donnerait une fonction de répartition décroissante, résultat absurde à détecter immédiatement.

Il vient alors, en utilisant $P(U = e^{-\lambda x}) = 0$,

$$F_X(x) = P(U \geq e^{-\lambda x}) = 1 - F_U(e^{-\lambda x})$$

Comme $x > 0$ et $\lambda > 0$, on a $-\lambda x < 0$, donc $e^{-\lambda x} \in]0, 1[$ et $F_U(e^{-\lambda x}) = e^{-\lambda x}$. Finalement

$$F_X(x) = 1 - e^{-\lambda x}$$

On reconnaît la fonction de répartition de la loi exponentielle de paramètre λ , d'où

$$-\frac{\ln U}{\lambda} \hookrightarrow \mathcal{E}(\lambda)$$

□

C'est la recette de simulation d'une loi exponentielle : tirer un nombre au hasard entre 0 et 1, en prendre le logarithme, changer le signe et diviser par λ .

3. Posons $X = F^{-1}(U)$. La fonction F étant continue et strictement croissante de $\mathbb{R}$ dans $]0, 1[$, elle réalise une bijection de $\mathbb{R}$ sur $]0, 1[$, et sa réciproque F^{-1} est une bijection de $]0, 1[$ sur $\mathbb{R}$, elle aussi continue et **strictement croissante**. Comme U prend ses valeurs dans $]0, 1[$, la quantité $F^{-1}(U)$ est bien définie.

Soit $x \in \mathbb{R}$. La stricte croissance de F permet d'écrire l'équivalence

$$F^{-1}(U) \leq x \iff U \leq F(x)$$

Justifions-la, car c'est tout l'exercice. Si $F^{-1}(U) \leq x$, on applique F , strictement croissante donc croissante, aux deux membres : $F(F^{-1}(U)) \leq F(x)$, c'est-à-dire $U \leq F(x)$. Réciproquement, si $U \leq F(x)$, on applique F^{-1} , elle aussi croissante : $F^{-1}(U) \leq F^{-1}(F(x)) = x$. L'équivalence est établie.

Par conséquent

$$F_X(x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F_U(F(x))$$

Or $F(x)$ appartient à $]0, 1[$ par hypothèse sur F , donc $F_U(F(x)) = F(x)$. Ainsi, pour tout réel x ,

$$F_X(x) = F(x)$$

La variable $F^{-1}(U)$ admet donc bien F pour fonction de répartition. $\square$

C'est la **méthode d'inversion**, et son intérêt pratique est considérable. Un ordinateur ne sait produire qu'une seule chose, un nombre pseudo-aléatoire uniforme sur $]0, 1[$. Ce théorème affirme qu'à partir de cette unique brique, et pourvu que l'on sache inverser F , on peut simuler **n'importe quelle** loi à densité. Les questions **2.**, **4.** et **5.** en sont trois instances : dans chacune, on résout l'équation $F(x) = u$ d'inconnue x , et la solution est la formule de simulation.

4. Vérifions d'abord que F a bien les propriétés attendues sur son support. Sur $[1, +\infty[$, la fonction $x \mapsto 1 - \frac{1}{x^2}$ est dérivable, de dérivée $\frac{2}{x^3}$, strictement positive : elle est strictement croissante. Elle vaut 0 en $x = 1$ et tend vers 1 en $+\infty$. Elle réalise donc une bijection strictement croissante de $[1, +\infty[$ sur $[0, 1[$.

Inversion. Soit $u \in]0, 1[$. On résout $F(x) = u$ d'inconnue $x \geq 1$:

$$1 - \frac{1}{x^2} = u \iff \frac{1}{x^2} = 1 - u \iff x^2 = \frac{1}{1 - u}$$

Comme $u < 1$, le nombre $1 - u$ est strictement positif et le quotient est bien défini. Comme de plus $x \geq 1 > 0$, on prend la racine positive :

$$x = \frac{1}{\sqrt{1 - u}}$$

La question **3.** ne s'applique pas telle quelle, puisque la fonction F n'est pas strictement croissante sur $\mathbb{R}$ tout entier : elle est nulle sur $]-\infty, 1[$. Mais sa restriction au support $[1, +\infty[$ est bien une bijection strictement croissante sur $[0, 1[$, et c'est la seule chose qui a servi dans la démonstration. La méthode d'inversion donne donc une première réponse : la variable

$$Y = \frac{1}{\sqrt{1 - U}}$$

a pour fonction de répartition F .

Simplification. La question **1.** a établi que $1 - U$ suit la même loi que U . Par conséquent, la variable $\frac{1}{\sqrt{U}}$ a la même loi que $\frac{1}{\sqrt{1 - U}}$, et l'on peut retenir la formule plus simple

$$X = \frac{1}{\sqrt{U}}$$

Vérification directe. Contrôlons ce résultat sans invoquer la question **3.**, ce qui est toujours prudent. Comme $U \in]0, 1[$, on a $\sqrt{U} \in]0, 1[$ et donc $\frac{1}{\sqrt{U}} > 1$: le support est correct. Soit $x \geq 1$. La fonction inverse étant **strictement décroissante** sur $]0, +\infty[$, l'inégalité se retourne :

$$\begin{aligned} \left(\frac{1}{\sqrt{U}} \leq x \right) &= \left(\sqrt{U} \geq \frac{1}{x} \right) \\ &= \left(U \geq \frac{1}{x^2} \right) \end{aligned}$$

le second passage utilisant la stricte croissance de la fonction carré sur les positifs. Comme $x \geq 1$, on a $\frac{1}{x^2} \in]0, 1[$, donc

$$P\left(U \geq \frac{1}{x^2}\right) = 1 - F_U\left(\frac{1}{x^2}\right) = 1 - \frac{1}{x^2}$$

et pour $x < 1$ la probabilité est nulle. On retrouve exactement F . $\square$

5. On cherche une variable de loi $\mathcal{U}([a, b])$, dont la fonction de répartition est $G(x) = \frac{x-a}{b-a}$ pour $x \in [a, b]$.

Inversion. Soit $u \in]0, 1[$. On résout $G(x) = u$:

$$\frac{x-a}{b-a} = u \iff x-a = u(b-a) \iff x = a + (b-a)u$$

la multiplication par $b-a > 0$ étant licite. La méthode d'inversion donne donc

$$X = a + (b-a)U$$

Vérification directe. Il s'agit d'une transformation affine de coefficient $b-a > 0$. Soit $x \in \mathbb{R}$. Comme $b-a > 0$, la division conserve le sens de l'inégalité :

$$F_X(x) = P(a + (b-a)U \leq x) = P\left(U \leq \frac{x-a}{b-a}\right) = F_U\left(\frac{x-a}{b-a}\right)$$

Si $x < a$, le quotient est négatif et $F_X(x) = 0$. Si $x > b$, le quotient dépasse 1 et $F_X(x) = 1$. Si $a \leq x \leq b$, le quotient appartient à $[0, 1]$ et

$$F_X(x) = \frac{x-a}{b-a}$$

C'est bien la fonction de répartition de la loi $\mathcal{U}([a, b])$, donc

$$a + (b-a)U \hookrightarrow \mathcal{U}([a, b])$$

□

On vérifie la cohérence sur les cas extrêmes : quand U est proche de 0, X est proche de a ; quand U est proche de 1, X est proche de b . La formule dilate l'intervalle $]0, 1[$ d'un facteur $b-a$, puis le translate de a . On contrôle aussi par l'espérance : $E(X) = a + (b-a)E(U) = a + \frac{b-a}{2} = \frac{a+b}{2}$, et par la variance, $V(X) = (b-a)^2 V(U) = \frac{(b-a)^2}{12}$, deux valeurs conformes au cours.

Méthode à retenir. Simuler une loi de fonction de répartition F à partir d'une uniforme se fait en deux gestes : on pose l'équation $F(x) = u$, on la résout en x , et la solution $x = F^{-1}(u)$ fournit directement la formule $X = F^{-1}(U)$. Dans la rédaction, tout le soin porte sur les **inégalités** : multiplier par un nombre négatif ou composer par une fonction décroissante (inverse, opposé) retourne le sens, tandis que composer par une fonction strictement croissante (exponentielle, logarithme, carré sur les positifs, racine) le conserve. Un dernier contrôle coûte peu et rapporte gros : la fonction de répartition obtenue doit être croissante, nulle avant le support et de limite 1 en $+\infty$.

Erreur classique. Écrire $X = \frac{\ln U}{\lambda}$ en oubliant le signe. Comme $U < 1$, le logarithme est négatif, la variable obtenue serait négative, ce qui est impossible pour une loi exponentielle : le contrôle du signe du support suffit à détecter la faute avant même de calculer. Seconde erreur, à la question 3. : croire que $F^{-1}(U)$ aurait pour fonction de répartition F^{-1} . C'est le contraire, et le retenir sous forme d'énoncé permet de ne plus se tromper : **appliquer la réciproque de F à une uniforme fabrique la loi de fonction de répartition F .**

Exercice 19 ★★ Somme de deux variables uniformes indépendantes

Notions : Somme de variables indépendantes : produit de convolution et stabilité, Indépendance de deux variables aléatoires réelles, lemme des coalitions, Loi uniforme sur un segment

Les variables aléatoires X et Y sont définies sur un même espace probabilisé, indépendantes, et suivent toutes deux la loi uniforme sur $[0, 1]$: $X \hookrightarrow \mathcal{U}([0, 1])$ et $Y \hookrightarrow \mathcal{U}([0, 1])$. On pose $S = X + Y$.

1. Déterminer $S(\Omega)$. Sans effectuer aucun calcul de densité, expliquer pourquoi S ne suit pas la loi uniforme sur $S(\Omega)$.
2. À l'aide du produit de convolution, déterminer une densité f_S de S . On détaillera très soigneusement la recherche du domaine d'intégration.
3. Vérifier que la fonction obtenue est bien une densité de probabilité.
4. Calculer $E(S)$ et $V(S)$ de deux façons : d'abord par les propriétés de l'espérance et de la variance, puis directement à partir de f_S .
5. Calculer $P(S \leq \frac{1}{2})$ et $P(S \geq \frac{3}{2})$.

Corrigé

1. Les variables X et Y prennent leurs valeurs dans $[0, 1]$, donc leur somme prend ses valeurs dans $[0, 2]$:

$$S(\Omega) = [0, 2]$$

Pourquoi S n'est pas uniforme. Une valeur de S proche de 0 est doublement exigeante : il faut à la fois que X soit proche de 0 et que Y le soit, puisque les deux variables sont positives. Il en va de même pour une valeur proche de 2, qui réclame X et Y toutes deux proches de 1. En revanche, une valeur de S proche de 1 s'obtient d'une infinité de manières : X peut valoir n'importe quoi dans $[0, 1]$, il suffit que Y soit proche de $1 - X$. Les valeurs centrales sont donc bien plus faciles à atteindre que les valeurs extrêmes, ce qui interdit à S d'être uniforme, puisqu'une loi uniforme donne le même poids à deux intervalles de même longueur.

Cette intuition se transforme en preuve en deux lignes. Comme $X \geq 0$ et $Y \geq 0$, l'événement $(S \leq \frac{1}{10})$ est inclus dans $(X \leq \frac{1}{10}) \cap (Y \leq \frac{1}{10})$, donc par indépendance

$$P(S \leq \frac{1}{10}) \leq P(X \leq \frac{1}{10}) P(Y \leq \frac{1}{10}) = \frac{1}{10} \times \frac{1}{10} = \frac{1}{100}$$

Or si S suivait la loi $\mathcal{U}([0, 2])$, on aurait $P(S \leq \frac{1}{10}) = \frac{1/10}{2} = \frac{1}{20}$. Comme $\frac{1}{100} < \frac{1}{20}$, la loi de S n'est pas $\mathcal{U}([0, 2])$. $\square$

2. Une densité de X et de Y est la fonction f définie par $f(u) = 1$ si $u \in [0, 1]$ et $f(u) = 0$ sinon. Les variables X et Y étant indépendantes et à densité, une densité de leur somme est donnée par le produit de convolution (le programme dispense de justifier la légitimité de cette formule) :

$$f_S(s) = \int_{-\infty}^{+\infty} f(t) f(s-t) dt$$

Recherche du domaine d'intégration. Tout repose sur cette étape. La fonction intégrée, $t \mapsto f(t)f(s-t)$, ne prend que les valeurs 0 et 1. Elle vaut 1 exactement lorsque les **deux** conditions suivantes sont réalisées :

$$0 \leq t \leq 1 \quad \text{et} \quad 0 \leq s-t \leq 1$$

La première condition porte sur t et se lit directement : $t \in [0, 1]$. La seconde doit être traduite en une condition sur t , et c'est là que se joue tout l'exercice. De $0 \leq s-t$ on tire $t \leq s$; de $s-t \leq 1$ on tire $t \geq s-1$. La seconde condition équivaut donc à $t \in [s-1, s]$.

L'ensemble des t qui contribuent à l'intégrale est ainsi

$$[0, 1] \cap [s - 1, s] = [\max(0, s - 1), \min(1, s)]$$

et l'intégrale, intégrale de la fonction constante 1 sur cet intervalle, vaut simplement sa **longueur** (elle vaut 0 si l'intervalle est vide). Il reste à discuter selon la position de s .

Premier cas : $s < 0$. Alors $\min(1, s) = s < 0 = \max(0, s - 1)$: l'intervalle est vide, donc $f_S(s) = 0$. C'est cohérent avec $S \geq 0$.

Deuxième cas : $0 \leq s \leq 1$. Alors $s - 1 \leq 0$, donc $\max(0, s - 1) = 0$, et $s \leq 1$, donc $\min(1, s) = s$. Le domaine est $[0, s]$, de longueur s :

$$f_S(s) = \int_0^s 1 \, dt = s$$

Troisième cas : $1 < s \leq 2$. Alors $s - 1 > 0$, donc $\max(0, s - 1) = s - 1$, et $s > 1$, donc $\min(1, s) = 1$. Le domaine est $[s - 1, 1]$, de longueur $1 - (s - 1) = 2 - s$:

$$f_S(s) = \int_{s-1}^1 1 \, dt = 2 - s$$

Quatrième cas : $s > 2$. Alors $\max(0, s - 1) = s - 1 > 1 = \min(1, s)$: l'intervalle est vide et $f_S(s) = 0$.

En résumé, une densité de S est la fonction f_S définie par

$$f_S(s) = s \text{ sur } [0, 1], \quad f_S(s) = 2 - s \text{ sur }]1, 2], \quad f_S(s) = 0 \text{ ailleurs}$$

Son graphe est un triangle isocèle de base $[0, 2]$ et de hauteur 1, de sommet au point d'abscisse 1. On dit d'ailleurs que S suit une **loi triangulaire**. La forme du graphe confirme exactement l'intuition de la question 1. : les valeurs proches de 1 sont les plus probables, celles proches de 0 et de 2 les plus rares.

3. Trois points sont à vérifier.

Positivité. Sur $[0, 1]$, $f_S(s) = s \geq 0$. Sur $]1, 2]$, $f_S(s) = 2 - s \geq 0$ car $s \leq 2$. Ailleurs f_S est nulle. Donc $f_S \geq 0$ sur $\mathbb{R}$.

Continuité. La fonction f_S est affine par morceaux, donc continue sur chacun des quatre intervalles de la définition. Aux points de raccord, les valeurs coïncident : en 0, $0 = 0$; en 1, $s = 1$ et $2 - s = 1$; en 2, $2 - s = 0$. La fonction f_S est donc continue sur $\mathbb{R}$ tout entier, ce qui est plus que suffisant (une densité doit être continue sauf éventuellement en un nombre fini de points).

Intégrale égale à 1. La fonction étant nulle hors de $[0, 2]$, l'intégrale $\int_{-\infty}^{+\infty} f_S(s) \, ds$ converge et se réduit à une intégrale sur un segment :

$$\begin{aligned} \int_{-\infty}^{+\infty} f_S(s) \, ds &= \int_0^1 s \, ds + \int_1^2 (2 - s) \, ds \\ &= \left[\frac{s^2}{2} \right]_0^1 + \left[2s - \frac{s^2}{2} \right]_1^2 \\ &= \frac{1}{2} + \left((4 - 2) - \left(2 - \frac{1}{2} \right) \right) \\ &= \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

On peut aussi conclure sans calcul : l'intégrale est l'aire du triangle, soit $\frac{1}{2} \times 2 \times 1 = 1$. La fonction f_S est bien une densité de probabilité. $\square$

4. Première méthode : par les propriétés générales. Pour $X \hookrightarrow \mathcal{U}([0, 1])$, le cours donne

$$E(X) = \frac{0 + 1}{2} = \frac{1}{2} \quad \text{et} \quad V(X) = \frac{(1 - 0)^2}{12} = \frac{1}{12}$$

et il en va de même pour Y . Ces deux espérances existent, donc par linéarité

$$E(S) = E(X) + E(Y) = \frac{1}{2} + \frac{1}{2} = 1$$

Pour la variance, la formule générale s'écrit $V(X + Y) = V(X) + V(Y) + 2 \operatorname{Cov}(X, Y)$. Comme X et Y sont **indépendantes**, leur covariance est nulle, et seulement pour cette raison :

$$V(S) = V(X) + V(Y) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$$

Seconde méthode : par la densité. La fonction f_S étant nulle hors du segment $[0, 2]$, les intégrales à considérer sont des intégrales sur un segment de fonctions continues : elles convergent, et convergent absolument puisque les fonctions intégrées sont positives. L'espérance et le moment d'ordre 2 existent donc.

$$\begin{aligned} E(S) &= \int_0^1 s \times s \, ds + \int_1^2 s(2-s) \, ds \\ &= \left[\frac{s^3}{3} \right]_0^1 + \left[s^2 - \frac{s^3}{3} \right]_1^2 \\ &= \frac{1}{3} + \left(\left(4 - \frac{8}{3} \right) - \left(1 - \frac{1}{3} \right) \right) \\ &= \frac{1}{3} + \frac{4}{3} - \frac{2}{3} = 1 \end{aligned}$$

Le théorème de transfert appliqué à $s \mapsto s^2$ donne ensuite

$$\begin{aligned} E(S^2) &= \int_0^1 s^2 \times s \, ds + \int_1^2 s^2(2-s) \, ds \\ &= \left[\frac{s^4}{4} \right]_0^1 + \left[\frac{2s^3}{3} - \frac{s^4}{4} \right]_1^2 \\ &= \frac{1}{4} + \left(\left(\frac{16}{3} - 4 \right) - \left(\frac{2}{3} - \frac{1}{4} \right) \right) \\ &= \frac{1}{4} + \frac{4}{3} - \frac{2}{3} + \frac{1}{4} = \frac{1}{2} + \frac{2}{3} = \frac{7}{6} \end{aligned}$$

et la formule de König-Huygens conclut :

$$V(S) = E(S^2) - (E(S))^2 = \frac{7}{6} - 1 = \frac{1}{6}$$

Les deux méthodes concordent. La première est immédiate et doit toujours être privilégiée quand on cherche seulement $E(S)$ et $V(S)$; la seconde ne sert qu'à contrôler la densité obtenue en question 2., ce qu'elle fait très bien ici.

5. Comme $\frac{1}{2} \in [0, 1]$, l'événement $(S \leq \frac{1}{2})$ ne fait intervenir que le premier morceau de la densité :

$$P\left(S \leq \frac{1}{2}\right) = \int_0^{1/2} s \, ds = \left[\frac{s^2}{2} \right]_0^{1/2} = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$$

De même, comme $\frac{3}{2} \in [1, 2]$, seul le second morceau intervient :

$$P\left(S \geq \frac{3}{2}\right) = \int_{3/2}^2 (2-s) \, ds = \left[2s - \frac{s^2}{2} \right]_{3/2}^2 = (4-2) - \left(3 - \frac{9}{8} \right) = 2 - \frac{15}{8} = \frac{1}{8}$$

Les deux résultats sont égaux, ce qui n'a rien d'un hasard : le graphe de f_S est symétrique par rapport à la droite d'équation $s = 1$, donc les deux événements ont des probabilités égales. On

peut aussi voir les deux nombres comme des aires de triangles rectangles isocèles de côté $\frac{1}{2}$, soit $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$.

Enfin, ce calcul chiffre l'écart avec la loi uniforme : sous $\mathcal{U}([0, 2])$ on aurait trouvé $P(S \leq \frac{1}{2}) = \frac{1}{4}$, soit le double. La somme concentre bien la masse au centre.

Méthode à retenir. Pour la somme de deux variables à densité **indépendantes**, on écrit le produit de convolution $f_S(s) = \int_{-\infty}^{+\infty} f_X(t)f_Y(s-t) dt$, puis on consacre l'essentiel du travail à la détermination du domaine d'intégration. La règle est toujours la même : écrire les deux systèmes de contraintes, l'un venant de $f_X(t)$, l'autre de $f_Y(s-t)$, traduire le second en un encadrement de t , intersecter, et discuter selon s . Une fois le domaine trouvé, le calcul de l'intégrale est en général très court.

Erreur classique. Ne conserver que la contrainte $0 \leq t \leq 1$ et oublier celle qui vient de $f_Y(s-t)$. On obtient alors $f_S(s) = 1$ pour tout s , ce qui n'est même pas une densité. Deuxième erreur, de nature différente : croire que la somme de deux variables uniformes indépendantes est uniforme. Aucune stabilité de ce type n'existe pour la loi uniforme, contrairement aux lois normales, gamma, de Poisson et binomiales de même paramètre p .

Exercice 20 ★★ Stabilité de la loi normale par somme

Notions : Somme de variables indépendantes : produit de convolution et stabilité, Loi normale, centrage-réduction et lecture de la table, Indépendance de deux variables aléatoires réelles, lemme des coalitions, Covariance et variance d'une somme dans le cas général

Les variables aléatoires X et Y sont indépendantes, avec $X \hookrightarrow \mathcal{N}(3, 4)$ et $Y \hookrightarrow \mathcal{N}(-1, 9)$.

On rappelle que Φ désigne la fonction de répartition de la loi $\mathcal{N}(0, 1)$, et on donne les valeurs suivantes.

u	0,83	1	1,11	1,5
$\Phi(u)$	0,7967	0,8413	0,8665	0,9332

- Déterminer la loi de $X + Y$.
- Déterminer la loi de $2X - Y$, en détaillant le calcul de la variance et en indiquant précisément où intervient l'hypothèse d'indépendance. En déduire $P(2X - Y \leq 14,5)$.
- Calculer $P(X + Y \leq 5)$.
- Calculer $P(X > Y)$ en introduisant la variable $D = X - Y$.
- Dans cette question, et dans cette seule question, on remplace la seconde variable par X elle-même : on pose $T = X + X$. Déterminer la loi de T , puis $P(T \leq 10)$. Que donnerait la formule « la variance de la somme est la somme des variances » appliquée à T , et pourquoi est-elle en défaut ici ?

Corrigé

1. Le théorème de stabilité de la loi normale, admis par le programme, s'énonce ainsi : si $U \hookrightarrow \mathcal{N}(m, \sigma^2)$ et $V \hookrightarrow \mathcal{N}(m', \sigma'^2)$ sont **indépendantes**, alors

$$U + V \hookrightarrow \mathcal{N}(m + m', \sigma^2 + \sigma'^2)$$

Autrement dit : les espérances s'ajoutent, les **variances** s'ajoutent (surtout pas les écarts-types). Ici X et Y sont indépendantes, d'espérances 3 et -1 , de variances 4 et 9. Donc

$$X + Y \hookrightarrow \mathcal{N}(2, 13)$$

L'écart-type de $X + Y$ vaut $\sqrt{13} \approx 3,606$. On notera qu'il est strictement inférieur à $2 + 3 = 5$, somme des écarts-types : ce sont bien les variances qui s'ajoutent.

2. On procède en deux temps.

Étape 1 : les transformées affines. Le cours donne la loi d'une transformée affine d'une variable normale. Comme $X \hookrightarrow \mathcal{N}(3, 4)$,

$$2X \hookrightarrow \mathcal{N}(2 \times 3, 2^2 \times 4) = \mathcal{N}(6, 16)$$

et comme $Y \hookrightarrow \mathcal{N}(-1, 9)$,

$$-Y \hookrightarrow \mathcal{N}((-1) \times (-1), (-1)^2 \times 9) = \mathcal{N}(1, 9)$$

Étape 2 : la somme. Les variables $2X$ et $-Y$ sont des fonctions de X et de Y respectivement ; comme X et Y sont indépendantes, $2X$ et $-Y$ le sont aussi (c'est le lemme des coalitions, admis, appliqué aux deux coalitions $\{X\}$ et $\{Y\}$). Le théorème de stabilité s'applique donc à $2X + (-Y) = 2X - Y$:

$$2X - Y \hookrightarrow \mathcal{N}(7, 25)$$

Le détail du calcul de la variance. La formule générale de la variance d'une combinaison s'écrit

$$V(aX + bY) = a^2V(X) + b^2V(Y) + 2ab \operatorname{Cov}(X, Y)$$

Avec $a = 2$ et $b = -1$:

$$V(2X - Y) = 4 \times 4 + 1 \times 9 + 2 \times 2 \times (-1) \times \operatorname{Cov}(X, Y) = 16 + 9 - 4 \operatorname{Cov}(X, Y)$$

C'est ici, et seulement ici, que sert l'indépendance : elle entraîne $\operatorname{Cov}(X, Y) = 0$, et le terme rectangle disparaît. Il reste $V(2X - Y) = 25$, d'écart-type 5. On observe au passage que le signe $-$ de $2X - Y$ ne diminue pas la variance : les carrés a^2 et b^2 sont positifs, et le terme croisé est nul. Une différence de variables indépendantes est **plus** dispersée que chacune d'elles, jamais moins.

Calcul de la probabilité. On centre et on réduit : si $W = 2X - Y \hookrightarrow \mathcal{N}(7, 25)$, alors $\frac{W-7}{5} \hookrightarrow \mathcal{N}(0, 1)$. Donc

$$P(W \leq 14,5) = P\left(\frac{W-7}{5} \leq \frac{14,5-7}{5}\right) = \Phi\left(\frac{7,5}{5}\right) = \Phi(1,5) = 0,9332$$

3. D'après la question 1., $S = X + Y \hookrightarrow \mathcal{N}(2, 13)$, donc $\frac{S-2}{\sqrt{13}} \hookrightarrow \mathcal{N}(0, 1)$ et

$$P(S \leq 5) = P\left(\frac{S-2}{\sqrt{13}} \leq \frac{5-2}{\sqrt{13}}\right) = \Phi\left(\frac{3}{\sqrt{13}}\right)$$

Or $\sqrt{13} \approx 3,6056$, donc $\frac{3}{\sqrt{13}} \approx 0,8321$, que la table arrondit à 0,83. D'où

$$P(X + Y \leq 5) \approx \Phi(0,83) = 0,7967$$

Il y a donc à peu près 80 % de chances que la somme ne dépasse pas 5.

4. L'événement $(X > Y)$ s'écrit $(X - Y > 0)$: on est ramené à une seule variable, ce qui est tout l'intérêt de la manoeuvre. Posons $D = X - Y$. Comme en question 2., $-Y \hookrightarrow \mathcal{N}(1, 9)$ est indépendante de X , donc par stabilité

$$D \hookrightarrow \mathcal{N}(3 + 1, 4 + 9) = \mathcal{N}(4, 13)$$

On centre et on réduit, en posant $Z = \frac{D-4}{\sqrt{13}} \hookrightarrow \mathcal{N}(0, 1)$:

$$\begin{aligned} P(X > Y) &= P(D > 0) = P\left(Z > \frac{0-4}{\sqrt{13}}\right) \\ &= P\left(Z > -\frac{4}{\sqrt{13}}\right) \\ &= 1 - \Phi\left(-\frac{4}{\sqrt{13}}\right) \\ &= \Phi\left(\frac{4}{\sqrt{13}}\right) \end{aligned}$$

la dernière égalité venant de la relation $\Phi(-u) = 1 - \Phi(u)$, conséquence de la parité de la densité φ . Comme $\frac{4}{\sqrt{13}} \approx 1,1094$, arrondi à 1,11 :

$$P(X > Y) \approx \Phi(1,11) = 0,8665$$

Ce résultat est plausible : X a une espérance de 3 et Y une espérance de -1 , donc X dépasse Y dans la grande majorité des cas, sans que ce soit certain pour autant.

5. Ici la seconde variable n'est plus indépendante de la première : c'est la première. On a $T = X + X = 2X$, et la loi d'une transformée affine donne directement

$$T = 2X \hookrightarrow \mathcal{N}(6, 16)$$

d'écart-type 4. Par centrage-réduction,

$$P(T \leq 10) = P\left(\frac{T-6}{4} \leq \frac{10-6}{4}\right) = \Phi(1) = 0,8413$$

La formule fautive. Si l'on appliquait aveuglément « la variance d'une somme est la somme des variances », on écrirait

$$V(T) \llcorner = \llcorner V(X) + V(X) = 4 + 4 = 8$$

et on conclurait à tort que $T \hookrightarrow \mathcal{N}(6, 8)$. L'espérance, elle, serait juste : la linéarité de l'espérance ne demande **aucune** hypothèse d'indépendance. C'est bien la variance, et elle seule, qui est en défaut.

