242 UTILISATION EN PROBABILITES DE LA TRANSFORMEE DE FOURIER OU DE LAPLACE ET DU PRODUIT DE CONVOLUTION

Introduction : Les transformées de Fourier et de Laplace étendent les propriétés de fonctions génératrices, définies pour des variables aléatoires (v.a) à valeurs entières positives, au cas des v.a. réelles (v.a.r) quelconques, à savoir : la caractérisation de la loi, de l'indépendance, de la convergence en loi et le calcul des moments.

I°) CARACTERISATION DE LA LOI D'UN VECTEUR ALEATOIRE

1.1. Transformée de Fourier

Déf : Soit X un $\vec{v.a}$ à valeurs dans $\mathbb{R}^d$, de loi $\mathbb{P}_X$. La fonction caractéristique φ_X de X (appelée aussi transformée de Fourier), est la transformée de Fourier $\widehat{\mathbb{P}_X}$ de sa loi, définie sur $\mathbb{R}^d$ par :

$$\boxed{\forall t \in \mathbb{R}^d,\ \varphi_X(t) = \mathbb{E}[e^{i\langle t,X\rangle}] = \widehat{\mathbb{P}_X}(t)}$$ [BL p 65]

Thm d'injectivité : Si X et Y sont deux $\vec{v.a}$ de lois $\mathbb{P}_X$ et $\mathbb{P}_Y$ tels que $\varphi_X = \varphi_Y$ alors $\mathbb{P}_X = \mathbb{P}_Y$. [BL p. 66]

Puisque la transformée de Fourier caractérise la loi, il est souhaitable d'avoir une formule d'inversion permettant d'obtenir effectivement la loi à partir de la fonction caractéristique :

Thm (formule d'inversion de Fourier) : Soit φ_X une fonction caractéristique intégrable d'un $\vec{v.a}$ X. Alors X admet une densité continue bornée f_X donnée par :

$$\boxed{\forall x \in \mathbb{R}^d,\ f_X(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle t,x\rangle}\varphi_X(t)\,dt}$$

[BL. p 69]

→ application : des fonctions caractéristiques peuvent coïncider sur un intervalle sans être égales.

Ex de Stoyanov : Soit $\Phi : \mathbb{R} \longrightarrow [0,1]$, $t \longmapsto \begin{cases} 1-|t| & \text{si } 0 \le |t| \le 1 \\ 0 & \text{sinon} \end{cases}$

Soit X une v.a de densité Φ.

On a $\Phi = \hat{\mu}$ où $\mu = f.\lambda$ avec $f(x) = \frac{1}{\sqrt{2\pi}}\varphi_X(x)$

Soient Y et Z 2 v.a telles que Y soit de densité f et

$$\mathbb{P}_Z = \frac{1}{2}\delta_0 + \sum_{k\in\mathbb{Z}} \frac{2}{(2k-1)^2\pi^2}\delta_{(2k-1)\pi}$$

Alors $\varphi_Y(t) = \varphi_Z(t)$, $\forall t \in [-1;1]$ et pourtant Y et Z ne sont pas égales. [OuV p. 212]

Thm (calcul des moments) : Soit X une v.a.r. On suppose que X admet un moment d'ordre $n \in \mathbb{N}^*$ (i.e $\mathbb{E}(|X|^n) < \infty$), alors φ_X est n fois dérivable, de dérivée k-ième ($k \le n$) :

$$\varphi^{(k)}(t) = i^k \int x^k e^{itx}\,d\mathbb{P}_X(x) = i^k\,\mathbb{E}(X^k e^{itX})$$

En particulier, $\varphi^{(k)}(0) = i^k\,\mathbb{E}(X^k)$ [BL. p 68]

NB : Ce théorème se généralise au cas de $\vec{v.a}$.

⚠ La réciproque est fausse :

contre-ex : Soit $\mathbb{P}$ la mesure de probabilité sur $\mathbb{Z}$ définie par :

$$\mathbb{P} = \sum_{n\ge 2} \frac{c}{n^2\ln n}(\delta_n + \delta_{-n})$$ (où c est choisie de façon à ce que $\mathbb{P}$ soit une probabilité).

