

CHAPITRE 13

Compléments sur les variables aléatoires à densité

Il n'y a rien de plus triste qu'une vie sans hasard.

HONORÉ DE BALZAC (1799-1850)

1

Loi du maximum, loi du minimum

Soient X, Y des variables aléatoires *indépendantes* à densité définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. Notons f_X et f_Y des densités respectivement de X et Y . On définit les applications

$$Z: \begin{cases} \Omega & \rightarrow \mathbb{R} \\ \omega & \mapsto \max(X(\omega); Y(\omega)) \end{cases} \quad \text{et} \quad T: \begin{cases} \Omega & \rightarrow \mathbb{R} \\ \omega & \mapsto \min(X(\omega); Y(\omega)) \end{cases}$$

Noter simplement $Z = \max(X, Y)$ et $T = \min(X, Y)$. Les applications Z et T sont des variables aléatoires.

- Pour tout $t \in \mathbb{R}$,

$$\begin{aligned} F_Z(t) &= \mathbf{P}(Z \leq t) \\ &= \mathbf{P}([X \leq t] \cap [Y \leq t]) \\ &= \mathbf{P}([X \leq t]) \cdot \mathbf{P}([Y \leq t]) \quad \text{Par indépendance} \\ F_Z(t) &= F_X(t) \cdot F_Y(t). \end{aligned}$$

Par produit, F_Z est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points (noté D). Z est une variable aléatoire à densité. Une densité est donnée par dérivation, pour $t \in \mathbb{R} \setminus D$

$$f_Z(t) = F_Z'(t) = F_X'(t) \cdot F_Y(t) + F_X(t) \cdot F_Y'(t) = f_X(t)F_Y(t) + f_Y(t)F_X(t).$$

La dernière égalité s'étend à tout réel t .

- Le calcul est similaire pour le minimum en rajoutant le passage au complémentaire. Pour tout $t \in \mathbb{R}$,

$$\begin{aligned} 1 - F_T(t) &= \mathbf{P}(T > t) \\ &= \mathbf{P}([X > t] \cap [Y > t]) \\ &= \mathbf{P}([X > t]) \cdot \mathbf{P}([Y > t]) \quad \text{Par indépendance} \\ 1 - F_T(t) &= (1 - F_X(t)) \times (1 - F_Y(t)). \end{aligned}$$

Par produit, F_T est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points (noté D_T). T est une variable aléatoire à densité. Une densité est donnée par dérivation, pour $t \in \mathbb{R} \setminus D_T$

$$-f_T(t) = (1 - F_T)'(t) = (1 - F_X)'(t) \cdot (1 - F_Y(t)) + (1 - F_X(t)) \cdot (1 - F_Y)'(t) = -f_X(t)(1 - F_Y(t)) - f_Y(t)(1 - F_X(t)).$$

Finalement, on peut poser pour tout réel t

$$f_T(t) = f_X(t)(1 - F_Y(t)) + f_Y(t)(1 - F_X(t)).$$

Le calcul précédent se généralise facilement au cas de n variables aléatoires indépendantes à densité $(X_1, X_2, \dots, X_n)$. Avec

$$Z = \max(X_1, \dots, X_n) \quad \text{et} \quad T = \min(X_1, \dots, X_n),$$

on a
$$F_Z(t) = \prod_{i=1}^n F_{X_i}(t) \quad \text{et} \quad 1 - F_T(t) = \prod_{i=1}^n (1 - F_{X_i}(t)).$$

Les variables Z et T restent des variables à densité. Retenons aussi la simplification lorsque les variables X_i suivent la même loi (ainsi, $F_{X_i} = F_{X_1}$)

on a
$$F_Z(t) = F_{X_1}(t)^n \quad \text{et} \quad 1 - F_T(t) = (1 - F_{X_1}(t))^n.$$

Une densité pour Z et pour T sont données par

$$\forall t \in \mathbb{R}, \quad f_Z(t) = n f_{X_1}(t) F_{X_1}(t)^{n-1} \quad \text{et} \quad f_T(t) = n f_{X_1}(t) (1 - F_{X_1}(t))^{n-1}.$$

Exercice 1



✧ Soient X, Y des variables aléatoires indépendantes suivant respectivement des lois exponentielles de paramètres λ et μ .

1. Donner la loi de $T = \min(X, Y)$.
2. En déduire l'existence et le calcul des espérances de $\min(X, Y)$ et $\max(X, Y)$.
3. Généraliser le résultat de la première question avec n variables.

p. 16

VAD1

Exercice 2



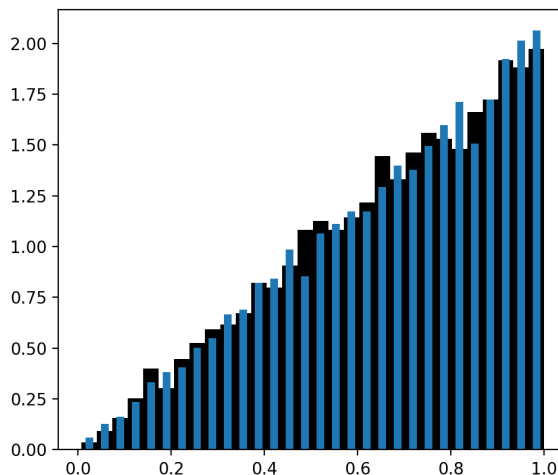
1. Que permet de conjecturer le programme et les résultats suivants?
2. Prouver votre conjecture puis généraliser le résultat obtenu? Comment interpréter graphiquement ce résultat?

p. 17

VAD2

Editeur

```
def simu1():
    return max(rd.random(2))
def simu2():
    return np.sqrt(rd.random())
m=5000
Ech1=np.zeros(m), Ech2=np.zeros(m)
for i in range(m):
    Ech1[i]=simu1()
    Ech2[i]=simu2()
plt.hist(Ech1,30,density=True,color='black')
plt.hist(Ech2,30,density=True,rwidth=0.5)
plt.show()
```



2

Loi d'une somme de variables aléatoires à densité

2.1

Calcul d'un produit de convolution

Soient f, g deux fonctions continues sur $\mathbb{R}$ (sauf éventuellement en un nombre fini de points). Pour tout réel x , on définit, sous-réserve d'existence, $f * g(x)$ par

$$f * g(x) = \int_{-\infty}^{+\infty} f(x-t)g(t) dt.$$

Cela définit alors une nouvelle fonction $x \in \mathbb{R} \mapsto f * g(x)$. On parle alors de produit de convolution.