Pourquoi. La formule exacte est toujours

$$V(X + Y) = V(X) + V(Y) + 2 \operatorname{Cov}(X, Y)$$

Le terme $2 \operatorname{Cov}(X, Y)$ ne s'annule que si la covariance est nulle, ce que l'indépendance garantit. Ici la seconde variable est X elle-même, et $\operatorname{Cov}(X, X) = V(X) = 4$: c'est la covariance maximale possible, celle d'une variable avec elle-même. La formule donne alors

$$V(T) = 4 + 4 + 2 \times 4 = 16$$

ce qui est bien la valeur obtenue par la transformation affine, $V(2X) = 2^2 V(X) = 16$. Les deux calculs corrects concordent, et l'écart avec la valeur fautive 8 est considérable : un facteur 2 sur la variance, donc un facteur $\sqrt{2}$ sur l'écart-type.

L'erreur se voit aussi sur la probabilité. Avec la variance fautive 8, l'écart-type serait $2\sqrt{2} \approx 2,83$, et on évaluerait $P(T \leq 10)$ en $\Phi\left(\frac{4}{2\sqrt{2}}\right) = \Phi(\sqrt{2})$, c'est-à-dire en un point strictement plus grand que 1. Comme Φ est strictement croissante, on obtiendrait une probabilité strictement supérieure à la vraie valeur 0,8413.

Méthode à retenir. Pour une combinaison $aX + bY$ de variables normales **indépendantes** : la loi reste normale (stabilité), l'espérance se calcule par linéarité, et la variance par $a^2 V(X) + b^2 V(Y)$. Ensuite, toute probabilité s'obtient par centrage-réduction, en se ramenant à Φ et en utilisant $\Phi(-u) = 1 - \Phi(u)$ pour les arguments négatifs. Une inégalité entre deux variables, du type $(X > Y)$, se traite toujours en introduisant la différence $D = X - Y$: on passe d'un couple à une seule variable, et le problème devient un calcul de table.

Erreur classique. Ajouter les écarts-types au lieu des variances : ici on écrirait $\sigma = 2 + 3 = 5$ au lieu de $\sqrt{13} \approx 3,61$. Autre erreur, plus grave car silencieuse : appliquer la stabilité sans vérifier l'indépendance. La question **5.** montre ce qu'il en coûte. Enfin, pour $2X - Y$, ne pas oublier d'élever 2 au carré : la variance est multipliée par 4, pas par 2.

Exercice 21 ★★ Covariance d'une variable et de son carré

Notions : Covariance et variance d'une somme dans le cas général, Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Loi uniforme sur un segment

Partie A

La variable aléatoire X suit la loi $\mathcal{U}([0, 1])$ et on pose $Y = X^2$.

1. Calculer $E(X)$, $E(X^2)$, $E(X^3)$ et $E(X^4)$.
2. En déduire $\text{Cov}(X, Y)$.
3. Calculer $V(X)$ et $V(Y)$, puis le coefficient de corrélation linéaire $\rho(X, Y)$. Interpréter la valeur obtenue.
4. Calculer $V(X + Y)$ à l'aide de la formule générale de la variance d'une somme, et contrôler le résultat par un calcul direct.

Partie B

La variable aléatoire Z suit la loi $\mathcal{U}([-1, 1])$ et on pose $W = Z^2$.

5. Montrer que $\text{Cov}(Z, W) = 0$.
6. Montrer que Z et W ne sont pourtant pas indépendantes.
7. Quelle conclusion générale faut-il retenir des questions 5. et 6.? Qu'en est-il de l'implication réciproque?

Corrigé**Partie A**

1. Une densité de X est la fonction f valant 1 sur $[0, 1]$ et 0 ailleurs. Pour tout entier $k \geq 1$, le théorème de transfert appliqué à la fonction continue $x \mapsto x^k$ donne

$$E(X^k) = \int_{-\infty}^{+\infty} x^k f(x) dx = \int_0^1 x^k dx$$

Cette intégrale est celle d'une fonction continue sur un **segment** : elle converge, et la convergence est absolue puisque la fonction intégrée est positive sur $[0, 1]$. Tous les moments de X existent donc, sans aucune discussion, et

$$E(X^k) = \left[\frac{x^{k+1}}{k+1} \right]_0^1 = \frac{1}{k+1}$$

a. $E(X) = \frac{1}{2}$

c. $E(X^3) = \frac{1}{4}$

b. $E(X^2) = \frac{1}{3}$

d. $E(X^4) = \frac{1}{5}$

La valeur $E(X) = \frac{1}{2}$ est bien celle du cours pour la loi $\mathcal{U}([0, 1])$, à savoir $\frac{a+b}{2}$.

2. La formule de König-Huygens pour la covariance s'écrit

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Le point à ne pas manquer est que le produit XY se simplifie : $XY = X \times X^2 = X^3$. Aucune loi conjointe n'est donc nécessaire, tout se ramène à des moments de X :

$$E(XY) = E(X^3) = \frac{1}{4} \quad \text{et} \quad E(Y) = E(X^2) = \frac{1}{3}$$

D'où

$$\text{Cov}(X, Y) = \frac{1}{4} - \frac{1}{2} \times \frac{1}{3} = \frac{1}{4} - \frac{1}{6} = \frac{3}{12} - \frac{2}{12} = \frac{1}{12}$$

La covariance est strictement positive, ce qui était prévisible : sur $[0, 1]$ la fonction $x \mapsto x^2$ est croissante, donc X et Y varient dans le même sens.

3. Par König-Huygens appliquée à la variance,

$$V(X) = E(X^2) - (E(X))^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

ce qui est bien la valeur du cours $\frac{(b-a)^2}{12} = \frac{1}{12}$. Pour Y , on utilise $Y^2 = X^4$:

$$V(Y) = E(Y^2) - (E(Y))^2 = E(X^4) - (E(X^2))^2 = \frac{1}{5} - \frac{1}{9} = \frac{9-5}{45} = \frac{4}{45}$$

Ces deux variances sont strictement positives, donc les écarts-types ne sont pas nuls et le coefficient de corrélation linéaire est bien défini :

$$\sigma(X) = \sqrt{\frac{1}{12}} = \frac{1}{2\sqrt{3}} \quad \text{et} \quad \sigma(Y) = \sqrt{\frac{4}{45}} = \frac{2}{3\sqrt{5}}$$

Le produit des écarts-types vaut

$$\sigma(X) \sigma(Y) = \frac{1}{2\sqrt{3}} \times \frac{2}{3\sqrt{5}} = \frac{2}{6\sqrt{15}} = \frac{1}{3\sqrt{15}}$$

d'où

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X) \sigma(Y)} = \frac{1/12}{1/(3\sqrt{15})} = \frac{3\sqrt{15}}{12} = \frac{\sqrt{15}}{4} \approx 0,968$$

On vérifie au passage que $|\rho(X, Y)| \leq 1$, comme l'exige le cours.

Interprétation. La valeur est très proche de 1, ce qui traduit une liaison croissante extrêmement forte. Et pour cause : la connaissance de X détermine **complètement** Y , puisque $Y = X^2$. Il n'y a aucun aléa résiduel. Pourtant $\rho(X, Y) \neq 1$, et ce n'est pas une imprécision de calcul. Le coefficient de corrélation ne mesure pas la dépendance en général, il mesure uniquement la qualité d'un ajustement **affine**. Or la relation entre X et Y est parabolique, pas affine : sur $[0, 1]$ la parabole est presque une droite, d'où ρ très proche de 1, mais pas tout à fait. Le cas d'égalité $|\rho| = 1$ est réservé aux relations de la forme $Y = aX + b$.

4. La formule générale de la variance d'une somme donne

$$\begin{aligned} V(X + Y) &= V(X) + V(Y) + 2\text{Cov}(X, Y) \\ &= \frac{1}{12} + \frac{4}{45} + 2 \times \frac{1}{12} \\ &= \frac{1}{12} + \frac{4}{45} + \frac{1}{6} \end{aligned}$$

On réduit au dénominateur commun 180 :

$$V(X + Y) = \frac{15}{180} + \frac{16}{180} + \frac{30}{180} = \frac{61}{180} \approx 0,339$$

Contrôle par le calcul direct. Posons $S = X + Y = X + X^2$. Par linéarité de l'espérance,

$$E(S) = E(X) + E(X^2) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$$

Par ailleurs $S^2 = (X + X^2)^2 = X^2 + 2X^3 + X^4$, donc par linéarité à nouveau

$$E(S^2) = E(X^2) + 2E(X^3) + E(X^4) = \frac{1}{3} + \frac{2}{4} + \frac{1}{5} = \frac{10 + 15 + 6}{30} = \frac{31}{30}$$

et donc

$$V(S) = \frac{31}{30} - \left(\frac{5}{6}\right)^2 = \frac{31}{30} - \frac{25}{36} = \frac{186}{180} - \frac{125}{180} = \frac{61}{180}$$

Les deux méthodes donnent la même valeur, ce qui valide en particulier la covariance trouvée en question 2. On remarquera que $V(X) + V(Y) = \frac{1}{12} + \frac{4}{45} = \frac{31}{180}$ seulement : oublier le terme de covariance reviendrait ici à diviser la variance par presque deux.

Partie B

5. Une densité de $Z \hookrightarrow \mathcal{U}([-1, 1])$ est la fonction g valant $\frac{1}{2}$ sur $[-1, 1]$ et 0 ailleurs. Le théorème de transfert donne, pour tout entier $k \geq 1$,

$$E(Z^k) = \int_{-1}^1 \frac{z^k}{2} dz$$

intégrale d'une fonction continue sur un segment, donc convergente. En particulier

$$E(Z) = \int_{-1}^1 \frac{z}{2} dz = \left[\frac{z^2}{4} \right]_{-1}^1 = \frac{1}{4} - \frac{1}{4} = 0$$

$$E(Z^3) = \int_{-1}^1 \frac{z^3}{2} dz = \left[\frac{z^4}{8} \right]_{-1}^1 = \frac{1}{8} - \frac{1}{8} = 0$$

Ces deux nullités s'expliquent d'un mot : on intègre une fonction **impaire** sur un intervalle symétrique par rapport à l'origine. C'est le seul ressort de la question.

Comme $ZW = Z \times Z^2 = Z^3$, la formule de König-Huygens donne

$$\text{Cov}(Z, W) = E(ZW) - E(Z)E(W) = E(Z^3) - 0 \times E(W) = 0 - 0 = 0$$

La covariance de Z et de son carré est donc **nulle**. $\square$

Géométriquement, le nuage des points (z, z^2) pour $z \in [-1, 1]$ est une parabole symétrique par rapport à l'axe des ordonnées : elle ne présente aucune tendance croissante ni décroissante, la meilleure droite d'ajustement est horizontale. La covariance, qui ne détecte que les tendances affines, ne voit rien.

6. Les variables Z et W sont pourtant très fortement liées, puisque W est entièrement déterminée par Z . Voici deux rédactions possibles, la première étant la plus parlante.

Par deux événements qui n'en font qu'un. Considérons les deux événements

$$A = (|Z| \leq \frac{1}{2}) \quad \text{et} \quad B = (W \leq \frac{1}{4})$$

Le premier ne dépend que de Z , le second ne dépend que de W . Or, pour un réel z , l'inégalité $z^2 \leq \frac{1}{4}$ équivaut à $|z| \leq \frac{1}{2}$. Ces deux événements sont donc **le même événement**, d'où $A \cap B = A$. Par ailleurs

$$P(A) = P(-\frac{1}{2} \leq Z \leq \frac{1}{2}) = \int_{-1/2}^{1/2} \frac{1}{2} dz = \frac{1}{2} \quad \text{et} \quad P(B) = P(A) = \frac{1}{2}$$

Si Z et W étaient indépendantes, on aurait $P(A \cap B) = P(A)P(B) = \frac{1}{4}$. Or $P(A \cap B) = P(A) = \frac{1}{2} \neq \frac{1}{4}$. Les variables Z et W ne sont donc pas indépendantes. $\square$

Par la définition, avec des événements de la forme $(Z \leq z)$ et $(W \leq w)$. Prenons $z = \frac{1}{2}$ et $w = \frac{1}{4}$. Comme $(W \leq \frac{1}{4}) = (-\frac{1}{2} \leq Z \leq \frac{1}{2})$, l'intersection des deux événements est

$$(Z \leq \frac{1}{2}) \cap (W \leq \frac{1}{4}) = (-\frac{1}{2} \leq Z \leq \frac{1}{2})$$

de probabilité $\frac{1}{2}$. D'autre part

$$P(Z \leq \frac{1}{2}) = \int_{-1}^{1/2} \frac{1}{2} dz = \frac{1}{2} \times \frac{3}{2} = \frac{3}{4} \quad \text{et} \quad P(W \leq \frac{1}{4}) = \frac{1}{2}$$

Le produit vaut $\frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$, différent de $\frac{1}{2}$. La caractérisation de l'indépendance $P(Z \leq z, W \leq w) = P(Z \leq z)P(W \leq w)$ est donc mise en défaut pour ce couple de valeurs, ce qui suffit à conclure. $\square$

7. La conclusion à retenir tient en une phrase : **une covariance nulle n'entraîne pas l'indépendance**. Les variables Z et W de la partie B en fournissent l'exemple le plus économique : leur covariance est nulle, et pourtant chacune détermine entièrement l'autre à un signe près.

L'implication **réciroque**, elle, est vraie et figure au cours : si Z et W sont indépendantes et admettent chacune une variance, alors $E(ZW) = E(Z)E(W)$, donc $\text{Cov}(Z, W) = 0$. On dispose donc d'une flèche, et d'une seule :

$$\text{indépendance} \implies \text{covariance nulle}, \quad \text{mais la réciproque est fausse}$$

L'usage pratique de cette flèche est presque toujours par contraposée : une covariance **non nulle** prouve la non-indépendance. En revanche, une covariance nulle ne prouve rien, et il faut alors revenir à la définition pour trancher, comme en question 6.

Méthode à retenir. Dès que la seconde variable est une fonction de la première, du type $Y = g(X)$, le produit se simplifie en $XY = Xg(X)$: la covariance se calcule uniquement avec des moments de X , obtenus par transfert, sans jamais avoir besoin de la loi de Y . Pour prouver une **non-indépendance**, deux voies : montrer que la covariance est non nulle (le plus rapide quand ça marche), ou exhiber un couple de valeurs (z, w) pour lequel $P(Z \leq z, W \leq w) \neq P(Z \leq z)P(W \leq w)$ (la seule voie possible quand la covariance est nulle).

Erreur classique. Conclure de $\text{Cov}(Z, W) = 0$ que Z et W sont indépendantes, puis écrire $V(Z + W) = V(Z) + V(W)$ en invoquant l'indépendance. Cette dernière égalité est ici vraie, mais pour la bonne raison, qui est la nullité de la covariance, jamais l'indépendance. Autre confusion à éviter : croire qu'un coefficient de corrélation proche de 1 signifie « Y est presque égale à X ». En partie A, $\rho \approx 0,968$ alors que $Y = X^2$ vaut par exemple 0,01 quand X vaut 0,1.

Exercice 22 ★★ Minimum et maximum de deux durées de vie

Notions : Minimum et maximum d'un échantillon de variables indépendantes, Loi exponentielle et absence de mémoire, Indépendance de deux variables aléatoires réelles, lemme des coalitions

Deux composants électroniques fonctionnent côte à côte. Leurs durées de vie, exprimées en années, sont modélisées par deux variables aléatoires indépendantes X et Y , avec $X \hookrightarrow \mathcal{E}(\lambda)$ et $Y \hookrightarrow \mathcal{E}(\mu)$, où $\lambda > 0$ et $\mu > 0$. On pose

$$M = \min(X, Y) \quad \text{et} \quad N = \max(X, Y)$$

1. Montrer que pour tout réel t , $P(M > t) = P(X > t) P(Y > t)$. En déduire la loi de M .
2. Déterminer la fonction de répartition de N , puis une densité de N , et enfin montrer que

$$E(N) = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}$$

3. Justifier que $M + N = X + Y$, et vérifier la cohérence des résultats des questions 1. et 2.
4. Application numérique : $\lambda = \frac{1}{3}$ et $\mu = \frac{1}{6}$. Calculer $E(X)$, $E(Y)$, $E(M)$ et $E(N)$.
5. Les deux composants sont montés dans un même appareil. Interpréter M et N selon que le montage est en série ou en parallèle.

Corrigé

1. Soit t un réel. Le minimum de deux nombres dépasse t si et seulement si **chacun** des deux dépasse t . Autrement dit, en termes d'événements,

$$(M > t) = (X > t) \cap (Y > t)$$

Les variables X et Y étant indépendantes, les événements $(X > t)$ et $(Y > t)$ le sont également, d'où

$$P(M > t) = P(X > t) P(Y > t)$$

Loi de M . Distinguons deux cas selon le signe de t .

Cas $t < 0$. Une durée de vie exponentielle est positive, donc $P(X > t) = P(Y > t) = 1$ et $P(M > t) = 1$, c'est-à-dire $F_M(t) = 0$.

Cas $t \geq 0$. Le cours donne la fonction de répartition d'une loi exponentielle : $F_X(t) = 1 - e^{-\lambda t}$, donc $P(X > t) = e^{-\lambda t}$, et de même $P(Y > t) = e^{-\mu t}$. Par conséquent

$$P(M > t) = e^{-\lambda t} \times e^{-\mu t} = e^{-(\lambda + \mu)t}$$

d'où

$$F_M(t) = 1 - P(M > t) = 1 - e^{-(\lambda + \mu)t}$$

La fonction F_M est donc nulle sur $]-\infty, 0[$ et vaut $1 - e^{-(\lambda + \mu)t}$ sur $[0, +\infty[$: c'est exactement la fonction de répartition de la loi exponentielle de paramètre $\lambda + \mu$. Comme la loi d'une variable est caractérisée par sa fonction de répartition,

$$M \hookrightarrow \mathcal{E}(\lambda + \mu) \quad \square$$

Le résultat mérite d'être médité : le minimum de deux variables exponentielles indépendantes est **encore** exponentielle, et les paramètres s'ajoutent. En langage de fiabilité, les taux de défaillance s'additionnent. En particulier $E(M) = \frac{1}{\lambda + \mu}$, qui est plus petit que $\frac{1}{\lambda}$ et que $\frac{1}{\mu}$: le premier des deux composants à lâcher tombe en panne plus tôt que chacun des deux pris isolément, ce qui est rassurant pour le bon sens.

2. Le maximum de deux nombres est inférieur ou égal à t si et seulement si **chacun** des deux l'est :

$$(N \leq t) = (X \leq t) \cap (Y \leq t)$$

Par indépendance, $F_N(t) = F_X(t) F_Y(t)$. Ainsi $F_N(t) = 0$ pour $t < 0$, et pour $t \geq 0$

$$F_N(t) = (1 - e^{-\lambda t})(1 - e^{-\mu t}) = 1 - e^{-\lambda t} - e^{-\mu t} + e^{-(\lambda+\mu)t}$$

N est bien à densité. La fonction F_N est continue sur $]-\infty, 0[$ et sur $[0, +\infty[$, et les deux expressions se raccordent en 0, où elles valent toutes deux 0. Elle est donc continue sur $\mathbb{R}$. Elle est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement en 0, seul point de raccord. La variable N est donc à densité, et on obtient une densité en dérivant là où c'est possible.

Pour $t > 0$,

$$f_N(t) = \lambda e^{-\lambda t} + \mu e^{-\mu t} - (\lambda + \mu) e^{-(\lambda+\mu)t}$$

et $f_N(t) = 0$ pour $t < 0$; en 0, on pose par exemple $f_N(0) = 0$, le choix étant sans importance. Cette fonction est positive, comme dérivée de la fonction croissante F_N .

Cette écriture est très commode : la densité de N est la **somme des densités** de $\mathcal{E}(\lambda)$ et de $\mathcal{E}(\mu)$, **moins** celle de $\mathcal{E}(\lambda + \mu)$.

Calcul de $E(N)$. Par définition,

$$E(N) = \int_0^{+\infty} t f_N(t) dt = \int_0^{+\infty} t (\lambda e^{-\lambda t} + \mu e^{-\mu t} - (\lambda + \mu) e^{-(\lambda+\mu)t}) dt$$

Pour tout $\alpha > 0$, l'intégrale $\int_0^{+\infty} t \alpha e^{-\alpha t} dt$ converge et vaut $\frac{1}{\alpha}$: c'est exactement l'espérance de la loi $\mathcal{E}(\alpha)$, résultat du cours. Les trois intégrales obtenues en développant convergent donc, respectivement vers $\frac{1}{\lambda}$, $\frac{1}{\mu}$ et $\frac{1}{\lambda+\mu}$. Leur combinaison linéaire converge, ce qui prouve l'existence de $E(N)$; et comme la fonction intégrée est positive, cette convergence est absolue. Il vient

$$E(N) = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu} \quad \square$$

3. Pour une issue ω donnée, les deux nombres $M(\omega)$ et $N(\omega)$ sont les deux nombres $X(\omega)$ et $Y(\omega)$, simplement rangés dans l'ordre croissant. Le plus petit est appelé $M(\omega)$, le plus grand $N(\omega)$, mais l'ensemble des deux valeurs est le même. Leur somme est donc inchangée :

$$M + N = X + Y$$

et, accessoirement, $MN = XY$. Toutes ces variables admettant une espérance, la linéarité donne

$$E(M) + E(N) = E(X) + E(Y) = \frac{1}{\lambda} + \frac{1}{\mu}$$

Vérification. Avec les résultats des questions **1.** et **2.** :

$$E(M) + E(N) = \frac{1}{\lambda + \mu} + \left(\frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu} \right) = \frac{1}{\lambda} + \frac{1}{\mu}$$

Les termes en $\frac{1}{\lambda+\mu}$ se compensent exactement, et l'identité est vérifiée. Ce contrôle est précieux : il ne coûte rien et il valide d'un coup la loi du minimum et l'espérance du maximum. C'est d'ailleurs de cette identité que l'on peut retrouver $E(N)$ si on l'a oubliée, une fois connue la loi de M .

4. Avec $\lambda = \frac{1}{3}$ et $\mu = \frac{1}{6}$, les espérances des durées de vie individuelles sont

$$E(X) = \frac{1}{\lambda} = 3 \quad \text{et} \quad E(Y) = \frac{1}{\mu} = 6$$

soit trois ans pour le premier composant et six ans pour le second. Le paramètre du minimum vaut

$$\lambda + \mu = \frac{1}{3} + \frac{1}{6} = \frac{2}{6} + \frac{1}{6} = \frac{1}{2}$$

donc $M \hookrightarrow \mathcal{E}(\frac{1}{2})$ et

$$E(M) = \frac{1}{\lambda + \mu} = 2$$

Puis

$$E(N) = 3 + 6 - 2 = 7$$

Le contrôle de la question **3.** est immédiat : $E(M) + E(N) = 2 + 7 = 9 = 3 + 6 = E(X) + E(Y)$. Les ordres de grandeur sont éloquentes. La première panne survient en moyenne au bout de 2 ans, soit plus tôt que la durée de vie moyenne du plus fragile des deux composants (3 ans). La dernière panne survient en moyenne au bout de 7 ans, soit plus tard que la durée de vie moyenne du plus robuste (6 ans).

5. Tout dépend de la façon dont les deux composants sont montés.

Montage en série. Le courant traverse successivement les deux composants : l'appareil ne fonctionne que si les **deux** fonctionnent. Il tombe donc en panne dès la première défaillance, et sa durée de vie est $M = \min(X, Y)$. Ici, l'appareil dure en moyenne 2 ans, c'est-à-dire moins longtemps que n'importe lequel de ses composants. Multiplier les composants en série, c'est multiplier les points de rupture.

Montage en parallèle. Les deux composants sont redondants : l'appareil fonctionne tant qu'**au moins un** des deux fonctionne. Il ne tombe en panne qu'à la seconde défaillance, et sa durée de vie est $N = \max(X, Y)$. Ici, l'appareil dure en moyenne 7 ans.

Le contraste est saisissant : les mêmes composants, montés différemment, donnent une durée de vie moyenne de 2 ans ou de 7 ans, soit un rapport de 3,5. C'est tout l'intérêt de la redondance en ingénierie.

Une dernière remarque, mathématique celle-là. Le minimum reste une variable exponentielle, donc il conserve la propriété d'absence de mémoire : un système en série encore en marche après 5 ans a la même loi de durée de vie résiduelle qu'un système neuf. Le maximum, lui, n'est **pas** exponentiel, comme le montre l'expression de F_N : un système en parallèle qui a survécu 5 ans a probablement déjà perdu un de ses deux composants, il est donc bel et bien usé. La redondance ne se comporte pas comme un composant neuf.

Méthode à retenir. Pour un minimum, on passe systématiquement par la **fonction de survie** $P(M > t)$, car $(M > t)$ se factorise en une intersection. Pour un maximum, on passe par la **fonction de répartition** $F_N(t) = P(N \leq t)$, pour la même raison. Dans les deux cas, l'indépendance transforme la probabilité d'une intersection en produit, et il ne reste qu'à dériver pour obtenir la densité. Le résultat à connaître par coeur : le minimum de variables exponentielles indépendantes est exponentielle, de paramètre la somme des paramètres.

Erreur classique. Écrire $P(N \leq t) = P(X \leq t) + P(Y \leq t)$, ou traiter le maximum comme le minimum en factorisant $(N > t)$. L'événement $(N > t)$ signifie « au moins un des deux dépasse t » : c'est une **réunion**, pas une intersection, et sa probabilité n'est pas un produit. Autre erreur, très répandue : croire que $\max(X, Y)$ suit une loi exponentielle parce que $\min(X, Y)$ en suit une.

Exercice 23 ★★ Somme de deux temps d'attente exponentiels

Notions : Somme de variables indépendantes : produit de convolution et stabilité, Loi exponentielle et absence de mémoire, Lois gamma à un paramètre et fonction Gamma d'Euler

Un guichet traite les clients l'un après l'autre. Le temps de traitement du premier client est modélisé par une variable $X \hookrightarrow \mathcal{E}(\lambda)$, celui du second par une variable $Y \hookrightarrow \mathcal{E}(\lambda)$, avec $\lambda > 0$, les deux variables étant indépendantes. Le temps total est $S = X + Y$.

1. Déterminer une densité de S par le produit de convolution. On détaillera la recherche du domaine d'intégration.
2. Vérifier que la fonction obtenue est bien une densité.
3. Calculer $E(S)$ et $V(S)$ de deux façons.
4. Montrer que $\lambda S \hookrightarrow \gamma(2)$.
5. Établir que, pour tout $t \geq 0$, $P(S > t) = e^{-\lambda t}(1 + \lambda t)$. Comparer à $P(X > t)$ et commenter : la variable S possède-t-elle la propriété d'absence de mémoire ?
6. Application numérique avec $\lambda = \frac{1}{10}$, le temps étant compté en minutes : calculer $P(S > 20)$.

On donne les valeurs approchées suivantes : $e^{-1} \approx 0,3679$ et $e^{-2} \approx 0,1353$.

Corrigé

1. Une densité commune à X et Y est la fonction f définie par $f(u) = \lambda e^{-\lambda u}$ si $u > 0$ et $f(u) = 0$ si $u \leq 0$. Les deux variables étant indépendantes et à densité, le produit de convolution fournit une densité de leur somme (sa légitimité est admise) :

$$f_S(s) = \int_{-\infty}^{+\infty} f(t) f(s-t) dt$$

Recherche du domaine d'intégration. La fonction intégrée est nulle sauf si les **deux** facteurs sont non nuls, c'est-à-dire si

$$t > 0 \quad \text{et} \quad s - t > 0$$

La seconde condition se traduit en $t < s$. Le domaine est donc l'intervalle

$$]0, +\infty[\cap]-\infty, s[=]0, s[$$

Cet intervalle est vide dès que $s \leq 0$, et c'est bien ce qu'il faut : la somme de deux durées positives est positive.

Cas $s \leq 0$. La fonction intégrée est nulle partout, donc $f_S(s) = 0$.

Cas $s > 0$. On intègre sur le segment $]0, s[$:

$$\begin{aligned} f_S(s) &= \int_0^s \lambda e^{-\lambda t} \times \lambda e^{-\lambda(s-t)} dt \\ &= \lambda^2 \int_0^s e^{-\lambda t - \lambda s + \lambda t} dt \\ &= \lambda^2 e^{-\lambda s} \int_0^s 1 dt \\ &= \lambda^2 s e^{-\lambda s} \end{aligned}$$

Le mécanisme est remarquable : dans le produit des deux exponentielles, les termes en t se compensent exactement, et il ne reste qu'une constante à intégrer. C'est ce qui rend ce calcul si court une fois le domaine trouvé.

En résumé, une densité de S est

$$f_S(s) = \lambda^2 s e^{-\lambda s} \text{ sur }]0, +\infty[, \quad f_S(s) = 0 \text{ sur }]-\infty, 0]$$

2. Positivité. Pour $s > 0$, les trois facteurs λ^2 , s et $e^{-\lambda s}$ sont strictement positifs, donc $f_S(s) > 0$. Ailleurs f_S est nulle.

Continuité. La fonction est continue sur $]-\infty, 0]$ et sur $]0, +\infty[$ comme produit de fonctions usuelles. En 0, la limite à droite vaut $\lambda^2 \times 0 \times 1 = 0$, qui est aussi la valeur à gauche : f_S est en fait continue sur $\mathbb{R}$ tout entier.