$\mathbb{P}$ n'admet pas de moment d'ordre 1 mais sa fonction caractéristique est dérivable en 0.

On a néanmoins la réciproque si φ est $2n$-fois dérivable en 0. [BL p. 68]

Quelques exemples de fonctions caractéristiques : [BL p. 66-67]

1. Si $X \sim \mathcal{B}(n,p)$, $\varphi_X(t) = (pe^{it} + 1 - p)^n$.
2. Si $X \sim \mathcal{P}(\lambda)$, $\varphi_X(t) = \exp(\lambda(e^{it}-1))$
3. Si $X \sim \mathcal{U}(a,b)$, $\varphi_X(t) = \dfrac{e^{itb} - e^{ita}}{it(b-a)}$
4. Si $X \sim \mathcal{C}(a)$ (i.e $f_X(x) = \frac{a}{\pi(a^2+x^2)}$), $\varphi_X(t) = e^{-a|t|}$
5. Si $X \sim \mathcal{N}(m,\sigma^2)$, $\varphi_X(t) = \exp(itm - \frac{1}{2}\sigma^2t^2)$
6. Si $X \sim \mathcal{N}_d(m,\Sigma^2)$, $\varphi_X(t) = e^{i\langle t,m\rangle - \frac{1}{2}t'\Sigma^2 t}$ où t' désigne la transposée de t.

1.2. Transformée de Laplace

Déf : Soit X un $\vec{v.a}$ à valeurs dans $\mathbb{R}^d$. On appelle transformée de Laplace (ou fonction génératrice des moments) la fonction :

$$\boxed{L_X(t) = \mathbb{E}[e^{\langle t,X\rangle}]}$$ définie pour les valeurs de t pour lesquelles $e^{\langle t,X\rangle}$ est intégrable. [BL. p 70]

Thm d'unicité : Deux $\vec{v.a}$ ayant la même transformée de Laplace sur un voisinage de 0 ont même loi. [BL p. 70]

Prop : Soit X une v.a.r telle que e^{tX} est intégrable pour t dans un intervalle contenant 0. Alors la transformée de Laplace L_X est définie sur un intervalle contenant 0. De plus, elle est analytique sur un voisinage de 0 et :

$$L_X(t) = \sum_{n\in\mathbb{N}} \frac{t^n}{n!}\mathbb{E}[X^n]$$ pour tout t dans ce voisinage.

En particulier :

$$\boxed{(L_X)^{(n)}(0) = \mathbb{E}[X^n]}$$ [BL p. 70-71]

→ application : caractérisation de la loi de v.a.r bornées.

Thm des moments : Soient X, Y deux v.a.r à valeurs dans un intervalle borné $[a,b]$. Si $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ pour tout $k \in \mathbb{N}$, alors X et Y ont même loi. [BL. p 70]

⚠ En général, une loi n'est pas caractérisée par ses moments.

Contre-ex : Soient $X \sim \mathcal{N}(0,1)$, $Z = e^X$ et Y v.a.r à densité.

$$f_Y(t) = f_Z(t)(1 + \sin(2\pi \ln t)).$$

$\forall t > 0$, Y et Z ont mêmes moments mais pas la même loi. [BL p. 73]

Inégalité de Chernoff : soient $X_1, \dots, X_n$ des v.a indépendantes, de même loi μ telles que $L\mu(t)$ existe dans un voisinage de 0.
On pose $H(r) = \sup_\lambda (\lambda(r+m) - \ln L\mu(\lambda))$: on l'appelle la transformation de Cramer.
Alors : $\mathbb{P}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - m \geq r\right) \leq \exp(-H(r))$

II) TRANSFORMEE DE FOURIER, PRODUIT DE CONVOLUTION ET INDEPENDANCE

2.1. caractérisation de l'indépendance.

