263 Variables aléatoires à densité. Exemples et Applications

Cadre : On se place sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. X désigne une variable ou un vecteur aléatoire réel (v.a).

I - Variables aléatoires à densité

1) Définitions, propriétés et lois usuelles

Def 1 : X est à densité si $\mathbb{P}_X \ll leb$. Il existe alors une unique fonction mesurable positive f_X telle que :

$$\forall A \in \mathcal{B}(\mathbb{R}),\ \mathbb{P}_X(A) = \int_A f_X \, dleb$$

Ex 1 : Si $f = \mathbb{1}_{[0,1]}$, on dit que X suit une loi uniforme sur $[0,1]$

Def 2 : Soit X de densité f_X. La fonction de répartition de X est définie sur $\mathbb{R}$ par : $F_X(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^{t} f_X(u)\,du$

Prop 1 : 1) La donnée de f_X (resp. F_X) détermine la loi de X.
2) Une v.a. à densité ne possède pas d'atome : $\forall x \in \mathbb{R}, \mathbb{P}(X=x)=0$
3) F_X est continue, dérivable à droite et en tout point de continuité de f_X

Def 3 (Inverse généralisé) : Soit X v.a. à densité. On appelle inverse généralisé de F_X la fonction F_X^- définie par :
$\forall u \in [0,1],\ F_X^-(u) = \inf\{x,\ F(x) \geq u\}$
Si F_X est bijective, $F_X^- = F_X^{-1}$

Prop 2 : Soit U de loi uniforme sur $[0,1]$ et X de fonction de répartition F. Alors $F^-(U)$ a même loi que X

Def 4 : On appelle espérance de X la quantité $\mathbb{E}[X] = \int_X x f_X(x)\,dx$
Soit $\Phi : \mathbb{R} \to \mathbb{R}$ une fonction mesurable. Alors on a :
$$\mathbb{E}[\Phi(X)] = \int_{\mathbb{R}} \Phi(x) f_X(x)\,dx$$

Def 5 : On dit que X possède un moment d'ordre $n \geq 0$ si $\mathbb{E}[|X|^n] < \infty$

Prop 3 (Changement de variable) Soient X de densité f_X, $\varphi : \mathbb{R} \to \mathbb{R}$ un $\mathcal{C}^1$-difféomorphisme. Alors $Y = \varphi(X)$ admet pour densité $f_Y(y) = |J_{\varphi^{-1}}(y)| f_X(\varphi^{-1}(y))$, $\forall y \in \mathbb{R}$

2) Lois usuelles

Loi de X	Densité f_X	$\mathbb{E}[X]$	Variance σ^2
Uniforme sur $[a,b]$ $\mathcal{U}([a;b])$	$\frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normale : $\mathcal{N}(m, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	m	σ^2
Exponentielle : $\mathcal{E}(\lambda)$ $\lambda > 0$	$\lambda e^{-\lambda x}\mathbb{1}_{x>0}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma : $\Gamma(\alpha, \lambda)$ $\alpha, \lambda \in \mathbb{R}_+^*$	$\frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x} x^{\alpha-1}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$

Applic° 1 : Etude de la loi Gamma $\boxed{\text{DEV}}$

Ex 2 : Si X est de Cauchy, ie de densité $f_X(x) = \frac{1}{\pi} \times \frac{1}{(x-a)^2 + b^2}$, alors X ne possède aucun moment.

Applic° 2 : Simuler une loi exponentielle ou de Cauchy à partir d'une loi uniforme

3) Couple aléatoires, indépendance

Def 6 : Soit (X,Y) un couple aléatoire à valeurs dans $\mathbb{R}^{m+n}$ de densité $f_{(X,Y)}$. Alors la loi marginale de X est donnée par :
$\forall x \in \mathbb{R}^m,\ f_X(x) = \int_{\mathbb{R}^n} f_{(X,Y)}(x,y)\,dy$

Prop 4 : Soient $X_1, \ldots, X_n$ des v.a.r. à densité. Elles sont indépendantes si et seulement si la densité du vecteur $(X_1, \ldots, X_n)$ est le produit des densités des X_i.

