263

VARIABLES ALÉATOIRES À DENSITÉ. EXEMPLES ET APPLICATIONS.

Cadre : Toutes les variables aléatoires sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^n$, $n \geq 1$.

I. Variables aléatoires à densité

1. Définitions et premières propriétés [OUV 1]

Déf 1. S'il existe une fonction f_X définie sur $\mathbb{R}^n$, à valeurs ≥ 0, Riemann-intégrable, telle que $\int_{\mathbb{R}^n} f_X(x)\,dx = 1$, et que $\forall A$ pavé de $\mathbb{R}^n$ on ait :
$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \int_A f_X(x)\,dx$, p. 184
Alors f_X est appelée densité de la v.a X.

Rq : f_X détermine entièrement la loi de X.

Déf 2. Si $x \mapsto |x| f_X(x)$ est Riemann-intégrable sur $\mathbb{R}^n$, X ~~admet une espérance~~, notée $\mathbb{E}[X]$ et définie par :

$$\mathbb{E}[X] = \int_{\mathbb{R}^n} x f_X(x)\,dx$$ p. 190, 191

- Si $\mathbb{E}[X^2] < +\infty$, on définit la variance de X, notée σ_X^2 par :

$$\sigma_X^2 = \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Déf 3. Si $n = 1$, l'application $F : \mathbb{R} \to [0,1]$ définie par :
$\forall x \in \mathbb{R} \quad F_X(x) = \mathbb{P}(X \leq x)$ p. 184
est appelée fonction de répartition de la v.a X.

Prop 4. Ici $n = 1$. Soit F_X la fonction de répartition de X.
(i) F_X détermine la loi de X, c'est une fonction croissante.
(ii) $\forall a, b \in \mathbb{R}$ tq $a < b$, on a $F_X(b) - F_X(a) = \mathbb{P}(a < X \leq b)$
et $\forall x \in \mathbb{R}$, $\lim_{x \to +\infty} F_X(x) = 1$ et $\lim_{x \to -\infty} F_X(x) = 0$. p. 185
(iii) Si X est à densité f_X alors $\forall x \in \mathbb{R}$ $\mathbb{P}(X = x) = 0$,
F_X est continue sur $\mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f_X(u)\,du$.
et F_X est dérivable en tout point de continuité de f_X.

2. Lois usuelles [OUV 2] p. 28

- **Loi uniforme sur $[a,b]$**, notée $\mathcal{U}[a,b]$. (où $a < b$)
  - densité : $f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$
  - espérance : $\frac{a+b}{2}$ - variance : $\frac{(b-a)^2}{12}$

App 5. (Simulation de lois) Soit X une v.a de fonction de répartition F_X. On définit $\forall t \in \mathbb{R}$, $G(t) = \inf(x \mid F(x) \geq t)$, la fonction pseudo-inverse de F_X. p. 23
Soit $Y \sim \mathcal{U}[0,1]$ alors $G(Y)$ admet F_X comme fonction de répartition.

- **Loi normale**, notée $\mathcal{N}(m, \sigma^2)$ (où $m \in \mathbb{R}$, $\sigma^2 > 0$)
  - densité : $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right)$
  - espérance : m - variance : σ^2

Ex 6. Cela permet par exemple de modéliser la répartition des notes d'une classe autour de la moyenne m.

- **Loi exponentielle**, notée $\mathcal{E}(\lambda)$ (où $\lambda \in \mathbb{R}_+^*$)
  - densité : $\mathbb{1}_{\mathbb{R}^+}(x) \lambda e^{-\lambda x}$
  - espérance : $\frac{1}{\lambda}$ - variance : $\frac{1}{\lambda^2}$

Ex 7. Cela permet de modéliser une durée de vie grâce à la propriété : $\mathbb{P}(X > t+s \mid X > s) = \mathbb{P}(X > t)$ qui caractérise la loi exponentielle.