Exercice 3



◆ Conditions suffisantes d'existence et propriétés

Soient f, g et h , trois fonctions continues sur $\mathbb{R}$ telles que f soit bornée sur $\mathbb{R}$ et $\int_{-\infty}^{+\infty} g(t) dt$, $\int_{-\infty}^{+\infty} h(t) dt$ soient absolument convergentes.

p. 17

1. Montrer que $f * g$ est bien définie sur $\mathbb{R}$.
2. Vérifier que $f * g = g * f$ et pour tout $\lambda \in \mathbb{R}$, $f * (g + \lambda h) = (f * g) + \lambda(f * h)$.

VAD3

Exemples. Nous en donnons deux. Chacun vérifie la condition de la question 1 de l'exercice précédent, le produit de convolution sera ainsi toujours bien défini.

• Exemple 1 avec la gaussienne.

Posons pour $t \in \mathbb{R}$, $f(t) = e^{-t^2/2}$. Soit $x \in \mathbb{R}$, calculons $f * f(x)$. Dans un premier temps, pour $t \in \mathbb{R}$

$$\begin{aligned} f(x-t)f(t) &= e^{-(x-t)^2/2} \cdot e^{-t^2/2} = e^{-((x-t)^2+t^2)/2} = e^{-(t^2-tx+\frac{x^2}{2})} = e^{-((t-x/2)^2+x^2/4)} \\ &= e^{-x^2/4} \cdot e^{-(t-x/2)^2}. \end{aligned}$$

On a donc

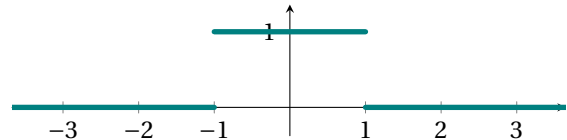
$$f * f(x) = \int_{-\infty}^{+\infty} f(x-t)f(t) dt = e^{-x^2/4} \int_{-\infty}^{+\infty} e^{-(t-x/2)^2} dt.$$

À l'aide du changement de variable affine $u = t - x/2$, il vient

$$f * f(x) = e^{-x^2/4} \int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{\pi} e^{-x^2/4}.$$

• Exemple 2 avec une fonction à support borné. Soit c définie sur $\mathbb{R}$ par

$$c(t) = \begin{cases} 1 & \text{si } t \in [-1; 1] \\ 0 & \text{sinon.} \end{cases}$$



Soit $x \in \mathbb{R}$. Calculons

$$c * c(x) = \int_{-\infty}^{+\infty} c(x-t)c(t) dt.$$

Tout d'abord, simplifions les bornes de l'intégrale.

$$c(x-t)c(t) \neq 0 \iff \begin{cases} -1 \leq t \leq 1 \\ -1 \leq t-x \leq 1 \end{cases} \iff \begin{cases} -1 \leq t \leq 1 \\ 1-x \leq t \leq 1+x \end{cases}.$$

Distinguons plusieurs cas :

- Si $x \leq -2$, alors les conditions précédentes ne peuvent être vérifiées simultanément : $c(x-t)c(t) = 0$ pour tout t et $c * c(x) = 0$.
- Si $x \geq 2$, on a de même, $c * c(x) = 0$.
- Si $x \in [-2; 0]$, on a $x-1 \leq -1$ et $x+1 \leq 1$. Les conditions se simplifient en $-1 \leq t \leq x+1$ et

$$c * c(x) = \int_{-1}^{x+1} c(x-t)c(t) dt = \int_{-1}^{x+1} 1 du = x+2.$$

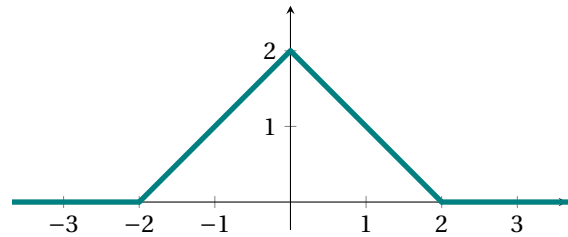
- Si $x \in [0; 2]$, on a maintenant $x-1 \leq t \leq 1$ et

$$c * c(x) = \int_{x-1}^1 c(x-t)c(t) dt = 2-x.$$

Finalement :

$$c * c(x) = \begin{cases} 0 & \text{si } x \notin [-2; 2] \\ 2+x & \text{si } x \in [-2; 0] \\ 2-x & \text{si } x \in [0; 2]. \end{cases}$$

On obtient un graphe en "triangle".



Exercice 4



◆◆◆ On définit les fonctions f , g et h sur $\mathbb{R}$ par

$$f(t) = \frac{1}{1+t^2}, \quad g(t) = \begin{cases} e^{-t} & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad h(t) = \begin{cases} 1 & \text{si } t \in [0; 1] \\ 0 & \text{sinon.} \end{cases}$$

p. 17

Calculer $f * h$, $g * g$, $g * h$.

VAD4

2.2 Le théorème de sommation

THÉORÈME

loi d'une somme

Soient X et Y deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbf{P})$, de densités respectives f_X et f_Y .

Si $\rightarrow$ Les variables X et Y sont indépendantes,

$\rightarrow$ La fonction h définie sur $\mathbb{R}$ par

$$h(x) = \int_{-\infty}^{+\infty} f_X(t) f_Y(x-t) dt$$

est bien définie et continue sauf en un nombre fini de points.

Alors

la variable $X+Y$ est à densité et h est une densité.

Exemples. Reprenons les deux premiers calculs de la section précédente.

- Soient X et Y deux variables aléatoires indépendantes dont la loi est une loi normale centrée réduite. Une densité est donnée par

$$\forall t \in \mathbb{R}, \quad f_X(t) = f_Y(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} = \frac{1}{\sqrt{2\pi}} f(t).$$

Le premier calcul justifie l'existence du produit de convolution $f_X * f_Y$ avec

$$\forall x \in \mathbb{R}, \quad f_X * f_Y(x) = \frac{1}{2\pi} f * f(x) = \frac{1}{2\pi} \sqrt{\pi} e^{-x^2/4} = \frac{1}{\sqrt{2} \cdot \sqrt{2\pi}} e^{-x^2/(2 \cdot \sqrt{2}^2)}.$$

On constate alors que $X+Y$ est encore une variable à densité qui suit encore une loi normale $\mathcal{N}(0; 2)$.

