Probabilités
Université Clermont Auvergne
IUT Clermont Auvergne
2A BUT Info, 2024-2025

A. El Ouni
A. Khaldi
C. Samir
A. Wohrer

TD6 – Variables aléatoires continues (2è partie)

Exercice 1. Dans une colonie de marmottes de Sibérie orientale, on s'intéresse à la durée de vie X d'une marmotte au hasard. On suppose que X suit une loi exponentielle, de paramètre $\lambda = 0.5$ an^{-1}.

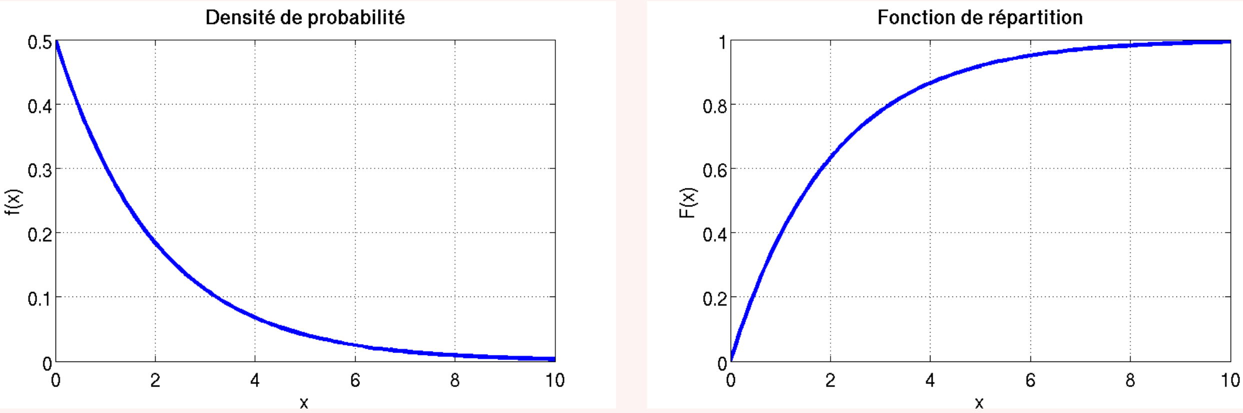
1. Calculez et représentez graphiquement la densité de probabilité $f(x)$ correspondante, puis la fonction de répartition $F(x)$.

Solution: Un calcul d'intégrale permet de retrouver l'expression de la **fonction de répartition d'une variable exponentielle :**

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_0^x \lambda e^{-\lambda u} du \\ &= \big[- e^{-\lambda u} \big]_0^x \\ &= 1 - e^{-\lambda x} \end{aligned}$$

Cette formule est à connaître, et il faut aussi savoir la retrouver facilement !

On aboutit ainsi aux graphes pour la densité de probabilité et la fonction de répartition (ici, pour la valeur $\lambda = 0.5$) :



2. Quelle est l'espérance de vie d'une marmotte dans la colonie ?

Solution: C'est un résultat de cours pour la loi exponentielle : $E(X) = 1/\lambda = 2$ ans.
Remarque : Ce résultat se démontre par un calcul d'intégrale. Il faut calculer

$$E(X) = \int_0^\infty x.\big(\lambda e^{-\lambda x}\big) dx$$

On effectue l'intégration par parties suivantes :

$$\boxed{\begin{array}{ccc} \lambda e^{-\lambda x} & \longleftarrow & -e^{-\lambda x} \\ x & \longrightarrow & 1 \end{array}}$$

et donc

$$\begin{aligned}
E(X) &= \big[-xe^{-\lambda x}\big]_0^\infty - \int_0^\infty 1.\big(-e^{-\lambda x}\big)dx \\
&= 0 - \left[\frac{1}{\lambda}e^{-\lambda x}\right]_0^\infty \\
&= \frac{1}{\lambda}
\end{aligned}$$

3. Quel âge est atteint par exactement 50% des marmottes de la colonie ?

Solution: Ce calcul revient à caractériser la **médiane** de la distribution exponentielle. On cherche le nombre q tel que

$$F(q) = \tfrac{1}{2}$$

ce qui donne, après résolution : $q = \frac{1}{\lambda}.\ln(2)$ (en se souvenant que $\ln(1/2) = -\ln(2)$). Soit ici : $q \simeq 1.39$ ans.