- **Loi Gamma**, notée $\Gamma(n, \lambda)$ (où $n > 0$, $\lambda > 0$)
  - densité : $\mathbb{1}_{\mathbb{R}^+}(x) \frac{\lambda^n}{\Gamma(n)} e^{-\lambda x} x^{n-1}$
- **Loi du chi-deux**, notée $\chi^2(n)$ (où $n \in \mathbb{N}^*$)

On a $\chi^2(n) \sim \Gamma(\frac{n}{2}, \frac{1}{2})$.
  - espérance : n - variance : $2n$.

Rq : $\chi^2(n)$ a la même loi qu'une somme de n carrés de v.a gaussiennes centrées réduites.
- $\mathcal{E}(\lambda) \sim \Gamma(1, \lambda)$.

3. Opérations sur les densités

Prop 8. Soit $X = (X_1, X_2)$ une v.a à valeurs dans $\mathbb{R}^2$ de densité f_X. Alors X_1 et X_2 admettent des densités respectives f_{X_1} et f_{X_2} données par : [OUV 1] p. 184

$$f_{X_1}(x_1) = \int_{\mathbb{R}} f_X(x_1, x_2)\,dx_2 \quad \text{et} \quad f_{X_2}(x_2) = \int_{\mathbb{R}} f_X(x_1, x_2)\,dx_1$$

Rq: On peut généraliser cette proposition à $\mathbb{R}^n$.

Ex 9: Soit $X=(X_1,X_2)$ de densité $f_X(x_1,x_2)=\frac{1}{2\pi}e^{-\frac{(x_1^2+x_2^2)}{2}}$

Alors $X_1 \sim \mathcal{N}(0,1)$ et X_2 aussi. [OUV1] p.188

Prop 9: Ici $n=1$. Soit X de densité f_X et $g:\mathbb{R}\to\mathbb{R}$ strictement monotone et dérivable. Alors $Y=g(X)$ admet une densité f_Y donnée par :

$$f_Y(y)=\begin{cases} f_X(g^{-1}(y))\,|(g^{-1})'(y)| & \text{si } y\in g(\mathbb{R}) \\ 0 & \text{sinon.}\end{cases}$$

[OUV1] p.189

Prop 10: (Critère d'indépendance) Soient X_1 et X_2 deux v.a à valeurs dans $\mathbb{R}^{d_1}$ et $\mathbb{R}^{d_2}$. Alors. [OUV2] p.43

(i) ($X_1 \perp\!\!\!\perp X_2$ et de densités resp. f_{X_1} et f_{X_2})

$\Rightarrow$ ((X_1,X_2) à densité et $f_{(X_1,X_2)}(x_1,x_2)=f_{X_1}(x_1)f_{X_2}(x_2)$)

(ii) ((X_1,X_2) admet une densité $f_X(x_1,x_2)=f_1(x_1)f_2(x_2)$ où f_1 et f_2 sont intégrables positives)

$\Rightarrow$ (f_1 et f_2 sont à un facteur positif près les densités de X_1 et X_2 et $X_1 \perp\!\!\!\perp X_2$)

App 11: $S\sim\mathcal{E}(\frac{1}{2})$ et $\Theta\sim\mathcal{U}[0,2\pi]$ indépendantes alors $X=\sqrt{S}\cos\Theta$ et $Y=\sqrt{S}\sin\Theta$ sont deux v.a réelles indépendantes de même loi $\mathcal{N}(0,1)$. [OUV2] p.65

Prop 12: Soient $X_1 \perp\!\!\!\perp X_2$ deux v.a à valeurs dans $\mathbb{R}^n$.

Alors : $\mathbb{P}_{X_1+X_2}=\mathbb{P}_{X_1}*\mathbb{P}_{X_2}$ [OUV2] p.61,62

- Si de plus, X_1 et X_2 sont à densités (f_{X_1} et f_{X_2})

Alors X_1+X_2 admet une densité $f_{X_1+X_2}$ donnée par :

$$\forall y\in\mathbb{R}^n,\quad f_{X_1+X_2}(y)=\int_{\mathbb{R}^n} f_{X_1}(x_1)f_{X_2}(y-x_1)\,d\lambda_n(x_1)$$

