PROBABILITÉS ET STATISTIQUES 1

I. Lois continues (SOLUTION)

Guillaume Euvrard

Exercice 1 **Densité**

Dans chacun des cas ci-dessous, discuter en fonction de $k \in \mathbb{R}$ si la fonction f est une densité.

En cas de réponse positive : considérons une variable aléatoire X de densité f.

a. Déterminer sa fonction de répartition.

b. Donner un intervalle de prédiction bilatéral à 95%, c'est-à-dire un intervalle $[a, b]$ tel que

$$P(X \in [a,b]) = 0.95, \qquad P(X < a) = 0.025 \qquad \text{et} \qquad P(X > b) = 0.025$$

❶ $f(x) = \begin{cases} \frac{k}{x} & \text{si } x \geqslant 1 \\ 0 & \text{sinon} \end{cases}$ ❷ $f(x) = \begin{cases} \frac{k}{x^2} & \text{si } x \geqslant 1 \\ 0 & \text{sinon} \end{cases}$ ✪❸ $f(x) = \begin{cases} \frac{k}{x^3} & \text{si } x \geqslant 2 \\ 0 & \text{sinon} \end{cases}$

❹ $f(x) = k\,e^{-2x}$ ❺ $f(x) = \begin{cases} k\,e^{-2x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$ ❻ $f(x) = \begin{cases} k & \text{si } x \in [A, B] \\ 0 & \text{sinon} \end{cases}$

Solution :

De façon générale, tout au long de ce module, il faut essayer le plus possible de représenter graphiquement les densités. Quand on parle d'intégrales, beaucoup d'étudiants pensent en terme de primitives plutôt que d'aires. S'ils restent dans cette façon de penser, ils ne comprendront rien aux lois continues.

Une fonction f est une densité si :

- elle est positive
- $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge et vaut 1.

Quand ces propriétés sont vérifiées, alors la fonction de répartition est $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t$

Ce qui donne, pour les fonctions proposées :

❶ L'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ diverge pour tout $k \neq 0$ car $\displaystyle\int_{1}^{+\infty} f(t)\,\mathrm{d}t = \int_{1}^{+\infty} \frac{k}{t}\,\mathrm{d}t$ diverge.
Pour $k = 0$, l'intégrale converge mais ne vaut pas 1.

❷ $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_{1}^{+\infty} \frac{k}{t^2}\,\mathrm{d}t = \left[-\frac{k}{t}\right]_1^{+\infty} = k$. Donc l'intégrale converge, elle vaut 1 ssi $k = 1$.
La fonction de répartition est alors

- si $x < 1$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = 0$.
- si $x \geq 1$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = \underbrace{\int_{-\infty}^{1} f(t)\,\mathrm{d}t}_{0} + \int_{1}^{x} f(t)\,\mathrm{d}t = 0 + \left[-\frac{1}{t}\right]_1^{x} = 1 - \frac{1}{x}$.

Donc, $F(x) = \begin{cases} 1 - \frac{1}{x} & \text{si } x \geqslant 1 \\ 0 & \text{sinon} \end{cases}$.

En posant $F(a) = 0.025$ et $1 - F(b) = 0.025$, on trouve

- $F(a) = 0.025 \Longrightarrow 1 - \frac{1}{a} = 0.025 \Longrightarrow a = \frac{40}{39}$
- $F(b) = 0.975 \Longrightarrow 1 - \frac{1}{b} = 0.975 \Longrightarrow b = 40.$

❸ $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_2^{+\infty} \frac{k}{t^3}\,\mathrm{d}t = \left[-\frac{k}{2t^2}\right]_2^{+\infty} = \frac{k}{8}$. Donc l'intégrale converge, elle vaut 1 ssi $k = 8$.

La fonction de répartition est alors $F(x) = \begin{cases} \left[-\frac{8}{2t^2}\right]_2^x = 1 - \frac{4}{x^2} & \text{si } x \geqslant 2 \\ 0 & \text{sinon} \end{cases}$

En posant $F(a) = 0.025$ et $1 - F(b) = 0.025$, on trouve $a = \sqrt{\frac{160}{39}}$ et $b = 4\sqrt{10}$.

❹ L'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ diverge pour tout $k \neq 0$ car $\displaystyle\int_{-\infty}^{0} f(t)\,\mathrm{d}t = \int_{-\infty}^{0} e^{-2t}\,\mathrm{d}t$ diverge.
Pour $k = 0$, l'intégrale converge mais ne vaut pas 1.