Applic° 4 : (Méthode Delta) : Soit $(X_n)_n$ une suite de v.a.r. On suppose qu'il existe θ, σ tels que $\sqrt{n}(X_n-\theta) \Rightarrow \mathcal{N}(0,\sigma^2)$. Soit $g : \mathbb{R} \to \mathbb{R}$ dérivable en θ telle que $g'(\theta) \neq 0$. Alors :

$$\sqrt{n}\left(g(X_n)-g(\theta)\right) \Rightarrow \mathcal{N}\left(0, g'(\theta)^2\sigma^2\right)$$

2) Vecteurs gaussiens

Soit $(X_1,..,X_d)$ un vecteur aléatoire. On note $\mathbb{E}[X] = {}^t(\mathbb{E}[X_1],..,\mathbb{E}[X_d])$ et $\Gamma = \mathbb{E}\left[(X-\mathbb{E}[X])\,{}^t(X-\mathbb{E}[X])\right]$ la matrice de covariance

Si $X \in \mathbb{R}^n$ de moyenne m et de covariance Γ, alors AX est de moyenne Am et de matrice de covariance $A\Gamma\,{}^tA$, pour $A \in \mathcal{M}_n(\mathbb{R})$

Def 9 : Un v.a. de $\mathbb{R}^d$ est un vecteur gaussien si toute combinaison linéaire de ses composantes est une v.a.r. gaussienne :
$\forall a \in \mathbb{R}^d$, ${}^taX \sim \mathcal{N}(m,\sigma^2)$

Si $m = \mathbb{E}[X]$ et Γ est la matrice de covariance de X, on note $X \sim \mathcal{N}(m,\Gamma)$

Rmq 2 : Les composantes d'un vecteur gaussien sont des v.a. gaussiennes, la réciproque est fausse.

C-ex 1 : Si $X \sim \mathcal{R}(1/2)$ et $Y \sim \mathcal{N}(0,1)$, alors Y et XY sont gaussiennes mais (Y, XY) n'est pas un vecteur gaussien.

Prop 8 : Soit X v.a. gaussien de moyenne m et covariance Γ. Alors $\forall u \in \mathbb{R}^d$, $\varphi_X(u) = \exp\left(i\,{}^tum - \frac{{}^tu\Gamma u}{2}\right)$

Thm 3 : $\mathcal{N}(m,\Gamma)$ admet une densité si et seulement si Γ est inversible et alors : $\forall x \in \mathbb{R}^d$, la densité est donnée par :

$$f(x) = \frac{1}{\sqrt{(2\pi)^d|\Gamma|}}\exp\left(-\frac{1}{2}\,{}^t(x-m)\Gamma^{-1}(x-m)\right)$$

Thm 4 (Cochran) Soit $X \sim \mathcal{N}(m,\sigma^2 I_n)$. Soit $\mathbb{R}^n = E_1 \oplus \cdots \oplus E_p$ une décompos° en sous espaces orthogonaux de dimensions $d_1,..,d_p$. Soit π_k la matrice de projection orthogonale sur E_k et $Y_k = \pi_k X$. Alors : 1) $Y_1,..,Y_p$ sont des v.a. gaussiens indépendants et $Y_k \sim \mathcal{N}(\pi_k m, \sigma^2 \pi_k)$

2) Les v.a. $\|Y_k - \pi_k m\|^2$ sont indépendantes et $\frac{\|Y_k-\pi_k m\|^2}{\sigma^2} \sim \chi^2(d_k)$

Thm 5 (Fisher) Soit $(X_1,..,X_n)$ un n-échantillon de loi $\mathcal{N}(m,\sigma^2)$. On pose $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$ et $\widehat{S_n^2} = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2$. Alors les v.a. $\overline{X_n}$ et $\widehat{S_n^2}$ sont indépendantes et

$$\begin{cases} \overline{X_n} \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right) \\ \frac{(n-1)\widehat{S_n^2}}{\sigma^2} \sim \chi^2(n-1) \end{cases}$$

Applic° 5 : Calcul d'un intervalle de confiance exact pour σ^2 avec un n-échantillon de loi $\mathcal{N}(m,\sigma^2)$, σ^2 inconnu.