Intégrale égale à 1. Soit $A > 0$. Une intégration par parties sur le segment $[0, A]$, avec $u(s) = s$ et $v'(s) = e^{-\lambda s}$, donc $v(s) = -\frac{1}{\lambda}e^{-\lambda s}$, donne

$$\begin{aligned} \int_0^A \lambda^2 s e^{-\lambda s} ds &= \lambda^2 \left(\left[-\frac{s}{\lambda} e^{-\lambda s} \right]_0^A + \frac{1}{\lambda} \int_0^A e^{-\lambda s} ds \right) \\ &= \lambda^2 \left(-\frac{A}{\lambda} e^{-\lambda A} + \frac{1}{\lambda} \times \frac{1 - e^{-\lambda A}}{\lambda} \right) \\ &= -\lambda A e^{-\lambda A} + 1 - e^{-\lambda A} \end{aligned}$$

Par croissances comparées, $A e^{-\lambda A} \rightarrow 0$ quand $A \rightarrow +\infty$, et $e^{-\lambda A} \rightarrow 0$. L'intégrale $\int_0^{+\infty} f_S(s) ds$ converge donc, et vaut 1. La fonction f_S est bien une densité de probabilité. $\square$

3. Première méthode : par les propriétés générales. Le cours donne $E(X) = E(Y) = \frac{1}{\lambda}$ et $V(X) = V(Y) = \frac{1}{\lambda^2}$. Par linéarité de l'espérance,

$$E(S) = E(X) + E(Y) = \frac{2}{\lambda}$$

et, puisque X et Y sont indépendantes, leur covariance est nulle, donc

$$V(S) = V(X) + V(Y) = \frac{2}{\lambda^2}$$

Seconde méthode : par la densité. Introduisons, pour $n \in \mathbb{N}$, l'intégrale

$$I_n = \int_0^{+\infty} s^n e^{-\lambda s} ds$$

Elle converge : la fonction intégrée est continue sur $[0, +\infty[$, positive, et par croissances comparées $s^n e^{-\lambda s} = o\left(\frac{1}{s^2}\right)$ au voisinage de $+\infty$, ce qui assure la convergence par comparaison avec une intégrale de Riemann convergente. La positivité de l'intégrande fait que cette convergence est de plus absolue.

On a immédiatement $I_0 = \frac{1}{\lambda}$. Pour $n \geq 1$, une intégration par parties sur $[0, A]$, avec $u(s) = s^n$ et $v(s) = -\frac{1}{\lambda}e^{-\lambda s}$, puis un passage à la limite, donne

$$I_n = \underbrace{\left[-\frac{s^n}{\lambda} e^{-\lambda s} \right]_0^{+\infty}}_{=0} + \frac{n}{\lambda} \int_0^{+\infty} s^{n-1} e^{-\lambda s} ds = \frac{n}{\lambda} I_{n-1}$$

le crochet étant nul par croissances comparées. Une récurrence immédiate donne alors

$$I_n = \frac{n!}{\lambda^{n+1}}$$

En particulier $I_2 = \frac{2}{\lambda^3}$ et $I_3 = \frac{6}{\lambda^4}$. Il vient

$$E(S) = \int_0^{+\infty} s \times \lambda^2 s e^{-\lambda s} ds = \lambda^2 I_2 = \lambda^2 \times \frac{2}{\lambda^3} = \frac{2}{\lambda}$$

puis, par le théorème de transfert appliqué à $s \mapsto s^2$,

$$E(S^2) = \lambda^2 I_3 = \lambda^2 \times \frac{6}{\lambda^4} = \frac{6}{\lambda^2}$$

et enfin, par König-Huygens,

$$V(S) = \frac{6}{\lambda^2} - \left(\frac{2}{\lambda}\right)^2 = \frac{6}{\lambda^2} - \frac{4}{\lambda^2} = \frac{2}{\lambda^2}$$

Les deux méthodes concordent, ce qui valide la densité obtenue en question 1.

4. Comme $\lambda > 0$, la transformation $s \mapsto \lambda s$ est affine strictement croissante. Posons $T = \lambda S$ et calculons sa fonction de répartition. Pour tout réel t ,

$$F_T(t) = P(\lambda S \leq t) = P\left(S \leq \frac{t}{\lambda}\right) = F_S\left(\frac{t}{\lambda}\right)$$

La fonction F_S étant continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf éventuellement en 0, il en va de même pour F_T , donc T est à densité, et on dérive :

$$f_T(t) = \frac{1}{\lambda} f_S\left(\frac{t}{\lambda}\right)$$

Pour $t \leq 0$, cela donne $f_T(t) = 0$. Pour $t > 0$,

$$f_T(t) = \frac{1}{\lambda} \times \lambda^2 \times \frac{t}{\lambda} \times e^{-\lambda \times t/\lambda} = t e^{-t}$$

Or la densité de la loi $\gamma(\nu)$ est $x \mapsto \frac{x^{\nu-1} e^{-x}}{\Gamma(\nu)}$ sur $]0, +\infty[$, nulle ailleurs. Pour $\nu = 2$, comme $\Gamma(2) = 1! = 1$, cette densité vaut $x e^{-x}$ sur $]0, +\infty[$. C'est exactement f_T , donc

$$\lambda S \hookrightarrow \gamma(2) \quad \square$$

Lecture par la stabilité. On pouvait prévoir ce résultat sans aucun calcul. Le cours indique que $\lambda X \hookrightarrow \gamma(1)$ et $\lambda Y \hookrightarrow \gamma(1)$, ces deux variables étant indépendantes puisque X et Y le sont. La stabilité des lois gamma par somme de variables indépendantes, admise, donne alors $\lambda X + \lambda Y \hookrightarrow \gamma(1 + 1) = \gamma(2)$, et $\lambda X + \lambda Y = \lambda S$. Le calcul de convolution mené aux questions 1. et 4. est donc une démonstration directe de ce cas particulier de la stabilité.

5. Soit $t \geq 0$. On calcule la probabilité de survie en intégrant la densité :

$$P(S > t) = \int_t^{+\infty} \lambda^2 s e^{-\lambda s} ds$$

Une intégration par parties sur $[t, A]$, avec les mêmes fonctions qu'en question 2., donne

$$\begin{aligned} \int_t^A \lambda^2 s e^{-\lambda s} ds &= \lambda^2 \left(\left[-\frac{s}{\lambda} e^{-\lambda s} \right]_t^A + \frac{1}{\lambda} \int_t^A e^{-\lambda s} ds \right) \\ &= \lambda^2 \left(\frac{t}{\lambda} e^{-\lambda t} - \frac{A}{\lambda} e^{-\lambda A} + \frac{1}{\lambda} \times \frac{e^{-\lambda t} - e^{-\lambda A}}{\lambda} \right) \\ &= \lambda t e^{-\lambda t} - \lambda A e^{-\lambda A} + e^{-\lambda t} - e^{-\lambda A} \end{aligned}$$

En faisant tendre A vers $+\infty$, les deux termes en A tendent vers 0 par croissances comparées, d'où

$$P(S > t) = e^{-\lambda t} (1 + \lambda t) \quad \square$$

Comparaison avec $P(X > t) = e^{-\lambda t}$. Le rapport des deux vaut

$$\frac{P(S > t)}{P(X > t)} = 1 + \lambda t$$

Il est strictement supérieur à 1 dès que $t > 0$, et il croît linéairement. C'est logique : attendre deux traitements prend plus de temps qu'un seul, et l'écart se creuse à mesure que t grandit.

Absence de mémoire. La loi de S n'est **pas** exponentielle, ne serait-ce que parce que sa densité s'annule en 0 alors qu'une densité exponentielle y vaut λ . Elle ne peut donc pas posséder l'absence de mémoire, qui caractérise la loi exponentielle parmi les lois à densité positives. On peut le voir directement : pour $t \geq 0$ et $h > 0$,

$$P_{(S>t)}(S > t + h) = \frac{P(S > t + h)}{P(S > t)} = e^{-\lambda h} \times \frac{1 + \lambda t + \lambda h}{1 + \lambda t}$$

Cette quantité dépend de t , alors que l'absence de mémoire exigerait qu'elle n'en dépende pas. La lecture concrète est limpide : si l'on attend depuis longtemps, il est très probable que le premier client soit déjà traité, donc l'attente restante est plus courte qu'au départ. La somme, contrairement à chacun de ses termes, garde la mémoire du temps écoulé.

6. Avec $\lambda = \frac{1}{10}$ et $t = 20$, on a $\lambda t = 2$, donc

$$P(S > 20) = e^{-2} (1 + 2) = 3e^{-2} \approx 3 \times 0,1353 \approx 0,406$$

Il y a donc environ 40,6 % de chances que le traitement des deux clients dépasse 20 minutes. À titre de comparaison, un seul client dépasserait 20 minutes avec la probabilité $e^{-2} \approx 0,135$, soit trois fois moins.

On notera enfin que 20 minutes est précisément l'espérance de S , puisque $E(S) = \frac{2}{\lambda} = 20$. La probabilité de dépasser la moyenne vaut donc ici 0,406, contre $e^{-1} \approx 0,368$ pour une seule attente exponentielle : les deux lois sont dissymétriques, mais pas de la même façon.

Méthode à retenir. Pour la somme de deux variables exponentielles de **même** paramètre, le domaine de convolution est $]0, s[$ et le calcul se réduit à intégrer une constante, parce que les exponentielles se recombinent. Le résultat, la densité $\lambda^2 s e^{-\lambda s}$, est un classique absolu de la filière : il faut savoir le retrouver en trois lignes. Pour tout ce qui touche aux moments, l'outil est la suite $I_n = \int_0^{+\infty} s^n e^{-\lambda s} ds = \frac{n!}{\lambda^{n+1}}$, établie par récurrence avec une intégration par parties.

Erreur classique. Écrire le domaine d'intégration $]0, +\infty[$ en oubliant la contrainte $s - t > 0$: on obtient une intégrale divergente ou une densité fautive. Autre erreur, conceptuelle : croire que la somme de deux exponentielles est exponentielle, par analogie avec le minimum, qui l'est. Il n'y a stabilité par somme que pour les lois normales, gamma, de Poisson et binomiales de même paramètre p ; la loi exponentielle $\mathcal{E}(\lambda)$ n'est pas de la partie, et c'est justement la loi gamma qui prend le relais.

Exercice 24 ★★★ Durée de vie d'un système en série

Notions : Minimum et maximum d'un échantillon de variables indépendantes, Loi exponentielle et absence de mémoire, n-uplets de variables aléatoires, indépendance mutuelle, échantillon

Un système est constitué de n composants, avec $n \geq 2$. Pour $i \in \llbracket 1, n \rrbracket$, la durée de vie du composant numéro i , comptée en heures, est une variable aléatoire $X_i \hookrightarrow \mathcal{E}(\lambda_i)$ avec $\lambda_i > 0$. Les variables $X_1, \dots, X_n$ sont mutuellement indépendantes.

Le système est monté **en série** : il fonctionne tant que tous ses composants fonctionnent. Sa durée de vie est donc

$$M = \min(X_1, \dots, X_n)$$

1. Montrer que $M \hookrightarrow \mathcal{E}\left(\sum_{i=1}^n \lambda_i\right)$.
2. On suppose désormais que tous les composants sont identiques : $\lambda_i = \lambda$ pour tout i . Donner la loi de M et $E(M)$. Commenter l'évolution de la fiabilité du système quand n augmente.
3. On prend $\lambda = \frac{1}{20\,000}$. Déterminer le nombre maximal de composants que le système peut comporter pour que $P(M > 1000) \geq 0,9$.
4. On monte maintenant les n composants identiques **en parallèle** : le système fonctionne tant qu'au moins un composant fonctionne, et sa durée de vie est $N = \max(X_1, \dots, X_n)$. Déterminer la fonction de répartition de N , puis montrer que

$$E(N) = \frac{1}{\lambda} \sum_{k=1}^n \frac{1}{k}$$

5. Comparer chiffrément les deux montages pour $n = 3$ et $\lambda = \frac{1}{20\,000}$.

On donne les valeurs approchées suivantes : $\ln(0,9) \approx -0,1054$, $e^{-0,05} \approx 0,9512$ et $e^{-0,15} \approx 0,8607$.

Corrigé

1. Soit t un réel. Le minimum de n nombres dépasse t si et seulement si chacun d'eux dépasse t , donc

$$(M > t) = \bigcap_{i=1}^n (X_i > t)$$

Les variables $X_1, \dots, X_n$ étant **mutuellement** indépendantes, la probabilité de cette intersection est le produit des probabilités :

$$P(M > t) = \prod_{i=1}^n P(X_i > t)$$

Cas $t < 0$. Chaque X_i est positive, donc chaque facteur vaut 1 et $P(M > t) = 1$, c'est-à-dire $F_M(t) = 0$.

Cas $t \geq 0$. Le cours donne $P(X_i > t) = e^{-\lambda_i t}$, donc

$$P(M > t) = \prod_{i=1}^n e^{-\lambda_i t} = e^{-(\sum_{i=1}^n \lambda_i)t}$$

en utilisant que le produit des exponentielles est l'exponentielle de la somme. Posons $\Lambda = \sum_{i=1}^n \lambda_i$, qui est strictement positif comme somme de réels strictement positifs. On obtient

$$F_M(t) = 1 - e^{-\Lambda t} \quad \text{pour } t \geq 0, \quad F_M(t) = 0 \quad \text{pour } t < 0$$

C'est exactement la fonction de répartition de la loi $\mathcal{E}(\Lambda)$, et la loi d'une variable est caractérisée par sa fonction de répartition. Donc

$$M \hookrightarrow \mathcal{E}\left(\sum_{i=1}^n \lambda_i\right) \quad \square$$

Le paramètre d'une loi exponentielle s'interprète en fiabilité comme un **taux de défaillance**. Le résultat se retient donc ainsi : dans un montage en série, les taux de défaillance s'additionnent.

2. Si tous les λ_i valent λ , la somme vaut $n\lambda$, donc

$$M \hookrightarrow \mathcal{E}(n\lambda) \quad \text{et} \quad E(M) = \frac{1}{n\lambda}$$

Commentaire. La durée de vie moyenne du système est inversement proportionnelle au nombre de composants. Doubler le nombre de composants divise par deux l'espérance de vie ; passer de 1 à 10 composants la divise par 10. La chute est encore plus brutale si l'on regarde les probabilités de survie : pour $t \geq 0$,

$$P(M > t) = e^{-n\lambda t} = (e^{-\lambda t})^n = (P(X_1 > t))^n$$

La fiabilité du système à l'instant t est la puissance n -ième de celle d'un composant. Comme $e^{-\lambda t} < 1$, cette puissance tend vers 0 quand n grandit, et vite. Un système en série est aussi fragile que la conjonction de toutes ses fragilités : c'est pour cela qu'un montage en série se dimensionne toujours avec des composants nettement plus fiables que ce que l'on exige du système.

3. Avec $\lambda = \frac{1}{20000}$ et $t = 1000$, on a $\lambda t = \frac{1000}{20000} = \frac{1}{20}$, donc

$$P(M > 1000) = e^{-n\lambda \times 1000} = e^{-n/20}$$

L'inéquation à résoudre est donc

$$e^{-n/20} \geq 0,9$$

La fonction logarithme népérien étant strictement croissante sur $]0, +\infty[$, elle conserve le sens de l'inégalité :

$$-\frac{n}{20} \geq \ln(0,9) \quad \Longleftrightarrow \quad \frac{n}{20} \leq -\ln(0,9) \quad \Longleftrightarrow \quad n \leq -20 \ln(0,9)$$

Attention au changement de sens lors de la multiplication par -1 . Numériquement, $\ln(0,9) \approx -0,10536$, donc

$$-20 \ln(0,9) \approx 2,107$$

Comme n est un entier, la condition s'écrit $n \leq 2$. Le système ne peut donc comporter **au plus que 2 composants**.

Vérification. Pour $n = 2$, $P(M > 1000) = e^{-0,1} \approx 0,905 \geq 0,9$: la contrainte est respectée. Pour $n = 3$, $P(M > 1000) = e^{-0,15} \approx 0,861 < 0,9$: elle ne l'est plus. La réponse est bien 2.

Le résultat est frappant. Chaque composant a une durée de vie moyenne de 20 000 heures, soit plus de deux ans en fonctionnement continu, et pourtant on ne peut pas en assembler trois en série si l'on exige 90 % de survie à seulement 1000 heures.

4. Le maximum de n nombres est inférieur ou égal à t si et seulement si chacun d'eux l'est :

$$(N \leq t) = \bigcap_{i=1}^n (X_i \leq t)$$

Par indépendance mutuelle, $F_N(t) = \prod_{i=1}^n F_{X_i}(t)$, soit, les composants étant identiques,

$$F_N(t) = (1 - e^{-\lambda t})^n \quad \text{pour } t \geq 0, \quad F_N(t) = 0 \quad \text{pour } t < 0$$

Cette fonction est continue sur $\mathbb{R}$ (les deux expressions valent 0 en 0) et de classe $\mathcal{C}^1$ sauf éventuellement en 0 : la variable N est à densité, et pour $t > 0$

$$f_N(t) = n\lambda e^{-\lambda t} (1 - e^{-\lambda t})^{n-1}$$

la densité étant nulle pour $t < 0$.

Une majoration utile. Pour $t \geq 0$, posons $a = 1 - e^{-\lambda t} \in [0, 1[$. L'identité classique

$$1 - a^n = (1 - a)(1 + a + \dots + a^{n-1})$$

et la majoration de chacun des n termes de la parenthèse par 1 donnent

$$1 - F_N(t) = 1 - a^n \leq n(1 - a) = n e^{-\lambda t}$$

Existence de $E(N)$. De $f_N(t) \leq n\lambda e^{-\lambda t}$ on tire $0 \leq t f_N(t) \leq n\lambda t e^{-\lambda t}$, et l'intégrale $\int_0^{+\infty} t e^{-\lambda t} dt$ converge (elle vaut $\frac{1}{\lambda^2}$). Par comparaison de fonctions positives, $E(N)$ existe, et l'intégrale converge absolument.

Première transformation : une intégration par parties. Sur $[0, A]$, prenons $u(t) = t$ et pour primitive de f_N la fonction $v(t) = F_N(t) - 1$, choisie de manière à tendre vers 0 en $+\infty$. Alors

$$\int_0^A t f_N(t) dt = \left[t(F_N(t) - 1) \right]_0^A + \int_0^A (1 - F_N(t)) dt = -A(1 - F_N(A)) + \int_0^A (1 - F_N(t)) dt$$

D'après la majoration précédente, $0 \leq A(1 - F_N(A)) \leq nA e^{-\lambda A}$, qui tend vers 0 par croissances comparées. En faisant tendre A vers $+\infty$:

$$E(N) = \int_0^{+\infty} (1 - (1 - e^{-\lambda t})^n) dt$$

Seconde transformation : un changement de variable. Posons $u = 1 - e^{-\lambda t}$. L'application $t \mapsto 1 - e^{-\lambda t}$ est de classe $\mathcal{C}^1$ et strictement croissante sur $[0, A]$, à valeurs dans $[0, B]$ avec $B = 1 - e^{-\lambda A}$. On a $du = \lambda e^{-\lambda t} dt$, et comme $e^{-\lambda t} = 1 - u$, il vient $dt = \frac{du}{\lambda(1-u)}$. Donc

$$\int_0^A (1 - (1 - e^{-\lambda t})^n) dt = \frac{1}{\lambda} \int_0^B \frac{1 - u^n}{1 - u} du$$

Or, pour $u \neq 1$, l'identité rappelée plus haut donne exactement

$$\frac{1 - u^n}{1 - u} = 1 + u + u^2 + \dots + u^{n-1} = \sum_{k=0}^{n-1} u^k$$

Le quotient, qui semblait problématique près de $u = 1$, est donc en réalité un simple polynôme. On intègre terme à terme cette somme **finie** :

$$\frac{1}{\lambda} \int_0^B \sum_{k=0}^{n-1} u^k du = \frac{1}{\lambda} \sum_{k=0}^{n-1} \frac{B^{k+1}}{k+1}$$

Quand A tend vers $+\infty$, B tend vers 1, et cette somme finie de fonctions continues tend vers

$$E(N) = \frac{1}{\lambda} \sum_{k=0}^{n-1} \frac{1}{k+1} = \frac{1}{\lambda} \sum_{k=1}^n \frac{1}{k} \quad \square$$

Contrôle sur les petits cas. Pour $n = 2$, la formule donne $E(N) = \frac{1}{\lambda} (1 + \frac{1}{2}) = \frac{3}{2\lambda}$. C'est bien ce que fournit la formule $E(N) = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}$ obtenue pour deux composants, quand on y fait $\mu = \lambda : \frac{1}{\lambda} + \frac{1}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda}$. Pour $n = 3$, la formule donne $\frac{11}{6\lambda}$.

Commentaire. La somme $\sum_{k=1}^n \frac{1}{k}$ croît, mais très lentement : elle vaut 1, puis 1,5, puis environ 1,83, 2,08, 2,28 pour n allant de 1 à 5. Ajouter un composant redondant améliore donc la durée de vie moyenne, mais de moins en moins : le passage de 1 à 2 composants la multiplie par 1,5, celui de 4 à 5 ne la multiplie que par 1,10 environ. La redondance a des rendements décroissants, ce que tout ingénieur sait, et que le calcul quantifie.

5. Prenons $n = 3$ et $\lambda = \frac{1}{20\,000}$, donc $\frac{1}{\lambda} = 20\,000$ heures pour un composant seul.

Espérances. En série, $E(M) = \frac{1}{3\lambda} = \frac{20\,000}{3} \approx 6667$ heures. En parallèle, $E(N) = \frac{11}{6\lambda} = \frac{11 \times 20\,000}{6} \approx 36\,667$ heures.

Probabilités de survie à 1000 heures. On a déjà $P(M > 1000) = e^{-0,15} \approx 0,861$. Pour le montage en parallèle, avec $e^{-0,05} \approx 0,9512$:

$$P(N > 1000) = 1 - (1 - e^{-0,05})^3 \approx 1 - (0,0488)^3 \approx 1 - 0,000116 \approx 0,99988$$

Montage	Espérance de vie (heures)	$P(\text{survie à } 1000 \text{ h})$
Composant seul	20 000	0,951
Série ($n = 3$)	6667	0,861
Parallèle ($n = 3$)	36 667	0,99988

L'écart est spectaculaire. Avec exactement les mêmes trois composants, la durée de vie moyenne passe de 6667 à 36 667 heures selon le montage, soit un rapport $\frac{11}{2} = 5,5$. Sur la probabilité de panne à 1000 heures, l'écart est plus net encore : environ 14 % de risque en série, contre un peu plus d'un cas sur 10 000 en parallèle.

Méthode à retenir. Minimum : on passe par $P(M > t)$, l'événement se factorise en intersection et l'indépendance mutuelle transforme la probabilité en produit. Maximum : on passe par $F_N(t)$, pour la même raison. Le minimum de variables exponentielles indépendantes est exponentielle de paramètre la somme des paramètres, résultat à connaître. Pour une espérance de maximum, la voie la plus sûre est l'intégration par parties qui ramène $\int_0^{+\infty} t f_N(t) dt$ à $\int_0^{+\infty} (1 - F_N(t)) dt$, suivie du changement de variable $u = 1 - e^{-\lambda t}$ qui fait apparaître une somme géométrique.

Erreur classique. Traiter le maximum comme le minimum en écrivant $P(N > t) = \prod P(X_i > t)$. L'événement $(N > t)$ est une réunion, pas une intersection. Autre erreur, dans la question 3. : oublier de changer le sens de l'inégalité en multipliant par -20 , ce qui conduirait à $n \geq 2,1$ et à une conclusion exactement inverse. Enfin, dans la question 4., croire que $\frac{1-u^n}{1-u}$ pose un problème en $u = 1$: c'est un polynôme, il n'y a aucune singularité.

Exercice 25 ★★★ Maximum d'un échantillon uniforme

Notions : Minimum et maximum d'un échantillon de variables indépendantes, Loi uniforme sur un segment, n-uplets de variables aléatoires, indépendance mutuelle, échantillon, Espérance d'une variable à densité : existence, théorème de transfert, linéarité

Soit $\theta > 0$ et n un entier supérieur ou égal à 1. Les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes et suivent toutes la loi $\mathcal{U}([0, \theta])$: elles forment un échantillon de taille n de cette loi.

On pose

$$M_n = \max(X_1, \dots, X_n) \quad \text{et} \quad m_n = \min(X_1, \dots, X_n)$$

1. Déterminer la fonction de répartition de M_n , puis une densité de M_n .
2. Calculer $E(M_n)$ et $V(M_n)$.
3. Déterminer le réel a_n tel que $E(a_n M_n) = \theta$, puis calculer $V(a_n M_n)$.
4. Déterminer la fonction de répartition, une densité et l'espérance de m_n .
5. Étudier les limites de $E(M_n)$, $V(M_n)$ et $E(m_n)$ quand n tend vers $+\infty$, et commenter.

Corrigé

1. Commençons par la fonction de répartition commune des X_i . Pour $X \hookrightarrow \mathcal{U}([0, \theta])$, le cours donne

$$F_X(x) = 0 \quad \text{si } x < 0, \quad F_X(x) = \frac{x}{\theta} \quad \text{si } 0 \leq x \leq \theta, \quad F_X(x) = 1 \quad \text{si } x > \theta$$

Le maximum de n nombres est inférieur ou égal à x si et seulement si chacun d'eux l'est, donc

$$(M_n \leq x) = \bigcap_{i=1}^n (X_i \leq x)$$

Les variables étant mutuellement indépendantes, la probabilité de l'intersection est le produit des probabilités, et comme elles ont toutes la même loi, ces n probabilités sont égales :

$$F_{M_n}(x) = \prod_{i=1}^n F_{X_i}(x) = (F_X(x))^n$$

D'où

$$F_{M_n}(x) = 0 \quad \text{si } x < 0, \quad F_{M_n}(x) = \left(\frac{x}{\theta}\right)^n \quad \text{si } 0 \leq x \leq \theta, \quad F_{M_n}(x) = 1 \quad \text{si } x > \theta$$

M_n est à densité. La fonction F_{M_n} est continue sur chacun des trois intervalles, et les raccords se font sans saut : en 0 les deux expressions valent 0, en θ elles valent toutes deux 1. Elle est donc continue sur $\mathbb{R}$, et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement en 0 et en θ , soit en un nombre fini de points. La variable M_n est donc à densité, et on obtient une densité en dérivant :

$$f_{M_n}(x) = \frac{n x^{n-1}}{\theta^n} \quad \text{si } 0 < x < \theta, \quad f_{M_n}(x) = 0 \quad \text{sinon}$$

2. La densité étant nulle hors de $[0, \theta]$, toutes les intégrales à venir sont des intégrales de fonctions continues sur un **segment** : elles convergent, et absolument puisque les fonctions intégrées sont

positives sur $[0, \theta]$. L'espérance et la variance de M_n existent donc sans discussion.

$$\begin{aligned} E(M_n) &= \int_0^\theta x \times \frac{n x^{n-1}}{\theta^n} dx \\ &= \frac{n}{\theta^n} \int_0^\theta x^n dx \\ &= \frac{n}{\theta^n} \times \frac{\theta^{n+1}}{n+1} \\ &= \frac{n\theta}{n+1} \end{aligned}$$

De même, par le théorème de transfert appliqué à $x \mapsto x^2$,

$$E(M_n^2) = \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx = \frac{n}{\theta^n} \times \frac{\theta^{n+2}}{n+2} = \frac{n\theta^2}{n+2}$$

La formule de König-Huygens donne alors

$$V(M_n) = \frac{n\theta^2}{n+2} - \left(\frac{n\theta}{n+1} \right)^2 = n\theta^2 \left(\frac{1}{n+2} - \frac{n}{(n+1)^2} \right)$$

On réduit la parenthèse au dénominateur commun $(n+2)(n+1)^2$, en développant soigneusement le numérateur :

$$(n+1)^2 - n(n+2) = (n^2 + 2n + 1) - (n^2 + 2n) = 1$$

La simplification est spectaculaire : il ne reste que 1. Donc

$$V(M_n) = \frac{n\theta^2}{(n+1)^2(n+2)}$$

Contrôle pour $n = 1$. Alors $M_1 = X_1$, et les formules donnent $E(M_1) = \frac{\theta}{2}$ et $V(M_1) = \frac{\theta^2}{4 \times 3} = \frac{\theta^2}{12}$, qui sont bien l'espérance et la variance de $\mathcal{U}([0, \theta])$.

3. Cherchons a_n réel tel que $E(a_n M_n) = \theta$. Par linéarité de l'espérance,

$$E(a_n M_n) = a_n E(M_n) = a_n \times \frac{n\theta}{n+1}$$

Cette quantité vaut θ si et seulement si $a_n \frac{n}{n+1} = 1$, c'est-à-dire

$$a_n = \frac{n+1}{n}$$

Ce coefficient est strictement supérieur à 1, ce qui est cohérent : M_n est un maximum de valeurs toutes inférieures à θ , donc $M_n < \theta$ à coup sûr, et il faut nécessairement l'agrandir pour retomber sur θ en moyenne.