App 13: Soient $X_1,\dots,X_n$ indépendantes de lois respectives $\Gamma(a_1,\lambda),\dots,\Gamma(a_n,\lambda)$ alors :

$$X_1+X_2+\dots+X_n \sim \Gamma(a_1+\dots+a_n,\lambda)$$

[OUV2] p.62

4. Fonction caractéristique et densité

Déf 14: On appelle fonction caractéristique de la v.a X l'application φ_X donnée par :

$$\forall t\in\mathbb{R}^n,\quad \varphi_X(t)=\int_{\mathbb{R}^n}\exp(i\langle x,t\rangle)\,d\mathbb{P}_X(x) = \mathbb{E}[e^{i\langle X,t\rangle}]$$

[OUV2] p.191

Rq: Si X est à densité, φ_X est la transformée de Fourier de f_X (car $d\mathbb{P}_X(x)=f_X(x)dx$)

Thm 15: La fonction caractéristique caractérise la loi, ie: $\varphi_X=\varphi_Y \iff \mathbb{P}_X=\mathbb{P}_Y$. [OUV2] p.197

Ex 16: - Si $X\sim\mathcal{N}(m,\sigma^2)$ alors $\varphi_X(t)=\exp(imt-\frac{t^2\sigma^2}{2})$

- Si $X\sim\mathcal{E}(\lambda)$ alors $\varphi_X(t)=\frac{1}{1-\frac{it}{\lambda}}$ [OUV2] p.28

Prop 17: Soit X une v.a telle que φ_X soit Lebesgue intégrable. Alors $\mathbb{P}_X$ est absolument continue par rapport à la mesure de Lebesgue et sa densité est donnée λ_n-pp par : $\forall x\in\mathbb{R}^n,\ f_X(x)=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\varphi_X(t)\exp(-i\langle x,t\rangle)\,dt$ [OUV2] p.199

Prop 18: (Critère d'indépendance) Soit $X=(X_1,X_2)$ à valeurs dans $\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$ alors : [OUV2] p.200

$$(X_1 \perp\!\!\!\perp X_2) \iff (\forall (t_1,t_2)\in\mathbb{R}^{d_1\times d_2},\ \varphi_{(X_1,X_2)}(t_1,t_2)=\varphi_{X_1}(t_1)\varphi_{X_2}(t_2))$$

Ex 19: X_1, X_2 iid de fonction caractéristique $\varphi(t)=\frac{1}{1+t^2}$ (loi de Laplace). On pose $Y_1=X_1-X_2$ et $Y_2=X_1+X_2$. Alors Y_1 et Y_2 ont m̂ fonction caractéristique mais ne sont pas indépendantes. [OUV2] p.201

II. Vecteurs gaussiens

1. Définitions et premières propriétés [CAD]

Déf 20: Un vecteur aléatoire X est gaussien si et seulement si toute combinaison linéaire de ses composantes est une v.a réelle gaussienne. Si X à valeurs dans $\mathbb{R}^n$ on note $\mathcal{N}_n(M,\Sigma)$ où $M=\mathbb{E}[X]$ p.182

et $\Sigma=(\mathrm{Cov}(X_i,X_j))_{\substack{1\le i\le n\\1\le j\le n}}$

DVPT 1
[OUV2]
p. 271

De plus, si Σ est inversible alors X admet une densité qui vaut : $f_X(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(x-m)^T \Sigma^{-1}(x-m)\right)$

Thm 21 Soient X et Y deux v.a réelles iid et L^2 telles que $X+Y \perp\!\!\!\perp X-Y$ alors X et Y sont des v.a gaussiennes.