3) Estimation

Cadre : Soient $X_1,..,X_n$ un n-échantillon de loi P_θ et admettant une densité f_θ, où θ est inconnu.

Def 10 : La fonction $\theta \mapsto L_{X_1,..,X_n}(\theta) = \prod_{i=1}^n f_\theta(X_i)$ s'appelle la fonction de vraisemblance du modèle

Si L admet un unique maximum atteint en $\hat{\theta}$, on appelle alors $\hat{\theta}$ l'estimateur du maximum de vraisemblance (EMV)

Rmq 3 : On utilise $\hat{\theta}$ pour estimer la valeur de θ

Applic° 6 : Soit $(X_1,..,X_n)$ un n-échantillon de loi $\mathcal{E}(\theta)$, θ inconnu. Alors l'EMV fournit $\hat{\theta} = \frac{1}{\overline{X_n}}$ où $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$

Prop5: Soient $X \perp\!\!\!\perp Y$, à valeurs dans $\mathbb{R}^d$. Alors $X+Y$ a une densité égale à $f_{X+Y}(z) = \int_{\mathbb{R}^d} f_X(x) f_Y(z-x)dx$

Ex3: Si $X \sim \Gamma(a,\lambda)$ et $Y \sim \Gamma(b,\lambda)$ alors $X+Y \sim \Gamma(a+b,\lambda)$

Def7: Si $X_1, \dots, X_n$ sont n variables indépendantes de loi normale $\mathcal{N}(0,1)$, alors $Z = \sum_{i=1}^{n} X_i^2$ suit une loi $\Gamma(\frac{1}{2}, \frac{1}{2})$ appelée loi du χ^2

* La somme de n v.a. indépendantes exponentielles $\mathcal{E}(\lambda)$ est une loi $\Gamma(n,\lambda)$, dite d'Erlang.

4) Fonctions caractéristiques

Def8: * Pour X à valeurs dans $\mathbb{R}^d$, on définit sa transformée de Laplace par : $\forall t \in \mathbb{R}^d$, $L_X(t) = \mathbb{E}[e^{\langle t,X\rangle}] \in \mathbb{R}^+ \cup \{+\infty\}$

* De même, on définit sa fonction caractéristique :

$$\varphi_X(t) = \mathbb{E}[e^{i\langle t,X\rangle}] = \int_{\mathbb{R}} e^{i\langle t,x\rangle} f_X(x)\,dx$$

Prop6: 1) La fonction caractéristique caractérise la loi

2) Si X est à densité et $\int_{\mathbb{R}} |\varphi_X(t)|dt < \infty$, alors $f_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_X(t)dt$

Rmq1: Soit $\varphi : \mathbb{R} \to \mathbb{C}$ continue telle que $\varphi(0)=1$ et $\int_{\mathbb{R}} |\varphi(t)|dt < \infty$

Si $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi(t)dt$ est positive et intégrable sur $\mathbb{R}$, alors φ est la fonction caractéristique de la loi de densité f.

Prop7: Soient X et Y deux v.a.

1) $X \perp\!\!\!\perp Y \iff \forall t,s,\ \varphi_{(X,Y)}(t,s) = \varphi_X(t)\varphi_Y(s)$

2) $X \perp\!\!\!\perp Y \Rightarrow \varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

⚠ La réciproque est fausse.