❺ $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_0^{+\infty} k\,e^{-2t}\,\mathrm{d}t = \left[-\frac{k}{2}\,e^{-2t}\right]_0^{+\infty} = \frac{k}{2}$. L'intégrale converge, elle vaut 1 ssi $k = 2$.
La fonction de répartition est alors

- si $x < 0$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = 0.$
- si $x \geq 0$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = \underbrace{\int_{-\infty}^{0} f(t)\,\mathrm{d}t}_{0} + \int_0^x f(t)\,\mathrm{d}t = \left[-e^{-2t}\right]_0^x = 1 - e^{-2x}.$

Donc, $F(x) = \begin{cases} 1 - e^{-2x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$

En posant $F(a) = 0.025$ et $1 - F(b) = 0.025$, on trouve

- $F(a) = 0.025 \Longrightarrow 1 - e^{-2a} = 0.025 \Longrightarrow a = \frac{\ln(0.975)}{-2}$
- $F(b) = 0.975 \Longrightarrow 1 - e^{-2b} = 0.975 \Longrightarrow b = \frac{\ln(0.025)}{-2}.$

❻ $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_A^B k\,\mathrm{d}t = [kt]_A^B = k(B - A)$. L'intégrale converge, elle vaut 1 ssi $k = \frac{1}{B-A}$.
La fonction de répartition est alors

- si $x < A$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = 0.$
- si $A \leq x \leq B$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = \underbrace{\int_{-\infty}^{A} f(t)\,\mathrm{d}t}_{0} + \int_A^x f(t)\,\mathrm{d}t = \left[\frac{t}{B-A}\right]_A^x = \frac{x-A}{B-A}.$
- si $x > B$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = \underbrace{\int_{-\infty}^{A} f(t)\,\mathrm{d}t}_{0} + \underbrace{\int_A^B f(t)\,\mathrm{d}t}_{1} + \underbrace{\int_B^x f(t)\,\mathrm{d}t}_{0} = 1.$

Donc $F(x) = \begin{cases} 0 & \text{si } x \leqslant A \\ \frac{x-A}{B-A} & \text{si } A \leqslant x \leqslant B \\ 1 & \text{si } B \leqslant x \end{cases}$

En posant $F(a) = 0.025$ et $1 - F(b) = 0.025$, on trouve $a = A + \frac{B-A}{40}$ et $b = B - \frac{B-A}{40}$.

Exercice 2 **✫ Changement de variable**

Soit X une variable aléatoire admettant une densité f.

❶ Écrire la fonction de répartition de X. Notons F cette fonction.

❷ Exprimer la fonction de répartition de $2X + 1$ au moyen de F.
En déduire une densité de $2X + 1$.

❷ Exprimer la fonction de répartition de $-2X + 1$ au moyen de F.
En déduire une densité de $-2X + 1$.

Solution :

❶ $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t$. Donc $F' = f$, sauf peut-être en un nombre fini de valeurs.

❷ Soit $Y = 2X + 1$ et G sa fonction de répartition. Alors

$$G(x) = P(2X + 1 \leqslant x) = P\left(X \leqslant \frac{x-1}{2}\right) = F\left(\frac{x-1}{2}\right)$$

D'où la densité $g(x) = G'(x) = \frac{1}{2} f\left(\frac{x-1}{2}\right)$

❸ Soit $Z = -2X + 1$ et H sa fonction de répartition. Alors

$$H(x) = P(-2X + 1 \leqslant x) = P\left(X \geqslant \frac{1-x}{2}\right) = 1 - F\left(\frac{1-x}{2}\right)$$

D'où la densité $h(x) = H'(x) = \frac{1}{2} f\left(\frac{1-x}{2}\right)$

Exercice 3 — **Loi uniforme**

Soit $(a, b) \in \mathbb{R}^2$ tel que $a < b$ et X une variable aléatoire admettant la densité

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon} \end{cases}$$

On dit que X suit une loi uniforme sur $[a, b]$ et on note : $X \rightsquigarrow \text{Unif}\,(a, b)$.

❶ Quelle est la fonction de répartition de X ? (voire Exo.1 Q.❻)

❷ Soit $Z = \dfrac{X-a}{b-a}$. Déterminer la fonction de répartition de Z puis sa densité.