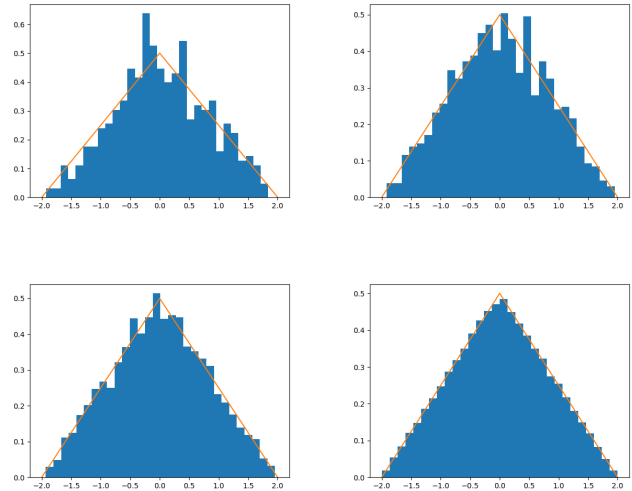
- Le deuxième calcul traduit le fait que la somme de deux lois uniformes sur $[-1; 1]$ est encore une variable aléatoire à densité dont une densité est donnée par

$$h(x) = \begin{cases} 0 & \text{si } x \notin [-2; 2] \\ (2+x)/4 & \text{si } x \in [-2; 0] \\ (2-x)/4 & \text{si } x \in [0; 2]. \end{cases}$$

Vérification expérimentale avec Python

```
m=500
# puis, 1000, 5000 et 10^4
# taille de l'échantillon
x=-1+2*np.random.rand(m)
y=-1+2*np.random.rand(m)
# simulation d'une loi uniforme
# continue sur [-1;1]

plt.hist(x+y,30,density=True)
# tracé de l'histogramme
plt.plot([-2,0,2],[0,0.5,0])
# tracé de la densité
plt.show()
# affichage
```



Remarque. La somme de deux variables aléatoires à densité n'est pas toujours une variable aléatoire à densité. Il suffit de considérer $X + (-X)$ pour s'en convaincre.

Exercice 5



◆◆ Probabilité de collision

Soient X et Y deux variables aléatoires indépendantes à densité bornée.
Que dire de la probabilité $P(X = Y)$?

p. 17

VAD5

[Un lien vers une explication en vidéo.](#)

3

Application aux lois usuelles

3.1 Stabilité par somme des lois γ

On rappelle les définitions des fonctions Γ et β

$$\Gamma: x \in \mathbb{R}_*^+ \mapsto \int_0^{+\infty} t^{x-1} e^{-t} dt \quad \text{et} \quad \beta: (x, y) \in \mathbb{R}^2 \mapsto \int_0^1 t^{x-1} (1-t)^{y-1} dt.$$

Les propriétés de ces fonctions et les liens entre elles sont étudiées à l'exercice ??, p. ??.

THÉORÈME

somme de deux lois γ

Soient $v_1, v_2 \in \mathbb{R}_*^+$ et X_1, X_2 deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbf{P})$.

Si $\left\{ \begin{array}{l} \rightarrow X_1 \text{ et } X_2 \text{ sont indépendantes.} \\ \rightarrow X_1 \hookrightarrow \gamma(v_1) \text{ et } X_2 \hookrightarrow \gamma(v_2). \end{array} \right.$

Alors $X_1 + X_2 \hookrightarrow \gamma(v_1 + v_2).$

Preuve. Notons $\alpha_1 = \Gamma(v_1)^{-1}$ et $\alpha_2 = \Gamma(v_2)^{-1}$ pour alléger les notations. Considérons de plus les densités de probabilité :

$$f(t) = \begin{cases} 0 & \text{si } t < 0 \\ \alpha_1 t^{v_1-1} e^{-t} & \text{si } t \geq 0 \end{cases} \quad \text{et} \quad g(t) = \begin{cases} 0 & \text{si } t < 0 \\ \alpha_2 t^{v_2-1} e^{-t} & \text{si } t \geq 0 \end{cases}$$

Soient $x \in \mathbb{R}$. Justifions l'existence et calculons $f * g(x)$. Soit $t \in \mathbb{R}$.

$$f(x-t)g(t) \neq 0 \iff \begin{cases} x-t \geq 0 \\ t \geq 0 \end{cases} \iff 0 \leq t \leq x.$$

- Si $x < 0$, on a toujours $f(x-t)g(t) = 0$ et par suite

$$f * g(x) = \int_{-\infty}^{+\infty} f(x-t)g(t) dt = 0.$$

- Si $x \geq 0$. On peut réduire l'intervalle d'intégration à $[0, x]$. $f * g(x)$ est alors bien défini en tant qu'intégrale d'une fonction continue sur un segment.

$$\begin{aligned} f * g(x) &= \int_{-\infty}^{+\infty} f(x-t)g(t) dt = \int_0^x f(x-t)g(t) dt \\ &= \int_0^x \alpha_1 (x-t)^{\nu_1-1} e^{-(x-t)} \cdot \alpha_2 t^{\nu_2-1} e^{-t} dt \\ f * g(x) &= \alpha_1 \alpha_2 e^{-x} \int_0^x (x-t)^{\nu_1-1} t^{\nu_2-1} dt. \end{aligned}$$

Effectuons le changement de variable $u = \frac{t}{x}$ de classe $\mathcal{C}^1$, strictement croissant de $[0; x]$ dans $[0; 1]$.

$$\begin{aligned} \int_0^x (x-t)^{\nu_1-1} t^{\nu_2-1} dt &= \int_0^1 (x-xu)^{\nu_1-1} (xu)^{\nu_2-1} \cdot (x du) \\ &= x^{\nu_1+\nu_2-1} \int_0^1 (1-u)^{\nu_1-1} u^{\nu_2-1} du = \beta(\nu_1, \nu_2) x^{\nu_1+\nu_2-1}. \end{aligned}$$

On obtient pour $x \in \mathbb{R}_*^+$,

$$f * g(x) = \underbrace{\alpha_1 \alpha_2 \beta(\nu_1, \nu_2)}_{\text{indépendant de } x} \cdot x^{\nu_1+\nu_2-1} e^{-x}$$

À une constante près, $f * g$ est une densité d'une loi $\gamma(\nu_1 + \nu_2)$.

$$f * g(x) = \alpha_1 \alpha_2 \beta(\nu_1, \nu_2) \times \Gamma(\nu_1 + \nu_2) \cdot \frac{1}{\Gamma(\nu_1 + \nu_2)} x^{\nu_1+\nu_2-1} e^{-x}.$$

Or, d'après le théorème sur les sommes, $f * g$ est une densité de la variable $X + Y$. Donc

$$1 = \int_{-\infty}^{+\infty} f * g(t) dt = \alpha_1 \alpha_2 \beta(\nu_1, \nu_2) \Gamma(\nu_1 + \nu_2)$$

On en déduit que

$$\alpha_1 \alpha_2 \beta(\nu_1, \nu_2) \cdot \Gamma(\nu_1 + \nu_2) = 1$$

et $X_1 + X_2$ suit une loi $\gamma(\nu_1 + \nu_2)$.