Ex4: * Si $X \sim \mathcal{N}(m,\sigma^2)$, $\varphi_X(t) = e^{itm - \frac{\sigma^2 t^2}{2}}$ donc l'espérance et la variance caractérisent une variable gaussienne

* Si $X \sim \mathcal{E}(\lambda)$, $\varphi_X(t) = \left(1 - \frac{it}{\lambda}\right)^{-1}$

II - Applications

1) Utilisation de la convergence

Thm1 (de Lévy) Soit $(X_n)_n$ une suite de v.a. dans $\mathbb{R}^d$ telle que $\lim_{n\to\infty} \varphi_{X_n} = \varphi$, avec φ continue en 0. Alors φ est la fonct° caractéristique d'une v.a. X et $(X_n)_n$ converge en loi vers X

On note $X_n \underset{n\to\infty}{\Longrightarrow} X$

Lme1 (de Scheffé) : Soit $(X_n)_n$ suite de v.a., dans $\mathbb{R}^d$, à densités notées f_n. Si $(f_n)_n$ converge presque sûrement vers une fonction f telle que $\int_{\mathbb{R}^d} f\,d\lambda_d = 1$, alors $(X_n)_n$ converge en loi vers une v.a. X de densité f. De plus, on a :

$$\lim_{n\to\infty} \sup_{A\in\mathcal{B}(\mathbb{R}^d)} \left|\mathbb{P}(X_n\in A) - \int_A f\,d\lambda_d\right| = 0.$$

Lme2 (de Slutsky) Soient $(X_n)_n$ et $(Y_n)_n$ des suites de v.a.r. sur $(\Omega,\mathcal{A},\mathbb{P})$ telles que $\begin{cases} X_n \Rightarrow X \\ Y_n \Rightarrow a, \text{ constante.}\end{cases}$ Alors $(X_n,Y_n) \Rightarrow (X,a)$

Thm2 (limite central) : Soit $(X_n)_n$ une suite de v.a.r indépendantes, de même loi et possédant un moment d'ordre 2. On note : $\mu = \mathbb{E}[X_1]$, $\sigma^2 = Var(X_1)$ et $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Alors :

$$\sqrt{n}(\overline{X}_n - \mu) \underset{n\to\infty}{\Longrightarrow} \mathcal{N}(0,\sigma^2) \qquad \boxed{\text{DEV}}$$

Applic° 3 : Calcul d'intervalle de confiance asymptotique (ICA)

48 Etude de la loi Γ

Références Cottrell-Duhamel, Chanabol-Ruchs

Soit $\lambda > 0$ et $a > 0$. La loi $\Gamma(a, \lambda)$ est une loi dont la densité est

$$f : x \mapsto \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbb{1}_{x \geq 0}(x)$$

Soit X une variable aléatoire dont f est la densité.

⋆ **Espérance** : Par un changement de variables $x = \lambda y$, on a

$$\mathbb{E}(X) = \frac{\lambda^a}{\Gamma(a)} \int_0^{+\infty} e^{-\lambda x} x^a dx = \frac{\lambda^a}{\lambda\Gamma(a)} \int_0^{+\infty} e^{-y} \left(\frac{y}{\lambda}\right)^a dy = \frac{\lambda^a}{\Gamma(a)} \frac{\Gamma(a+1)}{\lambda^{a+1}} = \frac{a}{\lambda}$$

⋆ **Variance** : Par le même changement de variables, on a

$$\mathbb{E}(X^2) = \frac{\lambda^a}{\lambda^{a+2}\Gamma(a)} \int_0^{\infty} e^{-y} y^{a+1} dy = \frac{\Gamma(a+2)}{\lambda^2\Gamma(a)} = \frac{a(a+1)}{\lambda^2}$$

Ainsi, $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X) = \frac{a}{\lambda^2}$.