Prop : une famille $(X_1, \dots, X_n)$ de v.a.r est indépendante (ssi) $\forall (t_1, \dots, t_n) \in \mathbb{R}^n$, $\varphi_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \prod_{i=1}^{n} \varphi_{X_i}(t_i)$ [BL p.83]

→ application aux v.a gaussiennes

Prop : Soit $X = (X_1, \dots, X_d)$ un vecteur gaussien à valeurs dans $\mathbb{R}^d$. Les v.a.r X_i sont indépendantes (ssi) elles sont non corrélées 2 à 2 (ce qui est équivalent à dire que la matrice de covariance est diagonale) [BL p.107]

⚠ Il faut supposer le vecteur X gaussien et non pas seulement chacune de ses variables aléatoires X_i gaussiennes.

Contre-ex : soient $X \sim \mathcal{N}(0,1)$, $\varepsilon \sim \frac{1}{2}(\delta_1 + \delta_{-1})$ avec X indépendant de ε. et $Z = \varepsilon X$. X et Z sont non corrélées, gaussiennes mais ne sont clairement pas indépendantes. [BL p112]

2.2. Loi et transformée de Fourier de la somme de deux v.a.r indépendantes

Déf : soient μ et ν deux probabilités sur $\mathbb{R}^d$. On appelle produit de convolution de μ et ν, et on note $\mu * \nu$, la probabilité sur $\mathbb{R}^d$ définie par : $(\mu * \nu)(B) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \mathbb{1}_B(x+y)\, \mu(dx)\nu(dy)$ pour tout borélien B de $\mathbb{R}^d$. [Sato p.8]

Prop : La loi de la somme de deux v.a.r indépendantes X et Y est le produit de convolution $\mathbb{P}_X * \mathbb{P}_Y$ de leur loi. [BL p.89]

Prop : si X admet une densité f_X et Y est une v.a.r indépendante de X, la somme X+Y admet une densité f_{X+Y} donnée par :
$$f_{X+Y}(t) = \int_{\mathbb{R}} f_X(t-y)\, \mathbb{P}_Y(dy) \quad \text{pour presque tout } t \in \mathbb{R}$$
[Brémaud p.123]

NB : En particulier, si X et Y sont des v.a.r à densités indépendantes alors la densité de X+Y est le produit de convolution usuel de leur densité. [Brémaud p.123]

Prop : soient X et Y deux v.a.r ou vectorielles indépendantes. Alors $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$, $\forall t \in \mathbb{R}$. [BL p.91]

NB : on a le résultat analogue pour la transformée de Laplace.
⚠ La réciproque est fausse.

Contre-ex : soit X une v.a suivant une loi de Cauchy $\mathcal{C}(1)$. Alors $\varphi_X(t) = e^{-|t|}$ $\forall t \in \mathbb{R}$ donc $\varphi_{X+X} = (\varphi_X)^2$ et pourtant X n'est pas indépendante d'elle-même (car les seules v.a indépendantes d'elles-mêmes sont les v.a constantes presque sûrement.

2.3. Lois infiniment divisibles (Lois I.D) [Sato p.31-39]

Déf : Soit X un v.a de loi μ. On dit que μ est infiniment divisible si : $\forall n$, $\exists \mu_n$ mesure de probabilité telle que $\mu = \mu_n^n$ où on a noté $\mu_n^n := \underbrace{\mu_n * \dots * \mu_n}_{n \text{ fois}} = \mu^{*n}$
Ceci est équivalent à dire que : $\forall n \in \mathbb{N}^*$, $\exists X_1, \dots, X_n$ i.i.d telles que $X \overset{\mathcal{L}}{=} X_1 + \dots + X_n$.

Il est clair que μ est I.D (ssi) sa fonction caractéristique est, pour tout entier $n \geq 1$, la puissance n-ième d'une fonction caractéristique

Ex : ① $\mathcal{N}(m, \sigma^2) = \mathcal{N}\left(\frac{m}{n}, \frac{\sigma^2}{n}\right)^{*n}$ est I.D

② $\mathcal{P}(\lambda) = \mathcal{P}\left(\frac{\lambda}{n}\right)^{*n}$ est I.D

③ $\Gamma(p, \lambda) = \Gamma\left(\frac{p}{n}, \lambda\right)^{*n}$ est I.D

④ $\mathcal{C}(a) = \left(\mathcal{C}\left(\frac{a}{n}\right)\right)^{*n}$ est I.D

Prop : Les probabilités à support fini, autres que les Dirac, ne peuvent pas être I.D

Ex : La loi binomiale n'est pas I.D

Prop : si μ est I.D, alors $\hat{\mu}$ ne s'annule jamais.
⚠ Réciproque fausse : prendre la loi binomiale.