❸ Dans le cadre d'un code informatique, vous avez besoin de tirer une variable aléatoire uniformément distribuée entre 25 et 30.
Supposons que vous disposez d'une fonction `rand()` qui retourne un nombre aléatoire uniformément distribué entre 0 et 1. Comment pouvez-vous l'utiliser ?

Solution :

❶ Déjà fait, $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = \begin{cases} 0 & \text{si } x \leqslant a \\ \frac{x-a}{b-a} & \text{si } a \leqslant x \leqslant b \\ 1 & \text{si } b \leqslant x \end{cases}$

❷ Soit G la fonction de répartition de Z. Alors $G(z) = P\left(\frac{X-a}{b-a} \leqslant z\right) = P(X \leqslant a + z(b-a))$.

D'où $G(z) = F(\underbrace{a + z(b-a)}_{x}) = \begin{cases} 0 & \text{si } a + z(b-a) \leqslant a \\ \frac{a+z(b-a)-a}{b-a} & \text{si } a \leqslant a + z(b-a) \leqslant b \\ 1 & \text{si } b \leqslant a + z(b-a) \end{cases} = \begin{cases} 0 & \text{si } z \leqslant 0 \\ z & \text{si } 0 \leqslant z \leqslant 1 \\ 1 & \text{si } 1 \leqslant z \end{cases}$

Donc une densité de Z est $g(z) = G'(z) = \begin{cases} 1 & \text{si } 0 \leqslant z \leqslant 1 \\ 0 & \text{sinon} \end{cases}$ Ainsi, $Z \rightsquigarrow \text{Unif}\,(0, 1)$.

❸ On définit $Z = \frac{X-25}{30-25} \rightsquigarrow \text{Unif}\,(0, 1)$. Ainsi, la fonction `rand()` tire Z.
On peut donc tirer X par l'instruction : `x = 25 + (30-25)*rand();`

Exercice 4 **Loi exponentielle**

Un serveur reçoit des requêtes de clients. On définit la variable aléatoire

$$T = \text{«Temps d'attente avant la prochaine requête»}$$

❶ Dans un premier temps, on suppose qu'il existe $\lambda > 0$ tel qu'une densité de T est

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

On dit alors que T suit une loi exponentielle de paramètre λ : $T \rightsquigarrow \operatorname{Exp}(\lambda)$.

(a) Vérifier que f est bien une densité et donner la fonction de répartition de T.

(b) Trouver $a \in \mathbb{R}$ tel que $P(T \leqslant a) = 0.95$.

(c) Soit $t_0 > 0$. Pour tout $\Delta t \in \mathbb{R}$, calculer la probabilité conditionnelle

$$P(T > t_0 + \Delta t \,|\, T > t_0)\,.$$

Comparer avec $P(T > \Delta t)$.
On dit alors que la variable T est «sans mémoire». Expliquer cette expression.

❷ **(Bonus)** On suppose maintenant que la loi de T est **sans mémoire**. De plus, T ne peut prendre que des valeurs positives et on suppose sa fonction de répartition continue et dérivable sur $\mathbb{R}^+$.

(a) Soit $t \in \mathbb{R}^+$. Traduire au moyen de la fonction de répartition F l'hypothèse :

$$\forall\, \Delta t \in \mathbb{R}_+^*, \quad P(T \leqslant t + \Delta t \,|\, T > t) = P(T \leqslant \Delta t)$$

(b) En divisant cette relation par Δt et en faisant tendre Δt vers 0, trouver une équation différentielle que F satisfait.

(c) En déduire que X suit une loi exponentielle.

Solution :

❶ Si $f(t) = \begin{cases} \lambda\, e^{-\lambda t} & \text{si } t \geqslant 0 \\ 0 & \text{sinon} \end{cases}$

(a) f est bien positive. De plus,

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_0^{+\infty} \lambda\, e^{-\lambda t}\,\mathrm{d}t = \left[-e^{-\lambda t}\right]_0^{+\infty} = 1$$

La fonction de répartition est

- si $x < 0$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = 0$.
- si $x \geq 0$ alors $F_X(x) = \displaystyle\int_{-\infty}^{x} f(t)\,\mathrm{d}t = \underbrace{\int_{-\infty}^{0} f(t)\,\mathrm{d}t}_{0} + \int_0^x f(t)\,\mathrm{d}t = \left[-e^{-\lambda t}\right]_0^x = 1 - e^{-\lambda x}$.