Prop 22 (Caractérisation de l'indépendance) Soit X gaussien $\mathcal{N}_n(m, \Sigma)$. Les composantes de X sont des v.a indépendantes ssi Σ est diagonale. p. 182

2. Projection de vecteurs gaussiens [CAD]

Thm 23 : (Cochran) Soit $\sigma > 0$, $X \sim \mathcal{N}_n(0, \sigma^2 I_n)$ et $V_1, \dots, V_p$ des sev orthogonaux de dim resp. $r_1, \dots, r_p$ tq $\mathbb{R}^n = V_1 \oplus \dots \oplus V_p$. Alors les projections orthogonales $\Pi_1, \dots, \Pi_p$ de X sur $V_1, \dots, V_p$ sont des vecteurs gaussiens indépendants et p. 97

$\forall i \in 1, \dots, p$: $\frac{1}{\sigma^2} \|\Pi_i\|^2 \sim \chi^2(r_i)$

App 24 : En considérant $V = \mathrm{Vect}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$ on en déduit que $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ et $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ sont indépendants et que $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$. p. 99

III - Utilisation des lois à densités en statistiques

1. Maximum de vraisemblance [CAD]

Dans toute cette partie $(\mathcal{H}^n, \{P_\theta\}_{\theta \in \Theta})$ est un modèle statistique dominé par une mesure σ-finie μ, avec $\mathcal{H} \subset \mathbb{R}^k$ et $\Theta \subset \mathbb{R}^d$.
Soit $(X_1, \dots, X_n) \sim P_\theta$.

Déf 25 : La vraisemblance du modèle est l'application :
$L_n : \mathcal{H}^n \times \Theta \to \mathbb{R}_+$ tq $\forall \theta \in \Theta$ $dP_\theta = L_n(\cdot\,;\theta)\, d\mu$.
En particulier, si L est la vraisemblance du modèle $(\mathcal{H}, \{Q_\theta\}_{\theta \in \Theta})$ dominé par ν et que $\forall \theta \in \Theta$, $P_\theta = Q_\theta^{\otimes n}$ alors p. 45

$$L_n : \mathcal{H}^n \times \Theta \to \mathbb{R}$$
$$(x_1, \dots, x_n, \theta) \mapsto \prod_{i=1}^n L(x_i, \theta)$$

Rq : Lorsqu'il s'agit de variables aléatoires à densité f on a
$$L_n(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i)$$

Prop 26 : Pour tout $\theta \in \Theta$, $L_n(\cdot\,;\theta) > 0$ P_θ - p.s. p. 44

Ex 27 : Pour le modèle statistique $(\mathbb{R}^n, \{\mathcal{N}(m, \sigma^2)^{\otimes n}\}_{m \in \mathbb{R}, \sigma > 0})$ on a $L_n(x_1, \dots, x_n; m, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(\frac{-\sum_{i=1}^n (x_i - m)^2}{2\sigma^2}\right)$ p. 44

Déf 28 : Un estimateur du maximum de vraisemblance est un estimateur $\hat{\theta}$ qui vérifie :
$$L_n(\cdot\,;\hat{\theta}) = \sup_{\theta \in \Theta} L_n(\cdot\,;\theta)$$ p. 66

Rq : Lorsque $P_\theta = Q_\theta^{\otimes n}$ on calcule l'EMV en maximisant la log-vraisemblance : $\ln L_n(x_1, \dots, x_n; \theta) = \sum_{i=1}^n \ln L(x_i; \theta)$, p. 41

App 29 : $X_1, \dots, X_n$ iid de loi $\mathcal{U}[0, \theta]$, $\theta > 0$ et $\hat{\theta}_n$ l'EMV du paramètre θ. Alors on a :

DVPT 2 [CAD]

1) $\hat{\theta}_n = \max_{1 \le i \le n} X_i$
2) Biais : $\mathbb{E}[\hat{\theta}_n] = \frac{n}{n+1}\theta$
3) Forte consistance : $\hat{\theta}_n \xrightarrow{p.s} \theta$
4) vitesse et loi limite : $n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} -\mathcal{E}(1/\theta)$
5) Risque quadratique : $\mathbb{E}[(\hat{\theta}_n - \theta)^2] = \frac{2\theta^2}{(n+1)(n+2)}$

2. Théorème central limite (TCL) [CAD]

Thm 30 : Soient $X_1, \dots, X_n$ iid et L^2 alors
$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - m\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$ p. 149
où $m = \mathbb{E}[X_1]$ et $\sigma^2 = \mathrm{Var}(X_1)$.

