Lois à densité usuelles

1. Variables à densité

On appelle *densité de probabilité* toute fonction f définie sur $\mathbb{R}$, à valeurs positives ou nulles, telle que :

$$\boxed{\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1}\ .$$

Les densités effectivement utilisées en modélisation sont des fonctions continues sur $\mathbb{R}$, privé au maximum d'un nombre fini de points, mais rien n'interdit d'utiliser des fonctions plus générales. [1]

Définition

On dit qu'une variable aléatoire X est absolument continue (ou « **à densité** ») s'il existe une densité de probabilité f telle que, pour tout intervalle non vide I de $\mathbb{R}$, la probabilité que X appartienne à I soit égale à l'intégrale de f sur I : [2]

$$\boxed{P[X \in I] = \int_I f(x)\,\mathrm{d}x}\ .$$

On dit alors que X admet f pour densité ou que f est *une* densité de X. [3]

Propriétés de la fonction de répartition d'une variable à densité

Soit $F : x \longmapsto P[X \leq x]$ la fonction de répartition d'une variable aléatoire X de densité f. Alors F est continue sur $\mathbb{R}$ et, pour tout point de continuité x de f, la fonction F est dérivable en x et admet $f(x)$ pour dérivée en x.

Caractérisation des variables à densité

Soit X une variable aléatoire réelle, de fonction de répartition F.

Soit f la fonction définie par : $\quad f(x) = \begin{cases} F'(x) & \text{si } F \text{ est dérivable en } x \\ 0 & \text{sinon} \end{cases}$.

Alors X est absolument continue si et seulement si f est une densité de probabilité.

Dans ce cas, la fonction f est une densité de X. [4]

---

1. Il convient alors de disposer d'une notion d'intégrale adaptée à ce contexte général, l'intégrale de Lebesgue par exemple.
2. La formule donnant les probabilités liées à X s'étend alors à toutes les parties boréliennes de $\mathbb{R}$

$$\forall B \in \mathcal{B}(\mathbb{R}), \quad P[X \in B] = \int_B f(x)\,\mathrm{d}x\ .$$

3. Les fonctions positives qui coïncident presque sûrement avec f sont aussi des densités de X. Il n'y a donc pas unicité de la densité f de X, mais les intégrales $\int_I f(x)\,\mathrm{d}x$ ne dépendent pas de la « version » choisie pour f.

   De même, l'adhérence de l'ensemble des points où une densité de X est strictement positive ne dépend pas de la densité choisie : c'est le plus petit ensemble fermé auquel la variable X appartient presque sûrement, appelé le *support* de X.
4. Une condition suffisante pour que X soit absolument continue est que sa fonction de répartition soit à la fois continue sur $\mathbb{R}$, et de classe C^1 sur $\mathbb{R}$ privé éventuellemet d'un nombre fini de points.

2. **Lois uniformes**

On dit que X suit la loi uniforme sur le segment $[a,b]$, notée $\mathcal{U}[a,b]$ (où $a < b$) si elle admet pour densité :

$$\boxed{f : x \longmapsto \begin{cases} \dfrac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases}}$$

a) Fonction de répartition

$$F_{\mathcal{U}[a,b]}(x) = \begin{cases} 0 & \text{si } x < a \\ \dfrac{x-a}{b-a} & \text{si } a \leq x < b \\ 1 & \text{si} b \leq x \end{cases}$$

b) Moyenne, variance, asymétrie, aplatissement

- $E(X) = \dfrac{a+b}{2}$
- $V(X) = \dfrac{(b-a)^2}{12}$
- $\gamma_1(X) = 0$
- $\gamma_2(X) = \dfrac{6}{5}$.

c) Fonction génératrice des moments

$$M_{\mathcal{U}[a,b]}(t) = E(\mathrm{e}^{tX}) = \begin{cases} \dfrac{\mathrm{e}^{tb} - \mathrm{e}^{ta}}{t(b-a)} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}$$

d) Simulation Excel de $\mathcal{U}[a,b]$

$$x = a + (b-a) * \text{ALEA}()$$

e) Utilisation

La loi uniforme continue est utilisée pour modéliser un aléa « équiréparti » ou, dans un cadre bayésien, comme distribution a priori non informative.