4. a) Une marmotte vient de naître. Quelle probabilité qu'elle atteigne l'âge de 4 ans ?
   b) Une marmotte a 10 ans (pile). Quelle probabilité qu'elle atteigne l'âge de 14 ans ?
   c) Plus généralement, expliquez ce que représente la quantité $P(X > x_0 + t \mid X > x_0)$, et ce qu'elle vaut dans notre modèle. Dans quelles circonstances est-ce réaliste ?

Solution:

a) De manière générale, on a la formule de "survie au temps t" suivante :

$$P(X \geq t) = 1 - F(t) = \exp^{-\lambda.t}$$

Ici, il vient donc $P(X \geq 4) = \exp(-\frac{1}{2}.4) \simeq 0.14$.

b) Il faut calculer la **probabilité conditionnelle** suivante :

$$\begin{aligned}
P(X \geq 14|X \geq 10) &= \frac{P(X \geq 14 \text{ ET } X \geq 10)}{P(X \geq 10)} \\
&= \frac{P(X \geq 14)}{P(X \geq 10)} \\
&= \frac{\exp(-\lambda.14)}{\exp(-\lambda.10)} \\
&= \exp(-\lambda.4)
\end{aligned}$$

ce qui fait donc 0.14 environ, exactement comme à la question précédente.

c) Ce résultat est un peu étonnant ! Une marmotte de 10 ans (donc très âgée pour cette colonie) a autant de chances de survivre encore 4 années, qu'une marmotte qui vient de naître ! En réalité le calcul précédent est valide pour n'importe quel âge considéré. On a :

$$P(X > x_0 + t \mid X > x_0) \ = e^{-\lambda . t}$$

soit avec des mots : *Quel que soit l'âge x_0 d'une marmotte, sa probabilité de survivre encore t années est égale à $e^{-\lambda t}$.*

Clairement, cette situation est un peu fausse dans la réalité : une marmotte très agée est plus fragile qu'une marmotte jeune, et a donc davantage de chances de mourir. Toutefois, on peut considérer que ce phénomène de vieillissement est négligeable si *les causes de mort des marmottes sont principalement extérieures*. Ce modèle est donc justifié dans la mesure où les marmottes constituent une espèce de type « proie » : il y a un 'réservoir' de marmottes qui meurent, un peu au hasard, sous l'action des prédateurs. En moyenne cette mort survient au bout de 2 ans, mais une marmotte très chanceuse peut, en théorie, vivre sans limite d'âge.

Exercice 2. On considère une cible circulaire, de rayon $R = 1$. Deux joueurs A et B s'affrontent au tir à l'arc, le vainqueur étant celui dont la flèche finit le plus près du centre. On introduit les variables aléatoires : $X_A =$ distance entre la flèche du joueur A et le centre, et de même $X_B =$ distance entre la flèche du joueur B et le centre.

1. Le joueur A tire sa flèche de façon *uniforme* dans toute la cible, c'est-à-dire que toutes les régions de la cible ont la même chance de recevoir la flèche.
   a) Soit un nombre $0 \leq x \leq 1$. Quelle est l'aire d'un disque de rayon x ?
   b) Déduisez-en la fonction de répartition $F_A(x) = P(X_A \leq x)$. Tracez-la.
   c) Déduisez-en la densité de probabilité $f_A(x)$ pour la variable X_A.
   d) Quelle est la probabilité que le joueur A tombe à une distance $x \leq 1/2$ du centre ?