App 31 : (Intervalle de confiance asymptotique)
$(\mathbb{R}^n, \{Q_\theta^{\otimes n}\}_{\theta \in]-1,1[})$ où Q_θ désigne la loi sur $\mathbb{R}$ de densité $\frac{1}{2}(1 + \theta x)\mathbb{1}_{]-1,1[}(x)$ p. 41
$\hat{\theta} = 3\overline{X}_n$ est un estimateur de θ sans biais et
$$I_{1-\alpha} = \left[\hat{\theta} - q\sqrt{(3 - \hat{\theta}^2)/n}\,,\ \hat{\theta} + q\sqrt{(3 - \hat{\theta}^2)/n}\right]$$
(où q est le quantile d'une $\mathcal{N}(0,1)$ d'ordre $1 - \frac{\alpha}{2}$) est un intervalle de confiance asymptotique pour θ au niveau $1 - \alpha$.

Estimateur du maximum de vraisemblance pour $\mathcal{U}([0,\theta])$, $\theta>0$

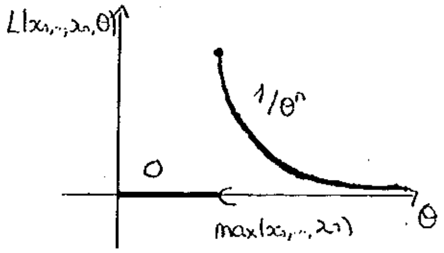
__Théorème__

Soient $X_1,\dots,X_n$ iid de loi $\mathcal{U}([0,\theta])$ avec $\theta>0$. Notons $\hat{\theta}_n$ l'estimateur du maximum de vraisemblance du paramètre θ. Alors on a :

1) $\hat{\theta}_n = \max\limits_{i\in[\![1,n]\!]}(X_i)$
2) Le biais de l'estimateur est $E[\hat{\theta}_n] = \frac{n}{n+1}\theta$
3) L'estimateur est fortement consistant : $\hat{\theta}_n \xrightarrow{p.s} \theta$
4) Vitesse et loi limite : $n(\hat{\theta}_n-\theta) \xrightarrow{\mathcal{L}} -\mathcal{E}\left(\frac{1}{\theta}\right)$
5) Le risque quadratique : $E[(\hat{\theta}_n-\theta)^2] = \frac{2\theta^2}{(n+1)(n+2)}$

1) La vraisemblance des observations $(x_1,\dots,x_n)$ est :

$$L(x_1,\dots,x_n,\theta) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(x_i) = \begin{cases} 0 & \text{si } \theta<\max(x_1,\dots,x_n) \\ \frac{1}{\theta^n} & \text{si } \theta\geq \max(x_1,\dots,x_n)\end{cases}$$



d'où l'estimateur cherché est $\hat{\theta}_n = \max\limits_{i\in[\![1,n]\!]}(X_i)$

2) Calculons la loi de $\hat{\theta}_n$.

Soit $t\geq 0$:

$F_{\hat{\theta}_n}(t) = \mathbb{P}(\hat{\theta}_n\leq t) = \mathbb{P}(X_i\leq t, \forall i\in[\![1,n]\!]) = \mathbb{P}(X_i\leq t)^n$ car les $(X_i)_i$ sont iid

donc $F_{\hat{\theta}_n}(t) = \begin{cases} \left(\frac{t}{\theta}\right)^n & \text{si } t\leq\theta \\ 1 & \text{sinon}\end{cases}$

d'où la densité de $\hat{\theta}_n$ est $f_{\hat{\theta}_n}(t) = \frac{n}{\theta}\left(\frac{t}{\theta}\right)^{n-1}\mathbb{1}_{[0,\theta]}(t)$

On peut désormais calculer le biais :

$$E[\hat{\theta}_n] = \int_0^\theta t f_{\hat{\theta}_n}(t)\,dt = \frac{n}{\theta^n}\int_0^\theta t^n\,dt = \frac{n}{\theta^n}\,\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta$$

donc $\hat{\theta}_n$ est un estimateur biaisé.