⋆ **Transformée de Laplace** : L'intégrale $\int_0^{+\infty} e^{-(\lambda-t)x} x^{a-1} dx$ est convergente si et seulement si $\lambda > t$. Dans ce cas, par le changement de variable $y = (\lambda - t)x$, on a

$$\mathbb{E}(e^{tX}) = \frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} e^{-(\lambda-t)x} x^{a-1} dx = \frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} e^{-y} \frac{y^{a-1}}{(\lambda - t)^a} dy = \frac{\lambda^a}{(\lambda - t)} \frac{\Gamma(a)}{\Gamma(a)} = \left(\frac{\lambda}{\lambda - t}\right)^a$$

⋆ **Fonction caractéristique** : On pose $\phi(t) = \mathbb{E}(e^{itX}) = \frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} e^{-(\lambda-it)x} x^{a-1} dx$.

Soit $D = \{z \in \mathbb{C} \mid \mathrm{Re}(z) < \lambda\}$. Pour tout $z \in D$, on pose:

$$\psi(z) = \frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} e^{-(\lambda-z)x} x^{a-1} dx$$

La fonction ψ est bien définie sur D car pour tout $z \in D$, on a:

$$|e^{-(\lambda-z)x} x^{a-1}| \leq x^{a-1} e^{-(\lambda - \mathrm{Re}(z))x}$$

où la fonction majorante est intégrable sur $\mathbb{R}^+$.

Soit $z \in D \cap \mathbb{R}$. On a

$$\frac{\lambda^a}{\Gamma(a)} \int_0^{+\infty} e^{-(\lambda-z)x} x^{a-1} dx = \frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} e^{-y} \frac{y^{a-1}}{(\lambda - z)^a} dy = \frac{\lambda^a}{(\lambda - z)^a} dy = \frac{\lambda^a}{(\lambda - z)^a} \frac{\Gamma(a)}{\Gamma(a)} = \left(\frac{\lambda}{\lambda - z}\right)^a$$

Soit $z \in D$. On pose $h : z \mapsto \frac{\lambda^a}{(\lambda - z)^a}$. Si log désigne la détermination principale du logarithme, on a $h(z) = \lambda^a \exp(a \log(\lambda - z))$. Il vient que h est holomorphe sur D.

Assertion La fonction ψ est holomorphe sur D. Soit $K \subset\subset D$ un compact. Il existe $\epsilon > 0$ tel que $K \subset D_\epsilon := \{z \in \mathbb{C} \mid \mathrm{Re}(z) < \lambda - \epsilon\}$. Posons

$$g(x, z) = e^{-(\lambda-z)x} x^{a-1}$$

C'est une fonction $\mathcal{C}^1$ et

$$\left|\frac{\partial g}{\partial z}(x, z)\right| = |x^{a-1} e^{-(\lambda-z)x}| \leq x^a e^{-\epsilon x}$$

qui est intégrable sur $\mathbb{R}^+$ et indépendante de z. Par le théorème d'holomorphie des intégrales à paramètres, il vient que ψ est holomorphe sur tout compact de D donc sur D. Par le théorème des zéros isolés, on a $h = \psi$ sur D. Donc finalement,

$$\phi(t) = \psi(it) = h(it) = \frac{\lambda^a}{(\lambda - it)^a}$$

⋆ **Somme de v.a. exponentielle indépendantes** Soit $X_1, \dots, X_n$ des variables indépendantes exponentielle de même loi $\mathcal{E}(\lambda)$. Alors $S_n = X_1 + \cdots + X_n$ suit une loi $\Gamma(n, \lambda)$ (dite loi d'Erlang). En effet, on calcule la transformée de Laplace de S_n. On a par indépendance

$$L_{S_n}(t) = \mathbb{E}(e^{tS_n}) = \prod_{k=1}^{n} \mathbb{E}(e^{tX_i}) = \mathbb{E}(e^{tX_1})^n = \left(\frac{\lambda}{\lambda - t}\right)^n$$

pour $\lambda > t$.