Solution:

a) Si $C(x)$ est le disque de rayon x, son aire est égale à πx^2.

b) Il s'agit ici d'une distribution uniforme *dans la cible 2D*. Autrement dit, pour toute 'sous-région' C dans la cible,

$$P(\text{flèche} \in C) = \frac{\text{Aire de } C}{\text{Aire totale de la cible}}$$

L'aire totale de la cible est $\pi R^2 = \pi$ (car $R = 1$). De même, notons $C(x)$ le disque de rayon x, dont l'aire est égale à πx^2. On peut donc écrire

$$P(X_A \leq x) = P(\text{flèche} \in C(x)) = \frac{\pi x^2}{\pi} = x^2.$$

Autrement dit, on vient de calculer la **fonction de répartion** de la variable X_A :

$$F_A(x) = x^2.$$

Comme toute fonction de répartition, il s'agit d'une fonction croissante, de 0 (lorsque $x = 0$) à 1 (lorsque $x = 1$).

c) On en déduit, en dérivant, la **densité de probabilité** correspondante :

$$f_A(x) = F_A'(x) = 2x.$$

d) Par définition de la fonction de répartition,

$$P(X_A \leq \tfrac{1}{2}) = F_A(\tfrac{1}{2}) = (1/2)^2 = \frac{1}{4} \quad = 25\%.$$

2. Le joueur B, quant à lui, a la fonction de répartition suivante :

$$F_B(x) = \begin{cases} 0 & \text{pour } x \leq 0, \\ \sqrt{x} & \text{pour } 0 < x < 1, \\ 1 & \text{pour } x \geq 1. \end{cases}$$

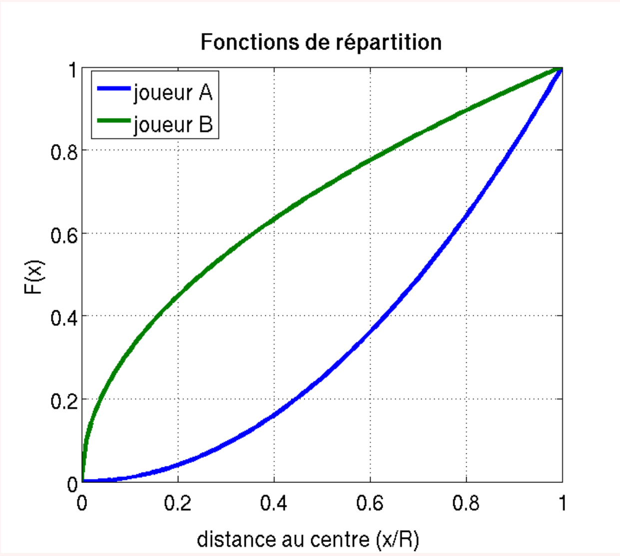
a) Tracez $F_B(x)$, dans la même figure que $F_A(x)$. Visuellement, lequel des deux joueurs semble le plus à même de gagner le concours ?

b) Calculez et tracez la densité de probabilité $f_B(x)$.

c) Quelle est la probabilité que le joueur B tombe à une distance $x \leq 1/2$ du centre ?

Solution:

a) Voici le graphe des deux fonctions de répartition :