Pour la variance, on utilise $V(aX) = a^2 V(X)$:

$$\begin{aligned} V(a_n M_n) &= \left(\frac{n+1}{n} \right)^2 \times \frac{n\theta^2}{(n+1)^2(n+2)} \\ &= \frac{(n+1)^2}{n^2} \times \frac{n\theta^2}{(n+1)^2(n+2)} \\ &= \frac{\theta^2}{n(n+2)} \end{aligned}$$

Les facteurs $(n+1)^2$ se simplifient et l'expression devient très simple.

4. Le minimum de n nombres dépasse x si et seulement si chacun d'eux dépasse x , donc

$$(m_n > x) = \bigcap_{i=1}^n (X_i > x) \quad \text{d'où} \quad P(m_n > x) = (1 - F_X(x))^n$$

par indépendance mutuelle et identité des lois. Pour $0 \leq x \leq \theta$, cela donne $P(m_n > x) = (1 - \frac{x}{\theta})^n$, d'où

$$F_{m_n}(x) = 0 \quad \text{si } x < 0, \quad F_{m_n}(x) = 1 - \left(1 - \frac{x}{\theta}\right)^n \quad \text{si } 0 \leq x \leq \theta, \quad F_{m_n}(x) = 1 \quad \text{si } x > \theta$$

Comme pour le maximum, cette fonction est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf éventuellement en 0 et en θ : la variable m_n est à densité, et en dérivant,

$$f_{m_n}(x) = \frac{n}{\theta} \left(1 - \frac{x}{\theta}\right)^{n-1} \quad \text{si } 0 < x < \theta, \quad f_{m_n}(x) = 0 \quad \text{sinon}$$

Espérance. À nouveau l'intégrale porte sur un segment, donc elle converge. On effectue le changement de variable $u = 1 - \frac{x}{\theta}$, de classe $\mathcal{C}^1$ et strictement décroissant, avec $x = \theta(1 - u)$ et $dx = -\theta du$; les bornes $x = 0$ et $x = \theta$ deviennent $u = 1$ et $u = 0$.

$$\begin{aligned} E(m_n) &= \int_0^\theta x \times \frac{n}{\theta} \left(1 - \frac{x}{\theta}\right)^{n-1} dx \\ &= \int_1^0 \theta(1 - u) \times \frac{n}{\theta} u^{n-1} \times (-\theta) du \\ &= n\theta \int_0^1 (u^{n-1} - u^n) du \\ &= n\theta \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= n\theta \times \frac{1}{n(n+1)} = \frac{\theta}{n+1} \end{aligned}$$

Contrôle par symétrie. Posons $Y_i = \theta - X_i$. Chaque Y_i suit encore la loi $\mathcal{U}([0, \theta])$ (c'est une transformation affine qui échange les deux bornes du segment), et ces variables sont mutuellement indépendantes. Retourner l'échantillon échange minimum et maximum :

$$\max(Y_1, \dots, Y_n) = \theta - m_n$$

En prenant l'espérance, $\frac{n\theta}{n+1} = \theta - E(m_n)$, d'où $E(m_n) = \theta - \frac{n\theta}{n+1} = \frac{\theta}{n+1}$. On retrouve bien le même résultat. On note au passage la jolie relation

$$E(m_n) + E(M_n) = \frac{\theta}{n+1} + \frac{n\theta}{n+1} = \theta$$

5. Reprenons les trois quantités calculées.

Espérance du maximum. En écrivant $\frac{n}{n+1} = \frac{1}{1+1/n}$, on obtient

$$\lim_{n \rightarrow +\infty} E(M_n) = \lim_{n \rightarrow +\infty} \frac{n\theta}{n+1} = \theta$$

et plus précisément l'écart à θ vaut $\theta - E(M_n) = \frac{\theta}{n+1}$, qui décroît comme $\frac{1}{n}$.

Variance du maximum. Le numérateur est de degré 1 en n , le dénominateur de degré 3, donc

$$V(M_n) = \frac{n\theta^2}{(n+1)^2(n+2)} \underset{n \rightarrow +\infty}{\sim} \frac{\theta^2}{n^2} \quad \text{donc} \quad \lim_{n \rightarrow +\infty} V(M_n) = 0$$

L'écart-type est donc de l'ordre de $\frac{\theta}{\sqrt{n}}$.

Espérance du minimum. $E(m_n) = \frac{\theta}{n+1} \rightarrow 0$.

Commentaire. Les trois résultats racontent la même histoire. Quand on tire de plus en plus de valeurs uniformément dans $[0, \theta]$, le plus petit tirage se rapproche de la borne 0 et le plus grand se rapproche de la borne θ , tandis que la dispersion du plus grand s'écroule. On peut le chiffrer directement sur les probabilités : pour tout n ,

$$P(M_n \geq 0,9\theta) = 1 - F_{M_n}(0,9\theta) = 1 - (0,9)^n$$

n	10	50	100
$P(M_n \geq 0,9\theta)$	0,651	0,995	0,99997

Avec cent tirages, il est pratiquement certain que le plus grand dépasse 90 % de θ . C'est très intuitif : pour que le maximum reste sous $0,9\theta$, il faut que **toutes** les valeurs tirées y restent, ce qui devient de plus en plus improbable à mesure que n grandit, chaque tirage supplémentaire multipliant la probabilité par 0,9.

Enfin, la question 3. prend tout son sens. La variable M_n reste toujours strictement inférieure à θ , donc sa moyenne aussi. En la multipliant par $a_n = \frac{n+1}{n}$, on corrige exactement ce décalage et on retombe en moyenne sur θ . Le prix à payer est une dispersion un peu plus grande, mais $V(a_n M_n) = \frac{\theta^2}{n(n+2)}$ tend elle aussi vers 0.

Méthode à retenir. Pour un maximum, on part toujours de $F_{M_n}(x) = (F_X(x))^n$; pour un minimum, de $P(m_n > x) = (1 - F_X(x))^n$. On dérive ensuite pour obtenir la densité, après avoir vérifié la continuité de la fonction de répartition et son caractère $\mathcal{C}^1$ sauf en un nombre fini de points. Le cas uniforme est le plus simple de tous, puisque $F_X(x) = \frac{x}{\theta}$ transforme la puissance n -ième en un simple monôme.

Erreur classique. Écrire $F_{M_n}(x) = (\frac{x}{\theta})^n$ pour tout réel x , en oubliant que cette expression n'est valable que sur $[0, \theta]$: sans les deux morceaux extrêmes, on obtient une fonction non bornée et on calcule des intégrales sur $\mathbb{R}$ tout entier qui divergent. Autre erreur, dans le calcul de la variance : oublier d'élever $E(M_n)$ au carré, ou tenter de simplifier $\frac{n}{n+2} - \frac{n^2}{(n+1)^2}$ sans réduire au même dénominateur. La simplification du numérateur en 1 est le signe que le calcul est juste.

Exercice 26 ★★ Le carré d'une variable normale centrée réduite

Notions : Loi d'une transformée : transformation affine et méthode de la fonction de répartition, Loi normale, centrage-réduction et lecture de la table, Lois gamma à un paramètre et fonction Gamma d'Euler

La variable aléatoire Z suit la loi normale centrée réduite $\mathcal{N}(0, 1)$. On note φ sa densité et Φ sa fonction de répartition, et on pose $Y = Z^2$.

1. Déterminer $Y(\Omega)$, puis montrer que pour tout $y > 0$,

$$F_Y(y) = 2\Phi(\sqrt{y}) - 1$$

2. En déduire que Y est une variable à densité et qu'une densité de Y est donnée par

$$f_Y(y) = \frac{e^{-y/2}}{\sqrt{2\pi y}} \text{ sur }]0, +\infty[, \quad f_Y(y) = 0 \text{ ailleurs}$$

3. Montrer que $T = \frac{Y}{2}$ suit la loi $\gamma(\frac{1}{2})$.
4. En écrivant que la densité de T est d'intégrale 1, retrouver la valeur $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
5. Calculer $E(Y)$, puis $E(Z^4)$ à l'aide d'une intégration par parties, et en déduire $V(Y)$.

Corrigé

1. La variable Z prend toutes les valeurs réelles, donc $Y = Z^2$ prend toutes les valeurs positives ou nulles :

$$Y(\Omega) = [0, +\infty[$$

On peut préciser tout de suite que $P(Y = 0) = P(Z = 0) = 0$, puisque Z est une variable à densité : la valeur 0 est atteinte, mais avec une probabilité nulle.

Calcul de F_Y . Soit y un réel.

Cas $y < 0$. L'événement $(Z^2 \leq y)$ est impossible car un carré est positif, donc $F_Y(y) = 0$.

Cas $y = 0$. $F_Y(0) = P(Z^2 \leq 0) = P(Z = 0) = 0$.

Cas $y > 0$. Pour un réel z , l'inégalité $z^2 \leq y$ équivaut à $|z| \leq \sqrt{y}$, c'est-à-dire à $-\sqrt{y} \leq z \leq \sqrt{y}$. C'est le point clé de l'exercice : passer au carré n'est pas une opération injective, et l'événement se traduit par un **encadrement**, pas par une seule inégalité. Donc

$$F_Y(y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$

Il reste à utiliser la parité. La densité φ de la loi $\mathcal{N}(0, 1)$ est une fonction **paire**, car $\varphi(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$ ne dépend que de z^2 . La courbe est donc symétrique par rapport à l'axe des ordonnées, ce qui se traduit sur les aires par la relation du cours

$$\Phi(-u) = 1 - \Phi(u) \quad \text{pour tout réel } u$$

Appliquée à $u = \sqrt{y}$, elle donne

$$F_Y(y) = \Phi(\sqrt{y}) - (1 - \Phi(\sqrt{y})) = 2\Phi(\sqrt{y}) - 1 \quad \square$$

2. *Continuité de F_Y .* Sur $]-\infty, 0]$, F_Y est nulle donc continue. Sur $]0, +\infty[$, elle est continue comme composée et combinaison de fonctions continues. En 0, la limite à droite vaut

$$\lim_{y \rightarrow 0^+} (2\Phi(\sqrt{y}) - 1) = 2\Phi(0) - 1 = 2 \times \frac{1}{2} - 1 = 0$$

puisque $\Phi(0) = \frac{1}{2}$ par symétrie de la loi centrée. Cette limite coïncide avec $F_Y(0) = 0$: la fonction F_Y est continue sur $\mathbb{R}$.

Caractère $\mathcal{C}^1$. Sur $] -\infty, 0[$, F_Y est nulle donc de classe $\mathcal{C}^1$. Sur $] 0, +\infty[$, la fonction $y \mapsto \sqrt{y}$ est de classe $\mathcal{C}^1$, et Φ l'est aussi puisque $\Phi' = \varphi$ est continue sur $\mathbb{R}$; la composée est donc de classe $\mathcal{C}^1$. Seul le point 0 échappe éventuellement à la règle, ce qui est autorisé : la fonction de répartition est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf en un point. La variable Y est donc à densité.

Calcul de la densité. Pour $y > 0$, en dérivant la composée,

$$f_Y(y) = F'_Y(y) = 2 \times \varphi(\sqrt{y}) \times \frac{1}{2\sqrt{y}} = \frac{\varphi(\sqrt{y})}{\sqrt{y}}$$

Or $\varphi(\sqrt{y}) = \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} = \frac{1}{\sqrt{2\pi}} e^{-y/2}$, donc

$$f_Y(y) = \frac{e^{-y/2}}{\sqrt{2\pi} \sqrt{y}} = \frac{e^{-y/2}}{\sqrt{2\pi y}} \quad \square$$

et $f_Y(y) = 0$ pour $y < 0$, la valeur en 0 étant sans importance.

Remarque. Cette densité tend vers $+\infty$ quand y tend vers 0 par valeurs supérieures. Ce n'est nullement contradictoire : une densité n'a aucune obligation d'être bornée, elle doit seulement être positive et d'intégrale 1. Le phénomène s'explique très bien : les valeurs de Z s'accumulent autour de 0, où la densité φ est maximale, et l'élévation au carré écrase encore ces valeurs vers 0, ce qui concentre énormément de masse près de l'origine.

3. Posons $T = \frac{Y}{2}$. La transformation $y \mapsto \frac{y}{2}$ est affine strictement croissante, donc pour tout réel t

$$F_T(t) = P\left(\frac{Y}{2} \leq t\right) = P(Y \leq 2t) = F_Y(2t)$$

La fonction F_T hérite de F_Y sa continuité sur $\mathbb{R}$ et son caractère $\mathcal{C}^1$ sauf en 0 : la variable T est à densité, et en dérivant, $f_T(t) = 2 f_Y(2t)$. Ainsi $f_T(t) = 0$ pour $t < 0$, et pour $t > 0$

$$f_T(t) = 2 \times \frac{e^{-2t/2}}{\sqrt{2\pi} \times 2t} = \frac{2e^{-t}}{\sqrt{4\pi t}} = \frac{2e^{-t}}{2\sqrt{\pi t}} = \frac{e^{-t}}{\sqrt{\pi t}}$$

On écrit ce résultat sous une forme qui parle :

$$f_T(t) = \frac{1}{\sqrt{\pi}} t^{-1/2} e^{-t} = \frac{1}{\sqrt{\pi}} t^{\frac{1}{2}-1} e^{-t} \quad \text{sur }]0, +\infty[$$

Or la densité de la loi $\gamma(\nu)$ est, par définition,

$$x \mapsto \frac{x^{\nu-1} e^{-x}}{\Gamma(\nu)} \quad \text{sur }]0, +\infty[, \quad 0 \text{ ailleurs}$$

Pour $\nu = \frac{1}{2}$, cette densité vaut $\frac{t^{-1/2} e^{-t}}{\Gamma(\frac{1}{2})}$. Comme le cours donne $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, les deux expressions coïncident exactement, donc

$$T = \frac{Z^2}{2} \hookrightarrow \gamma\left(\frac{1}{2}\right) \quad \square$$

4. La question **3.** a produit la densité de T par un calcul qui n'utilise à aucun moment la valeur de $\Gamma(\frac{1}{2})$: on n'y a manipulé que φ , une dérivation et une simplification de racines carrées. On peut donc s'en servir pour **retrouver** cette valeur, sans raisonnement circulaire.

Rappelons d'abord que l'intégrale définissant $\Gamma(\frac{1}{2})$ converge. En effet

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} t^{-1/2} e^{-t} dt$$

est impropre aux deux bornes. En 0, $t^{-1/2} e^{-t} \sim t^{-1/2}$, et l'intégrale de Riemann $\int_0^1 \frac{dt}{t^{1/2}}$ converge car l'exposant $\frac{1}{2}$ est strictement inférieur à 1. En $+\infty$, par croissances comparées, $t^{-1/2} e^{-t} =$

$o\left(\frac{1}{t^2}\right)$, et l'intégrale de Riemann correspondante converge. Les fonctions étant positives, ces comparaisons sont légitimes.

Maintenant, f_T est une densité de probabilité, donc son intégrale sur $\mathbb{R}$ vaut 1. Comme f_T est nulle sur $]-\infty, 0]$,

$$1 = \int_0^{+\infty} f_T(t) dt = \frac{1}{\sqrt{\pi}} \int_0^{+\infty} t^{-1/2} e^{-t} dt = \frac{\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi}}$$

d'où

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad \square$$

Cette valeur, qui peut sembler mystérieuse, provient donc en droite ligne de la constante de normalisation $\frac{1}{\sqrt{2\pi}}$ de la loi normale, elle-même issue de l'intégrale de Gauss $\int_{-\infty}^{+\infty} e^{-t^2/2} dt = \sqrt{2\pi}$, admise par le programme. Les deux résultats sont deux visages du même fait.

5. Espérance de Y . Aucun calcul d'intégrale n'est nécessaire. La variable Z suit la loi $\mathcal{N}(0, 1)$, donc $E(Z) = 0$ et $V(Z) = 1$. La formule de König-Huygens donne

$$V(Z) = E(Z^2) - (E(Z))^2 \quad \text{soit} \quad 1 = E(Z^2) - 0$$

d'où

$$E(Y) = E(Z^2) = 1$$

Moment d'ordre 4 de Z . Par le théorème de transfert appliqué à $z \mapsto z^4$,

$$E(Z^4) = \int_{-\infty}^{+\infty} z^4 \varphi(z) dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z^4 e^{-z^2/2} dz$$

Convergence. La fonction intégrée est continue et positive sur $\mathbb{R}$, et par croissances comparées $z^4 e^{-z^2/2} = o\left(\frac{1}{z^2}\right)$ au voisinage de $+\infty$ comme de $-\infty$. Par comparaison avec une intégrale de Riemann convergente, l'intégrale converge, et absolument puisque l'intégrande est positif. Le moment d'ordre 4 existe donc.

Calcul par intégration par parties. Tout repose sur une identité que l'on doit reconnaître immédiatement :

$$\varphi'(z) = -z \varphi(z)$$

qui se vérifie en dérivant $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$. On écrit alors $z^4 \varphi(z) = z^3 \times (z \varphi(z))$ et on intègre par parties sur un segment $[-A, A]$, avec

$$u(z) = z^3, \quad v'(z) = z \varphi(z) = -\varphi'(z), \quad \text{donc} \quad v(z) = -\varphi(z)$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[-A, A]$, l'intégration par parties est licite :

$$\int_{-A}^A z^4 \varphi(z) dz = \left[-z^3 \varphi(z) \right]_{-A}^A + 3 \int_{-A}^A z^2 \varphi(z) dz$$

Évaluons le crochet, en utilisant la parité de φ , qui donne $\varphi(-A) = \varphi(A)$:

$$\left[-z^3 \varphi(z) \right]_{-A}^A = -A^3 \varphi(A) + (-A)^3 \varphi(-A) = -A^3 \varphi(A) - A^3 \varphi(A) = -2A^3 \varphi(A)$$

Or $A^3 \varphi(A) = \frac{1}{\sqrt{2\pi}} A^3 e^{-A^2/2} \rightarrow 0$ quand $A \rightarrow +\infty$, par croissances comparées (l'exponentielle en $-A^2/2$ l'emporte largement sur la puissance A^3). Le crochet tend donc vers 0.

Quant à la seconde intégrale, elle tend vers $\int_{-\infty}^{+\infty} z^2 \varphi(z) dz = E(Z^2) = 1$. En faisant tendre A vers $+\infty$:

$$E(Z^4) = 0 + 3 \times 1 = 3$$

Variance de Y . Comme $Y^2 = Z^4$, le moment d'ordre 2 de Y existe et vaut 3. Par König-Huygens,

$$V(Y) = E(Y^2) - (E(Y))^2 = E(Z^4) - 1^2 = 3 - 1 = 2$$

Contrôle par la loi gamma. On sait qu'une variable de loi $\gamma(\nu)$ a pour espérance ν et pour variance ν . Avec $T \hookrightarrow \gamma(\frac{1}{2})$ et $Y = 2T$, la linéarité de l'espérance et la formule $V(aX) = a^2V(X)$ donnent

$$E(Y) = 2 \times \frac{1}{2} = 1 \quad \text{et} \quad V(Y) = 2^2 \times \frac{1}{2} = 2$$

Les deux valeurs coïncident avec celles obtenues directement, ce qui valide à la fois le calcul de $E(Z^4)$ et l'identification de la loi de T .

Méthode à retenir. La loi d'une transformée $g(X)$ se cherche **toujours** par la fonction de répartition : on écrit $F_{g(X)}(y) = P(g(X) \leq y)$, on traduit l'inégalité en une condition sur X , on exprime le résultat à l'aide de F_X , puis on dérive. Quand g n'est pas injective, comme ici avec le carré, la traduction produit un encadrement, donc une **différence** de deux valeurs de F_X . Deuxième réflexe, propre à la loi normale : $\varphi' = -z\varphi$ est l'outil universel pour calculer les moments par intégration par parties, et la relation $\Phi(-u) = 1 - \Phi(u)$ sert dès qu'un argument négatif apparaît.

Erreur classique. Écrire $F_Y(y) = P(Z \leq \sqrt{y}) = \Phi(\sqrt{y})$ en oubliant la branche négative. On obtient alors une « densité » deux fois trop petite, dont l'intégrale vaut $\frac{1}{2}$. Autre erreur, de calcul cette fois : dériver $\Phi(\sqrt{y})$ en $\varphi(\sqrt{y})$ sans le facteur $\frac{1}{2\sqrt{y}}$ venant de la dérivée de la racine.

Exercice 27 ★★★ Moments d'une loi gamma

Notions : Lois gamma à un paramètre et fonction Gamma d'Euler, Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Variance, écart-type et moments d'une variable à densité

Soit $\nu > 0$. La variable aléatoire X suit la loi $\gamma(\nu)$, c'est-à-dire qu'une densité de X est

$$f(x) = \frac{x^{\nu-1} e^{-x}}{\Gamma(\nu)} \quad \text{sur }]0, +\infty[, \quad f(x) = 0 \quad \text{ailleurs}$$

On rappelle que $\Gamma(a) = \int_0^{+\infty} t^{a-1} e^{-t} dt$ pour tout $a > 0$, et que $\Gamma(a+1) = a\Gamma(a)$.

1. Montrer que pour tout entier $k \geq 1$, le moment $E(X^k)$ existe et vaut

$$E(X^k) = \frac{\Gamma(\nu+k)}{\Gamma(\nu)} = \nu(\nu+1)\cdots(\nu+k-1)$$

2. En déduire $E(X)$, $E(X^2)$ et $V(X)$.

3. Montrer que $E\left(\frac{1}{X}\right)$ existe si et seulement si $\nu > 1$, et qu'elle vaut alors $\frac{1}{\nu-1}$.

4. Montrer que $E(e^{tX})$ existe si et seulement si $t < 1$, et qu'elle vaut alors $(1-t)^{-\nu}$.

5. Vérifier la cohérence de tous ces résultats pour $\nu = 1$.

Corrigé

1. Soit k un entier supérieur ou égal à 1. La fonction $x \mapsto x^k$ est continue sur $]0, +\infty[$, donc le théorème de transfert, admis par le programme, s'applique : le moment $E(X^k)$ existe si et seulement si l'intégrale ci-dessous converge absolument, et il vaut alors

$$E(X^k) = \int_0^{+\infty} x^k \frac{x^{\nu-1} e^{-x}}{\Gamma(\nu)} dx = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^{\nu+k-1} e^{-x} dx$$

Convergence. L'intégrande est continu et **positif** sur $]0, +\infty[$, donc la convergence de l'intégrale équivaut à sa convergence absolue : il n'y a qu'un seul point à examiner.

En 0 : comme $k \geq 1$ et $\nu > 0$, l'exposant vérifie $\nu+k-1 \geq \nu > 0$, donc $x^{\nu+k-1} e^{-x} \rightarrow 0$ quand x tend vers 0 par valeurs supérieures. L'intégrande se prolonge par continuité en 0, il n'y a donc aucun problème à cette borne.

En $+\infty$: par croissances comparées, $x^{\nu+k-1} e^{-x} = o\left(\frac{1}{x^2}\right)$, et l'intégrale de Riemann $\int_1^{+\infty} \frac{dx}{x^2}$ converge. Par comparaison de fonctions positives, l'intégrale converge.

Le moment $E(X^k)$ existe donc, et en posant $a = \nu+k > 0$ on reconnaît exactement la définition de $\Gamma(a)$:

$$E(X^k) = \frac{\Gamma(\nu+k)}{\Gamma(\nu)}$$

Forme développée. Montrons par récurrence sur $k \geq 1$ la propriété

$$\mathcal{P}(k) : \quad \Gamma(\nu+k) = \nu(\nu+1)\cdots(\nu+k-1)\Gamma(\nu)$$

Initialisation. Pour $k = 1$, la relation fonctionnelle $\Gamma(a+1) = a\Gamma(a)$ appliquée à $a = \nu$ donne $\Gamma(\nu+1) = \nu\Gamma(\nu)$, ce qui est bien $\mathcal{P}(1)$.

Hérédité. Supposons $\mathcal{P}(k)$ vraie pour un entier $k \geq 1$. En appliquant la relation fonctionnelle à $a = \nu+k$, qui est bien strictement positif :

$$\Gamma(\nu+k+1) = (\nu+k)\Gamma(\nu+k) = (\nu+k) \times \nu(\nu+1)\cdots(\nu+k-1)\Gamma(\nu)$$

ce qui est exactement $\mathcal{P}(k+1)$, le dernier facteur du produit étant $\nu + (k+1) - 1 = \nu + k$. La propriété est donc vraie pour tout $k \geq 1$. Comme $\Gamma(\nu) \neq 0$, on peut simplifier :

$$E(X^k) = \nu(\nu+1) \cdots (\nu+k-1) \quad \square$$

Ce produit compte exactement k facteurs, allant de ν à $\nu+k-1$.

2. Il suffit de spécialiser.

Pour $k=1$, le produit se réduit à son unique facteur :

$$E(X) = \nu$$

Pour $k=2$, il comporte deux facteurs :

$$E(X^2) = \nu(\nu+1)$$

La formule de König-Huygens donne alors

$$V(X) = E(X^2) - (E(X))^2 = \nu(\nu+1) - \nu^2 = \nu^2 + \nu - \nu^2 = \nu$$

On retiendra ce fait remarquable : pour une loi $\gamma(\nu)$, l'espérance et la variance sont **égales**, toutes deux égales au paramètre. L'écart-type vaut donc $\sqrt{\nu}$.

3. La fonction $x \mapsto \frac{1}{x}$ est continue sur $]0, +\infty[$, qui contient toutes les valeurs de X ayant une probabilité non nulle de se réaliser. Le théorème de transfert donne

$$E\left(\frac{1}{X}\right) = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} \frac{1}{x} \times x^{\nu-1} e^{-x} dx = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^{\nu-2} e^{-x} dx$$

sous réserve de convergence absolue, qui équivaut ici encore à la convergence, l'intégrande étant positif. L'intégrale est impropre aux deux bornes, et cette fois la borne 0 est le point délicat.

En $+\infty$. Par croissances comparées, $x^{\nu-2} e^{-x} = o\left(\frac{1}{x^2}\right)$ quel que soit ν . L'intégrale $\int_1^{+\infty} x^{\nu-2} e^{-x} dx$ converge donc **toujours**.

En 0. Comme $e^{-x} \rightarrow 1$ lorsque $x \rightarrow 0^+$, on a l'équivalence de fonctions positives

$$x^{\nu-2} e^{-x} \underset{x \rightarrow 0^+}{\sim} x^{\nu-2} = \frac{1}{x^{2-\nu}}$$

L'intégrale de Riemann $\int_0^1 \frac{dx}{x^\alpha}$ converge si et seulement si $\alpha < 1$. Ici $\alpha = 2 - \nu$, donc la condition s'écrit

$$2 - \nu < 1 \quad \Longleftrightarrow \quad \nu > 1$$

Par le critère d'équivalence pour les fonctions positives, $\int_0^1 x^{\nu-2} e^{-x} dx$ converge si et seulement si $\nu > 1$. Comme la borne $+\infty$ ne pose jamais problème,

$$E\left(\frac{1}{X}\right) \text{ existe} \quad \Longleftrightarrow \quad \nu > 1$$

Calcul de la valeur. Supposons $\nu > 1$. Alors $\nu - 1 > 0$, donc $\Gamma(\nu - 1)$ est bien défini, et en écrivant $x^{\nu-2} = x^{(\nu-1)-1}$ on reconnaît

$$\int_0^{+\infty} x^{(\nu-1)-1} e^{-x} dx = \Gamma(\nu - 1)$$

D'où, en utilisant la relation fonctionnelle sous la forme $\Gamma(\nu) = \Gamma((\nu-1) + 1) = (\nu-1)\Gamma(\nu-1)$:

$$E\left(\frac{1}{X}\right) = \frac{\Gamma(\nu-1)}{\Gamma(\nu)} = \frac{\Gamma(\nu-1)}{(\nu-1)\Gamma(\nu-1)} = \frac{1}{\nu-1} \quad \square$$

Interprétation de la divergence. Quand $\nu < 1$, la densité de X explose au voisinage de 0 ; pour $\nu = 1$ elle reste bornée, puisqu'elle vaut alors e^{-x} , mais la fonction $x \mapsto \frac{1}{x}$ n'en est pas moins non intégrable en 0. Dans les deux cas, la variable prend des valeurs très proches de 0 avec une probabilité trop grande pour que la moyenne de $\frac{1}{X}$, qui y devient énorme, reste finie. On notera aussi que $\frac{1}{\nu-1}$ tend vers $+\infty$ quand ν tend vers 1 par valeurs supérieures, ce qui est cohérent avec la disparition de l'espérance en $\nu = 1$. Enfin, cet exemple illustre une erreur de raisonnement fréquente : $E\left(\frac{1}{X}\right)$ n'a rien à voir avec $\frac{1}{E(X)} = \frac{1}{\nu}$, et les deux quantités diffèrent bel et bien.