3) Soit $\varepsilon>0$, on a

$$\mathbb{P}(|\hat{\theta}_n-\theta|\geq\varepsilon) = 1-\mathbb{P}(\theta-\varepsilon\leq\hat{\theta}_n\leq\theta+\varepsilon) = 1-(F_{\hat{\theta}_n}(\theta+\varepsilon)-F_{\hat{\theta}_n}(\theta-\varepsilon))$$
$$= 1-\left(1-\left(\frac{\theta-\varepsilon}{\theta}\right)^n\right) = \left(\frac{\theta-\varepsilon}{\theta}\right)^n = \left(1-\frac{\varepsilon}{\theta}\right)^n$$

or $1-\frac{\varepsilon}{\theta}<1$ donc $\sum\limits_n \mathbb{P}(|\hat{\theta}_n-\theta|\geq\varepsilon)$ converge

On en déduit par le lemme de Borel-Cantelli que $\hat{\theta}_n \xrightarrow{ps} \theta$.
L'estimateur est donc fortement consistant.

4) Soit $t \in \mathbb{R}^+$, on a

$$\mathbb{P}(n(\theta-\hat{\theta}_n) \geq t) = \mathbb{P}(\theta - \frac{t}{n} \geq \hat{\theta}_n) = \left(\frac{\theta - t/n}{\theta}\right)^n = \left(1-\frac{t}{n\theta}\right)^n = e^{n\ln(1-\frac{t}{n\theta})} \longrightarrow e^{-t/\theta}$$

donc $\mathbb{P}(n(\theta-\hat{\theta}_n) \leq t) \underset{n\to\infty}{\longrightarrow} 1 - e^{-t/\theta}$

d'où $n(\theta - \hat{\theta}_n) \xrightarrow{\mathcal{L}} \mathcal{E}(\frac{1}{\theta})$ et $n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} -\mathcal{E}(\frac{1}{\theta})$

5) Par le lemme de transfert, on a :

$$E[\hat{\theta}_n^2] = \int_{\mathbb{R}} t^2 f_{\hat{\theta}_n}(t)\,dt = \frac{n}{\theta^n}\int_0^\theta t^{n+1}\,dt = \frac{n}{\theta^n}\left[\frac{t^{n+2}}{n+2}\right]_0^\theta = \frac{n}{n+2}\theta^2$$

donc le risque quadratique est :

$$E[(\hat{\theta}_n - \theta)^2] = E[\hat{\theta}_n^2 - 2\theta\hat{\theta}_n + \theta^2] = E(\hat{\theta}_n^2) - 2\theta E[\hat{\theta}_n] + \theta^2$$

$$\text{"} \quad = \frac{n}{n+2}\theta^2 - \frac{2\theta^2 n}{n+1} + \theta^2$$

$$\text{"} \quad = \theta^2 \frac{n(n+1) - 2n(n+2) + (n+1)(n+2)}{(n+1)(n+2)}$$

$$\text{"} \quad = \frac{2\theta^2}{(n+1)(n+2)}$$

<u>Référence</u> : Cadre/Vial, "Statistique mathématique".

<u>Remarque</u> : C'était prévisible que l'estimateur serait de moyenne plus petite que θ car $\hat{\theta}_n$ prend des valeurs plus petites que θ.

Bernstein sur les lois normales

Caroline Robet & Maylis Varvenne

Théorème. *Soit X et Y deux variables aléatoires réelles indépendantes, identiquement distribuées de loi μ (de fonction caractéristique φ), L^2, telles que X+Y et X-Y soient indépendantes alors X et Y sont des variables aléatoires gaussiennes.*

Démonstration. On note $\gamma = \mu * \mu$.