Remarque. Ce calcul a permis de retrouver l'égalité

$$\forall \nu_1, \nu_2 \in \mathbb{R}_*^+, \quad \beta(\nu_1, \nu_2) = \int_0^1 (1-u)^{\nu_1-1} u^{\nu_2-1} du = \frac{\Gamma(\nu_1) \Gamma(\nu_2)}{\Gamma(\nu_1 + \nu_2)}.$$

COROLLAIRE

cas particulier des lois exponentielles

Soient $n \in \mathbb{N}$ et $X_1, \dots, X_n$, n variables aléatoires définies sur un même espace probabilisé.

- Si** $\left\{ \begin{array}{l} \rightarrow \text{Les variables } X_1, \dots, X_n \text{ sont mutuellement indépendantes.} \\ \rightarrow \text{Pour tout } i \in [1; n], X_i \hookrightarrow \mathcal{E}(1). \end{array} \right.$

Alors $X_1 + X_2 + \dots + X_n \hookrightarrow \gamma(n).$

Preuve. Il suffit de procéder par récurrence en rappelant que la loi exponentielle de paramètre 1 est aussi une loi $\gamma(1)$.

Exercice 6



◆ Cas général - loi d'Erlang

Soient $X_1, \dots, X_n$ des variables mutuellement indépendantes suivant une loi $\mathcal{E}(\lambda)$. Vérifier que la somme $X_1 + X_2 + \dots + X_n$ est à densité et préciser une densité.

p. 17

3.2 Stabilité par somme des lois normales

PROPOSITION

somme de deux lois normales

Soient $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_+^*$ et $X_1, X_2 : \Omega \rightarrow \mathbb{R}$ deux variables aléatoires.

Si $\left\{ \begin{array}{l} \rightarrow X_1 \text{ et } X_2 \text{ sont indépendantes.} \\ \rightarrow X_1 \hookrightarrow \mathcal{N}(m_1, \sigma_1^2) \text{ et } X_2 \hookrightarrow \mathcal{N}(m_2, \sigma_2^2). \end{array} \right.$

Alors $X_1 + X_2 \hookrightarrow \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2).$

Exercice 7



◆◆ Preuve

1. On suppose dans cette question que $m_1 = m_2 = 0$. Prouver l'énoncé en adaptant le premier exemple de la section et en admettant l'égalité suivante :

$$-\frac{(x-t)^2}{2\sigma_1^2} - \frac{t^2}{2\sigma_2^2} = -\frac{\sigma^2}{2\sigma_1^2\sigma_2^2} \left(t - \frac{\sigma_2^2}{\sigma^2} x \right)^2 - \frac{x^2}{2\sigma^2} \quad \text{où } \sigma = \sqrt{\sigma_1^2 + \sigma_2^2}. \quad \text{p. 18}$$

2. En déduire le cas général avec m_1 et m_2 non nécessairement nuls.

VAD7

Exercice 8



On considère X, Y deux variables aléatoires indépendantes de loi respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$. Soient $S = X + Y$ et $T = X - Y$.

1. Déterminer la loi de S puis celle de T .
2. On suppose que S et T sont indépendantes. Donner la loi de $S + T$. En déduire $\sigma_1 = \sigma_2$.

p. 18

VAD8

On en déduit par récurrence :

COROLLAIRE

somme de n lois normales

Soient $X_1, \dots, X_n$, n variables aléatoires indépendantes telles que pour tout $i \in \llbracket 1; n \rrbracket$, $X_i \hookrightarrow \mathcal{N}(m_i, \sigma_i^2)$. Alors

$$X_1 + \dots + X_n \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n m_i, \sum_{i=1}^n \sigma_i^2\right).$$

Exemple. En particulier, retenons que si $(X_i)_{i \in \llbracket 1; n \rrbracket}$ est un vecteur aléatoire composé de variables mutuellement indépendantes et de loi identique $\mathcal{N}(\mu, \sigma^2)$, alors

$$\overline{X}_n^* = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \hookrightarrow \mathcal{N}(0; 1).$$

Remarque. Nous avons vu que les lois normales sont stables par transformations affines, ainsi toute combinaison linéaire de variables aléatoires suivant des lois normales restent une loi normale.

! Attention. La condition d'indépendance est nécessaire. Par exemple, si $X \hookrightarrow \mathcal{N}(0, 1)$ et Y une loi uniforme sur $\{-1; 1\}$ alors $XY \hookrightarrow \mathcal{N}(0, 1)$, X et Y ne sont pas indépendantes. Par contre avec vérifie que $X + XY = (1 + Y)X$ ne suit même pas une loi normale. La variable n'est même pas à densité car $\mathbf{P}(X + XY = 0) = 1/2$.

4

Un exemple détaillé

Soient X, Y deux variables aléatoires indépendantes suivant une loi uniforme sur $]0; 1[$. On pose

$$Z = \frac{\min(X, Y)}{\max(X, Y)}.$$

Déterminons la loi de Z.

Simulation Python et conjecture

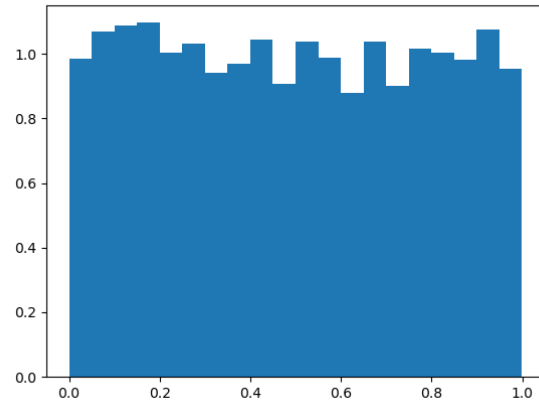
Editeur

```
import random as rd
import matplotlib.pyplot as plt

def simulation():
    X=rd.random()
    Y=rd.random()
    Z=min(X,Y)/max(X,Y)
    return Z

Ech=[]
for i in range(5000):
    Z=simulation()
    Ech.append(Z)

plt.clf()
plt.hist(Ech, bins=20, density=True)
plt.show()
```



De manière assez surprenante, la loi semble uniforme sur l'intervalle $[0; 1]$. Une preuve s'impose!