4. Soit t un réel. La fonction $x \mapsto e^{tx}$ est continue sur $]0, +\infty[$, et le théorème de transfert donne

$$E(e^{tX}) = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} e^{tx} x^{\nu-1} e^{-x} dx = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} x^{\nu-1} e^{-(1-t)x} dx$$

sous réserve de convergence, l'intégrande étant à nouveau positif. Tout dépend du signe de $1-t$. Cas $t \geq 1$. Alors $1-t \leq 0$, donc $-(1-t)x \geq 0$ et $e^{-(1-t)x} \geq 1$ pour tout $x > 0$. Il en résulte

$$x^{\nu-1} e^{-(1-t)x} \geq x^{\nu-1}$$

Or $\int_1^{+\infty} x^{\nu-1} dx$ diverge, puisqu'une intégrale de Riemann $\int_1^{+\infty} \frac{dx}{x^\alpha}$ ne converge que pour $\alpha > 1$, et qu'ici $\alpha = 1-\nu < 1$. Par comparaison de fonctions positives, l'intégrale diverge, donc $E(e^{tX})$ **n'existe pas**. C'est parfaitement intuitif : l'exponentielle e^{tx} croît alors au moins aussi vite que e^{-x} décroît, et le produit ne tend pas vers 0.

Cas $t < 1$. Alors $1-t > 0$. Effectuons le changement de variable $u = (1-t)x$, c'est-à-dire $x = \frac{u}{1-t}$, avec $dx = \frac{du}{1-t}$. L'application $x \mapsto (1-t)x$ est de classe $\mathcal{C}^1$ et strictement croissante, et elle réalise une bijection de $]0, +\infty[$ sur lui-même. Sur un segment $[\varepsilon, A]$ puis par passage à la limite,

$$\begin{aligned} \int_0^{+\infty} x^{\nu-1} e^{-(1-t)x} dx &= \int_0^{+\infty} \left(\frac{u}{1-t}\right)^{\nu-1} e^{-u} \frac{du}{1-t} \\ &= \frac{1}{(1-t)^{\nu-1}} \times \frac{1}{1-t} \int_0^{+\infty} u^{\nu-1} e^{-u} du \\ &= \frac{\Gamma(\nu)}{(1-t)^\nu} \end{aligned}$$

La dernière intégrale est exactement $\Gamma(\nu)$, qui converge puisque $\nu > 0$: le changement de variable transporte donc la convergence, et l'intégrale de départ converge. En divisant par $\Gamma(\nu)$:

$$E(e^{tX}) = (1-t)^{-\nu} \quad \text{pour tout } t < 1 \quad \square$$

On vérifie au passage que la valeur obtenue est cohérente en $t = 0$: elle vaut 1, ce qui est bien $E(e^0) = E(1) = 1$.

5. Pour $\nu = 1$, la densité de X devient

$$f(x) = \frac{x^0 e^{-x}}{\Gamma(1)} = e^{-x} \quad \text{sur }]0, +\infty[$$

puisque $\Gamma(1) = 0! = 1$. C'est exactement la densité de la loi exponentielle de paramètre 1, ce qui confirme l'identité $\gamma(1) = \mathcal{E}(1)$ annoncée par le cours. Les résultats des questions précédentes doivent donc redonner ceux, bien connus, de la loi $\mathcal{E}(1)$.

Quantité	Valeur en $\nu = 1$	Loi $\mathcal{E}(1)$
$E(X)$	1	$\frac{1}{\lambda} = 1$
$V(X)$	1	$\frac{1}{\lambda^2} = 1$
$E(X^k)$	$1 \times 2 \times \dots \times k = k!$	$\frac{k!}{\lambda^k} = k!$
$E\left(\frac{1}{X}\right)$	n'existe pas	n'existe pas
$E(e^{tX})$	$\frac{1}{1-t}$	$\frac{1}{1-t}$

Détaillons les deux dernières lignes, qui sont les plus instructives.

L'espérance de $\frac{1}{X}$. La condition $\nu > 1$ n'est pas remplie pour $\nu = 1$, donc l'espérance n'existe pas. On le retrouve directement : pour la loi $\mathcal{E}(1)$, l'intégrande vaut $\frac{e^{-x}}{x} \sim \frac{1}{x}$ au voisinage de 0, et $\int_0^1 \frac{dx}{x}$ diverge. Les deux approches concordent.

L'espérance de e^{tX} . Pour $\nu = 1$ la formule donne $(1-t)^{-1}$. Le calcul direct pour $\mathcal{E}(1)$ confirme, pour $t < 1$:

$$\int_0^{+\infty} e^{tx} e^{-x} dx = \int_0^{+\infty} e^{-(1-t)x} dx = \left[\frac{-e^{-(1-t)x}}{1-t} \right]_0^{+\infty} = \frac{1}{1-t}$$

Toutes les formules générales sont donc validées sur ce cas particulier. Cette vérification vaut d'être menée systématiquement : $\nu = 1$ est le seul paramètre pour lequel on connaît la loi sous un autre nom, c'est donc le seul point de contrôle indépendant dont on dispose.

Méthode à retenir. Tout se ramène à une reconnaissance de la fonction Γ : dès qu'une intégrale prend la forme $\int_0^{+\infty} x^{a-1} e^{-x} dx$ avec $a > 0$, elle vaut $\Gamma(a)$, et la relation $\Gamma(a+1) = a\Gamma(a)$ permet ensuite de simplifier les quotients de Γ . Pour l'existence d'un moment, la discipline est invariable : intégrande positif, donc convergence et convergence absolue coïncident ; on examine séparément chaque borne impropre, en comparant à une intégrale de Riemann. En $+\infty$, l'exponentielle règle toujours l'affaire par croissances comparées ; en 0, tout se joue sur l'exposant de x , avec le seuil 1.

Erreur classique. Écrire $E\left(\frac{1}{X}\right) = \frac{1}{E(X)}$, c'est-à-dire $\frac{1}{\nu}$ au lieu de $\frac{1}{\nu-1}$. L'espérance n'est pas compatible avec le passage à l'inverse, et l'écart est ici visible à l'œil nu : pour $\nu = 2$, l'espérance de $\frac{1}{X}$ vaut 1, alors que $\frac{1}{E(X)}$ vaut $\frac{1}{2}$. Autre erreur : oublier de discuter l'existence et calculer $E(e^{tX})$ pour tout t réel, alors que la formule $(1-t)^{-\nu}$ n'a aucun sens pour $t \geq 1$, ce que la divergence de l'intégrale confirme.

Exercice 28 ★★★ Comparer deux temps d'attente indépendants

Notions : Indépendance de deux variables aléatoires réelles, lemme des coalitions, Loi exponentielle et absence de mémoire, Espérance d'une variable à densité : existence, théorème de transfert, linéarité

Deux tâches sont lancées au même instant, dans deux files séparées et sans influence l'une sur l'autre. On note X la durée de la première et Y celle de la seconde, exprimées dans la même unité de temps. On suppose que $X \hookrightarrow \mathcal{E}(\lambda)$ et $Y \hookrightarrow \mathcal{E}(\mu)$, avec $\lambda > 0$ et $\mu > 0$, et que X et Y sont indépendantes.

L'objectif est de calculer la probabilité que la première tâche se termine avant la seconde.

On **admet** le résultat suivant, dont la justification sort du programme : si X et Y sont indépendantes et si Y admet une densité f_Y , alors

$$P(X < Y) = \int_{-\infty}^{+\infty} F_X(t) f_Y(t) dt$$

sous réserve de convergence de cette intégrale impropre. Cette formule sera utilisée telle quelle dans tout l'exercice, sans être redémontrée.

1. Rappeler l'expression de $F_X(t)$ et celle de $f_Y(t)$ pour tout réel t . En déduire que l'intégrale ci-dessus se réduit à une intégrale sur $[0, +\infty[$, justifier sa convergence, puis montrer que

$$P(X < Y) = \frac{\lambda}{\lambda + \mu}$$

2. Donner sans nouveau calcul la valeur de $P(Y < X)$. Vérifier que $P(X < Y) + P(Y < X) = 1$ et en déduire la valeur de $P(X = Y)$. Que vaut $P(X < Y)$ lorsque $\lambda = \mu$, et pourquoi était-ce prévisible ?
3. On pose $Z = 2Y$.
 - a. Déterminer la loi de Z par la méthode de la fonction de répartition.
 - b. Justifier que X et Z sont indépendantes.
 - c. En déduire la valeur de $P(X < 2Y)$.
4. Deux guichets traitent chacun un dossier, de façon indépendante. Le temps de traitement du premier guichet suit une loi exponentielle de durée moyenne 4 minutes, celui du second une loi exponentielle de durée moyenne 6 minutes. Calculer la probabilité que le premier guichet termine avant le second, puis la probabilité que le premier termine avant le double du temps mis par le second.
5. On appelle **taux** d'une loi exponentielle le paramètre λ , qui est l'inverse de la durée moyenne. Interpréter le résultat $\frac{\lambda}{\lambda + \mu}$ à l'aide de ce vocabulaire, puis le réécrire en fonction des seules durées moyennes $a = \frac{1}{\lambda}$ et $b = \frac{1}{\mu}$.

Corrigé

1. La stratégie est simple : on injecte dans la formule admise les deux expressions du cours, on repère où l'intégrande est nul, et on se ramène à des intégrales exponentielles élémentaires. La fonction de répartition d'une loi exponentielle de paramètre λ vaut $F_X(t) = 1 - e^{-\lambda t}$ pour $t \geq 0$, et $F_X(t) = 0$ pour $t < 0$. Une densité de Y est donnée par $f_Y(t) = \mu e^{-\mu t}$ pour $t \geq 0$, et $f_Y(t) = 0$ pour $t < 0$.

Pour $t < 0$, les deux facteurs sont nuls, donc le produit $F_X(t)f_Y(t)$ est nul. L'intégrale sur $\mathbb{R}$ se réduit donc à

$$P(X < Y) = \int_0^{+\infty} (1 - e^{-\lambda t}) \mu e^{-\mu t} dt$$

Convergence. L'intégrande est positif sur $[0, +\infty[$, et il est majoré par $f_Y(t)$ puisque $0 \leq F_X(t) \leq 1$. Or $\int_0^{+\infty} f_Y(t) dt$ converge et vaut 1. Par comparaison de fonctions positives, l'intégrale étudiée converge.

Calcul. Pour $A > 0$, on développe le produit et on sépare en deux intégrales, chacune convergente :

$$\begin{aligned} \int_0^A (1 - e^{-\lambda t}) \mu e^{-\mu t} dt &= \int_0^A \mu e^{-\mu t} dt - \int_0^A \mu e^{-(\lambda+\mu)t} dt \\ &= [-e^{-\mu t}]_0^A - \frac{\mu}{\lambda + \mu} [-e^{-(\lambda+\mu)t}]_0^A \\ &= (1 - e^{-\mu A}) - \frac{\mu}{\lambda + \mu} (1 - e^{-(\lambda+\mu)A}) \end{aligned}$$

Comme $\lambda > 0$ et $\mu > 0$, les deux exponentielles tendent vers 0 quand A tend vers $+\infty$. Il reste

$$P(X < Y) = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda + \mu - \mu}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}$$

2. Les hypothèses sont symétriques en X et Y : il suffit d'échanger les rôles de λ et de μ dans la question précédente, ce qui donne

$$P(Y < X) = \frac{\mu}{\lambda + \mu}$$

La somme vaut

$$P(X < Y) + P(Y < X) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} = \frac{\lambda + \mu}{\lambda + \mu} = 1$$

Les trois événements $(X < Y)$, $(X = Y)$ et $(Y < X)$ sont deux à deux incompatibles et leur réunion est l'univers tout entier. Leurs probabilités ont donc pour somme 1, d'où

$$P(X = Y) = 1 - (P(X < Y) + P(Y < X)) = 0$$

Ce résultat est cohérent avec le cours : pour des variables à densité, l'égalité exacte de deux durées est un événement de probabilité nulle, sans être pour autant impossible.

Si $\lambda = \mu$, alors $P(X < Y) = \frac{\lambda}{2\lambda} = \frac{1}{2}$. C'était prévisible : les deux durées suivent alors la même loi, aucune des deux tâches n'est avantagée, et la probabilité qu'elles soient exactement égales est nulle.

3. a. On applique la méthode de la fonction de répartition, qui est la seule méthode au programme pour transformer une variable à densité. La variable Y étant à valeurs positives, il en va de même de $Z = 2Y$, donc $F_Z(t) = 0$ pour $t < 0$. Pour $t \geq 0$:

$$F_Z(t) = P(2Y \leq t) = P\left(Y \leq \frac{t}{2}\right) = F_Y\left(\frac{t}{2}\right) = 1 - e^{-\mu t/2}$$

On reconnaît la fonction de répartition d'une loi exponentielle de paramètre $\frac{\mu}{2}$, donc

$$Z \hookrightarrow \mathcal{E}\left(\frac{\mu}{2}\right)$$

Ce résultat s'écrit aussi directement avec la règle du cours sur les transformations affines d'une loi exponentielle : multiplier une durée par 2 divise son taux par 2, et double bien la durée moyenne, qui passe de $\frac{1}{\mu}$ à $\frac{2}{\mu}$.

b. La variable Z ne dépend que de Y , puisque $Z = g(Y)$ avec $g : y \mapsto 2y$. Le lemme des coalitions, admis par le programme, assure alors que X et Z sont indépendantes, car X et Y le sont.

c. Les événements $(X < 2Y)$ et $(X < Z)$ sont le même événement. On applique donc la question 1. au couple (X, Z) , qui vérifie exactement les mêmes hypothèses avec μ remplacé par $\frac{\mu}{2}$:

$$P(X < 2Y) = P(X < Z) = \frac{\lambda}{\lambda + \frac{\mu}{2}} = \frac{2\lambda}{2\lambda + \mu}$$

la dernière égalité s'obtenant en multipliant numérateur et dénominateur par 2.

Ce résultat est plus grand que $\frac{\lambda}{\lambda + \mu}$, ce qui est rassurant : laisser au second guichet un temps deux fois plus long ne peut qu'augmenter les chances du premier de passer en tête.

4. Une loi exponentielle de paramètre λ a pour espérance $\frac{1}{\lambda}$. Une durée moyenne de 4 minutes donne donc $\lambda = \frac{1}{4}$, et une durée moyenne de 6 minutes donne $\mu = \frac{1}{6}$.

$$P(X < Y) = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{6}} = \frac{\frac{1}{4}}{\frac{3+2}{12}} = \frac{1}{4} \times \frac{12}{5} = \frac{3}{5} = 0,6$$

Le premier guichet termine avant le second avec une probabilité de 0,6, soit 60 % des cas. Avec la question 3. :

$$P(X < 2Y) = \frac{2 \times \frac{1}{4}}{2 \times \frac{1}{4} + \frac{1}{6}} = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{6}} = \frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4} = 0,75$$

5. Le paramètre λ est le taux de la première tâche : c'est le nombre moyen d'achèvements par unité de temps, autrement dit une mesure de sa rapidité. De même μ mesure la rapidité de la seconde. La formule

$$P(X < Y) = \frac{\lambda}{\lambda + \mu}$$

se lit alors ainsi : la probabilité que le premier gagne la course est **la part de sa rapidité dans la rapidité totale**. Deux tâches également rapides se partagent la course à parts égales, et une tâche deux fois plus rapide que l'autre gagne deux fois sur trois.

En fonction des durées moyennes $a = \frac{1}{\lambda}$ et $b = \frac{1}{\mu}$, on obtient

$$\frac{\lambda}{\lambda + \mu} = \frac{\frac{1}{a}}{\frac{1}{a} + \frac{1}{b}} = \frac{\frac{1}{a}}{\frac{a+b}{ab}} = \frac{b}{a+b}$$

La probabilité de gagner est donc la part de la durée moyenne de l'**autre** dans la somme des deux durées moyennes. Sur l'application numérique, cela redonne immédiatement $\frac{6}{4+6} = 0,6$, ce qui fournit un contrôle rapide du calcul de la question 4.

Méthode à retenir. Face à une comparaison de deux temps d'attente indépendants, on ne cherche jamais à décrire le couple (X, Y) : on applique la formule fournie, on réduit l'intégrale au domaine où l'intégrande est non nul, on justifie la convergence par comparaison avec une densité, et on intègre. Toute variante du type $P(X < 2Y)$, $P(X < Y + c)$ ou $P(3X < Y)$ se ramène au même calcul en posant une nouvelle variable dont on détermine d'abord la loi par la fonction de répartition, puis en invoquant le lemme des coalitions pour conserver l'indépendance.

Erreur classique. Intervertir les rôles de F et de f dans la formule admise. L'intégrale $\int_{-\infty}^{+\infty} f_X(t)F_Y(t) dt$ ne donne pas $P(X < Y)$ mais $P(Y < X)$, c'est-à-dire l'événement contraire : le résultat obtenu est alors $\frac{\mu}{\lambda + \mu}$, faux mais parfaitement vraisemblable, donc difficile à détecter. Le contrôle de la question 2., où les deux probabilités doivent avoir pour somme 1, et le test $\lambda = \mu$, qui doit donner $\frac{1}{2}$, permettent de repérer l'inversion. Autre erreur fréquente : croire que l'indépendance de X et Y entraîne à elle seule $P(X < Y) = \frac{1}{2}$, ce qui n'est vrai que si les deux variables suivent la même loi.

Exercice 29 ★★★ Moyenne d'un échantillon de variables normales

Notions : n-uplets de variables aléatoires, indépendance mutuelle, échantillon, Somme de variables indépendantes : produit de convolution et stabilité, Loi normale, centrage-réduction et lecture de la table, Covariance et variance d'une somme dans le cas général

Soit n un entier naturel non nul. On considère $X_1, X_2, \dots, X_n$ des variables aléatoires **mutuellement indépendantes**, suivant toutes la même loi $\mathcal{N}(m, \sigma^2)$, avec m réel et $\sigma > 0$. Un tel n -uplet est appelé échantillon de taille n de la loi $\mathcal{N}(m, \sigma^2)$.

On pose

$$S_n = \sum_{i=1}^n X_i \quad \text{et} \quad M_n = \frac{S_n}{n}$$

Les valeurs suivantes de la fonction de répartition Φ de la loi normale centrée réduite sont fournies :

$$\Phi(1) = 0,8413 \quad \Phi(1,96) = 0,9750 \quad \Phi(2) = 0,9772$$

1. Démontrer par récurrence sur n que $S_n \hookrightarrow \mathcal{N}(nm, n\sigma^2)$. On précisera à chaque étape le résultat du cours utilisé et l'endroit exact où sert l'indépendance.
2. Calculer $E(M_n)$ et $V(M_n)$. Pour chacun des deux calculs, indiquer si l'hypothèse d'indépendance est nécessaire.
3. En déduire la loi de M_n .
4. On prend $m = 100$, $\sigma = 15$ et $n = 25$. Calculer $P(97 \leq M_n \leq 103)$.
5. On garde $m = 100$ et $\sigma = 15$, mais n est maintenant inconnu. Déterminer le plus petit entier n tel que

$$P(|M_n - m| \leq 2) \geq 0,95$$

6. Exprimer l'écart-type de M_n en fonction de n . Combien faut-il multiplier la taille de l'échantillon pour diviser par 2 la dispersion de M_n autour de m ? Commenter le coût pratique de cette exigence.

Corrigé

1. Le résultat du cours à mobiliser est la **stabilité de la loi normale par somme de variables indépendantes**, admise par le programme : si $U \hookrightarrow \mathcal{N}(m_1, \sigma_1^2)$ et $V \hookrightarrow \mathcal{N}(m_2, \sigma_2^2)$ sont indépendantes, alors

$$U + V \hookrightarrow \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$$

Notons $\mathcal{H}_n$ la proposition « $S_n \hookrightarrow \mathcal{N}(nm, n\sigma^2)$ ».

Initialisation. Pour $n = 1$, on a $S_1 = X_1 \hookrightarrow \mathcal{N}(m, \sigma^2)$, ce qui est bien $\mathcal{N}(1 \times m, 1 \times \sigma^2)$. Donc $\mathcal{H}_1$ est vraie.

Hérédité. Supposons $\mathcal{H}_n$ vraie pour un entier $n \geq 1$. On écrit

$$S_{n+1} = S_n + X_{n+1}$$

La variable S_n ne dépend que de $X_1, \dots, X_n$, et X_{n+1} est indépendante de ce n -uplet. Le **lemme des coalitions**, admis par le programme, garantit donc que S_n et X_{n+1} sont indépendantes. C'est précisément ici que sert l'hypothèse d'indépendance mutuelle : sans elle, la stabilité ne s'applique pas.

La stabilité donne alors

$$S_{n+1} \hookrightarrow \mathcal{N}(nm + m, n\sigma^2 + \sigma^2) = \mathcal{N}((n+1)m, (n+1)\sigma^2)$$

ce qui est exactement $\mathcal{H}_{n+1}$.

Conclusion. Par principe de récurrence, pour tout entier $n \geq 1$, $S_n \hookrightarrow \mathcal{N}(nm, n\sigma^2)$. $\square$

2. L'espérance est **linéaire** dans tous les cas, sans aucune hypothèse d'indépendance :

$$E(M_n) = E\left(\frac{S_n}{n}\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \times nm = m$$

La moyenne de l'échantillon a donc la même espérance que chacune des variables qui la composent. Ce calcul resterait valable même si les X_i étaient liées.

Pour la variance, la situation est différente. La formule générale de la variance d'une somme s'écrit

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

Les variables étant indépendantes deux à deux, toutes les covariances sont nulles, et il reste $V(S_n) = n\sigma^2$. C'est **ici** que l'indépendance est indispensable. En utilisant ensuite $V(aX) = a^2V(X)$ avec $a = \frac{1}{n}$:

$$V(M_n) = \frac{1}{n^2} V(S_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

3. D'après la question **1.**, $S_n \hookrightarrow \mathcal{N}(nm, n\sigma^2)$. Or $M_n = \frac{1}{n}S_n$ est une transformation affine de S_n , et le cours affirme qu'une transformation affine non constante d'une variable normale suit encore une loi normale. Ses paramètres sont ceux calculés en question **2.**, donc

$$M_n \hookrightarrow \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$$

4. Avec $m = 100$, $\sigma = 15$ et $n = 25$, on obtient

$$V(M_n) = \frac{15^2}{25} = \frac{225}{25} = 9 \quad \text{donc} \quad \sigma(M_n) = 3$$

et $M_{25} \hookrightarrow \mathcal{N}(100, 9)$. Le réflexe est le **centrage-réduction** : on pose

$$T = \frac{M_{25} - 100}{3} \hookrightarrow \mathcal{N}(0, 1)$$

L'encadrement $97 \leq M_{25} \leq 103$ équivaut à $-3 \leq M_{25} - 100 \leq 3$, c'est-à-dire à $-1 \leq T \leq 1$.
Donc

$$P(97 \leq M_{25} \leq 103) = P(-1 \leq T \leq 1) = \Phi(1) - \Phi(-1)$$

La densité φ étant paire, on a $\Phi(-x) = 1 - \Phi(x)$, d'où

$$P(97 \leq M_{25} \leq 103) = 2\Phi(1) - 1 = 2 \times 0,8413 - 1 = 0,6826$$

La moyenne de 25 observations tombe donc à moins de 3 unités de m dans environ 68 % des cas.

5. La stratégie : traduire l'inégalité demandée en une condition sur Φ , puis utiliser la stricte croissance de Φ pour la transformer en une inégalité sur n .

D'après la question **3.**, $\sigma(M_n) = \frac{15}{\sqrt{n}}$, et la variable

$$T = \frac{M_n - 100}{\frac{15}{\sqrt{n}}} = \frac{\sqrt{n}(M_n - 100)}{15}$$

suit la loi $\mathcal{N}(0, 1)$. En posant $u = \frac{2\sqrt{n}}{15}$, l'événement $|M_n - 100| \leq 2$ s'écrit $|T| \leq u$, donc

$$P(|M_n - 100| \leq 2) = P(-u \leq T \leq u) = 2\Phi(u) - 1$$

La condition demandée devient

$$2\Phi(u) - 1 \geq 0,95 \iff \Phi(u) \geq 0,975$$

La table fournit $\Phi(1,96) = 0,9750$. Comme Φ est strictement croissante, cette condition équivaut à $u \geq 1,96$, soit

$$\frac{2\sqrt{n}}{15} \geq 1,96 \iff \sqrt{n} \geq \frac{1,96 \times 15}{2} = \frac{29,4}{2} = 14,7$$

La fonction carré étant croissante sur $[0, +\infty[$, on élève au carré :

$$n \geq 14,7^2 = 216,09$$

Le plus petit entier qui convient est donc

$$n = 217$$

Contrôle. Pour $n = 216$, on obtient $\sqrt{216} \approx 14,697$, donc $u \approx 1,9596 < 1,96$: la condition n'est pas atteinte. Pour $n = 217$, $\sqrt{217} \approx 14,731$ et $u \approx 1,9641 > 1,96$: elle l'est. La valeur 217 est donc bien la plus petite.

6. L'écart-type de M_n vaut

$$\sigma(M_n) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

Il décroît comme $\frac{1}{\sqrt{n}}$, et non comme $\frac{1}{n}$. Pour le diviser par 2, il faut donc que $\sqrt{n}$ soit multiplié par 2, c'est-à-dire que n **soit multiplié par 4**.

Sur l'exemple de la question **4.** : avec $n = 25$ observations, $\sigma(M_n) = 3$; il faut passer à $n = 100$ pour obtenir $\sigma(M_n) = 1,5$, puis à $n = 400$ pour descendre à 0,75. Le gain en précision est réel, mais son prix croît très vite : chaque division par 2 de la dispersion coûte quatre fois plus d'observations. C'est la raison pour laquelle, dans la pratique, on ne cherche pas à réduire indéfiniment la dispersion d'une moyenne.

Méthode à retenir. Dès qu'un énoncé propose une somme ou une moyenne de variables normales indépendantes, l'enchaînement est toujours le même : stabilité pour obtenir la loi de S_n , linéarité pour l'espérance, indépendance pour la variance, transformation affine pour passer de S_n à M_n , puis centrage-réduction pour tout calcul numérique. Retenir le résultat sous sa forme compacte : la moyenne d'un échantillon de taille n de la loi $\mathcal{N}(m, \sigma^2)$ suit $\mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$.

Erreur classique. Écrire $M_n \hookrightarrow \mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right)$ en confondant le second paramètre de la notation,

qui est la **variance** $\frac{\sigma^2}{n}$ et non l'écart-type. Cette confusion fausse tout le centrage-réduction qui suit. Deuxième erreur fréquente en question **5.** : oublier de repasser de $\sqrt{n}$ à n , ou arrondir 216,09 à 216 par réflexe d'arrondi, alors que l'inégalité est large et impose de prendre l'entier immédiatement **supérieur**.

Exercice 30 ★★★ Variance d'une somme de n variables

Notions : Covariance et variance d'une somme dans le cas général, n -uplets de variables aléatoires, indépendance mutuelle, échantillon, Indépendance de deux variables aléatoires réelles, lemme des coalitions

Soit n un entier supérieur ou égal à 2, et soient $X_1, X_2, \dots, X_n$ des variables aléatoires réelles admettant toutes une variance. On note $S = \sum_{i=1}^n X_i$.

On utilisera librement la linéarité de l'espérance, admise par le programme dans le cas général, ainsi que la formule de König-Huygens pour la covariance.

1. En posant $Y_i = X_i - E(X_i)$ pour tout i de $\llbracket 1, n \rrbracket$, démontrer la formule générale

$$V(S) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

2. Que devient cette formule lorsque les variables sont indépendantes deux à deux ? L'indépendance mutuelle est-elle nécessaire pour obtenir ce résultat ?
3. On suppose désormais que $V(X_i) = \sigma^2$ pour tout i , avec $\sigma > 0$, et que $\text{Cov}(X_i, X_j) = c$ pour tout couple (i, j) avec $i \neq j$.
 - a. Montrer que $V(S) = n\sigma^2 + n(n-1)c$.
 - b. En utilisant le fait qu'une variance est toujours positive, en déduire que nécessairement

$$c \geq -\frac{\sigma^2}{n-1}$$

- c. Exprimer le coefficient de corrélation $\rho(X_i, X_j)$ pour $i \neq j$, et retrouver la contrainte précédente sous la forme d'une minoration de ce coefficient.
4. Application numérique avec $n = 5$ et $\sigma^2 = 4$: calculer $V(S)$ pour $c = 1$, puis pour $c = -1$. Interpréter le second résultat.
 5. Soient X et Y deux variables admettant une variance, et a, b, c, d quatre réels.
 - a. Calculer $\text{Cov}(aX + b, cY + d)$ en fonction de $\text{Cov}(X, Y)$.
 - b. Calculer $V(X - Y)$ à l'aide de la formule de la question 1.
 - c. On suppose $a \neq 0$, $c \neq 0$, $V(X) > 0$ et $V(Y) > 0$. Montrer que

$$\rho(aX + b, cY + d) = \rho(X, Y) \quad \text{si } ac > 0, \quad \rho(aX + b, cY + d) = -\rho(X, Y) \quad \text{si } ac < 0$$

Corrigé

1. La stratégie consiste à se ramener à des variables centrées, pour lesquelles la variance et la covariance s'écrivent simplement comme des espérances de produits, puis à développer un carré de somme.

Posons $Y_i = X_i - E(X_i)$. Par linéarité de l'espérance, $E(Y_i) = 0$ pour tout i , et l'on dispose des deux identités

$$V(X_i) = E(Y_i^2) \quad \text{et} \quad \text{Cov}(X_i, X_j) = E(Y_i Y_j)$$

la seconde étant la formule de König-Huygens pour la covariance, écrite avec des variables centrées.