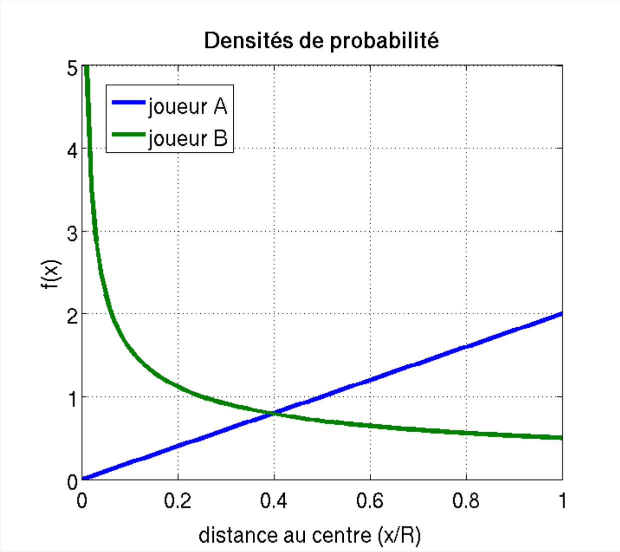


Il est clair que le joueur B a plus de chances de gagner le concours. En effet, pour tout $x \in [0, 1]$, on constate que $F_B(x) \geq F_A(x)$. Autrement dit, quel que soit le seuil x que l'on fixe, le joueur B a davantage de chances que le joueur A d'atterir à l'intérieur du disque de rayon x. Il est donc clairement plus adroit (en moyenne).

b) On peut calculer, en dérivant, la densité de probabilité pour le joueur B :

$$f_B(x) = F_B'(x) = \frac{1}{2\sqrt{x}}.$$

Voici les densités de probabilités pour les deux joueurs (qui correspondent donc aux fonctions dérivées du graphe précédent) :



Sous cette représentation aussi, il est clair que le joueur B tombe généralement plus près du centre que le joueur A.

c) Par définition de la fonction de répartition,

$$P(X_B \leq \tfrac{1}{2}) = F_B(\tfrac{1}{2})\frac{1}{\sqrt{2}} \quad \simeq 71\%.$$

Comme prévu, le joueur B a beaucoup plus de chances que le joueur A de tomber à distance $\leq (1/2)$ du centre.

3. Les deux joueurs tirent une flèche simultanément, et indépendamment l'un de l'autre.
   a) Supposons que le joueur A a tiré à distance « y » du centre. Quelle est la probabilité que le joueur B tombe plus proche que y ?
   b) Déduisez-en la probabilité que le joueur B gagne le concours.

Solution: La question précédente nous a montré que le joueur B est clairement meilleur (en moyenne) que le joueur A. Cependant, le joueur A peut toujours gagner le concours s'il a de la chance sur son lancer ! Dans cette question, on calcule la probabilité que cela se produise.

a) Si le joueur A a tiré à distance « y » du centre, le joueur B gagne le concours avec probabilité :

$$P(X_B \leq y) = F_B(y) = \sqrt{y}$$

On peut réécrire cela sous la forme d'une probabilité conditionnelle :

$$P(B \text{ gagne}|X_A = y) = F_B(y) = \sqrt{y}$$

b) Il s'agit à présent de moyenner la probabilité précédente, sur toutes les valeurs possibles prenables par X_A. Cela revient à appliquer une formule de la probabilité totale (comme la variable X_A est continue, on utilise une intégrale au-lieu d'une somme).

$$P(B \text{ gagne}) = \int_{y=0}^{1} P(B \text{ gagne et } X_A = y)\mathrm{d}y$$

$$= \int_{y=0}^{1} P(B \text{ gagne} | X_A = y) f_A(y) \mathrm{d}y$$
$$= \int_{y=0}^{1} F_B(y) f_A(y) \mathrm{d}y.$$

Il reste simplement à calculer cette intégrale. Dans notre cas, on obtient

$$P(B \text{ gagne}) = \int_{y=0}^{1} \sqrt{y}.2y\mathrm{d}y \quad = 2\int_{y=0}^{1} y^{(3/2)}\mathrm{d}y$$
$$= 2\left[\tfrac{2}{5}y^{(5/2)}\right]_0^1 \quad = \tfrac{4}{5}.$$

Ainsi, le joueur B a 4 chances sur 5 de gagner le concours (et le joueur A, 1 chance sur 5).