Donc $F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geqslant 0 \\ 0 & \text{sinon} \end{cases}$

(b) En résolvant $F(a) = 0.95$, on trouve $a = -\frac{\ln(0.05)}{\lambda} = \frac{\ln(20)}{\lambda}$. Un intervalle de prédiction à 95% est $[0, a]$.

(c) $P(T > t_0 + \Delta t \,|\, T > t_0) = \dfrac{P(T > t_0 + \Delta t \text{ et } T > t_0)}{P(T > t_0)}$

Cas $\Delta t \leqslant 0$ **:** $\quad P(T > t_0 + \Delta t \text{ et } T > t_0) = P(T > t_0)$

Donc $P(T > t_0 + \Delta t \,|\, T > t_0) = 1 = P(T > \Delta t)$. En effet $P(T > \Delta t) = 1 - F(\underbrace{\Delta t}_{<0}) = 1$

Cas $\Delta t > 0$ **:** $\quad P(T > t_0 + \Delta t \text{ et } T > t_0) = P(T > t_0 + \Delta t)$

Donc $P(T > t_0 + \Delta t \,|\, T > t_0) = \dfrac{e^{-\lambda(t_0+\Delta t)}}{e^{-\lambda t_0}} = e^{-\lambda \Delta t}$

Dans les deux cas, $P(T > t_0 + \Delta t \,|\, T > t_0) = P(T > \Delta t)$.
Et, par passage au complémentaire, $P(T \leqslant t_0 + \Delta t \,|\, T > t_0) = P(T \leqslant \Delta t)$.

L'expression «sans mémoire» veut dire que, si on est à la date t_0 et qu'on n'a pas vu de requête arriver, la loi de ΔT (le temps restant à attendre) est la même que la loi de T (le temps qu'il fallait attendre à l'instant initial).

❷ On suppose que la loi de T est sans mémoire, que T ne prend que des valeurs positives, que sa fonction de répartition F est dérivable sur $\mathbb{R}^+$.

(a) Pour tout $\Delta t \in \mathbb{R}_+^*$, $P(T \leqslant t + \Delta t \,|\, T > t) = \dfrac{P(t < T < T + \Delta t)}{P(T > t)} = \dfrac{F(t+\Delta t) - F(t)}{1 - F(t)}$

La propriété «sans mémoire» s'écrit donc : $\dfrac{F(t+\Delta t) - F(t)}{1 - F(t)} = F(\Delta t)$

Soit encore : $F(t + \Delta t) - F(t) = F(\Delta t)\Big(1 - F(t)\Big)$

(b) En divisant par Δt et en faisant tendre Δt vers 0, on obtient : $F'(t) = F'(0)\Big(1 - F(t)\Big)$

En effet, comme $F(0) = 0$, $\dfrac{F(\Delta t)}{\Delta t} = \dfrac{F(\Delta t) - F(0)}{\Delta t} \xrightarrow[\Delta t \to 0]{} F'(0)$

Comme $F'(0)$ ne dépend pas de t, c'est une constante. Appelons la λ. Alors F est solution de l'équation différentielle :

$$F' + \lambda F = \lambda$$

(c) Résolvons l'équation différentielle :
Équation homogène : $F_0(t) = ke^{-\lambda t}, \quad k \in \mathbb{R}$
Solution particulière : $F(t) = 1$ est une solution particulière.

Finalement, l'ensemble des solutions sur $\mathbb{R}^+$ est : $F(t) = 1 + ke^{-\lambda t}, \quad k \in \mathbb{R}$.

Comme de plus, $F(0) = 0$, on déduit que $k = -1$ et que : $\quad F(t) = 1 - e^{-\lambda t}$.

On peut faire remarquer que, en relâchant les hypothèses, la loi exponentielle n'est pas la seule loi sans mémoire. La loi géométrique :

$$X(\Omega) = \mathbb{N}^*, \qquad \forall n \in \mathbb{N}^*, P(X{=}n) = p\,(1-p)^{n-1} \quad \text{où } p \in \,]0,1[$$

est aussi sans mémoire.