Par linéarité, $E(S) = \sum_{i=1}^n E(X_i)$, donc

$$S - E(S) = \sum_{i=1}^n (X_i - E(X_i)) = \sum_{i=1}^n Y_i$$

Le développement du carré d'une somme finie donne

$$\left(\sum_{i=1}^n Y_i \right)^2 = \sum_{i=1}^n \sum_{j=1}^n Y_i Y_j = \sum_{i=1}^n Y_i^2 + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} Y_i Y_j$$

où l'on a simplement isolé les n termes diagonaux $i = j$. En prenant l'espérance et en utilisant sa linéarité, valable pour une somme finie :

$$\begin{aligned} V(S) &= E((S - E(S))^2) \\ &= \sum_{i=1}^n E(Y_i^2) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} E(Y_i Y_j) \\ &= \sum_{i=1}^n V(X_i) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \text{Cov}(X_i, X_j) \end{aligned}$$

Il reste à regrouper la dernière somme. La covariance est **symétrique** : $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$. Chaque paire d'indices distincts $\{i, j\}$ apparaît donc exactement deux fois, une fois sous la forme (i, j) et une fois sous la forme (j, i) , avec la même valeur. D'où

$$\sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \text{Cov}(X_i, X_j) = 2 \sum_{\substack{1 \leq i < j \leq n}} \text{Cov}(X_i, X_j)$$

et finalement

$$V(S) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j) \quad \square$$

2. Si les variables sont indépendantes deux à deux, alors pour tout couple (i, j) avec $i \neq j$, les variables X_i et X_j sont indépendantes, donc $\text{Cov}(X_i, X_j) = 0$. La double somme disparaît et il reste

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i)$$

L'indépendance **mutuelle n'est pas nécessaire** : la formule de la question 1. ne fait intervenir que des covariances de couples, donc seule l'indépendance deux à deux est utilisée. C'est une remarque importante, car l'indépendance deux à deux est strictement plus faible que l'indépendance mutuelle. En revanche, dès qu'on veut la **loi** de la somme, par exemple par produit de convolution ou par stabilité, l'indépendance mutuelle redevient indispensable.

3. a. Il y a n termes dans la première somme, tous égaux à σ^2 . La seconde somme porte sur les paires $\{i, j\}$ avec $i < j$, dont le nombre est

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

et chacune apporte la même covariance c . Donc

$$V(S) = n\sigma^2 + 2 \times \frac{n(n-1)}{2} \times c = n\sigma^2 + n(n-1)c$$

b. Une variance est l'espérance d'un carré, donc elle est toujours positive ou nulle. Ainsi

$$n\sigma^2 + n(n-1)c \geq 0$$

Comme $n > 0$, on peut diviser par n sans changer le sens de l'inégalité :

$$\sigma^2 + (n-1)c \geq 0 \iff (n-1)c \geq -\sigma^2$$

Enfin $n-1 > 0$ puisque $n \geq 2$, donc en divisant par $n-1$:

$$c \geq -\frac{\sigma^2}{n-1}$$

Autrement dit, des variables ne peuvent pas être « toutes très négativement corrélées entre elles » : plus elles sont nombreuses, plus la covariance commune est contrainte à rester proche de 0 par valeurs négatives. Il n'y a en revanche aucune borne supérieure issue de ce calcul, la positivité de la variance ne donnant qu'une minoration.

c. Toutes les variances valent $\sigma^2 > 0$, donc $\sigma(X_i) = \sigma$ pour tout i et, pour $i \neq j$,

$$\rho(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sigma(X_i)\sigma(X_j)} = \frac{c}{\sigma^2}$$

La contrainte de la question **b.** se divise par $\sigma^2 > 0$ et devient

$$\rho(X_i, X_j) \geq -\frac{1}{n-1}$$

Cette écriture est plus parlante : pour $n = 2$ on retrouve la borne -1 déjà connue, pour $n = 5$ la corrélation commune ne peut pas descendre en dessous de $-0,25$, et pour n grand elle ne peut presque plus être négative.

4. Avec $n = 5$ et $\sigma^2 = 4$, la formule de la question **3. a.** donne $V(S) = 5 \times 4 + 5 \times 4 \times c = 20 + 20c$.

a. Pour $c = 1$: $V(S) = 20 + 20 = 40$.

b. Pour $c = -1$: $V(S) = 20 - 20 = 0$.

Dans le premier cas, la variance de la somme, égale à 40, est le double de la somme des variances, qui vaut 20. Les variables varient « dans le même sens », et leurs fluctuations s'ajoutent au lieu de se compenser.

Le second cas est le cas limite. La borne de la question **3. b.** vaut ici

$$-\frac{\sigma^2}{n-1} = -\frac{4}{4} = -1$$

et $c = -1$ est donc la plus petite covariance commune possible. Elle donne $V(S) = 0$: la somme S est une variable de variance nulle, c'est-à-dire une variable **presque sûrement constante**,

égale à son espérance $\sum_{i=1}^5 E(X_i)$. Les cinq variables se compensent exactement. On rencontre cette situation, par exemple, avec des variables décrivant la répartition d'un total fixé entre cinq postes : si le total est imposé, ce qui est gagné sur l'un est perdu sur les autres.

5. a. La covariance est **bilinéaire**, et la covariance d'une variable avec une constante est nulle, puisqu'une constante ne fluctue pas. Donc

$$\begin{aligned} \text{Cov}(aX + b, cY + d) &= \text{Cov}(aX, cY + d) \\ &= \text{Cov}(aX, cY) \\ &= ac \text{Cov}(X, Y) \end{aligned}$$

Un décalage b ou d n'a donc aucun effet : seuls les coefficients multiplicatifs comptent.

b. On applique la formule de la question 1. avec $n = 2$ aux variables X et $-Y$:

$$V(X - Y) = V(X) + V(-Y) + 2 \operatorname{Cov}(X, -Y)$$

Or $V(-Y) = (-1)^2 V(Y) = V(Y)$ et, d'après la question a. avec $a = 1$, $b = 0$, $c = -1$, $d = 0$, on a $\operatorname{Cov}(X, -Y) = -\operatorname{Cov}(X, Y)$. D'où

$$V(X - Y) = V(X) + V(Y) - 2 \operatorname{Cov}(X, Y)$$

Le signe change devant la covariance, mais les deux variances s'**ajoutent** toujours.

c. Les écarts-types se transforment selon $\sigma(aX + b) = |a| \sigma(X)$, la valeur absolue étant indispensable car un écart-type est positif. Comme $a \neq 0$, $c \neq 0$ et $\sigma(X)\sigma(Y) > 0$, le dénominateur est non nul et

$$\rho(aX + b, cY + d) = \frac{ac \operatorname{Cov}(X, Y)}{|a| \sigma(X) \times |c| \sigma(Y)} = \frac{ac}{|ac|} \times \frac{\operatorname{Cov}(X, Y)}{\sigma(X)\sigma(Y)} = \frac{ac}{|ac|} \rho(X, Y)$$

Le facteur $\frac{ac}{|ac|}$ vaut 1 si $ac > 0$ et -1 si $ac < 0$, ce qui donne exactement les deux formules annoncées. $\square$

Interprétation : le coefficient de corrélation ne dépend ni de l'origine choisie, ni de l'unité de mesure, contrairement à la covariance qui, elle, est multipliée par ac . Changer des euros en centimes multiplie une covariance par 10 000 et laisse la corrélation inchangée. Le signe de ρ ne se retourne que si l'une des deux variables est renversée, par exemple en remplaçant une note par un nombre de fautes.

Méthode à retenir. La variance d'une somme n'est jamais la somme des variances tant que la nullité des covariances n'a pas été établie. La démonstration générale tient en trois gestes : centrer les variables, développer le carré de la somme en séparant les termes diagonaux, puis regrouper les couples symétriques par un facteur 2. Ce même schéma sert pour toutes les variantes, $V(X + Y)$, $V(X - Y)$, ou la variance d'une combinaison linéaire quelconque.

Erreur classique. Écrire $V(S) = \sum_i V(X_i)$ dès que l'énoncé parle de plusieurs variables, sans avoir vérifié l'indépendance deux à deux. Autre confusion fréquente : croire que le facteur 2 de la formule pourrait être remplacé par un comptage sur tous les couples ordonnés, ce qui reviendrait à compter chaque covariance quatre fois. Enfin, dans la question 5. c., oublier la valeur absolue dans $\sigma(aX + b) = |a| \sigma(X)$ conduit à annoncer ρ inchangé même lorsque $ac < 0$, alors qu'une corrélation positive devient négative.

Exercice 31 ★★★ Un paramètre choisi au hasard

Notions : Espérance conditionnelle et formule de l'espérance totale, Loi exponentielle et absence de mémoire, Fonction de répartition d'une variable aléatoire réelle quelconque, Variable aléatoire à densité : définition et reconnaissance d'une densité

Une usine fabrique des pièces à l'aide de trois machines. La machine 1 produit 20 % des pièces, la machine 2 en produit 30 % et la machine 3 en produit 50 %. La durée de vie, exprimée en années, d'une pièce issue de la machine i suit la loi $\mathcal{E}(\lambda_i)$, avec

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 4$$

On prélève une pièce au hasard dans la production et on note X sa durée de vie. Pour i dans $\llbracket 1, 3 \rrbracket$, on note M_i l'événement « la pièce provient de la machine i ». Les événements M_1, M_2, M_3 forment un système complet d'événements de probabilités non nulles.

1. À l'aide de la formule des probabilités totales appliquée à l'événement $(X \leq t)$, montrer que pour tout réel $t \geq 0$,

$$F_X(t) = 1 - (0,2 e^{-t} + 0,3 e^{-2t} + 0,5 e^{-4t})$$

et préciser la valeur de $F_X(t)$ pour $t < 0$.

2. Justifier que X est une variable à densité et déterminer une densité f_X de X . Vérifier que $\int_{-\infty}^{+\infty} f_X(t) dt = 1$.
3. Montrer que $E(X)$ existe et la calculer par intégration directe. On rappelle que pour $a > 0$, l'intégrale $\int_0^{+\infty} t e^{-at} dt$ converge et vaut $\frac{1}{a^2}$.
4. Montrer que $E(X^2)$ existe et la calculer, puis en déduire $V(X)$ et $\sigma(X)$. On rappelle que pour $a > 0$, l'intégrale $\int_0^{+\infty} t^2 e^{-at} dt$ converge et vaut $\frac{2}{a^3}$.
5. Une pièce prélevée au hasard a duré plus d'un an. Quelle est la probabilité qu'elle provienne de la machine 1 ? On donnera les trois probabilités conditionnelles $P_{(X>1)}(M_i)$ à trois décimales.
6. Calculer $E(X | M_i)$ pour chaque machine, puis vérifier la formule de l'espérance totale. Comparer $E(X)$ à la moyenne arithmétique des trois espérances conditionnelles et commenter.

On donne les valeurs approchées suivantes : $e^{-1} \approx 0,367879$, $e^{-2} \approx 0,135335$, $e^{-4} \approx 0,018316$ et $\sqrt{0,386875} \approx 0,622$.

Corrigé

1. La stratégie est celle de tout énoncé où un paramètre est tiré au hasard : on conditionne par la machine, on utilise la loi exponentielle correspondante, puis on recolle par la formule des probabilités totales.

Le système (M_1, M_2, M_3) est complet et chacun de ces événements est de probabilité non nulle. Pour tout réel t , la formule des probabilités totales donne

$$P(X \leq t) = \sum_{i=1}^3 P(M_i) P_{M_i}(X \leq t)$$

Sachant M_i , la durée de vie suit $\mathcal{E}(\lambda_i)$, donc pour $t \geq 0$ on a $P_{M_i}(X \leq t) = 1 - e^{-\lambda_i t}$. Ainsi,

pour $t \geq 0$:

$$\begin{aligned} F_X(t) &= 0,2(1 - e^{-t}) + 0,3(1 - e^{-2t}) + 0,5(1 - e^{-4t}) \\ &= (0,2 + 0,3 + 0,5) - (0,2e^{-t} + 0,3e^{-2t} + 0,5e^{-4t}) \\ &= 1 - (0,2e^{-t} + 0,3e^{-2t} + 0,5e^{-4t}) \end{aligned}$$

Pour $t < 0$, chaque probabilité conditionnelle $P_{M_i}(X \leq t)$ est nulle, car une loi exponentielle est à valeurs positives. Donc $F_X(t) = 0$ pour $t < 0$.

2. Il faut revenir à la définition du cours : X est à densité si F_X est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points.

Sur $]-\infty, 0[$, F_X est nulle, donc de classe $\mathcal{C}^1$. Sur $]0, +\infty[$, F_X est une combinaison linéaire de fonctions exponentielles, donc de classe $\mathcal{C}^1$. En 0, la limite à gauche vaut 0 et

$$F_X(0) = 1 - (0,2 + 0,3 + 0,5) = 0$$

donc F_X est continue en 0. Elle est donc continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf éventuellement en 0 : X est bien une variable à densité.

Une densité s'obtient en dérivant F_X là où c'est possible, et en attribuant une valeur arbitraire au point exceptionnel. Pour $t > 0$:

$$f_X(t) = F'_X(t) = 0,2e^{-t} + 0,6e^{-2t} + 2e^{-4t}$$

et $f_X(t) = 0$ pour $t \leq 0$. On remarque que ces coefficients sont exactement $0,2 \times \lambda_1$, $0,3 \times \lambda_2$ et $0,5 \times \lambda_3$: la densité de X est la moyenne pondérée des trois densités exponentielles, pondérée par les proportions de production.

Vérification. La fonction f_X est positive, continue sauf en 0, et pour $A > 0$:

$$\int_0^A f_X(t) dt = 0,2(1 - e^{-A}) + 0,3(1 - e^{-2A}) + 0,5(1 - e^{-4A}) \xrightarrow{A \rightarrow +\infty} 0,2 + 0,3 + 0,5 = 1$$

L'intégrale converge et vaut 1, ce qui confirme que f_X est bien une densité.

3. L'existence de $E(X)$ se justifie **avant** tout calcul. La fonction $t \mapsto tf_X(t)$ est nulle sur $]-\infty, 0[$ et positive sur $]0, +\infty[$: la convergence absolue équivaut donc à la convergence de $\int_0^{+\infty} tf_X(t) dt$. Or

$$tf_X(t) = 0,2te^{-t} + 0,6te^{-2t} + 2te^{-4t}$$

est une combinaison linéaire de trois fonctions dont les intégrales sur $[0, +\infty[$ convergent, d'après le rappel de l'énoncé. Donc $E(X)$ existe, et par linéarité de l'intégrale :

$$\begin{aligned} E(X) &= 0,2 \times \frac{1}{1^2} + 0,6 \times \frac{1}{2^2} + 2 \times \frac{1}{4^2} \\ &= 0,2 + 0,6 \times 0,25 + 2 \times 0,0625 \\ &= 0,2 + 0,15 + 0,125 \\ &= 0,475 \end{aligned}$$

La durée de vie moyenne d'une pièce prise au hasard est donc de 0,475 année, soit environ 5,7 mois.

4. Le même argument s'applique à $t \mapsto t^2 f_X(t)$, positive sur $]0, +\infty[$ et combinaison linéaire de fonctions $t \mapsto t^2 e^{-at}$ intégrables sur $[0, +\infty[$. Le moment d'ordre 2 existe, et le théorème de transfert donne

$$\begin{aligned} E(X^2) &= 0,2 \times \frac{2}{1^3} + 0,6 \times \frac{2}{2^3} + 2 \times \frac{2}{4^3} \\ &= 0,2 \times 2 + 0,6 \times 0,25 + 2 \times 0,03125 \\ &= 0,4 + 0,15 + 0,0625 \\ &= 0,6125 \end{aligned}$$

La formule de König-Huygens donne alors

$$V(X) = E(X^2) - (E(X))^2 = 0,6125 - 0,475^2 = 0,6125 - 0,225625 = 0,386875$$

et $\sigma(X) = \sqrt{0,386875} \approx 0,622$ année. L'écart-type est du même ordre que l'espérance, ce qui est caractéristique d'un mélange de lois exponentielles : la durée de vie est très dispersée.

5. Il s'agit d'une inversion de conditionnement, donc de la formule de Bayes. On calcule d'abord chaque terme du dénominateur, à savoir $P(M_i) P_{M_i}(X > 1) = P(M_i) e^{-\lambda_i}$:

$$\text{a. } 0,2 e^{-1} \approx 0,073576 \quad \text{b. } 0,3 e^{-2} \approx 0,040601 \quad \text{c. } 0,5 e^{-4} \approx 0,009158$$

en utilisant $e^{-1} \approx 0,367879$, $e^{-2} \approx 0,135335$ et $e^{-4} \approx 0,018316$. La formule des probabilités totales donne

$$P(X > 1) = 1 - F_X(1) = 0,2 e^{-1} + 0,3 e^{-2} + 0,5 e^{-4} \approx 0,123334$$

La formule de Bayes fournit alors

$$P_{(X>1)}(M_1) = \frac{P(M_1) P_{M_1}(X > 1)}{P(X > 1)} = \frac{0,073576}{0,123334} \approx 0,597$$

et de même

$$P_{(X>1)}(M_2) = \frac{0,040601}{0,123334} \approx 0,329 \quad P_{(X>1)}(M_3) = \frac{0,009158}{0,123334} \approx 0,074$$

La somme de ces trois valeurs vaut $0,597 + 0,329 + 0,074 = 1$, ce qui valide les calculs.

Le résultat est spectaculaire : avant toute information, la machine 1 n'expliquait que 20 % de la production ; sachant que la pièce a tenu plus d'un an, elle en explique près de 60 %. Une longue durée de vie est un indice fort en faveur de la machine la plus fiable, c'est-à-dire celle dont le paramètre λ est le plus petit.

6. Sachant M_i , la variable X suit $\mathcal{E}(\lambda_i)$, dont l'espérance vaut $\frac{1}{\lambda_i}$. Donc

$$E(X | M_1) = 1, \quad E(X | M_2) = \frac{1}{2} = 0,5, \quad E(X | M_3) = \frac{1}{4} = 0,25$$

La formule de l'espérance totale s'écrit

$$E(X) = \sum_{i=1}^3 P(M_i) E(X | M_i) = 0,2 \times 1 + 0,3 \times 0,5 + 0,5 \times 0,25 = 0,2 + 0,15 + 0,125 = 0,475$$

On retrouve exactement la valeur obtenue en question 3. par intégration directe, ce qui confirme les deux calculs. C'est le contrôle à faire systématiquement dans ce type d'exercice : l'intégration et le conditionnement doivent donner le même nombre.

La moyenne **arithmétique** des trois espérances conditionnelles, elle, vaut

$$\frac{1 + 0,5 + 0,25}{3} = \frac{1,75}{3} \approx 0,583$$

Elle est nettement supérieure à $E(X) = 0,475$. L'écart s'explique par les poids : la machine 3, dont les pièces ne durent en moyenne que 0,25 année, fournit à elle seule la moitié de la production, tandis que la machine 1, la plus fiable, n'en fournit qu'un cinquième. La moyenne réelle est donc tirée vers le bas par les machines rapides à défaillir. Traiter les trois machines à égalité reviendrait à surestimer de plus de 20 % la durée de vie moyenne du stock.

Méthode à retenir. Lorsqu'un paramètre est tiré au hasard, on ne cherche jamais la densité en premier : on écrit d'abord la **fonction de répartition** par la formule des probabilités totales

appliquée à $(X \leq t)$, puis on dérive. Les espérances se calculent ensuite soit par intégration de la densité, soit par la formule de l'espérance totale, et les deux voies doivent concorder. Toute question du type « sachant que la pièce a duré plus de t , d'où vient-elle ? » relève de la formule de Bayes, avec $P(X > t) = 1 - F_X(t)$ au dénominateur.

Erreur classique. Croire qu'un mélange de lois exponentielles est encore une loi exponentielle, et écrire $X \hookrightarrow \mathcal{E}(\bar{\lambda})$ avec $\bar{\lambda} = 0,2 \times 1 + 0,3 \times 2 + 0,5 \times 4 = 2,8$. Cela donnerait $E(X) = \frac{1}{2,8} \approx 0,357$, très loin de la vraie valeur 0,475. Le mélange porte sur les **fonctions de répartition**, pas sur les paramètres : ce sont les probabilités qui se moyennent, jamais les λ . Deuxième erreur : appliquer l'absence de mémoire à X . Chaque loi conditionnelle est sans vieillissement, mais le mélange, lui, vieillit, puisqu'observer une longue durée de vie modifie la probabilité de provenance de la pièce, comme le montre la question 5.

Exercice 32 ★★★★★ Caractérisation de la loi exponentielle par l'absence de mémoire

Notions : Loi exponentielle et absence de mémoire, Fonction de répartition d'une variable aléatoire réelle quelconque

Le cours établit qu'une variable de loi exponentielle est sans vieillissement. Cet exercice démontre la réciproque : c'est la **seule** loi à densité qui possède cette propriété.

Soit X une variable aléatoire à densité, à valeurs strictement positives, telle que

$$P(X > t) > 0 \quad \text{pour tout réel } t \geq 0$$

et vérifiant la propriété d'**absence de mémoire** : pour tous réels $s \geq 0$ et $t \geq 0$,

$$P_{(X>s)}(X > s+t) = P(X > t)$$

On pose, pour tout réel $t \geq 0$, $G(t) = P(X > t)$.

1. Premières propriétés de G .

a. Exprimer G à l'aide de F_X , puis justifier que G est continue sur $[0, +\infty[$, décroissante, que $G(0) = 1$ et que $\lim_{t \rightarrow +\infty} G(t) = 0$.

b. Montrer que l'inclusion $(X > s+t) \subset (X > s)$ est vraie pour $s \geq 0$ et $t \geq 0$, puis en déduire l'équation fonctionnelle

$$G(s+t) = G(s)G(t)$$

2. Résolution sur les rationnels.

a. Montrer par récurrence que pour tout entier naturel n , $G(n) = G(1)^n$.

b. Soit q un entier naturel non nul. Montrer, par une récurrence analogue, que $G\left(\frac{k}{q}\right) = G\left(\frac{1}{q}\right)^k$ pour tout entier naturel k . En prenant $k = q$, en déduire que $G\left(\frac{1}{q}\right) = G(1)^{1/q}$.

c. En déduire que $G(r) = G(1)^r$ pour tout rationnel positif r .

3. Passage aux réels.

a. Justifier que $0 < G(1) < 1$.

b. On **admet** que tout réel positif est limite d'une suite de rationnels positifs (densité de $\mathbb{Q}$ dans $\mathbb{R}$). En utilisant la continuité de G et celle de la fonction $x \mapsto G(1)^x$, montrer que

$$G(t) = G(1)^t \quad \text{pour tout réel } t \geq 0$$

4. **Conclusion.** On pose $\lambda = -\ln(G(1))$. Justifier que $\lambda > 0$, puis déterminer $F_X(t)$ pour tout réel t et conclure que $X \hookrightarrow \mathcal{E}(\lambda)$.

5. Énoncer en une phrase le théorème que l'on vient de démontrer, en le rapprochant du résultat du cours sur la loi exponentielle.

Corrigé

1. a. Par définition de la fonction de répartition, $(X > t)$ est l'événement contraire de $(X \leq t)$,

donc

$$G(t) = 1 - F_X(t)$$

Continuité. La variable X étant à densité, sa fonction de répartition F_X est continue sur $\mathbb{R}$. Donc G , qui en diffère par une constante et un signe, est continue sur $\mathbb{R}$, et en particulier sur $[0, +\infty[$.

Décroissance. F_X est croissante, donc $G = 1 - F_X$ est décroissante.

Valeur en 0. La variable X ne prend que des valeurs strictement positives, donc $P(X \leq 0) = 0$, c'est-à-dire $F_X(0) = 0$ et $G(0) = 1$.

Limite en $+\infty$. Le cours donne $\lim_{t \rightarrow +\infty} F_X(t) = 1$, donc $\lim_{t \rightarrow +\infty} G(t) = 0$.

b. Si ω est un événement élémentaire tel que $X(\omega) > s + t$, alors, comme $t \geq 0$, on a $s + t \geq s$ et donc $X(\omega) > s$. L'inclusion $(X > s + t) \subset (X > s)$ est donc établie, et par conséquent

$$(X > s + t) \cap (X > s) = (X > s + t)$$

Par hypothèse $P(X > s) = G(s) > 0$, la probabilité conditionnelle est donc bien définie et

$$P_{(X>s)}(X > s + t) = \frac{P((X > s + t) \cap (X > s))}{P(X > s)} = \frac{G(s + t)}{G(s)}$$

L'absence de mémoire s'écrit alors $\frac{G(s + t)}{G(s)} = G(t)$, et en multipliant par $G(s) \neq 0$:

$$G(s + t) = G(s) G(t)$$

2. a. Notons $\mathcal{H}_n$ la proposition « $G(n) = G(1)^n$ ».

Initialisation. Pour $n = 0$, $G(0) = 1$ d'après la question **1. a.**, et $G(1)^0 = 1$. Donc $\mathcal{H}_0$ est vraie.

Hérédité. Supposons $\mathcal{H}_n$ vraie pour un entier $n \geq 0$. L'équation fonctionnelle appliquée à $s = n$ et $t = 1$, tous deux positifs, donne

$$G(n + 1) = G(n) G(1) = G(1)^n \times G(1) = G(1)^{n+1}$$

d'où $\mathcal{H}_{n+1}$.

Conclusion. Pour tout entier naturel n , $G(n) = G(1)^n$.

b. Soit q un entier naturel non nul, fixé. La même récurrence, menée cette fois avec le pas $\frac{1}{q}$ au lieu de 1, donne le résultat. Précisément, pour $k = 0$ l'égalité est $G(0) = 1$, et si $G\left(\frac{k}{q}\right) = G\left(\frac{1}{q}\right)^k$, alors

$$G\left(\frac{k+1}{q}\right) = G\left(\frac{k}{q} + \frac{1}{q}\right) = G\left(\frac{k}{q}\right) G\left(\frac{1}{q}\right) = G\left(\frac{1}{q}\right)^{k+1}$$

Le cas particulier $k = q$ donne $\frac{k}{q} = 1$, donc

$$G(1) = G\left(\frac{1}{q}\right)^q$$

Le nombre $G\left(\frac{1}{q}\right)$ est strictement positif par hypothèse. La fonction $x \mapsto x^q$ étant strictement croissante et bijective de $]0, +\infty[$ sur $]0, +\infty[$, l'équation ci-dessus admet une unique solution strictement positive, à savoir la racine q -ième de $G(1)$:

$$G\left(\frac{1}{q}\right) = G(1)^{1/q}$$

Conclusion. Pour tout entier naturel non nul q et tout entier naturel k , $G\left(\frac{k}{q}\right) = G\left(\frac{1}{q}\right)^k$, et $G\left(\frac{1}{q}\right) = G(1)^{1/q}$.

c. Soit r un rationnel positif. On l'écrit $r = \frac{p}{q}$ avec p entier naturel et q entier naturel non nul. D'après **b.** appliqué avec $k = p$, puis d'après le résultat précédent :

$$G(r) = G\left(\frac{p}{q}\right) = G\left(\frac{1}{q}\right)^p = (G(1)^{1/q})^p = G(1)^{p/q} = G(1)^r$$

3. a. D'une part, $G(1) = P(X > 1) > 0$ par hypothèse. D'autre part, $G(1) \leq 1$ puisque c'est une probabilité. Il reste à écarter le cas $G(1) = 1$.

Supposons $G(1) = 1$. Alors, d'après la question **2. a.**, $G(n) = 1^n = 1$ pour tout entier naturel n . La suite $(G(n))$ serait donc constante égale à 1, alors que G tend vers 0 en $+\infty$, ce qui imposerait $\lim_{n \rightarrow +\infty} G(n) = 0$. Contradiction. Donc $G(1) \neq 1$ et finalement

$$0 < G(1) < 1$$

b. Soit t un réel positif. Le résultat admis fournit une suite (r_k) de rationnels positifs qui converge vers t .

D'un côté, G est continue en t , donc $G(r_k)$ tend vers $G(t)$.

De l'autre, la fonction $x \mapsto G(1)^x = e^{x \ln(G(1))}$ est définie et continue sur $\mathbb{R}$, car $G(1) > 0$; donc $G(1)^{r_k}$ tend vers $G(1)^t$.

Or, pour tout k , la question **2. c.** donne $G(r_k) = G(1)^{r_k}$. Ces deux suites sont donc égales terme à terme, et par unicité de la limite d'une suite convergente :

$$G(t) = G(1)^t$$

4. Comme $0 < G(1) < 1$, on a $\ln(G(1)) < 0$, donc

$$\lambda = -\ln(G(1)) > 0$$

Pour tout réel $t \geq 0$:

$$G(t) = G(1)^t = e^{t \ln(G(1))} = e^{-\lambda t}$$

d'où

$$F_X(t) = 1 - G(t) = 1 - e^{-\lambda t} \quad \text{pour } t \geq 0$$

Pour $t < 0$, la variable X étant à valeurs strictement positives, $F_X(t) = P(X \leq t) = 0$.