Exercice 9



◆ Une première approche?

Expliquer pourquoi le schéma de preuve suivant ne convient pas.

- Étape 1. Trouver, en reprenant la partie 1, la loi de $U = \max(X, Y)$ et $V = \min(X, Y)$
- Étape 2. En déduire, en passant par la fonction de répartition, la loi de $\ln(U)$ et $-\ln(V)$.
- Étape 3. Appliquer le théorème de sommation pour obtenir la loi de

$$W = \ln(U/V) = \ln(U) + (-\ln(V)).$$

- Étape 4. Conclure, de nouveau avec la fonction de répartition, en obtenant la loi

$$Z = \exp(W).$$

p. 18

VAD9

Calcul de la loi

- Appliquons le logarithme afin d'avoir des sommes.

$$\begin{aligned} \ln(Z) &= \ln(\min\{X, Y\}) - \ln(\max\{X, Y\}) \\ &= \min\{\ln(X), \ln(Y)\} - \max\{\ln(X), \ln(Y)\} \quad \text{car } \ln \text{ est croissante} \\ &= -|\ln(X) - \ln(Y)|. \end{aligned}$$

- Donnons une densité de $\ln(Y)$ à valeurs dans $\mathbb{R}^-$. Soit $t \in \mathbb{R}^-$

$$\begin{aligned} F_{\ln(Y)}(t) &= \mathbf{P}(\ln(Y) \leq t) \\ &= \mathbf{P}(Y \leq e^t) \quad \text{car l'exponentielle est croissante} \\ &= F_Y(e^t) = e^t \quad \text{car } e^t \in]0; 1] \text{ et } F_Y(x) = x \text{ si } x \in [0; 1]. \end{aligned}$$

et $F_{\ln(Y)}(t) = 0$ si $t \in \mathbb{R}_*^+$. Par dérivation, une densité est

$$g(t) = \begin{cases} e^t & \text{si } t \leq 0 \\ 0 & \text{sinon.} \end{cases}$$

En adaptant le résultat précédent, on constate que X est bien à densité avec une densité f donnée par

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

Autrement dit

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

- Soient $x, t \in \mathbb{R}$.

$$f(x-t)g(t) \neq 0 \iff \begin{cases} x-t \geq 0 \\ t \leq 0 \end{cases} \iff \begin{cases} t \leq x \\ t \leq 0. \end{cases}$$

Distinguons deux cas.

→ Si $x \in \mathbb{R}^+$, la condition devient $t \leq 0$ et le produit de convolution va être convergent

$$\int_{-\infty}^{+\infty} f(x-t)g(t)dt = \int_{-\infty}^0 f(x-t)g(t)dt = \int_{-\infty}^0 e^{-(x-t)}e^t dt = e^{-x} \int_{-\infty}^0 e^{2t} dt = \frac{1}{2}e^{-x}.$$

→ Pour $x \in \mathbb{R}^-$, on a maintenant

$$\begin{aligned} h(x) &= \int_{-\infty}^{+\infty} f(x-t)g(t)dt = \int_{-\infty}^x f(x-t)g(t)dt = \int_{-\infty}^x e^{-(x-t)}e^t dt \\ &= e^{-x} \int_{-\infty}^x e^{2t} dt = e^{-x} \left[\frac{1}{2}e^{2t} \right]_{-\infty}^x = \frac{1}{2}e^x. \end{aligned}$$

En résumé, on a convergence pour tout $x \in \mathbb{R}$ et

$$f * g(x) = \frac{1}{2}e^{-|x|}$$

De plus, par le lemme de coalitions, les variables $\ln(X)$ et $-\ln(Y)$ sont indépendantes. Ensuite par le théorème de sommation $\ln(X) - \ln(Y)$ est une variable à densité et une densité est donnée par $h = f * g$.

Vocabulaire. On retrouve une loi de Laplace (voir exercice ??, p.??).

- Justifions maintenant que z est à densité et déterminons une densité. Soit H , la fonction de répartition de $\ln(X) - \ln(Y)$. Soit $t \in [0; 1]$

$$\begin{aligned} F_Z(t) &= \mathbf{P}(Z \leq t) \\ &= \mathbf{P}(-\ln(Z) \geq -\ln(t)) \\ &= \mathbf{P}(|\ln(X) - \ln(Y)| \geq -\ln(t)) \\ &= 1 - \mathbf{P}(|\ln(X) - \ln(Y)| \leq -\ln(t)) \\ &= 1 - \mathbf{P}(-\ln(t) > \ln X - \ln Y > \ln t) \\ &= 1 - \mathbf{P}(-\ln(t) \geq \ln X - \ln Y > \ln t) \\ F_Z(t) &= 1 - (H(-\ln(t)) - H(\ln(t))). \end{aligned}$$

Puis par dérivation, une densité est donnée sur $[0; 1]$ par

$$f_Z(t) = \frac{1}{t}H'(-\ln(t)) + \frac{1}{t}H'(\ln(t)) = \frac{1}{t}h(-\ln(t)) + \frac{1}{t}h(\ln(t)).$$

Après simplifications, on trouve simplement $f_Z(t) = 1$. Au final

$$f_Z(t) = \begin{cases} 0 & \text{si } t \notin [0; 1] \\ 1 & \text{sinon.} \end{cases}$$

C'est l'expression d'une densité d'une loi uniforme sur $[0, 1]$:

$$Z \hookrightarrow \mathcal{U}([0; 1]).$$

Exercice 10



Les questions sont indépendantes.

1. Soient X, Y deux variables aléatoires indépendantes de loi $\mathcal{U}([0; 1])$. Donner la loi de

$$T = -\ln(\max\{X, Y\}).$$

p. 18

2. Reprendre le problème avec trois variables aléatoires indépendantes X, Y, T suivant des lois uniformes sur $[0; 1]$ et

$$Z = \frac{\min(X, Y, T)}{\max(X, Y, T)}.$$

VAD10



Exercices



Compléments sur les variables à densité

Exercice 11. ♦ Anticiper la réponse de la machine à la suite des commandes suivantes :

VAD11

» Solution p. 19

Editeur

```
import numpy.random as rd
C=0
m=5000
for i in range(m):
    if rd.normal(0,1)<rd.normal(0,1):
        C+=1
print(C/m)
```

Exercice 12. ♦

VAD12

Soient $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi normale centrée réduite et $\alpha \in \mathbb{R} \setminus \{\pm 1; 0\}$. On définit

$$Y_0 = X_0 \quad \text{et} \quad \forall n \in \mathbb{N}, \quad Y_{n+1} = \alpha Y_n + X_{n+1}.$$

1. Calculer pour tout $n \in \mathbb{N}$, $E(Y_n)$ et $V(Y_n)$.
2. Justifier que pour tout $n \in \mathbb{N}^*$, la variable Y_n suit une loi normale dont on précisera les paramètres.
3. Calculer pour tout $p \in \mathbb{N}^*$, $\text{Cov}(Y_n, Y_{n+p})$.