3. **Lois normales**

On dit que Z suit la loi normale standard, notée $\mathcal{N}(0,1)$, si elle admet pour densité :

$$\boxed{\varphi : x \longmapsto \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}}$$

La fonction de répartition $\Phi : x \longmapsto \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \mathrm{e}^{-t^2/2} \, \mathrm{d}t$ de Z et sa fonction réciproque Φ^{-1} sont disponibles sous Excel :

$$\begin{cases} \Phi(x) = \text{LOI.NORMALE.STANDARD}\,(x) \\ \Phi^{-1}(p) = \text{LOI.NORMALE.STANDARD.INVERSE}\,(p) \end{cases}$$

Exemples

- $\Phi(1) = 1 - \Phi(-1) = \text{LOI.NORMALE.STANDARD}\,(1) \simeq 0,841$
- $\Phi(2) = 1 - \Phi(-2) = \text{LOI.NORMALE.STANDARD}\,(2) \simeq 0,977$
- $\Phi(3) = 1 - \Phi(-3) = \text{LOI.NORMALE.STANDARD}\,(3) \simeq 0,999$

- $\Phi^{-1}(0,95) = \text{LOI.NORMALE.STANDARD.INVERSE}\,(0,95) \simeq 1,645$

d'où : $P[-1,645 \leq Z \leq +1,645] \simeq 90\%$

- $\Phi^{-1}(0,975) = \text{LOI.NORMALE.STANDARD.INVERSE}\,(0,975) \simeq 1,960$

d'où : $P[-1,96 \leq Z \leq +1,96] \simeq 95\%$

- $\Phi^{-1}(0,995) = \text{LOI.NORMALE.STANDARD.INVERSE}\,(0,995) \simeq 2,576$

d'où : $P[-2,576 \leq Z \leq +2,576] \simeq 99\%$

On dit que X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$ ($\sigma > 0$) si $Z = \dfrac{X-\mu}{\sigma}$ suit la loi normale standard. La variable aléatoire X admet alors pour densité

$$f_X : x \longmapsto \frac{1}{\sigma}\,\varphi(\frac{x-\mu}{\sigma}) = \boxed{\frac{1}{\sigma\sqrt{2\pi}}\ \mathrm{e}^{-\dfrac{(x-\mu)^2}{2\sigma^2}}}$$

et pour fonction de répartition

$$F_X : x \longmapsto \Phi(\frac{x-\mu}{\sigma})\ .$$

Sous Excel :

$$\begin{cases} f_X(x) = \text{LOI.NORMALE}\,(x;\mu;\sigma;FAUX) \\ F_X(x) = \text{LOI.NORMALE}\,(x;\mu;\sigma;VRAI) \\ F_X^{-1}(p) = \text{LOI.NORMALE.INVERSE}\,(p;\mu;\sigma) \end{cases}$$

a) Moyenne, variance, asymétrie, aplatissement

- $E(X) = \mu$
- $V(X) = \sigma^2$
- $\gamma_1(X) = 0$
- $\gamma_2(X) = 0$

b) Fonction génératrice des moments

$$\begin{cases} M_{\mathcal{N}(0,1)}(x) = E(\mathrm{e}^{tZ}) = \exp(\dfrac{t^2}{2}) \\ M_{\mathcal{N}(\mu,\sigma^2)}(x) = E(\mathrm{e}^{tX}) = E(\mathrm{e}^{t(\mu+\sigma Z)}) = \exp(\mu t + \dfrac{\sigma^2 t^2}{2}) \end{cases}$$

c) Simulation Excel de $\mathcal{N}(\mu, \sigma^2)$

$$x = \text{LOI.NORMALE.INVERSE}\,(\text{ALEA}();\mu;\sigma)$$

d) Convolution

$$\mathcal{N}(\mu,\sigma^2) * \mathcal{N}(\nu,\tau^2) = \mathcal{N}(\mu+\nu, \sigma^2+\tau^2)$$

e) Utilisation

Chaque fois qu'une quantité aléatoire résulte d'un cumul d'aléas indépendants et de petite taille par rapport à leur valeur cumulée, on peut raisonnablement s'attendre à ce qu'une modélisation par une loi normale soit adaptée.