Contexte : Ce genre de calcul est très typique dans le domaine de la *détection du signal* (qui fut développée pendant la Seconde Guerre Mondiale, dans le cadre de la détection des signaux radar). Dans ce contexte, X_A représente un « bruit de fond » et X_B représente un « signal ». Le nombre $d = P(X_B > X_A)$ (calculé à la dernière question) permet de quantifier à quel point le signal « se détache » du bruit, et pourra donc être détecté par une méthode appropriée.

Exercice 3.
Un plateau circulaire est découpé en 8 zones angulaires identiques, numérotées de 1 à 8. On fait tourner une aiguille autour de l'axe du plateau.

1. Soit X la variable aléatoire qui désigne la zone sur laquelle l'aiguille s'arrête.
   (a) Expliquez pourquoi X est une v.a. discrète.
   (b) Donnez la loi de probabilité de X.
   (c) Trouvez la fonction de répartition $F(x)$ de la variable X, et dessinez son graphe.
   (d) Calculez l'espérance, la variance et l'écart-type de X.
2. Soit θ la variable aléatoire qui désigne l'angle auquel s'arrête l'aiguille (la position horizontale vers la droite correspondant à $\theta = 0$).
   (a) Expliquez pourquoi θ est une v.a. continue.
   (b) Quelles sont les valeurs possibles pour θ ?
   (c) Sans faire le calcul, donnez des valeurs pour
   $$P(0 \leq \theta \leq \pi), \quad P(\pi \leq \theta \leq 2\pi), \quad P(0 \leq \theta \leq 2\pi), \quad P(\frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2})$$
   (d) Proposer une fonction de densité pour θ.
   (e) Trouvez la fonction de répartition $F(\theta)$ de la variable θ, et dessinez son graphe.
   (f) Calculez l'espérance, la variance et l'écart type de θ.
3. Un joueur veut miser une grande somme. À votre avis, a-t-il plus de chance de gagner en misant sur une zone (valeur de X), ou sur une valeur de l'angle θ ? Justifiez.

Solution: Fait en classe.

Exercice 4.
Un système électronique fait intervenir deux composants de durées de vies respectives X et Y. Ces deux variables aléatoires suivent des lois exponentielles indépendantes de paramètres respectifs λ_1 et λ_2.

1. On suppose que les composants sont montés en série.
   (a) Expliquer pourquoi la durée de vie D du système s'exprime par : $D = \min(X, Y)$.
   (b) Soit un nombre $a > 0$. Exprimer $P(D > a)$ à l'aide des variables X et Y.
   (c) En déduire $F_D(a)$, la fonction de répartition de D. Quelle loi suit la variable D ?
   (d) Donner l'espérance $E(D)$. Comment ce nombre se compare-t-il à $E(X)$ et $E(Y)$?

Solution:

(a) Composants montés en série $\rightarrow$ le système est en panne dès que l'un de deux composants est en panne. Par suite, $D = \min(X, Y)$.

(b) Pour tout $a \geq 0$ on a :

$$\begin{array}{rcl} P(D > a) & = & P(\min(X, Y) > a) \\ & = & P(X > a \cap Y > a) \\ & = & P(X > a) \times P(Y > a) \end{array}$$

car X et Y sont indépendantes. D'autre part, on a :

$$\begin{array}{rcl} P(X > a) & = & 1 - P(X \leq a) \\ & = & 1 - F_X(a) \\ & = & 1 - (1 - e^{-\lambda_1 a}) \\ & = & e^{-\lambda_1 a} \end{array}$$

De même, $P(Y > a) = e^{-\lambda_2 a}$. Par conséquent,

$$\begin{array}{rcl} P(D > a) & = & e^{-\lambda_1 a} \times e^{-\lambda_2 a} \\ & = & e^{-(\lambda_1 + \lambda_2) a} \end{array}$$

(c) La fonction de répartition de D est :

$$\begin{array}{rcl} F_D(a) & = & P(D \leq a) \\ & = & 1 - P(D > a) \\ & = & 1 - e^{-(\lambda_1 + \lambda_2) a} \end{array}$$

Par conséquent, D suit la loi exponentielle de paramètre $\lambda_1 + \lambda_2$: $D \sim \mathcal{E}(\lambda_1 + \lambda_2)$.