Exercice 5 **Loi normale**

Soit Z une variable aléatoire admettant pour densité la fonction $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

❶ Exprimer sa fonction de répartition sous une forme intégrale. On note Φ cette fonction qu'on ne cherchera pas à expliciter.

❷ Expliquer pourquoi, pour tout $\beta > 0$, $\Phi(-\beta) = 1 - \Phi(\beta)$.
On admet que, pour $\beta = 1.96$, on a $\Phi(-\beta) = 1 - \Phi(\beta) = 0.025$.

❸ Soit $(m, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$ et $X = m + \sigma Z$. Exprimer la fonction de répartition de X à partir de Φ, puis la densité de X.

Donner un intervalle de prédiction bilatéral à 95% pour la variable X.

On dit que X suit une loi normale de paramètres m et σ^2 : $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$.

Solution :

❶ Il faut admettre que $\displaystyle\int_{-\infty}^{+\infty} \varphi(t)\,\mathrm{d}t = 1$.

On obtient $\Phi(x) = \displaystyle\int_{-\infty}^{x} \varphi(t)\,\mathrm{d}t$

❷ La fonction φ étant paire, on a

$$\Phi(\beta) + \Phi(-\beta) = \int_{-\infty}^{\beta} \varphi(t)\,\mathrm{d}t + \underbrace{\int_{-\infty}^{-\beta} \varphi(t)\,\mathrm{d}t}_{u=-t} = \int_{-\infty}^{\beta} \varphi(t)\,\mathrm{d}t + \int_{\beta}^{+\infty} \varphi(u)\,\mathrm{d}u = \int_{-\infty}^{+\infty} \varphi(t)\,\mathrm{d}t = 1$$

❸ La fonction de répartition de X est

$$F(x) = P(X \leqslant x) = P\left(Z \leqslant \frac{x-m}{\sigma}\right) = \Phi\left(\frac{x-m}{\sigma}\right)$$

Sa densité est donc

$$F'(x) = \frac{1}{\sigma}\,\Phi'\left(\frac{x-m}{\sigma}\right) = \frac{1}{\sigma}\,\varphi\left(\frac{x-m}{\sigma}\right)$$

Finalement, $\quad f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left(\dfrac{(x-m)^2}{2\sigma^2}\right)$

D'après le changement de variable, on a

$$P(a \leq Z \leq b) = P\left(a \leq \frac{X-m}{\sigma} \leq b\right) = P(m - a\sigma \leq X \leq m + b\sigma)$$

Pour un intervalle de prédiction : $P(Z \in [-1.96, 1.96]) = 95\%$, donc $P(X \in [m - 1.96\sigma, m + 1.96\sigma]) = 95\%$.

Exercice 6 **✮ Loi uniforme**

Soit une variable aléatoire X telle que $X \rightsquigarrow \mathrm{Unif}\,(-1, 2)$

❶ Déterminer la fonction de répartition de $|X|$.

❷ En déduire une densité de $|X|$.

Solution :

❶ Considérons d'abord la fonction de répartition de X : $F(x) = \begin{cases} 0 & \text{si } x \leqslant -1 \\ \frac{x+1}{3} & \text{si } -1 \leqslant x \leqslant 2 \\ 1 & \text{si } 2 \leqslant x \end{cases}$

La fonction de répartition de $|X|$ est $G(x) = P(|X| \leqslant x)$.

Si $x < 0$, $G(x) = 0$

Si $x \geqslant 0$, $G(x) = P(-x \leqslant X \leqslant x) = F(x) - F(-x)$.

Cas où $x \in [0,1]$: $G(x) = \frac{x+1}{3} - \frac{-x+1}{3} = \frac{2x}{3}$

Cas où $x \in [1,2]$: alors $F(-x) = 0$ et $G(x) = \frac{x+1}{3}$

Cas où $x > 2$: alors $F(-x) = 0$ et $F(x) = 1$ d'où : $G(x) = 1$.

Finalement, $G(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{2x}{3} & \text{si } x \in [0,1] \\ \frac{x+1}{3} & \text{si } x \in [1,2] \\ 1 & \text{si } x \geqslant 1 \end{cases}$

❷ Une densité g de G s'obtient en dérivant G : $g(x) = \begin{cases} \frac{2}{3} & \text{si } x \in [0,1[\\ \frac{1}{3} & \text{si } x \in]1,2] \\ 0 & \text{sinon} \end{cases}$