» Solution p. 19

Exercice 13. ♦♦

VAD13

Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, \mathbf{P})$ telles que $X \hookrightarrow \mathcal{N}(a, \sigma^2)$ et $Y \hookrightarrow \mathcal{N}(b, \sigma^2)$. Prouver l'équivalence

$$\mathbf{P}(X \leq Y) \geq \frac{1}{2} \iff a \leq b.$$

» Solution p. 19

Exercice 14. ♦♦

Soient f, g deux fonctions positives non nulles bornées, continues et telles que les intégrales $\int_{-\infty}^{+\infty} f(t) dt$ et $\int_{-\infty}^{+\infty} g(t) dt$ soient convergentes. Montrer que

$$\int_{-\infty}^{+\infty} (f * g)(t) dt = \left(\int_{-\infty}^{+\infty} f(t) dt \right) \left(\int_{-\infty}^{+\infty} g(t) dt \right).$$

» Solution p. 20

Exercice 15. ♦♦

D'après oraux HEC 2013 # VAD16

Soient X et Y deux variables aléatoires indépendantes, définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et de même loi $\mathcal{N}(0, 1)$. On note α , la probabilité que la matrice ci-dessous soit diagonalisable.

$$M = \begin{bmatrix} 0 & X & 0 \\ Y & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

1. Calculer $\mathbf{P}(X = Y)$ et $\mathbf{P}(XY > 0)$.
2. Calculer α .

>> Solution p. 20

Exercice 16. ♦♦ Soient X, Y , et Z trois variables aléatoires, indépendantes de même loi. On pose

VAD17

$$\forall a, b, c \in \mathbb{R}, \quad M(a, b, c) = \begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix}.$$

1. a) Justifier que si $a + b + c \neq 0$ alors $M(a, b, c)$ est diagonalisable.
b) Comparer les événements $[X + Y + Z \neq 0]$ et "La matrice $M(X, Y, Z)$ est diagonalisable". En déduire la probabilité que la matrice aléatoire $M(X, Y, Z)$ soit diagonalisable lorsque les trois variables suivent des lois normales.
2. Reprendre la question en supposant que $X + 1$, Y et Z suivent des lois de Poisson.

>> Solution p. 21

Loi d'une somme, d'un maximum, d'un produit, etc

Exercice 17. ♦ Soient n un entier ≥ 2 et $X_1, \dots, X_n$ des variables indépendantes de densité f :

VAD18

$$\begin{cases} f(x) = \frac{1}{x^2} & \text{si } x \in]1, +\infty[\\ f(x) = 0 & \text{sinon} \end{cases}$$

1. Étudier l'existence de $E(X_i)$ et $V(X_i)$. Les calculer si possible.
2. Soit $Y = \min_{1 \leq i \leq n} (X_i)$. Déterminer la loi de Y et étudier l'existence de $E(Y)$.
3. De même avec $Z = \max_{1 \leq i \leq n} (X_i)$.

>> Solution p. 21

Exercice 18. ♦ **Maximum et loi de Gumbel**

VAD19

On dit que la variable X suit la loi de Gumbel si sa fonction de répartition F_X est définie sur $\mathbb{R}$ par $F_X(x) = e^{-e^{-x}}$.

1. Soit $n \in \mathbb{N}^*$. On considère n variables aléatoires mutuellement indépendantes $X_1, \dots, X_n$ suivant chacune la loi de Gumbel. Montrer que la variable aléatoire Z définie par $Z = \max(X_1, \dots, X_n) - \ln n$ suit également la loi de Gumbel.
2. On considère maintenant une suite de variables aléatoires indépendantes $(Y_n)_{n \in \mathbb{N}^*}$ suivant chacune la loi exponentielle $\mathcal{E}(1)$. Posons $S_n = \max(Y_1, \dots, Y_n) - \ln n$. Montrer que l'on a, pour tout $x \geq 0$, la convergence suivante :

$$F_{S_n}(x) \xrightarrow{n \rightarrow \infty} F_X(x) \quad \text{où } X \text{ suit une loi de Gumbel.}$$

>> Solution p. 21

Exercice 19. ♦ **Loi d'un produit d'une variable discrète et d'une continue**

VAD20

Soient X et Y deux variables aléatoires indépendantes définies sur un même espace probabilisé. On suppose que X est une variable discrète à valeurs dans $\{-1; 1\}$ et que Y suit la loi normale centrée réduite $\mathcal{N}(0; 1)$. Donner la loi de $Z = XY$?

>> Solution p. 22

Exercice 20. ♦ **Exemples avec les lois uniformes**

D'après EDHEC # VAD21

On considère deux variables aléatoires X et Y , définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes et suivant la loi uniforme sur $[0, 1]$. On pose : $Z = X + Y$.

1. a) Déterminer une densité de Z .
b) Montrer que, pour tout $x \in]0, 1[$, les événements $(Z > 1)$ et $(1 - x < Z \leq 1 + x)$ sont indépendants.
2. On pose : $T = \max(X, Y)$. On admet que T est une variable aléatoire définie sur $(\Omega, \mathcal{A}, \mathbf{P})$.
a) Montrer que T est une variable aléatoire à densité, puis donner une densité de T .
b) En déduire que T possède une espérance et la déterminer.
3. On pose : $U = |X - Y|$ et on admet que U est une variable aléatoire définie sur $(\Omega, \mathcal{A}, \mathbf{P})$. Montrer que U est combinaison linéaire de Z et T , puis en déduire l'espérance de U .

>> Solution p. 22

Exercice 21. ♦ **Loi d'une différence de lois exponentielles**

D'après Ecricome # VAD22

Soient a et b deux réels strictement positifs, X et Y deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes et suivant chacune une loi exponentielle de paramètres respectifs a et b .