Thm : Soit $(\mu_n)_n$ une suite de lois I.D telle que $\mu_n \xrightarrow[n\to\infty]{} \mu$. Alors μ est infiniment divisible

Thm de Représentation de Lévy-Khintchine

(i) Si μ est une loi I.D sur $\mathbb{R}^d$, alors sa fonction caractéristique vérifie :
$$\hat{\mu}(\xi) = \exp\left[-\frac{1}{2}\langle \xi, A\xi\rangle + i\langle \gamma, \xi\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle \xi, x\rangle} - 1 - i\langle \xi, x\rangle \mathbb{1}_{\{|x|<1\}}(x)\right)\nu(dx)\right] \quad \forall \xi \in \mathbb{R}^d$$
où . A est une matrice $d \times d$ symétrique positive
. $\gamma \in \mathbb{R}^d$
. ν est une mesure sur $\mathbb{R}^d$ telle que :
$$\nu(\{0\}) = 0 \text{ et } \int_{\mathbb{R}^d} (|x|^2 \wedge 1)\,\nu(dx) < \infty$$

(ii) La représentation de μ donnée en (i) par A, ν et γ est unique.

La réciproque est vraie.

NB : si ν satisfait de plus la condition : $\int_{|x| \leq 1} |x|\,\nu(dx) < \infty$, alors :

$\hat{\mu}(\xi) = \exp\left(-\frac{1}{2}\langle \xi, A\xi\rangle + i\langle \gamma, \xi\rangle + \int_{\mathbb{R}^d} (e^{i\langle \xi, x\rangle} - 1)\nu(dx)\right)$

Ex : pour une loi de Poisson $\mathcal{P}(\lambda)$, $d=1$, $A=0$, $\gamma=0$ et $\nu = \lambda\delta_1$

2.4. Variables aléatoires gaussiennes et indépendance.

Thm : Soient X et Y deux v.a.r indépendantes de même loi et de carré intégrable telles que les v.a $X+Y$ et $X-Y$ soient indépendantes. Alors X et Y sont deux v.a gaussiennes. [REN p 301-305] [OUVT]

NB : Ce thm est encore vrai sans supposer X et Y équidistribuées ni de carré intégrable (→thm de Bernstein) [Ouv. p 271]

⚠ La réciproque est fausse sauf si $X \sim \mathcal{N}(m, \sigma^2)$ et $Y \sim \mathcal{N}(m', \sigma^2)$

III) ETUDE DU COMPORTEMENT ASYMPTOTIQUE

3.1. Lien entre convergence en loi et convergence des fonctions caractéristiques

Thm : Soit $(X_n)_n$ une suite de v.a à valeurs dans $\mathbb{R}^d$. Alors la suite $(X_n)_n$ converge en loi vers X (ssi) $\varphi_{X_n}(t) \xrightarrow[n\to\infty]{} \varphi_X(t)$, $\forall t \in \mathbb{R}$ [BL. p129]

NB : Il ne faut pas nécessairement que la limite soit une fonction caractéristique, il suffit qu'elle soit continue en 0 :

Thm de Paul-Lévy : si $(\varphi_{X_n})_n$ est une suite de fonctions caractéristiques qui converge simplement vers une fonction g continue en 0, alors g est la fonction caractéristique d'une v.a X et $X_n \xrightarrow{loi} X$ [COT p 106]

NB : on ne peut pas se passer de l'hypothèse de continuité en 0.

contre-ex : Soit pour tout $n \geq 1$, $X_n \sim \mathcal{N}(0, n^2)$. Alors $\varphi_{X_n}(t) = e^{-\frac{n^2t^2}{2}}$ qui converge simplement sur $\mathbb{R}$ vers $\varphi = \mathbb{1}_{\{0\}}$ qui n'est pas continue en 0 et n'est donc pas une fonction caractéristique. Donc $(X_n)_{n\geq 1}$ ne converge pas en loi.