(d) L'espérance de D est

$$E(D) = \frac{1}{\lambda_1 + \lambda_2}$$

Par comparaison, $E(X) = \frac{1}{\lambda_1}$, on a donc $E(D) < E(X)$. Ce qui est normal puisque, par définition D est toujours inférieur (ou égal) à X. Le même raisonnement tient également pour Y.

2. On suppose maintenant que les composants sont montés en parallèle, et que le système fonctionne tant qu'au moins un des composants fonctionne. On note S la durée de vie du système.
   (a) Expliquer pourquoi la durée de vie S du système s'exprime par : $S = \max(X, Y)$.
   (b) Calculer $F_S(a)$, la fonction de répartition de S, puis sa densité de probabilité $f_S(a)$.
   (c) Calculer l'espérance $E(S)$. Prouver que $E(S) > E(D)$.

Solution:

(a) Composants montés en parallèles $\rightarrow$ le système est en panne dès que les deux composants sont en panne. Par suite, $S = \max(X, Y)$.

(b) On a :

$$\begin{aligned} F_S(a) &= P(S \leq a) \\ &= P(\max(X, Y) \leq a) \\ &= P(X \leq a \cap Y \leq a) \\ &= P(X \leq a) \times P(Y \leq a) \\ &= (1 - e^{-\lambda_1 a}) \times (1 - e^{-\lambda_2 a}) \\ &= 1 - e^{-\lambda_1 a} - e^{-\lambda_2 a} + e^{-(\lambda_1 + \lambda_2)a} \end{aligned}$$

(c) La densité de probabilité de S notée $f_S(a)$ est la dérivée de la fonction de répartition $F_S(a)$. Par suite, $f_S(a) = \lambda_1 e^{-\lambda_1 a} + \lambda_2 e^{-\lambda_2 a} - (\lambda_1 + \lambda_2)e^{-(\lambda_1+\lambda_2)a}$.

(d) On remarque que la densité de probabilité de S s'écrit comme : $f_S(a) = f_X(a) + f_Y(a) - f_D(a)$ où $f_X(a)$, $f_Y(a)$ et $f_D(a)$ sont les densités de probabilité de X, Y et D. Ainsi, l'espérance de S est donnée par :

$$\begin{aligned} E(S) &= \int_0^{+\infty} a f_S(a) da \\ &= \int_0^{+\infty} a(f_X(a) + f_Y(a) - f_D(a)) da \\ &= \int_0^{+\infty} a f_X(a) da + \int_0^{+\infty} a f_Y(a) - \int_0^{+\infty} a f_D(a) da \\ &= E(X) + E(Y) - E(D) \\ &= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2} \end{aligned}$$

(e) On a :

$$\begin{aligned} E(S) - E(D) &= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{2}{\lambda_1 + \lambda_2} \\ &= \frac{\lambda_2(\lambda_1 + \lambda_2)}{\lambda_1\lambda_2(\lambda_1 + \lambda_2)} + \frac{\lambda_1(\lambda_1 + \lambda_2)}{\lambda_1\lambda_2(\lambda_1 + \lambda_2)} - \frac{2\lambda_1\lambda_2}{\lambda_1\lambda_2(\lambda_1 + \lambda_2)} \\ &= \frac{\lambda_1\lambda_2 + \lambda_2^2 + \lambda_1^2 + \lambda_1\lambda_2 - 2\lambda_1\lambda_2}{\lambda_1\lambda_2(\lambda_1 + \lambda_2)} \\ &= \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1\lambda_2(\lambda_1 + \lambda_2)} > 0 \end{aligned}$$

c'est à dire, $E(S) > E(D)$. Donc, la durée de vie en parallèle est plus longue que celle en série.