1. Déterminer la fonction de répartition, puis donner une densité de la variable $T = -X$.
2. Montrer que $Y - X$ admet une densité, notée h , définie par :

$$h(t) = \begin{cases} \frac{ab}{a+b} e^{-bt}, & \text{pour } t > 0 \\ \frac{ab}{a+b} e^{at}, & \text{pour } t \leq 0. \end{cases}$$

3. On considère la variable $Z = |Y - X|$.

- a) Montrer que pour tout $s \in [0, +\infty[$, $\mathbf{P}(Z \leq s) = 1 - \frac{be^{-as} + ae^{-bs}}{a+b}$.
- b) Vérifier que Z est une variable aléatoire à densité, puis donner une densité de Z .
- c) Montrer que Z admet une espérance et la calculer.

>> Solution p. 23

Exercice 22. ♦♦ Lois de Pareto et produit

VAD23

Pour $a \in \mathbb{R}_*^+$, on définit la fonction f_a sur $\mathbb{R}$ par : $f_a(t) = \begin{cases} \frac{a}{t^{a+1}} & \text{si } t \in [1; +\infty[\\ 0 & \text{sinon.} \end{cases}$

1. Vérifier que f_a est une densité de probabilité.
On dit qu'une variable X suit une loi de Pareto de paramètre a si X admet f_a comme densité.
2. Soient $X_1, \dots, X_n$, n variables aléatoires indépendantes qui suivent toutes une loi de Pareto de paramètre a . Justifier que le produit $X_1 \dots X_n$ est une variable à densité, et en donner une densité.

>> Solution p. 24

Exercice 23. ♦♦ Loi de la partie fractionnaire

VAD24

Soit X une variable aléatoire de loi exponentielle $\mathcal{E}(\lambda)$.

1. Donner la loi de $[X]$. Préciser son espérance et sa variance.
2. Démontrer que la variable $Y = X - [X]$ est à densité et donner une densité.

>> Solution p. 24

Exercice 24. ♦♦ Soient X, Y deux variables aléatoires indépendantes de la loi $\mathcal{E}(1)$.

VAD25

1. Soit $\alpha \in \mathbb{R}_*^+$. Vérifier que $X - \alpha Y$ est une variable à densité dont une densité est donnée par

$$f(x) = \begin{cases} \frac{e^{x/\alpha}}{\alpha+1} & \text{si } x < 0 \\ \frac{e^{-x}}{\alpha+1} & \text{si } x \geq 0 \end{cases}.$$

2. Soit $Z = \frac{X}{Y}$. Vérifier que Z est une variable aléatoire à densité et en déterminer une densité.
3. Déterminer la loi de $\frac{X}{X+Y}$.

>> Solution p. 25

Exercice 25. ♦♦ Somme de lois uniformes

VAD26

Soient $Z, X_1, \dots, X_n$, $n+1$ variables aléatoires mutuellement indépendantes de loi uniforme sur $[0; 1]$.

1. Justifier que $S_n = \sum_{i=1}^n X_i$ est une variable à densité et qu'il existe une densité f_n dont la restriction à $[0; 1]$ est donnée par :

$$\forall x \in [0; 1], \quad f_n(x) = \frac{x^{n-1}}{(n-1)!}.$$

2. Calculer $\mathbf{P}(Z > X_1 + X_2 + \dots + X_n)$.

>> Solution p. 25

Exercice 26. ♦♦ Un exemple sans l'hypothèse d'indépendance

VAD27

Soit X une variable aléatoire de loi uniforme sur $[0; 1]$.

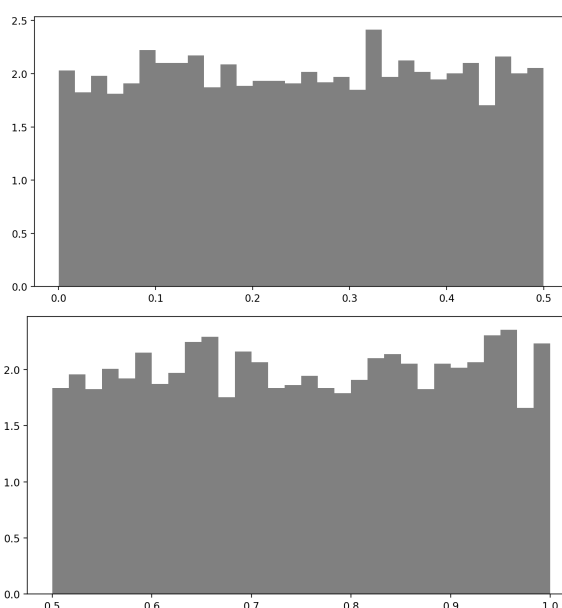
1. Déterminer la loi et l'espérance de $Y = \min(X, 1-X)$, ainsi que de $Z = \max(X, 1-X)$.
2. Vérifier que vos résultats sont en accord avec les simulations numériques suivantes :

Editeur

```
ech=np.zeros(5000)
for i in range(5000):
    x=np.random.rand()
    ech[i]=min(x,1-x)
plt.hist(ech,30,density=True)
plt.show()
```

Editeur

```
ech=np.zeros(5000)
for i in range(5000):
    x=np.random.rand()
    ech[i]=max(x,1-x)
plt.hist(ech,30,density=True)
plt.show()
```



3. Calculer, si elles existent, les espérances de X/Z et de X/Y .

Exercice 27. ♦♦♦ Minimum sur un nombre aléatoire de variables

VAD28

On considère une suite $(X_k)_{k \in \mathbb{N}^*}$ de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes, de même loi exponentielle de paramètre 1. On pose pour tout $n \in \mathbb{N}^*$, $U_n = \min(X_1, \dots, X_n)$.

- Donner la loi de U_n .
- Soit N une variable aléatoire qui suit une loi géométrique de paramètre $p \in]0, 1[$. On suppose que N est indépendant des X_k , pour tout $k \in \mathbb{N}^*$, et on pose :

$$U_N = \min(X_1, \dots, X_N)$$

i.e. $U_N(\omega) = \min(X_1(\omega), \dots, X_{N(\omega)}(\omega))$, pour tout $\omega \in \Omega$. On admet que U_N est une variable aléatoire.

Déterminer la loi de U_N .

- Soient $q \in]0, 1[$ et $I(q) = \int_0^1 \frac{\ln(t)}{(1-qt)^2} dt$.

- Déterminer deux réels a et b tels que : $\forall t \in]0, 1/q[, \frac{1}{t(1-qt)} = \frac{a}{t} + \frac{b}{1-qt}$.
- En déduire l'existence et la valeur de $I(q)$.
- Conclure en montrant que U_N a une espérance que l'on calculera.
Indication : on pourra utiliser le changement de variable $\varphi(t) = e^{-t}$.