La fonction F_X coïncide donc exactement avec la fonction de répartition de la loi exponentielle de paramètre λ . Comme la loi d'une variable aléatoire réelle est entièrement déterminée par sa fonction de répartition, on conclut

$$X \hookrightarrow \mathcal{E}(\lambda) \quad \square$$

5. Le théorème démontré s'énonce ainsi : **une variable aléatoire à densité, à valeurs strictement positives et de fonction de survie partout non nulle, est sans vieillissement si et seulement si elle suit une loi exponentielle.** Le sens direct est au cours, le sens réciproque vient d'être établi.

C'est ce qui donne à la loi exponentielle sa place particulière en fiabilité : dès qu'un composant ne garde aucune trace de son passé, sa durée de vie ne peut pas suivre autre chose qu'une loi exponentielle, et le seul choix qui reste au modélisateur est celui du paramètre λ . Réciproquement, un matériel qui s'use, dont la probabilité de tomber en panne augmente avec l'âge, ne peut pas être modélisé par une loi exponentielle.

Méthode à retenir. Face à une équation fonctionnelle du type $G(s+t) = G(s)G(t)$, l'itinéraire est toujours le même : les entiers d'abord par récurrence, puis les inverses d'entiers par extraction

d'une racine q -ième, puis les rationnels par combinaison des deux, et enfin les réels par continuité et densité de $\mathbb{Q}$ dans $\mathbb{R}$. Cette progression, des entiers vers les réels, est un raisonnement classique qui se réutilise tel quel dans d'autres contextes.

Erreur classique. Conclure directement de $G(s+t) = G(s)G(t)$ que G est une exponentielle, en invoquant « la seule fonction qui transforme les sommes en produits ». Cette affirmation est fausse sans hypothèse supplémentaire : il existe des solutions non continues de cette équation, et c'est **la continuité de G** qui force la forme exponentielle. Sauter l'étape rationnelle et l'argument de continuité, c'est sauter tout l'exercice. Autre maladresse fréquente : oublier de vérifier $G(1) \neq 1$, cas dans lequel λ vaudrait 0 et la loi exponentielle ne serait pas définie.

Exercice 33 ★★★★★ Espérance par l'intégrale de la queue

Notions : Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Fonction de répartition d'une variable aléatoire réelle quelconque, Loi exponentielle et absence de mémoire, Loi uniforme sur un segment

Soit X une variable aléatoire à densité, à valeurs positives, de densité f et de fonction de répartition F_X . On suppose donc $f(t) = 0$ et $F_X(t) = 0$ pour tout $t < 0$.

L'objectif est d'établir une formule qui donne $E(X)$ à partir de la seule fonction de répartition, sans jamais dériver ni utiliser la densité.

1. Soit $A > 0$. À l'aide d'une intégration par parties dans laquelle on choisit $t \mapsto -(1 - F_X(t))$ comme primitive de f , montrer que

$$\int_0^A t f(t) dt = -A(1 - F_X(A)) + \int_0^A (1 - F_X(t)) dt$$

2. Dans cette question, on suppose que $E(X)$ existe. On **admet** que si $E(X)$ existe, alors $A(1 - F_X(A))$ tend vers 0 quand A tend vers $+\infty$. En déduire que l'intégrale $\int_0^{+\infty} (1 - F_X(t)) dt$ converge et que

$$E(X) = \int_0^{+\infty} (1 - F_X(t)) dt$$

3. Retrouver par cette formule l'espérance de la loi $\mathcal{E}(\lambda)$ et celle de la loi $\mathcal{U}([0, a])$ avec $a > 0$.
4. Soient $U_1, \dots, U_n$ des variables mutuellement indépendantes, de même loi $\mathcal{U}([0, 1])$, et $M_n = \max(U_1, \dots, U_n)$.
 - a. Déterminer F_{M_n} .
 - b. Calculer $E(M_n)$ par la formule de la question 2.
 - c. Refaire le calcul par la densité de M_n et comparer la longueur des deux méthodes.
5. Soient $Y_1, \dots, Y_n$ mutuellement indépendantes, de même loi $\mathcal{E}(\lambda)$, et $m_n = \min(Y_1, \dots, Y_n)$. Calculer $P(m_n > t)$ pour $t \geq 0$, puis $E(m_n)$ par la formule de la question 2.
6. On considère enfin la fonction g définie par $g(t) = \frac{1}{t^2}$ pour $t \geq 1$ et $g(t) = 0$ pour $t < 1$.
 - a. Vérifier que g est une densité de probabilité. On note Z une variable admettant g pour densité ; elle est bien à valeurs positives.
 - b. Montrer que $E(Z)$ n'existe pas.
 - c. Déterminer $1 - F_Z(t)$ pour tout $t \geq 0$, montrer que l'intégrale $\int_0^{+\infty} (1 - F_Z(t)) dt$ diverge, et vérifier que l'hypothèse admise à la question 2. est ici en défaut.

Corrigé

1. La stratégie : la formule à démontrer fait apparaître $1 - F_X$, donc il faut faire jouer à cette quantité le rôle de la primitive de f dans une intégration par parties.

La fonction F_X est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sauf éventuellement en un nombre fini de points, avec $F'_X = f$ là où la dérivée existe. Posons

$$u(t) = t \quad \text{et} \quad v(t) = -(1 - F_X(t)) = F_X(t) - 1$$

Alors $u'(t) = 1$ et $v'(t) = f(t)$. Les fonctions u et v sont continues sur $[0, A]$ et de classe $\mathcal{C}^1$ sauf en un nombre fini de points ; l'intégration par parties est donc licite sur chacun des sous-

intervalles où elles sont de classe $\mathcal{C}^1$, et les termes de bord intermédiaires se compensent deux à deux par continuité de u et de v . On obtient

$$\begin{aligned}\int_0^A t f(t) dt &= \left[-t(1 - F_X(t)) \right]_0^A + \int_0^A (1 - F_X(t)) dt \\ &= -A(1 - F_X(A)) + 0 \times (1 - F_X(0)) + \int_0^A (1 - F_X(t)) dt \\ &= -A(1 - F_X(A)) + \int_0^A (1 - F_X(t)) dt\end{aligned}$$

le terme en $t = 0$ étant nul puisqu'il contient le facteur 0.

2. On isole l'intégrale cherchée dans l'égalité précédente :

$$\int_0^A (1 - F_X(t)) dt = \int_0^A t f(t) dt + A(1 - F_X(A))$$

Examinons les deux termes de droite quand A tend vers $+\infty$.

Comme X est à valeurs positives, la fonction $t \mapsto tf(t)$ est nulle sur $]-\infty, 0[$ et positive sur $[0, +\infty[$: la convergence absolue de l'intégrale équivaut à sa convergence. L'espérance $E(X)$ existant par hypothèse, le premier terme tend donc vers

$$\int_0^{+\infty} t f(t) dt = E(X)$$

Le second terme tend vers 0, c'est le résultat admis par l'énoncé.

Le membre de gauche admet donc une limite finie quand A tend vers $+\infty$. Autrement dit, l'intégrale impropre $\int_0^{+\infty} (1 - F_X(t)) dt$ converge, et sa valeur est

$$E(X) = \int_0^{+\infty} (1 - F_X(t)) dt \quad \square$$

La quantité $1 - F_X(t) = P(X > t)$ s'appelle la **queue** de la loi. La formule se lit donc : l'espérance d'une variable positive est l'aire sous sa courbe de survie.

3. *Loi exponentielle.* Pour $t \geq 0$, $1 - F_X(t) = e^{-\lambda t}$. Cette fonction est positive et, pour $A > 0$,

$$\int_0^A e^{-\lambda t} dt = \left[-\frac{1}{\lambda} e^{-\lambda t} \right]_0^A = \frac{1}{\lambda} (1 - e^{-\lambda A}) \xrightarrow{A \rightarrow +\infty} \frac{1}{\lambda}$$

L'intégrale converge, donc

$$E(X) = \frac{1}{\lambda}$$

Loi uniforme sur $[0, a]$. Ici $F_X(t) = \frac{t}{a}$ pour $t \in [0, a]$ et $F_X(t) = 1$ pour $t > a$. La queue est donc nulle au-delà de a , et l'intégrale n'est même pas impropre :

$$E(X) = \int_0^a \left(1 - \frac{t}{a}\right) dt = \left[t - \frac{t^2}{2a} \right]_0^a = a - \frac{a^2}{2a} = a - \frac{a}{2} = \frac{a}{2}$$

Les deux valeurs sont bien celles du cours.

4. a. Pour tout réel t , l'événement $(M_n \leq t)$ signifie que **toutes** les variables sont inférieures ou égales à t :

$$(M_n \leq t) = \bigcap_{i=1}^n (U_i \leq t)$$

Les U_i étant mutuellement indépendantes et de même loi,

$$F_{M_n}(t) = \prod_{i=1}^n P(U_i \leq t) = (F_U(t))^n$$

où F_U est la fonction de répartition commune. Comme $F_U(t) = t$ sur $[0, 1]$, on obtient $F_{M_n}(t) = t^n$ pour $t \in [0, 1]$, avec $F_{M_n}(t) = 0$ pour $t < 0$ et $F_{M_n}(t) = 1$ pour $t > 1$.

b. La variable M_n est à valeurs dans $[0, 1]$, donc positive, et sa queue est nulle au-delà de 1. La formule de la question **2.** donne alors

$$E(M_n) = \int_0^1 (1 - t^n) dt = \left[t - \frac{t^{n+1}}{n+1} \right]_0^1 = 1 - \frac{1}{n+1} = \frac{n+1-1}{n+1} = \frac{n}{n+1}$$

c. Par la densité, il faut d'abord dériver : $f_{M_n}(t) = n t^{n-1}$ sur $]0, 1[$, et 0 ailleurs. Puis

$$E(M_n) = \int_0^1 t \times n t^{n-1} dt = n \int_0^1 t^n dt = n \times \frac{1}{n+1} = \frac{n}{n+1}$$

Les deux méthodes donnent le même résultat, mais la seconde exige une étape supplémentaire, la dérivation de F_{M_n} , qui est précisément l'étape où les erreurs de calcul apparaissent. Dès qu'un maximum ou un minimum est en jeu, la fonction de répartition est immédiate alors que la densité demande un travail : la formule de la queue permet donc d'aller directement au but. Le résultat lui-même est parlant : le maximum de n tirages uniformes sur $[0, 1]$ vaut en moyenne $\frac{n}{n+1}$, valeur qui se rapproche de 1 à mesure que n grandit, tout en lui restant strictement inférieure.

5. Pour un **minimum**, c'est la queue qui est immédiate, et non la fonction de répartition. Pour $t \geq 0$, l'événement $(m_n > t)$ signifie que toutes les variables dépassent t :

$$P(m_n > t) = \prod_{i=1}^n P(Y_i > t) = (e^{-\lambda t})^n = e^{-n\lambda t}$$

Le minimum est à valeurs positives, la formule de la question **2.** s'applique :

$$E(m_n) = \int_0^{+\infty} e^{-n\lambda t} dt = \frac{1}{n\lambda}$$

en reprenant le calcul déjà fait à la question **3.** avec λ remplacé par $n\lambda$. On reconnaît d'ailleurs dans $t \mapsto e^{-n\lambda t}$ la queue d'une loi $\mathcal{E}(n\lambda)$, ce qui redonne le résultat du cours : le minimum de n variables exponentielles indépendantes de même paramètre λ suit la loi $\mathcal{E}(n\lambda)$. Avec n composants montés en série, la première panne survient en moyenne n fois plus tôt que la panne d'un composant isolé.

6. a. La fonction g est positive sur $\mathbb{R}$ et continue sauf en 1. Pour $A > 1$:

$$\int_1^A \frac{dt}{t^2} = \left[-\frac{1}{t} \right]_1^A = 1 - \frac{1}{A} \xrightarrow{A \rightarrow +\infty} 1$$

L'intégrale $\int_{-\infty}^{+\infty} g(t) dt$ converge et vaut 1 : g est bien une densité de probabilité. La variable Z associée est à valeurs dans $[1, +\infty[$, donc positive : aucune translation n'est nécessaire pour appliquer le cadre de l'exercice.

b. Étudions la convergence de $\int_1^{+\infty} t g(t) dt$. Pour $A > 1$:

$$\int_1^A t \times \frac{1}{t^2} dt = \int_1^A \frac{dt}{t} = \ln(A) \xrightarrow{A \rightarrow +\infty} +\infty$$

L'intégrale diverge, donc $E(Z)$ **n'existe pas**. C'est l'exemple de référence d'une variable parfaitement définie, à valeurs positives et de densité très simple, mais sans espérance.

c. La variable Z étant à valeurs supérieures ou égales à 1, on a $P(Z > t) = 1$ pour $0 \leq t < 1$. Pour $t \geq 1$, le calcul du a. donne

$$1 - F_Z(t) = \int_t^{+\infty} \frac{ds}{s^2} = \frac{1}{t}$$

L'intégrale se coupe donc en deux morceaux :

$$\int_0^A (1 - F_Z(t)) dt = \int_0^1 1 dt + \int_1^A \frac{dt}{t} = 1 + \ln(A) \xrightarrow{A \rightarrow +\infty} +\infty$$

L'intégrale de la queue diverge elle aussi. Les deux membres de la formule de la question 2. sont donc simultanément infinis : la formule ne se contredit pas, elle cesse simplement de s'appliquer, faute d'espérance.

On peut être plus précis sur la cause. Ici

$$A(1 - F_Z(A)) = A \times \frac{1}{A} = 1 \xrightarrow{A \rightarrow +\infty} 1 \neq 0$$

L'hypothèse admise en question 2. est donc en défaut, et c'est exactement là que le mécanisme se rompt. Cet exemple montre que cette hypothèse n'est pas une formalité technique : elle exprime que la queue de la loi doit décroître assez vite, et une queue en $\frac{1}{t}$ décroît trop lentement pour qu'une espérance existe.

Méthode à retenir. Pour une variable **positive**, l'espérance est l'intégrale de sa queue $P(X > t)$. Cette formule est la voie la plus courte chaque fois que la fonction de répartition est plus simple que la densité, ce qui est le cas systématique des maximums (où F est un produit) et des minimums (où c'est la queue qui est un produit). Le réflexe : devant $E(\max)$ ou $E(\min)$, ne pas dériver, intégrer la queue.

Erreur classique. Appliquer la formule à une variable qui prend des valeurs négatives. La démonstration utilise de façon essentielle que f et F_X sont nulles sur $]-\infty, 0[$; pour une variable de signe quelconque, la formule correcte comporte un second terme et n'est pas celle-ci. Autre erreur : conclure de la divergence de $\int_0^{+\infty} (1 - F_X(t)) dt$ que « la formule est fausse ». Elle n'est pas fausse : sa démonstration suppose l'existence de $E(X)$, et la question 6. montre que les deux intégrales divergent ensemble, ce qui est la meilleure cohérence que l'on puisse espérer.

Exercice 34 ★★★★★ Temps d'attente successifs et loi gamma

Notions : Somme de variables indépendantes : produit de convolution et stabilité, Lois gamma à un paramètre et fonction Gamma d'Euler, Loi exponentielle et absence de mémoire, n-uplets de variables aléatoires, indépendance mutuelle, échantillon

Une installation industrielle tombe en panne de façon répétée. On note T_1 le temps écoulé entre la mise en service et la première panne, et plus généralement, pour $k \geq 2$, on note T_k le temps écoulé entre la $(k-1)$ -ième et la k -ième panne. On suppose que les variables $T_1, T_2, \dots$ sont mutuellement indépendantes et suivent toutes la loi $\mathcal{E}(\lambda)$, avec $\lambda > 0$.

Pour tout entier $n \geq 1$, l'instant de la n -ième panne est

$$S_n = T_1 + T_2 + \dots + T_n$$

On rappelle que la loi $\gamma(\nu)$, avec $\nu > 0$, admet pour densité la fonction nulle sur $]-\infty, 0]$ et égale à $\frac{x^{\nu-1}e^{-x}}{\Gamma(\nu)}$ sur $]0, +\infty[$, que son espérance et sa variance valent toutes deux ν , et que $\Gamma(n) = (n-1)!$ pour tout entier $n \geq 1$.

1. Déterminer une densité de S_2 par le produit de convolution. On rappelle que la légitimité de ce produit n'est pas à démontrer.
2. Montrer par récurrence que, pour tout entier $n \geq 1$, une densité de S_n est donnée par

$$f_n(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} \quad \text{sur }]0, +\infty[, \quad f_n(t) = 0 \quad \text{sur }]-\infty, 0]$$

On précisera où intervient le lemme des coalitions.

3. En déduire que $\lambda S_n \hookrightarrow \gamma(n)$, puis retrouver $E(S_n)$ et $V(S_n)$. On contrôlera ces deux valeurs par un calcul direct sur la somme.
4. Montrer, pour tout réel $t \geq 0$ et tout entier $n \geq 1$,

$$P(S_n > t) = e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}$$

On pourra raisonner par récurrence sur n , en utilisant une intégration par parties dans l'hérédité.

5. Interpréter concrètement cette formule en termes de nombre de pannes survenues avant l'instant t , et reconnaître une loi discrète usuelle.
6. Application numérique : l'installation subit en moyenne une panne toutes les 100 heures, soit $\lambda = \frac{1}{100}$ panne par heure. Calculer $P(S_3 > 200)$ à trois décimales, et commenter le résultat.

On donne la valeur approchée $e^{-2} \approx 0,135335$.

Corrigé

1. Les variables T_1 et T_2 sont indépendantes et à densité, donc une densité de leur somme est donnée par le produit de convolution

$$f_{S_2}(t) = \int_{-\infty}^{+\infty} f_{T_1}(x) f_{T_2}(t-x) dx$$

Le réflexe est de repérer le domaine où l'intégrande est non nul. La densité $f_{T_1}(x) = \lambda e^{-\lambda x}$ n'est non nulle que pour $x > 0$, et $f_{T_2}(t-x) = \lambda e^{-\lambda(t-x)}$ n'est non nulle que pour $t-x > 0$, c'est-à-dire $x < t$.

Si $t \leq 0$, ces deux conditions sont incompatibles : l'intégrande est identiquement nul et $f_{S_2}(t) = 0$. C'est cohérent, une somme de deux durées positives est positive.

Si $t > 0$, l'intégrande n'est non nul que pour $x \in]0, t[$, et

$$\begin{aligned} f_{S_2}(t) &= \int_0^t \lambda e^{-\lambda x} \times \lambda e^{-\lambda(t-x)} dx \\ &= \lambda^2 e^{-\lambda t} \int_0^t e^{-\lambda x + \lambda x} dx \\ &= \lambda^2 e^{-\lambda t} \int_0^t dx \\ &= \lambda^2 t e^{-\lambda t} \end{aligned}$$

Le point remarquable est la simplification des exponentielles : leur produit ne dépend plus de x , ce qui rend l'intégrale triviale. C'est ce phénomène qui rend le calcul faisable pour tout n .

2. Notons $\mathcal{H}_n$ la proposition « la fonction f_n de l'énoncé est une densité de S_n ».

Initialisation. Pour $n = 1$, la formule donne $f_1(t) = \frac{\lambda^1 t^0 e^{-\lambda t}}{0!} = \lambda e^{-\lambda t}$ sur $]0, +\infty[$, ce qui est bien la densité de $T_1 \hookrightarrow \mathcal{E}(\lambda)$. Donc $\mathcal{H}_1$ est vraie. On vérifie au passage que $\mathcal{H}_2$ correspond bien au résultat de la question 1.

Hérédité. Supposons $\mathcal{H}_n$ vraie pour un entier $n \geq 1$. On écrit

$$S_{n+1} = S_n + T_{n+1}$$

La variable S_n ne dépend que de $T_1, \dots, T_n$, et T_{n+1} est indépendante de ce n -uplet. Le **lemme des coalitions**, admis par le programme, assure donc que S_n et T_{n+1} sont indépendantes : c'est exactement ce qui autorise l'emploi du produit de convolution à chaque étape.

Pour $t \leq 0$, la somme de $n+1$ durées positives étant positive, la densité est nulle. Pour $t > 0$, le même repérage de domaine qu'en question 1. donne $x \in]0, t[$, et

$$\begin{aligned} f_{n+1}(t) &= \int_0^t f_n(x) \lambda e^{-\lambda(t-x)} dx \\ &= \int_0^t \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \times \lambda e^{-\lambda t + \lambda x} dx \\ &= \frac{\lambda^{n+1} e^{-\lambda t}}{(n-1)!} \int_0^t x^{n-1} dx \\ &= \frac{\lambda^{n+1} e^{-\lambda t}}{(n-1)!} \times \frac{t^n}{n} \\ &= \frac{\lambda^{n+1} t^n e^{-\lambda t}}{n!} \end{aligned}$$

la dernière égalité utilisant $n \times (n-1)! = n!$. C'est bien la formule annoncée au rang $n+1$.

Conclusion. Pour tout entier $n \geq 1$, f_n est une densité de S_n . $\square$

3. On applique la méthode de la fonction de répartition à la transformation affine $U = \lambda S_n$, avec $\lambda > 0$. Pour $u \leq 0$, $F_U(u) = 0$. Pour $u > 0$:

$$F_U(u) = P(\lambda S_n \leq u) = P\left(S_n \leq \frac{u}{\lambda}\right) = F_{S_n}\left(\frac{u}{\lambda}\right)$$

La fonction F_U hérite de F_{S_n} sa continuité sur $\mathbb{R}$ et son caractère $\mathcal{C}^1$ sauf éventuellement en 0 : la variable U est donc à densité, et on l'obtient en dérivant. Une densité de U est

$$f_U(u) = \frac{1}{\lambda} f_n\left(\frac{u}{\lambda}\right) = \frac{1}{\lambda} \times \frac{\lambda^n \left(\frac{u}{\lambda}\right)^{n-1} e^{-u}}{(n-1)!} = \frac{\lambda^n \times \lambda^{-(n-1)} \times \lambda^{-1} u^{n-1} e^{-u}}{(n-1)!} = \frac{u^{n-1} e^{-u}}{(n-1)!}$$

Comme $(n-1)! = \Gamma(n)$, on reconnaît exactement la densité de la loi $\gamma(n)$ rappelée dans l'énoncé :

$$\lambda S_n \hookrightarrow \gamma(n)$$

L'espérance et la variance d'une loi $\gamma(\nu)$ valant ν , on a $E(\lambda S_n) = n$ et $V(\lambda S_n) = n$. Les règles $E(aX) = aE(X)$ et $V(aX) = a^2V(X)$ donnent alors

$$E(S_n) = \frac{n}{\lambda} \quad \text{et} \quad V(S_n) = \frac{n}{\lambda^2}$$

Contrôle direct. Chaque T_k a pour espérance $\frac{1}{\lambda}$ et pour variance $\frac{1}{\lambda^2}$. Par linéarité de l'espérance,

$$E(S_n) = \sum_{k=1}^n E(T_k) = \frac{n}{\lambda}$$

et, les variables étant indépendantes donc de covariances deux à deux nulles,

$$V(S_n) = \sum_{k=1}^n V(T_k) = \frac{n}{\lambda^2}$$

Les deux voies concordent. Le résultat est intuitif : si une panne survient en moyenne toutes les $\frac{1}{\lambda}$ unités de temps, la n -ième arrive en moyenne au bout de $\frac{n}{\lambda}$.

4. Notons $\mathcal{K}_n$ la proposition « pour tout $t \geq 0$, $P(S_n > t) = e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}$ ».

Initialisation. Pour $n = 1$, la somme se réduit au terme $k = 0$, qui vaut 1. Le membre de droite est donc $e^{-\lambda t}$, et le membre de gauche vaut $P(T_1 > t) = e^{-\lambda t}$ puisque $T_1 \hookrightarrow \mathcal{E}(\lambda)$. Donc $\mathcal{K}_1$ est vraie.

Hérédité. Supposons $\mathcal{K}_n$ vraie pour un entier $n \geq 1$. Par définition de la densité f_{n+1} ,

$$P(S_{n+1} > t) = \int_t^{+\infty} \frac{\lambda^{n+1} x^n e^{-\lambda x}}{n!} dx = \frac{\lambda^n}{n!} \int_t^{+\infty} x^n \lambda e^{-\lambda x} dx$$

On intègre par parties en posant $u(x) = x^n$ et $v(x) = -e^{-\lambda x}$, de sorte que $u'(x) = nx^{n-1}$ et $v'(x) = \lambda e^{-\lambda x}$. Ces deux fonctions sont de classe $\mathcal{C}^1$ sur $[t, +\infty[$. Par croissances comparées, $x^n e^{-\lambda x}$ tend vers 0 en $+\infty$, donc le terme de bord converge :

$$\begin{aligned} \int_t^{+\infty} x^n \lambda e^{-\lambda x} dx &= \left[-x^n e^{-\lambda x} \right]_t^{+\infty} + \int_t^{+\infty} n x^{n-1} e^{-\lambda x} dx \\ &= t^n e^{-\lambda t} + n \int_t^{+\infty} x^{n-1} e^{-\lambda x} dx \end{aligned}$$

En reportant, et en séparant les deux morceaux :

$$P(S_{n+1} > t) = \frac{\lambda^n t^n e^{-\lambda t}}{n!} + \frac{\lambda^n \times n}{n!} \int_t^{+\infty} x^{n-1} e^{-\lambda x} dx$$

Le second morceau se reconnaît : comme $\frac{n}{n!} = \frac{1}{(n-1)!}$, il vaut

$$\frac{\lambda^n}{(n-1)!} \int_t^{+\infty} x^{n-1} e^{-\lambda x} dx = \int_t^{+\infty} f_n(x) dx = P(S_n > t)$$

D'où, en utilisant l'hypothèse de récurrence,

$$P(S_{n+1} > t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} + e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} = e^{-\lambda t} \sum_{k=0}^n \frac{(\lambda t)^k}{k!}$$

le terme isolé étant exactement celui d'indice $k = n$. C'est $\mathcal{K}_{n+1}$.

Conclusion. La formule est établie pour tout entier $n \geq 1$. $\square$

5. Notons N_t le nombre de pannes survenues jusqu'à l'instant t . Dire que la n -ième panne n'a pas encore eu lieu à l'instant t , c'est dire qu'il y a eu au plus $n-1$ pannes avant t . Les événements $(S_n > t)$ et $(N_t \leq n-1)$ sont donc le même événement, et la formule de la question 4. s'écrit

$$P(N_t \leq n-1) = \sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Le terme général de cette somme est exactement $P(Y = k)$ pour une variable $Y \hookrightarrow \mathcal{P}(\lambda t)$. Autrement dit, le nombre de pannes survenues avant l'instant t suit une **loi de Poisson de paramètre λt** .

Le résultat est cohérent avec l'intuition : le paramètre λt est le nombre moyen de pannes sur une durée t , puisque $E(N_t) = \lambda t$ et que les pannes arrivent au rythme de λ par unité de temps. On voit ainsi apparaître les deux faces d'un même phénomène : compter les pannes conduit à une loi de Poisson, mesurer les temps d'attente entre pannes conduit à des lois exponentielles.

6. Avec $\lambda = \frac{1}{100}$ et $t = 200$, on obtient d'abord $\lambda t = \frac{200}{100} = 2$. La formule de la question 4. appliquée à $n = 3$ donne

$$P(S_3 > 200) = e^{-2} \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} \right) = e^{-2} (1 + 2 + 2) = 5e^{-2}$$

Avec $e^{-2} \approx 0,135335$:

$$P(S_3 > 200) \approx 5 \times 0,135335 = 0,676675 \approx 0,677$$

Commentaire. La troisième panne survient en moyenne à l'instant $E(S_3) = \frac{3}{\lambda} = 300$ heures. Il est donc naturel qu'à 200 heures elle n'ait pas encore eu lieu dans plus de deux cas sur trois. On peut relire ce nombre autrement : 0,677 est la probabilité d'observer au plus deux pannes en 200 heures, alors qu'on en attend deux en moyenne, ce qui est bien un résultat majoritaire mais pas écrasant.

Méthode à retenir. Une somme de n variables exponentielles indépendantes de même paramètre λ se traite en deux temps : convolution et récurrence pour la densité, puis normalisation par λ pour reconnaître une loi $\gamma(n)$, seule forme de la loi gamma disponible au programme. Toutes les caractéristiques (E , V , queue) s'en déduisent ensuite sans nouveau calcul d'intégrale. La formule de la queue $P(S_n > t)$ est à connaître, car elle transforme un problème de temps continu en un simple calcul de loi de Poisson.

Erreur classique. Écrire que S_n suit une loi exponentielle de paramètre $\frac{\lambda}{n}$, sous prétexte que son espérance vaut $\frac{n}{\lambda}$. C'est faux : une loi exponentielle a une densité maximale en 0, alors que f_n s'annule en 0 pour $n \geq 2$, la n -ième panne ne pouvant raisonnablement pas survenir immédiatement. Deuxième erreur, plus discrète : oublier de justifier l'indépendance de S_n et de T_{n+1} dans l'hérédité, alors que le produit de convolution ne s'applique qu'à des variables indépendantes. Le lemme des coalitions doit être cité, faute de quoi la récurrence n'a pas de fondement.