Utilité: définir des processus de Poisson. On pose $(T_k)_k$ une suite de variables al"attires indépendantes identiquement distribuées selon $\mathcal{E}(\lambda)$. Alors $\forall \geq 0, N(t) = \sum_{n=1}^{+\infty} \mathbb{1}_{S_n \leq t}$ définit un processus de Poisson.

⋆ **Loi du χ^2** Soit $Y \sim \mathcal{N}(0,1)$. Alors Y^2 suit la loi $\Gamma(1/2, 1/2)$ dite $\chi^2(1)$.

En effet, si $z \leq 0$, on a $(Y^2 < z) = 0$. Si $z > 0$, on a

$$\mathbb{P}(Y^2 < z) = \mathbb{P}(-\sqrt{z} < Y < \sqrt{z}) = \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2) = 2\int_0^{\sqrt{z}} \frac{1}{2\pi} \exp(-\frac{1}{2}u^2) du$$

On dérive ensuite cette application continue et de classe $\mathcal{C}^1$ par morceaux pour voir que Y^2 admet la densité

$$\frac{1}{\sqrt{2\pi}} z^{-1/2} e^{-z/2} \mathbb{1}_{]0,+\infty[}(z)$$

On reconnait la loi $\Gamma(1/2, 1/2)$. En utilisant le fait que la $\Gamma(a+b, \lambda) = \Gamma(a, \lambda) + \Gamma(b, \lambda)$, on prouve de même que si $Y_1, \dots, Y_n$ sont indépendantes de même lois normales $\mathcal{N}(0,1)$, alors $Z := \sum_{i=1}^{n} Y_i^2$ a la loi $\Gamma(n/2, 1/2)$.

Utilité: tests statistiques du χ^2 .

Application: Détermination d'un intervalle de confiance exact On a besoin tout d'abord du lemme suivant:

Lemme 48.1

Soit $Y \sim \Gamma(n, \gamma)$. Alors $\gamma Y \sim \Gamma(n, 1)$.

Démonstration Soit $h \in \mathcal{C}_c^0(\mathbb{R})$. On calcule $\mathbb{E}(h(\gamma Y)) = \int_{\mathbb{R}} h(\gamma y) \frac{\gamma^n}{(n-1)!} y^{n-1} e^{-\gamma y} dy = \int_{\mathbb{R}} h(x) \frac{x^{n-1}}{(n-1)!} e^{-x} dx$ en effectuant le changement de variables $\gamma y = x$. ■

Soit $(X_1, \dots, X_n)$ un n échantillon de loi $\mathcal{E}(\lambda)$ avec $\lambda > 0$ inconnu. On souhaite avoir un intervalle de confiance pour λ. Une étude du maximum de vraisemblance fournit que $\widehat{\lambda} := \frac{1}{\bar{X}_n}$ est un estimateur fortement consistant de λ. Par ailleurs, $n\bar{X} = \sum_{i=1}^{n} X_i$ est la somme de n variables exponentielles de même paramètres donc une variable de loi $\Gamma(n, \lambda)$. Par le lemme, $\lambda n \bar{X} \sim \Gamma(n, 1)$. A nouveau, $2n\lambda\bar{X} \sim \Gamma(n, 1/2) = \chi^2(2n)$. Par conséquent

$$\mathbb{P}(2\lambda\bar{X} \in [\chi_{\alpha/2}(2n), \chi_{1-\alpha/2}(2n)]) = 1 - \alpha$$

donc on obtient bien un intervalle de confiance exact pour λ de niveau $1 - \alpha$.

26 Théorème limite central

Références Zuily-Queffelec, *Eléments d'analyse pour l'agrégation* et Cadre-Vial, *Statistiques.*

C'est un théorème clé en théorie des probabilités. Il souligne le rôle central des variables gaussiennes, qui peuvent être vues comme le comportement global d'une multitude de petits phénomènes. En pratique, quand $n \geq 30$, le TLC fournit une bonne approximation de la situation. Notons qu'on peut estimer son erreur grâce à l'inégalité de Berry-Esseen.