» Solution p. 26

Exercice 28. ♦♦♦ Soit une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires indépendantes toutes de loi uniforme continue sur $[0; 1]$. # VAD29

On définit, pour tout $n \in \mathbb{N}^*$, les variables aléatoires S_n et T_n :

$$S_n = \sum_{j=1}^n X_j \quad \text{et} \quad T_n = S_n - \lfloor S_n \rfloor.$$

- Vérifier que S_2 est à densité et donner une densité.
- Montrer que T_2 suit la loi uniforme sur $[0; 1]$.
- On note h la fonction définie sur $\mathbb{R}$ par $h(x) = x - \lfloor x \rfloor$.
 - Soit $n \in \mathbb{N}^*$. Vérifier que $h(h(S_n) + X_{n+1}) = T_{n+1}$.
 - En déduire que T_n suit la loi uniforme sur $[0; 1]$.

» Solution p. 27

Problème 29. ♦♦♦ Lois de Cauchy

- Préliminaire sur la loi de Cauchy

Soit $a \in \mathbb{R}_+^*$. On définit sur $\mathbb{R}$ la fonction f_a par : $f_a(x) = \frac{a}{\pi(a^2 + x^2)}$.

- a) Montrer que f_a est une densité de probabilité.

Dans la suite, X est une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbf{P})$ admettant f_a pour densité. On dit alors que X suit une loi de Cauchy de paramètre a et on écrit $X \hookrightarrow \mathcal{C}(a)$.

- b) Donner la fonction de répartition de X .

- c) Soit $\lambda \in \mathbb{R}_*^+$. Reconnaître la loi de λX lorsque $X \hookrightarrow \mathcal{C}(a)$.

2. Quotient de lois normales centrées

Soient N et N' deux variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes, à valeurs dans $\mathbb{R}^*$, suivant une loi normale centrée réduite $\mathcal{N}(0; 1)$.

- a) Montrer que la variable aléatoire $Z = \ln |N|$ est une variable aléatoire à densité, et en déterminer une densité. Quelle est une densité de la variable aléatoire $-Z$?

- b) Montrer qu'une densité h de la variable aléatoire $\ln |N/N'|$ est donnée par : $\forall x \in \mathbb{R}, \quad h(x) = \frac{2}{\pi} \frac{e^x}{e^{2x} + 1}$.

- c) i) En déduire que la fonction de répartition de $U = |N/N'|$ est donnée par $F_U(x) = 2 \arctan(x) / \pi \cdot \mathbf{1}_{\mathbb{R}^+}(x)$.

- ii) Vérifier que pour tout $x \in \mathbb{R}, \mathbf{P}(N/N' = x) = 0$. En remarquant que N et $-N$ ont même loi, déduire que $G(x) + G(-x) = 1$ où G désigne la fonction de répartition de N/N' .

- iii) Conclure en montrant que $N/N' \hookrightarrow \mathcal{C}(1)$.

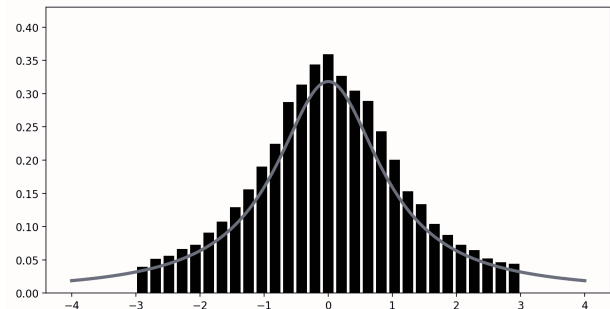
3. Proposer une fonction python qui simule une loi de Cauchy de paramètre 1, puis de paramètre $\lambda \in \mathbb{R}_*^+$.

4. À l'aide du code suivant, que peut-on conjecturer sur la loi de $1/X$ si $X \hookrightarrow \mathcal{C}(1)$? Prouver la conjecture.

>> Solution p. 28

Editeur

```
plt.clf()
X=np.zeros(50000)
for iter in range(50000):
    X[iter]=1/Cauchy()
plt.hist(X,np.linspace(-3,3,30),density=
    True,rwidth=0.8,color='k')
x=np.linspace(-4,4,200)
plt.plot(x,1/(np.pi*(1+x**2)),linewidth=3)
plt.show()
```



Compléments

Problème 30. ♦♦ Maximum de loi normale

D'après EDHEC 2006 #VADp2

On considère deux variables aléatoires X et Y , définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes et suivant toutes deux la loi normale centrée réduite (de densité notée φ et de fonction de répartition notée Φ).

On pose $Z = \max(X, Y)$ et l'on se propose de déterminer la loi de Z , ainsi que son espérance et sa variance.

1. a) Montrer que Z est une variable aléatoire à densité définie elle aussi sur $(\Omega, \mathcal{A}, \mathbf{P})$.

- b) Vérifier que Z admet pour densité la fonction f définie pour tout réel x par :

$$f(x) = 2\varphi(x)\Phi(x).$$

2. a) Rappeler la valeur de l'intégrale $\int_{-\infty}^{+\infty} e^{-t^2/2} dt$.

- b) En déduire la convergence et la valeur de $\int_{-\infty}^{+\infty} e^{-t^2} dt$.

- c) En remarquant que, pour tout réel $x, \varphi'(x) = -x\varphi(x)$, montrer, grâce à une intégration par parties, que :

$$\int_0^{+\infty} xf(x) dx = \frac{1}{\sqrt{2\pi}} + \frac{1}{\pi} \int_0^{+\infty} e^{-t^2} dt.$$

- d) Montrer de même que : $\int_{-\infty}^0 xf(x) dx = -\frac{1}{\sqrt{2\pi}} + \frac{1}{\pi} \int_{-\infty}^0 e^{-t^2} dt$. En déduire que Z a une espérance et donner sa valeur.

3. a) Montrer que X^2 et Z^2 suivent la même loi.

- b) Déterminer $\mathbf{E}(Z^2)$, puis donner la valeur de la variance de Z .

>> Solution p. 29

Table des matières

13 Compléments sur les variables aléatoires à densité	1
1 Loi du maximum, loi du minimum	1
2 Loi d'une somme de variables aléatoires à densité	2
2.1 Calcul d'un produit de convolution	2
2.2 Le théorème de sommation	4
3 Application aux lois usuelles	5
3.1 Stabilité par somme des lois γ	5
3.2 Stabilité par somme des lois normales	7
4 Un exemple détaillé	7