- **Mq $\forall (u,v) \in \mathbb{R}^2$, $\widehat{\gamma}(u+v)\widehat{\gamma}(u-v) = \widehat{\gamma}(u)^2|\widehat{\gamma}(v)|^2$ où $\widehat{\gamma}$ est la transformée de Fourier de γ.**
  Par indépendance des variables X+Y et X-Y, et des variables X et Y, on a sur les fonctions caractéristiques que $\forall (u,v) \in \mathbb{R}^2$:
  $$\varphi_{(X+Y,X-Y)}(u,v) = \varphi_{(X+Y)}(u)\varphi_{(X-Y)}(v) = \varphi(u)\varphi(u)\varphi(v)\overline{\varphi(v)} = \varphi(u)^2|\varphi(v)|^2.$$
  De plus, on a aussi pour tous réels u et v :
  $$\varphi_{(X+Y,X-Y)}(u,v) = \varphi_{(X,Y)}(u+v,u-v) = \varphi(u+v)\varphi(u-v),$$
  on a donc pour tous réels u et v,
  $$\varphi(u+v)\varphi(u-v) = \varphi(u)^2|\varphi(v)|^2,$$
  Or on a $\widehat{\gamma} = \varphi_{X+Y} = \varphi^2$, en mettant au carré l'équation précédente, on obtient alors la relation :
  $$\forall (u,v) \in \mathbb{R}^2, \widehat{\gamma}(u+v)\widehat{\gamma}(u-v) = \widehat{\gamma}(u)^2|\widehat{\gamma}(v)|^2 \quad (1)$$

- **Soit $\overline{\gamma}$ la probabilité définie pour tout borélien A par $\overline{\gamma}(A) = \gamma(-A)$ et $\delta = \gamma \star \overline{\gamma}$. Montrons que la transformée de Fourier $\widehat{\delta}$ de δ vérifie la relation : $\forall (u,v) \in \mathbb{R}^2$, $\widehat{\delta}(u+v)\widehat{\delta}(u-v) = (\widehat{\delta}(u))^2(\widehat{\delta}(v))^2$**
  Le théorème de transfert permet d'établir l'égalité sur les transformées de Fourier de γ et $\overline{\gamma}$:
  $$\forall t \in \mathbb{R}, \ \ \widehat{\overline{\gamma}}(t) = \overline{\widehat{\gamma}}(t).$$
  On a donc $\forall t \in \mathbb{R}$, $\widehat{\delta}(t) = \widehat{\gamma}(t)\widehat{\overline{\gamma}}(t) = \widehat{\gamma}(t)\overline{\widehat{\gamma}}(t) = |\widehat{\gamma}(t)|^2 \geqslant 0$. On prend alors le carré des modules dans la relation (1), pour obtenir que pour tous réels u et v,
  $$\widehat{\delta}(u+v)\widehat{\delta}(u-v) = (\widehat{\delta}(u))^2(\widehat{\delta}(v))^2 \quad (2)$$

- **Montrons que $G = \{t \in \mathbb{R} | \widehat{\delta}(t) \neq 0\}$ est un groupe, que $\widehat{\delta}$ ne s'annule jamais et déterminons $\widehat{\delta}$ puis $|\widehat{\gamma}|$.** Si u et v sont dans G, par la relation précédente on a $\widehat{\delta}(u+v) \neq 0$, $\widehat{\delta}(-u) \neq 0$ et $\widehat{\delta}(0) = 1$ donc G est un groupe. De plus, $\widehat{\delta}$ est continue donc G est ouvert car G$=\widehat{\delta}^{-1}(\mathbb{R}^{+*})$ d'où $G = \mathbb{R}$. Donc $\widehat{\delta}$ ne s'annule pas et est donc strictement positive.
  On pose alors pour tout réel t, $f(t) = -ln(\widehat{\delta}(t))$. La relation (2) donne pour tous réels u et v
  $$f(u+v) + f(u-v) = 2(f(u) + f(v)).$$
  Comme f est continue, positive et que $f(0) = 0$, il existe $a > 0$ tel que pour tout réel u, $f(u) = au^2$ (argument classique : sur les entiers, les rationnels puis par continuité sur les réels). Donc $\widehat{\delta}(u) = \exp(-au^2)$ et $\forall t \in R$
  $$|\widehat{\gamma}(t)| = \exp(-\frac{at^2}{2})$$