4. **Lois Gamma**

On dit que X suit la loi $\Gamma(\nu,\lambda)$, ou ν et λ sont deux réels strictement positifs, appelés respectivement *paramètre de forme* et *paramètre d'échelle*, si elle admet pour densité :

$$\boxed{f_X : t \longmapsto \begin{cases} \dfrac{1}{\Gamma(\nu)} \dfrac{t^{\nu-1}}{\lambda^\nu} \mathrm{e}^{-\dfrac{t}{\lambda}} & \text{si } t > 0 \\ 0 & \text{sinon} \end{cases}}$$

où $\Gamma(\nu) = \displaystyle\int_0^{+\infty} t^{\nu-1}\,\mathrm{e}^{-t}\mathrm{d}t.$ [5]

La densité f_X, la fonction de répartition F_X de X et sa fonction réciproque F_X^{-1} sont disponibles sous Excel :

$$\begin{cases} f_X(x) = \text{LOI.GAMMA}\,(x;\nu;\lambda;FAUX) \\ F_X(x) = \text{LOI.GAMMA}\,(x;\nu;\lambda;VRAI) \\ F_X^{-1}(p) = \text{LOI.GAMMA.INVERSE}\,(p;\nu;\lambda) \end{cases}$$

a) Cas particuliers

- $\Gamma(1,\lambda)$ est la loi exponentielle d'espérance λ, qui est la loi du temps d'attente de la première arrivée d'un processus de Poisson homogène d'intensité λ.
- $\Gamma(\frac{n}{2},2)$, appelée loi du khi-deux à n degrés de liberté, est la loi de la somme des carrés de n variables aléatoires indépendantes suivant chacune la loi normale standard.

b) Moyenne, variance, asymétrie, aplatissement

- $E(X) = \lambda\nu$
- $V(X) = \lambda\nu$
- $\gamma_1(X) = \dfrac{2}{\sqrt{\nu}}$
- $\gamma_2(X) = \dfrac{6}{\nu}$

b) Fonction génératrice des moments

$$M_{\Gamma(\nu,\lambda)}(x) = E(\mathrm{e}^{tX}) = \frac{1}{(1-\lambda t)^\nu}$$

c) Simulation Excel de $\mathcal{N}(\mu,\sigma^2)$

$$x = \text{LOI.GAMMA.INVERSE}\,(\text{ALEA}();\nu;\lambda)$$

d) Convolution

$$\Gamma(\mu,\lambda) * \Gamma(\nu,\lambda) = \Gamma(\mu+\nu,\lambda)$$

e) Utilisation

Les lois Γ peuvent être utilisées pour modéliser des variables unimodales à valeurs positives, comme des durées de vie ou des coûts de sinistres. En assurance IARD, elles permettent aussi de décrire la structure de risque d'un portefeuille hétérogène dans un but de tarification à posteriori.

---

5. La fonction gamma d'Euler, $\Gamma : x \longmapsto \displaystyle\int_0^{+\infty} t^{x-1}\,\mathrm{e}^{-t}\mathrm{d}t$ est définie sur $]0,+\infty[$ et vérifie les propriétés suivantes :
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
- Pour tout $x > 0$, $\Gamma(x+1) = x\,\Gamma(x)$
- Pour tout $n \in \mathbb{N}$, $\Gamma(n+1) = n!$
- Pour tout $k \in \mathbb{N}^*$ et tout $x > 0$, $\Gamma^{(k)}(x) = \displaystyle\int_0^{+\infty} \big(\ln(t)\big)^k\, t^{x-1}\,\mathrm{e}^{-t}\mathrm{d}t$.

On peut en calculer les valeurs numériques à l'aide de la fonction Excel « GAMMA ».