Exercice 35 ★★★★★ Une famille de densités à paramètre

Notions : Variable aléatoire à densité : définition et reconnaissance d'une densité, Espérance d'une variable à densité : existence, théorème de transfert, linéarité, Variance, écart-type et moments d'une variable à densité, Loi d'une transformée : transformation affine et méthode de la fonction de répartition

Pour tout réel $a > 0$, on définit la fonction f_a sur $\mathbb{R}$ par

$$f_a(x) = \frac{a}{x^{a+1}} \quad \text{si } x \geq 1, \quad f_a(x) = 0 \quad \text{si } x < 1$$

1. Montrer que f_a est une densité de probabilité, pour toute valeur de $a > 0$. Dans toute la suite, X désigne une variable aléatoire admettant f_a pour densité.
2. Déterminer la fonction de répartition F_a de X sur $\mathbb{R}$.
3. Montrer que $E(X)$ existe si et seulement si $a > 1$, et qu'alors

$$E(X) = \frac{a}{a-1}$$

4. Montrer que $V(X)$ existe si et seulement si $a > 2$, et qu'alors

$$V(X) = \frac{a}{(a-1)^2(a-2)}$$

5. On pose $Y = \ln(X)$.

a. Déterminer la loi de Y par la méthode de la fonction de répartition.

b. En déduire une méthode permettant de simuler X à partir d'une variable $U \hookrightarrow \mathcal{U}([0, 1])$. On rédigera la réponse en français et en mathématiques, sans écrire de programme.

6. Soit $b > 0$. Déterminer la loi de $Z = X^b$ et reconnaître une loi de la même famille.
7. On prend $a = 3$. Calculer $E(X)$, $V(X)$, $\sigma(X)$, la médiane de X (le réel μ tel que $F_3(\mu) = 0,5$) et $P(X > 2)$. Comparer la médiane et l'espérance, et commenter.

On donne les valeurs approchées suivantes : $2^{1/3} \approx 1,260$ et $\sqrt{3} \approx 1,7321$.

Corrigé

1. Trois points sont à vérifier : la positivité, la continuité sauf en un nombre fini de points, et l'intégrale égale à 1.

La fonction f_a est positive sur $\mathbb{R}$, puisque $a > 0$ et $x^{a+1} > 0$ pour $x \geq 1$. Elle est continue sur $]-\infty, 1[$ et sur $]1, +\infty[$, donc continue sur $\mathbb{R}$ sauf éventuellement en 1, ce qui est autorisé.

Reste l'intégrale. Elle est nulle sur $]-\infty, 1[$, et pour $A > 1$:

$$\int_1^A \frac{a}{x^{a+1}} dx = a \int_1^A x^{-a-1} dx = a \left[\frac{x^{-a}}{-a} \right]_1^A = [-x^{-a}]_1^A = 1 - A^{-a}$$

Comme $a > 0$, $A^{-a} = \frac{1}{A^a}$ tend vers 0 quand A tend vers $+\infty$. L'intégrale converge donc, et

$$\int_{-\infty}^{+\infty} f_a(x) dx = 1$$

La fonction f_a est bien une densité de probabilité, et ce pour toute valeur de $a > 0$. $\square$

2. Pour $x < 1$, la densité est nulle sur $]-\infty, x]$, donc $F_a(x) = 0$. Pour $x \geq 1$, le calcul précédent, mené jusqu'à x au lieu de A , donne directement

$$F_a(x) = \int_1^x \frac{a}{s^{a+1}} ds = 1 - x^{-a} = 1 - \frac{1}{x^a}$$

On contrôle la cohérence : $F_a(1) = 1 - 1 = 0$, la fonction est continue en 1, croissante, et tend vers 1 en $+\infty$.

3. L'existence de $E(X)$ demande la convergence absolue de $\int_{-\infty}^{+\infty} x f_a(x) dx$. Cette fonction étant nulle pour $x < 1$ et positive pour $x \geq 1$, la convergence absolue équivaut à la convergence de

$$\int_1^{+\infty} x \times \frac{a}{x^{a+1}} dx = a \int_1^{+\infty} \frac{dx}{x^a}$$

C'est une intégrale de Riemann en $+\infty$, d'exposant a : elle converge si et seulement si $a > 1$. Le paramètre a contrôle donc la vitesse à laquelle la densité s'écrase, et une densité qui décroît trop lentement interdit à la variable d'avoir une espérance.

Supposons $a > 1$. Pour $A > 1$:

$$a \int_1^A x^{-a} dx = a \left[\frac{x^{1-a}}{1-a} \right]_1^A = \frac{a}{1-a} (A^{1-a} - 1)$$

Comme $a > 1$, l'exposant $1 - a$ est strictement négatif, donc A^{1-a} tend vers 0. Il reste

$$E(X) = \frac{a}{1-a} \times (-1) = \frac{a}{a-1}$$

Cette valeur est toujours strictement supérieure à 1, ce qui est cohérent puisque X ne prend que des valeurs supérieures ou égales à 1.

4. On passe par le moment d'ordre 2. Le théorème de transfert appliqué à $x \mapsto x^2$ donne, sous réserve de convergence absolue,

$$E(X^2) = \int_1^{+\infty} x^2 \times \frac{a}{x^{a+1}} dx = a \int_1^{+\infty} \frac{dx}{x^{a-1}}$$

L'intégrande est positif : c'est encore une intégrale de Riemann en $+\infty$, d'exposant $a - 1$, qui converge si et seulement si $a - 1 > 1$, c'est-à-dire $a > 2$.

La variance existe si et seulement si le moment d'ordre 2 existe, donc si et seulement si $a > 2$. Il faut noter que la condition est **strictement plus forte** que celle de la question 3. : pour $1 < a \leq 2$, la variable a une espérance mais pas de variance.

Supposons $a > 2$. Le même calcul que ci-dessus, avec l'exposant $a - 1$ à la place de a , donne

$$E(X^2) = \frac{a}{a-2}$$

La formule de König-Huygens fournit alors

$$V(X) = \frac{a}{a-2} - \left(\frac{a}{a-1} \right)^2 = \frac{a}{a-2} - \frac{a^2}{(a-1)^2}$$

On réduit au même dénominateur $(a-2)(a-1)^2$:

$$\begin{aligned} V(X) &= \frac{a(a-1)^2 - a^2(a-2)}{(a-2)(a-1)^2} \\ &= \frac{a[(a-1)^2 - a(a-2)]}{(a-2)(a-1)^2} \\ &= \frac{a[a^2 - 2a + 1 - a^2 + 2a]}{(a-2)(a-1)^2} \\ &= \frac{a}{(a-1)^2(a-2)} \end{aligned}$$

Le crochet se réduit au seul terme 1, ce qui explique la simplicité de la formule finale. On vérifie que cette expression est bien strictement positive pour $a > 2$, comme doit l'être une variance.

5. a. La méthode de la fonction de répartition est la seule méthode au programme. La variable X étant à valeurs dans $[1, +\infty[$, la variable $Y = \ln(X)$ est à valeurs dans $[0, +\infty[$, donc $F_Y(y) = 0$ pour $y < 0$.

Pour $y \geq 0$, la fonction exponentielle étant strictement croissante :

$$F_Y(y) = P(\ln(X) \leq y) = P(X \leq e^y) = F_a(e^y)$$

Comme $y \geq 0$ entraîne $e^y \geq 1$, la question **2.** s'applique et

$$F_Y(y) = 1 - (e^y)^{-a} = 1 - e^{-ay}$$

On reconnaît la fonction de répartition de la loi exponentielle de paramètre a , d'où

$$Y = \ln(X) \hookrightarrow \mathcal{E}(a)$$

b. Le passage au logarithme ramène donc la loi étudiée à une loi exponentielle, dont on sait simuler une réalisation par la méthode d'inversion.

Soit $U \hookrightarrow \mathcal{U}([0, 1])$. On a $P(U = 0) = 0$, donc $\ln(U)$ est défini presque sûrement. Posons

$$V = -\frac{1}{a} \ln(U)$$

Pour $y \geq 0$, en utilisant la croissance de l'exponentielle puis la loi de U :

$$F_V(y) = P\left(-\frac{1}{a} \ln(U) \leq y\right) = P(\ln(U) \geq -ay) = P(U \geq e^{-ay}) = 1 - e^{-ay}$$

la dernière égalité venant de ce que e^{-ay} appartient à $]0, 1]$ et que U est uniforme sur $[0, 1]$. Donc $V \hookrightarrow \mathcal{E}(a)$, et V a la même loi que Y .

D'après la question **a.**, $X = e^Y$. On simule donc X en posant

$$X = e^V = e^{-\frac{1}{a} \ln(U)} = U^{-1/a}$$

Contrôle direct. Pour $x \geq 1$, on a $x^{-a} \in]0, 1]$ et

$$P(U^{-1/a} \leq x) = P(U \geq x^{-a}) = 1 - x^{-a} = F_a(x)$$

La variable $U^{-1/a}$ a donc bien la loi voulue. En résumé : tirer un nombre au hasard uniformément entre 0 et 1, puis l'élever à la puissance $-\frac{1}{a}$.

6. La variable X est à valeurs dans $[1, +\infty[$ et $b > 0$, donc $Z = X^b$ est aussi à valeurs dans $[1, +\infty[$: ainsi $F_Z(z) = 0$ pour $z < 1$.

Pour $z \geq 1$, la fonction $x \mapsto x^{1/b}$ étant strictement croissante sur $[0, +\infty[$:

$$F_Z(z) = P(X^b \leq z) = P(X \leq z^{1/b}) = F_a(z^{1/b}) = 1 - (z^{1/b})^{-a} = 1 - z^{-a/b}$$

C'est exactement la fonction de répartition obtenue en question **2.**, avec le paramètre $\frac{a}{b}$ à la place de a . Donc Z suit la loi de la même famille, de paramètre $\frac{a}{b}$, et admet pour densité

$$f_{a/b}(z) = \frac{a/b}{z^{a/b} + 1} \quad \text{pour } z \geq 1$$

La famille est donc **stable par élévation à une puissance strictement positive**, l'exposant venant simplement diviser le paramètre. Cette propriété est cohérente avec la question **5.** : élever

X à la puissance b revient à multiplier $\ln(X)$ par b , et multiplier par b une variable de loi $\mathcal{E}(a)$ donne une variable de loi $\mathcal{E}\left(\frac{a}{b}\right)$.

7. Pour $a = 3$, les conditions $a > 1$ et $a > 2$ sont satisfaites : l'espérance et la variance existent toutes les deux.

a. $E(X) = \frac{3}{3-1} = \frac{3}{2} = 1,5$

c. $\sigma(X) = \sqrt{0,75} = \frac{\sqrt{3}}{2} \approx 0,866$

b. $V(X) = \frac{3}{(3-1)^2(3-2)} = \frac{3}{4} = 0,75$

d. $P(X > 2) = 1 - F_3(2) = 2^{-3} = \frac{1}{8} = 0,125$

Médiane. Elle est solution de $F_3(\mu) = 0,5$ avec $\mu \geq 1$, soit

$$1 - \frac{1}{\mu^3} = 0,5 \iff \frac{1}{\mu^3} = 0,5 \iff \mu^3 = 2 \iff \mu = 2^{1/3} \approx 1,260$$

Comparaison. La médiane vaut environ 1,260 et l'espérance 1,5 : la médiane est nettement **inférieure** à l'espérance. Cela traduit une loi très dissymétrique. La moitié des valeurs sont inférieures à 1,26, mais la densité décroît assez lentement pour que les valeurs rares et grandes tirent la moyenne vers le haut. Ce comportement est typique des lois dites à queue lourde, que l'on rencontre pour modéliser des tailles de sinistres, des montants de patrimoine ou des durées de connexion : la moyenne y est mal représentative de l'individu ordinaire, et la médiane la décrit mieux.

Méthode à retenir. Sur une famille de densités à paramètre, l'ordre des questions est toujours le même : vérifier que c'est bien une densité, calculer F , puis discuter l'existence des moments **avant** de les calculer. La discussion se ramène ici à un critère de Riemann en $+\infty$, et il faut retenir que l'existence de $E(X)$ n'entraîne jamais celle de $V(X)$: chaque moment a sa propre condition, de plus en plus exigeante. Enfin, toute transformation $\ln(X)$, X^b , e^X se traite par la fonction de répartition, jamais en manipulant la densité.

Erreur classique. Annoncer $E(X) = \frac{a}{a-1}$ sans discuter le cas $a \leq 1$. La formule obtenue reste calculable, par exemple $\frac{0,5}{-0,5} = -1$ pour $a = 0,5$, et fournit une espérance négative pour une variable qui ne prend que des valeurs supérieures à 1 : l'absurdité du résultat signale que le calcul a été mené sans vérifier la convergence. Deuxième erreur : dériver F_Z « en dérivant la densité de X » au lieu d'appliquer la méthode de la fonction de répartition, ce qui fait perdre le facteur issu de la transformation.

Exercice 36 ★★★★★ Synthèse : fiabilité d'un parc de machines

Notions : Minimum et maximum d'un échantillon de variables indépendantes, Loi exponentielle et absence de mémoire, Somme de variables indépendantes : produit de convolution et stabilité, n-uplets de variables aléatoires, indépendance mutuelle, échantillon, Espérance conditionnelle et formule de l'espérance totale

Un atelier exploite des machines dont la durée de vie, exprimée en heures de fonctionnement, suit une loi exponentielle. Les quatre parties sont largement indépendantes les unes des autres.

On donne $e^{-1} \approx 0,367879$, $e^{-1,5} \approx 0,223130$, $e^{-2} \approx 0,135335$, $e^{-3} \approx 0,049787$, $e^{-4} \approx 0,018316$, $e^{-5} \approx 0,006738$ et $e^{-8/7} \approx 0,318907$.

Partie A. Une machine isolée.

La durée de vie X d'une machine suit la loi $\mathcal{E}(\lambda)$ avec $\lambda = \frac{1}{2000}$.

1. Donner $E(X)$ et $\sigma(X)$, et commenter le fait que ces deux nombres soient égaux.
2. Calculer $P(X > 3000)$ à trois décimales.
3. Une machine fonctionne déjà depuis 1000 heures. Calculer la probabilité qu'elle fonctionne encore 3000 heures de plus, c'est-à-dire $P_{(X>1000)}(X > 4000)$. Quelle propriété de la loi exponentielle ce résultat illustre-t-il, et quelle conséquence pratique cela a-t-il sur la politique de remplacement préventif de l'atelier ?

Partie B. Un parc de cinq machines.

L'atelier fait fonctionner simultanément 5 machines, dont les durées de vie $X_1, \dots, X_5$ sont mutuellement indépendantes et suivent toutes la loi $\mathcal{E}(\lambda)$ avec $\lambda = \frac{1}{2000}$. On pose

$$m = \min(X_1, \dots, X_5) \quad \text{et} \quad M = \max(X_1, \dots, X_5)$$

4. Déterminer la loi de m , instant de la première panne du parc, ainsi que $E(m)$.
5. Déterminer la fonction de répartition de M , instant de la dernière panne du parc, puis une densité de M .
6. Soit N le nombre de machines, parmi les cinq, dont la durée de vie dépasse 2000 heures. Déterminer la loi de N , puis calculer $P(N \geq 3)$ à trois décimales.

Partie C. Remplacements successifs.

Un unique poste de travail est équipé d'une machine. Dès qu'elle tombe en panne, elle est immédiatement remplacée par une machine neuve identique. Les durées de vie successives T_1, T_2, T_3 sont mutuellement indépendantes de loi $\mathcal{E}(\lambda)$ avec $\lambda = \frac{1}{2000}$, et l'on note $S_3 = T_1 + T_2 + T_3$ l'instant de la troisième panne.

L'énoncé fournit, pour des variables $T_1, \dots, T_n$ mutuellement indépendantes de loi $\mathcal{E}(\lambda)$ et $S_n = T_1 + \dots + T_n$, la formule

$$P(S_n > t) = e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} \quad (t \geq 0)$$

7. Déterminer une densité de S_2 par le produit de convolution, puis admettre ou retrouver que S_3 a pour densité $t \mapsto \frac{\lambda^3 t^2 e^{-\lambda t}}{2}$ sur $]0, +\infty[$.

8. Calculer $E(S_3)$ et $V(S_3)$.
9. Calculer $P(S_3 > 6000)$ à trois décimales, et comparer ce résultat à $E(S_3)$.

Partie D. Deux types de machines et coût.

Le parc contient en réalité deux types de machines : 70 % sont de type α , de durée de vie $\mathcal{E}\left(\frac{1}{2000}\right)$, et 30 % sont de type β , de durée de vie $\mathcal{E}\left(\frac{1}{3500}\right)$. On tire une machine au hasard dans le parc et on note D sa durée de vie, A l'événement « la machine est de type α » et B l'événement « la machine est de type β ».

On rappelle la formule de l'espérance totale : pour un système complet d'événements de probabilités non nulles (A, B) ,

$$E(D) = P(A) E(D | A) + P(B) E(D | B)$$

10. Calculer la durée de vie moyenne d'une machine tirée au hasard dans le parc.
11. Une machine tirée au hasard a dépassé 4000 heures de fonctionnement. Calculer la probabilité qu'elle soit de type β , à trois décimales, et commenter l'écart avec la proportion initiale de 30 %.

Corrigé

1. Pour une loi exponentielle de paramètre λ , le cours donne $E(X) = \frac{1}{\lambda}$ et $V(X) = \frac{1}{\lambda^2}$. Ici

$$E(X) = 2000 \text{ heures} \quad V(X) = 2000^2 = 4 \times 10^6 \quad \sigma(X) = \sqrt{V(X)} = 2000 \text{ heures}$$

L'égalité $E(X) = \sigma(X)$ n'est pas un hasard : elle est vraie pour toute loi exponentielle, puisque $\sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda}$. Elle signifie que la dispersion des durées de vie est du même ordre que la durée de vie moyenne, autrement dit que ce modèle décrit un matériel très imprévisible : certaines machines cassent très tôt, d'autres durent beaucoup plus longtemps que la moyenne.

2. On utilise directement la queue de la loi exponentielle :

$$P(X > 3000) = e^{-\lambda \times 3000} = e^{-3000/2000} = e^{-1,5} \approx 0,223$$

Un peu plus d'une machine sur cinq dépasse donc 3000 heures.

3. Comme $4000 > 1000$, on a l'inclusion $(X > 4000) \subset (X > 1000)$, donc

$$P_{(X > 1000)}(X > 4000) = \frac{P(X > 4000)}{P(X > 1000)} = \frac{e^{-4000/2000}}{e^{-1000/2000}} = \frac{e^{-2}}{e^{-0,5}} = e^{-1,5} \approx 0,223$$

On retrouve exactement la valeur de la question 2. Il s'agit de l'**absence de mémoire** de la loi exponentielle : la probabilité qu'une machine tienne encore 3000 heures est la même qu'elle soit neuve ou qu'elle ait déjà tourné 1000 heures. La machine ne vieillit pas, elle ne garde aucune trace de son passé.

La conséquence pratique est nette : dans ce modèle, un **remplacement préventif est sans intérêt**. Remplacer une machine en état de marche par une machine neuve ne change rien à sa loi de survie future, et coûte le prix d'une machine. Une politique préventive n'a de sens que pour un matériel qui s'use, ce que la loi exponentielle ne décrit pas.

4. Pour un **minimum**, c'est la queue qui se factorise. Pour $t \geq 0$, l'événement $(m > t)$ signifie que les cinq machines ont dépassé t :

$$P(m > t) = P\left(\bigcap_{i=1}^5 (X_i > t)\right) = \prod_{i=1}^5 P(X_i > t) = (e^{-\lambda t})^5 = e^{-5\lambda t}$$

la deuxième égalité utilisant l'indépendance mutuelle. On reconnaît la queue d'une loi exponentielle de paramètre 5λ , donc

$$m \hookrightarrow \mathcal{E}(5\lambda) = \mathcal{E}\left(\frac{5}{2000}\right) = \mathcal{E}\left(\frac{1}{400}\right)$$

d'où

$$E(m) = \frac{1}{5\lambda} = \frac{2000}{5} = 400 \text{ heures}$$

Le résultat mérite un commentaire : avec cinq machines en service, la première panne survient en moyenne au bout de 400 heures, soit cinq fois plus tôt qu'avec une machine seule. Multiplier les équipements multiplie les occasions de panne.

5. Pour un **maximum**, c'est la fonction de répartition qui se factorise. Pour $t \geq 0$, l'événement $(M \leq t)$ signifie que les cinq machines sont déjà tombées en panne :

$$F_M(t) = \prod_{i=1}^5 P(X_i \leq t) = (1 - e^{-\lambda t})^5$$

et $F_M(t) = 0$ pour $t < 0$. Cette fonction est continue sur $\mathbb{R}$, nulle en 0, croissante, de limite 1 en $+\infty$, et de classe $\mathcal{C}^1$ sauf éventuellement en 0 : M est bien une variable à densité. En dérivant pour $t > 0$:

$$f_M(t) = 5\lambda e^{-\lambda t} (1 - e^{-\lambda t})^4$$

et $f_M(t) = 0$ pour $t \leq 0$. Contrairement au minimum, le maximum ne suit **pas** une loi exponentielle : le parc entier vieillit, même si chaque machine prise isolément ne vieillit pas.

6. Pour chaque machine, considérons l'épreuve « la durée de vie dépasse 2000 heures ». Sa probabilité de succès est

$$p = P(X_i > 2000) = e^{-2000/2000} = e^{-1} \approx 0,367879$$

et elle est la même pour les cinq machines. Les durées de vie étant mutuellement indépendantes, les cinq épreuves le sont aussi. La variable N compte les succès parmi 5 épreuves identiques et indépendantes, donc

$$N \hookrightarrow \mathcal{B}(5, e^{-1})$$

En posant $q = 1 - e^{-1} \approx 0,632121$, on calcule les trois termes utiles :

$$P(N = 3) = \binom{5}{3} p^3 q^2 = 10 \times e^{-3} \times q^2 \approx 10 \times 0,049787 \times 0,399576 \approx 0,198937$$

$$P(N = 4) = \binom{5}{4} p^4 q = 5 \times e^{-4} \times q \approx 5 \times 0,018316 \times 0,632121 \approx 0,057888$$

$$P(N = 5) = p^5 = e^{-5} \approx 0,006738$$

Ces trois événements étant deux à deux incompatibles,

$$P(N \geq 3) = P(N = 3) + P(N = 4) + P(N = 5) \approx 0,263564 \approx 0,264$$

Contrôle d'ordre de grandeur. On a $E(N) = 5e^{-1} \approx 1,84$ machine : on attend donc moins de deux survivantes en moyenne, et il est cohérent qu'en obtenir au moins trois soit un événement minoritaire, de probabilité voisine de 0,26.

7. Les variables T_1 et T_2 sont indépendantes et à densité. Le produit de convolution donne, pour $t > 0$, l'intégrande n'étant non nul que pour $x \in]0, t[$:

$$f_{S_2}(t) = \int_0^t \lambda e^{-\lambda x} \times \lambda e^{-\lambda(t-x)} dx = \lambda^2 e^{-\lambda t} \int_0^t dx = \lambda^2 t e^{-\lambda t}$$

et $f_{S_2}(t) = 0$ pour $t \leq 0$, une somme de durées positives étant positive.

On recommence avec $S_3 = S_2 + T_3$. La variable S_2 ne dépend que de T_1 et T_2 , et T_3 est indépendante de ce couple : le lemme des coalitions assure l'indépendance de S_2 et T_3 , ce qui autorise une nouvelle convolution. Pour $t > 0$:

$$f_{S_3}(t) = \int_0^t \lambda^2 x e^{-\lambda x} \times \lambda e^{-\lambda(t-x)} dx = \lambda^3 e^{-\lambda t} \int_0^t x dx = \lambda^3 e^{-\lambda t} \times \frac{t^2}{2} = \frac{\lambda^3 t^2 e^{-\lambda t}}{2}$$

C'est bien la densité annoncée. Le mécanisme est toujours le même : les exponentielles se recombinent en $e^{-\lambda t}$, indépendant de x , et il ne reste qu'une intégrale de polynôme.

8. Le plus rapide est de passer par la somme, sans utiliser la densité. Chaque T_k a pour espérance $\frac{1}{\lambda} = 2000$ et pour variance $\frac{1}{\lambda^2} = 2000^2$.

Par linéarité de l'espérance, valable sans hypothèse :

$$E(S_3) = 3 \times 2000 = 6000 \text{ heures}$$

Les trois variables étant indépendantes, leurs covariances deux à deux sont nulles et la variance de la somme est la somme des variances :

$$V(S_3) = 3 \times 2000^2 = 12 \times 10^6 \quad \text{donc} \quad \sigma(S_3) = 2000\sqrt{3} \approx 3464 \text{ heures}$$

9. On applique la formule fournie avec $n = 3$ et $t = 6000$. On calcule d'abord

$$\lambda t = \frac{6000}{2000} = 3$$

puis

$$P(S_3 > 6000) = e^{-3} \left(\frac{3^0}{0!} + \frac{3^1}{1!} + \frac{3^2}{2!} \right) = e^{-3} (1 + 3 + 4,5) = 8,5 e^{-3}$$

Avec $e^{-3} \approx 0,049787$:

$$P(S_3 > 6000) \approx 8,5 \times 0,049787 = 0,423190 \approx 0,423$$

Comparaison. L'instant 6000 heures est exactement l'espérance de S_3 , calculée en question **8**. Pourtant la probabilité de dépasser cet instant n'est pas 0,5 mais 0,423 : la troisième panne survient avant sa date moyenne dans près de 58 % des cas. Cela rappelle qu'une espérance n'est pas une médiane, et que pour une loi dissymétrique comme celle-ci, la moyenne est tirée vers le haut par les scénarios longs.

10. Le couple (A, B) forme un système complet d'événements, avec $P(A) = 0,7$ et $P(B) = 0,3$, deux probabilités non nulles. Sachant A , la durée de vie suit $\mathcal{E}\left(\frac{1}{2000}\right)$, donc $E(D | A) = 2000$;

sachant B , elle suit $\mathcal{E}\left(\frac{1}{3500}\right)$, donc $E(D | B) = 3500$.

La formule de l'espérance totale donne

$$E(D) = 0,7 \times 2000 + 0,3 \times 3500 = 1400 + 1050 = 2450 \text{ heures}$$

La durée de vie moyenne du parc est donc de 2450 heures, valeur intermédiaire entre les deux durées de vie des deux types, mais plus proche de celle du type α , qui est majoritaire.

11. Il s'agit d'une inversion de conditionnement, donc de la formule de Bayes. On calcule d'abord les deux termes du dénominateur :

$$P(A) P_A(D > 4000) = 0,7 \times e^{-4000/2000} = 0,7 e^{-2} \approx 0,7 \times 0,135335 \approx 0,094735$$

$$P(B) P_B(D > 4000) = 0,3 \times e^{-4000/3500} = 0,3 e^{-8/7} \approx 0,3 \times 0,318907 \approx 0,095672$$

l'exposant venant de $\frac{4000}{3500} = \frac{8}{7}$. La formule des probabilités totales donne

$$P(D > 4000) = 0,7 e^{-2} + 0,3 e^{-8/7} \approx 0,190407$$

et la formule de Bayes fournit

$$P_{(D>4000)}(B) = \frac{0,3 e^{-8/7}}{0,7 e^{-2} + 0,3 e^{-8/7}} \approx \frac{0,095672}{0,190407} \approx 0,502$$

Commentaire. Avant toute information, une machine avait 30 % de chances d'être de type β . Sachant qu'elle a dépassé 4000 heures, cette probabilité passe à environ 50,2 % : les chances sont désormais équilibrées, alors que le type β était minoritaire dans le parc. Une longue durée de service est donc un indice sérieux en faveur du matériel le plus fiable. C'est le même mécanisme que dans toute inférence par Bayes : l'observation ne change pas la composition du parc, elle change ce que l'on peut raisonnablement croire d'une machine particulière compte tenu de ce que l'on a observé d'elle.

Méthode à retenir. Un problème de fiabilité se décompose toujours selon la structure du parc. Un montage où toutes les machines doivent tenir conduit à un **minimum**, donc à un produit de queues, et le minimum de lois exponentielles indépendantes est encore exponentiel, de paramètre égal à la somme des paramètres. Un montage où il suffit qu'une machine tienne conduit à un **maximum**, donc à un produit de fonctions de répartition, et le résultat n'est plus exponentiel. Un remplacement à l'identique conduit à une **somme**, donc à une convolution. Enfin, un parc hétérogène conduit à un **conditionnement**, donc à l'espérance totale et à Bayes.

Erreur classique. Traiter le maximum comme le minimum et écrire $M \hookrightarrow \mathcal{E}\left(\frac{\lambda}{5}\right)$ par symétrie avec la question 4. C'est faux : seule la queue du minimum se factorise, et la fonction de répartition du maximum n'est pas celle d'une exponentielle, comme le montre la question 5. Autre erreur, en partie D : moyenner les paramètres en écrivant $\bar{\lambda} = 0,7 \times \frac{1}{2000} + 0,3 \times \frac{1}{3500}$, puis annoncer $E(D) = \frac{1}{\bar{\lambda}} = \frac{14000}{6,1} \approx 2295$ heures. Ce sont les espérances qui se moyennent, jamais les paramètres, et l'écart avec la valeur exacte 2450 heures n'est pas négligeable.