- **On pose pour t réel, $g(t) = \frac{\widehat{\gamma}}{|\widehat{\gamma}|}$. Montrons que g vérifie la relation $\forall (u,v) \in \mathbb{R}^2,\ g^2(u+v) = g^2(u)g^2(v)$**

  On prend les modules dans la relation (1), on a donc pour tous réels u et v,

$$|\widehat{\gamma}(u+v)||\widehat{\gamma}(u-v)| = (|\widehat{\gamma}(u)|)^2(|\widehat{\gamma}(v)|)^2.$$

  On quotiente cette relation par la relation (1) pour obtenir

$$\forall (u,v) \in \mathbb{R}^2 g(u+v)g(u-v) = g^2(u);$$

  en échangeant u en v, on a aussi

$$\forall (u,v) \in \mathbb{R}^2 g(u+v)g(v-u) = g^2(v).$$

  On multiplie membre à membre ces égalités en tenant compte du fait que pour tout t, on a $g(-t) = \overline{g(t)}$ et $|g(t)| = 1$, d'où

$$\forall (u,v) \in \mathbb{R}^2,\ g^2(u+v) = g^2(u)g^2(v).$$

Lemme. *Soit Φ une application borélienne de $\mathbb{R}$ dans $\mathbb{C}$ telle que pour tout réel t, $|\Phi(t)| = 1$ et pour tous réels s et t, $\Phi(s+t) = \Phi(s)\Phi(t)$ alors $\exists c \in \mathbb{R}$ tel que $\forall t \in \mathbb{R},\ \Phi(t) = exp(ict)$*

Démonstration. On pose $h(x) = \int_0^x \Phi(t)dt$ si on avait h=0 alors on aurait $Re(\Phi)$ et $Im(\Phi)$ qui seraient nulles presque partout ce qui n'est pas le cas car $|\Phi(t)| = 1$ donc $\exists a \in \mathbb{R}$ tel que $h(a) \neq 0$. On a

$$h(x+a) - h(x) = \int_x^{x+a} \Phi(t)dt = \int_0^a \Phi(x+s)ds = \Phi(x)\int_0^a \Phi(s)ds = \Phi(x)h(a)$$

ce qui donne

$$\Phi(x) = \frac{h(x+a) - h(x)}{h(a)}.$$

La continuité de h entraîne la continuité de Φ ; h est donc dérivable et donc Φ est dérivable. En dérivant par rapport à s l'égalité $\Phi(s+t) = \Phi(s)\Phi(t)$, on a donc pour tout s et t, $\Phi'(s+t) = \Phi'(s)\Phi(t)$ et donc pour tout t $\Phi'(t) = \Phi'(0)\Phi(t)$. Si $\Phi'(0) = 0$, alors pour tout t $\Phi'(t) = 0$ et donc $\Phi(t) = 1$ par les hypothèses sur Φ. Si $\Phi'(0) \neq 0$, on a en évaluant la relation précédente en 0 que $\Phi(0) = 1$ d'où en posant $c = -i\Phi'(0)$, on a $\Phi(t) = \exp(ict)$. Puis comme $|\Phi(1)| = |\exp(ic)| = 1$, c est réel. □

- On applique ce lemme à g^2, il existe donc un réel m tel que pour tout t $g^2(t) = \exp(i2mt)$. Comme g(0)=1 et g est continue, on obtient $g(t) = \exp(imt)$. D'où, $\widehat{\gamma}(t) = g(t)|\widehat{\gamma}(t)| = \exp(imt - at^2/2)$ avec $a > 0$. Donc, par injectivité de la transformée de Fourier, γ est la probabilité gaussienne $\mathcal{N}(m,a)$. Donc la variable aléatoire X+Y est gaussienne. De même, on montre que X-Y est gaussienne. Ces variables aléatoires étant indépendantes la variable aléatoire (X+Y,X-Y) est aussi gaussienne ainsi les variables aléatoires X et Y sont gaussiennes, comme transformations linéaires de la variable aléatoire gaussienne (X+Y,X-Y)

  □

<u>**Référence**</u> : Ouvrard, Probabilité 2, ex 13.4 p280.