Théorème 26.1 (***Théorème limite central***)

Soit $(X_n)_n$ une suite de variables aléatoires i.i.d.. Notons $\overline{X_n}$ sa moyenne empirique (i.e. moyenne de Cesàro). Si $X_n \in L^2$, en notant σ^2 la variance de X_1, on a:

$$\sqrt{n}\left(\overline{X_n} - \mathbb{E}(X_1)\right) \Rightarrow \mathcal{N}(0, \sigma^2)$$

Démonstration On note $m = \mathbb{E}(X_1)$. Quitte à considérer les variables $Y_n := \frac{X_n - m}{\sigma}$, on peut supposer $m = 0$ et $\sigma = 1$. On cherche donc à montrer que $\sqrt{n}\overline{X_n} = \frac{\sum_{k=1}^n X_k}{\sqrt{n}}$ converge en loi vers $\mathcal{N}(0,1)$. On note $S_n = \sum_{k=1}^n X_k$. On rappelle que si $Z \sim \mathcal{N}(0,1)$, on a $\varphi_Z(t) = e^{-t^2/2}$. Ce résultat découle de la résolution d'une équation différentielle sur l'intégrale à paramètre $\varphi_Z(t)$. On va pour cela utiliser le théorème de Lévy, i.e. prouver que

$$\forall t \in \mathbb{R}, \lim_{n \to +\infty} \varphi_{S_n/\sqrt{n}}(t) = e^{-t^2/2}$$

D'une part, par indépendance et identique distribution des X_n, on a:

$$\forall t \in \mathbb{R}, \varphi_{S_n/\sqrt{n}}(t) = \varphi_{S_n}\left(\frac{t}{\sqrt{n}}\right) = \left(\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

D'autre part, on effectue un développement limité à l'ordre deux en l'origine:

$$\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right) = \varphi_{X_1}(0) + \frac{t}{\sqrt{n}}\varphi'_{X_1}(0) + \frac{t^2}{2n}\varphi''_{X_1}(0) + o\left(\frac{1}{n}\right)$$

Ce développement limité est licite car $X_1 \in L^2$ et donc $\varphi \in \mathcal{C}^2$ et on connait ses dérivées en fonctions des moments de X_1, à savoir:

$$\varphi'_{X_1}(0) = im = 0 \ , \ \varphi''_{X_1}(0) = -\sigma^2 = -1$$

On en déduit que:

$$\varphi_{X_1}\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)$$

On en déduit que:

$$\varphi_{S_n/\sqrt{n}}(t) = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

Pour conclure, on va utiliser le lemme suivant.

Lemme 26.1

Pour toute suite de complexes $(z_n)_n$ convergeant vers $z \in \mathbb{C}$, on a

$$\lim_{n \to +\infty} \left(1 + \frac{z_n}{n}\right)^n = e^z$$

Démonstration Par la formule du binôme de Newton, on a

$$e^{z_n} - \left(1 + \frac{z_n}{n}\right)^n = \sum_{k=0}^{+\infty} \frac{z_n^k}{k!} - \sum_{k=0}^{+\infty} \binom{n}{k} \frac{z_n^k}{n^k}$$

De plus, ont l'égalité:

$$\frac{1}{n^k}\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!n^k} = \frac{1}{k!}\prod_{j=0}^{k-1}\left(\frac{n-j}{n}\right) = \frac{1}{k!}\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right)$$

Ainsi, on a

$$e^{z_n} - \left(1+\frac{z_n}{n}\right)^n = \sum_{k=n+1}^{+\infty}\frac{z_n^k}{k!} + \sum_{k=0}^{n}\frac{z_n^k}{k!}\left(\underbrace{1-\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right)}_{\geq 0}\right)$$

D'où par passage au module:

$$\begin{aligned}
\left|e^{z_n} - \left(1+\frac{z_n}{n}\right)^n\right| &\leq \sum_{k=n+1}^{+\infty}\frac{|z_n^k|}{k!} + \sum_{k=0}^{n}\frac{|z_n|^k}{k!}\left(1-\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right)\right) \\
&\leq {}^{|z_n|} - \left(1+\frac{|z_n|}{n}\right)^n \\
&\leq e^{|z_n|} - e^{n\ln(1+|z_n|/n)} \\
&\leq e^{|z_n|} - e^{n(|z_n|/n-|z_n|^2/2n^2)} \\
&\leq e^{|z_n|}\left(1-e^{-|z_n|^2/2n}\right) \\
&\leq \frac{|z_n|^2}{2n}e^{|z_n|}
\end{aligned}$$

Les inégalités annoncées découlent d'études de fonctions.

Cela permet de conclure car:

$$e^z - \left(1+\frac{z_n}{n}\right)\Big| \leq |e^z - e^{z_n}| + \left|e^{z_n} - \left(1+\frac{z_n}{n}\right)\right| \leq |e^z - e^{z_n}| + \frac{|z_n|^2}{2n}e^{|z_n|}$$

Le terme de droite tend vers 0 par hypothèse $z_n \to z$, et donc la suite $|z_n|$ est bornée. ■

On applique le lemme. On a donc

$$\lim_{n\to+\infty}\varphi_{S_n/\sqrt{n}}(t) = \left(1-\frac{t^2}{2n}+o\left(\frac{1}{n}\right)\right)^n = e^{-t^2/2}$$

On reconnait la la fonction caractéristique d'une loi normale centrée-réduite. ■

Application: calcul d'un intervalle de confiance asymptotique Soit $(X_1,\dots,X_n)$ un n-échantillon de même loi dont la densité est donnée par

$$f : x \mapsto \ln\theta x\theta^{-x^2/2}\mathbb{1}_{\mathbb{R}^+}(x)$$

On cherche à donner un intervalle de confiance au niveau $1-\alpha$ pour θ.

- On calcule l'estimateur du maximum de vraisemblance:

$$\widehat{\theta} = \exp\left(\frac{2n}{\sum_{i=1}^n X_i}\right)$$

- On calcule les moments d'ordre 2 et d'ordre 4 de X_1: on trouve, par intégration par parties:

$$\mathbb{E}(X_1^2) = \frac{2}{\ln\theta} \text{ et } \mathbb{E}(X_1^4) = \frac{8}{\ln^2\theta}$$

  Et donc $\mathrm{Var}(X_1^2) = \frac{4}{(\ln\theta)^2}$.

- Par la loi des grands nombres, on a $\widehat{\theta} = g\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right) \rightarrow_{p.s} g(2/\ln\theta) = \theta$ où $g : x \mapsto e^{2/x}$.
- On applique le théorème central limite pour $(X_i^2)_i$ i.i.d.

$$\sqrt{n}\left(\frac{\sum_{i=1}^n X_i^2}{n} - \frac{2}{\ln\theta}\right) \Rightarrow \mathcal{N}\left(0, \frac{4}{\ln^2\theta}\right)$$

- Comme $g \in C^1(\mathbb{R}^{*+})$ on en déduit par la méthode δ que

$$\sqrt{n}(\widehat{\theta} - \theta)\mathcal{N}(0, (\theta\ln\theta)^2)$$

- Par le lemme de Slutsky, on a

$$\sqrt{n}\frac{(\widehat{\theta} - \theta)}{\widehat{\theta}\ln\widehat{\theta}} \Rightarrow \mathcal{N}(0,1)$$

  Par suite,

$$\left[\widehat{\theta} \pm \frac{q_{1-\alpha/2}\widehat{\theta}\ln\widehat{\theta}}{\sqrt{n}}\right]$$

  est un intervalle de confiance asymptotique de niveau $1-\alpha$ pour θ (où $q_{1-\alpha/2}$ désigne le quantile d'ordre $1-\alpha/2$ de la loi normale centrée réduite)