3.2. Théorèmes limites

Thm (loi faible des grands nombres) : Soit $(X_n)_n$ une suite de v.a.r i.i.d intégrables. On pose $S_n = X_1 + \dots + X_n$.
Alors : $\frac{S_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}[X_1]$ [B.L p140]

Théorème central limite : Soit $(X_n)_n$ une suite de v.a indépendantes, de carré intégrable, équidistribuées et à valeurs dans $\mathbb{R}$. On pose $\sigma^2 = Var(X)$ et $m = \mathbb{E}[X]$.
Alors $\sqrt{n}\left(\frac{S_n}{n} - m\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2)$ [BL p.145] [OUVT]

NB : Ce théorème se généralise à $\mathbb{R}^d$ en remplaçant σ^2 par C, la matrice de covariance de X. (TCL multidimensionnel)

→ application 1 : approximation de lois
- $\mathcal{P}(n)$ peut être approchée par une loi normale $\mathcal{N}(n, n)$
- $\Gamma(n, \lambda)$, en tant que somme de v.a de lois $\mathcal{E}(\lambda)$, peut être approchée par une loi $\mathcal{N}(\frac{n}{\lambda}, \frac{n}{\lambda^2})$

→ application 2 : intervalles de confiance. [Rivoirard].
Soit un modèle $(\mathcal{X}^n, (P_\theta)_{\theta\in\Theta})$

Principe : on a construit un estimateur de θ, ce qui nous permet d'estimer des quantités mais on aimerait savoir avec quelle précision. On souhaiterait notamment donner un intervalle autour de notre estimation tel qu'on soit sûr à 95 % que la vraie valeur soit dedans.

Déf : un intervalle de confiance pour le paramètre θ, de niveau de confiance $1-\alpha \in]0;1[$ est un intervalle qui a la probabilité $1-\alpha$ de contenir la vraie valeur du paramètre θ.

Ex : Soit $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (moyenne empirique) estimateur de la moyenne.

Si on cherche à localiser la moyenne m de la loi d'un échantillon $(X_1, \dots, X_n)$, on établit avec probabilité $1-\alpha$ que m se situe dans l'intervalle $[\overline{X_n} - a\frac{\sigma}{\sqrt{n}} ; \overline{X_n} + a\frac{\sigma}{\sqrt{n}}]$ où a est tel que $\mathbb{P}(|Z| \leq a) = 1-\alpha$ avec $Z \sim \mathcal{N}(0,1)$

NB : L'inégalité de Chernoff permet également de construire des intervalles de confiance pour m

Thm de Poisson : Soit pour tout $n \in \mathbb{N}^*$ une v.a de loi binomiale $\mathcal{B}(n, p_n)$. On suppose que $\lim_{n\to\infty} np_n = \lambda$, où $\lambda > 0$. Alors la suite $(X_n)_n$ converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$. [Ouv. p 310]

On peut généraliser ce thm :

Thm des évènements rares : Soit, pour tout $n \in \mathbb{N}^*$, une famille finie $\{A_{n,j} \mid 1 \leq j \leq M_n\}$ d'évènements indépendants. On pose $\mathbb{P}(A_{n,j}) = p_{n,j}$ et on note : $S_n = \sum_{j=1}^{M_n} \mathbb{1}_{A_{n,j}}$

On suppose que la suite de terme général M_n tend en croissant vers $+\infty$ et que :
$\max_{1\leq j\leq M_n} p_{n,j} \xrightarrow[n\to+\infty]{} 0$ et $\sum_{j=1}^{M_n} p_{n,j} \xrightarrow[n\to+\infty]{} \lambda$

Alors la suite $(S_n)_{n\in\mathbb{N}^*}$ converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$. [Ouv. p311]

Ref :
- Barbe-Ledoux
- Ouvrard 2
- Rényi, "calcul des probabilités"
- Sato "Lévy processes [...]"
- Rivoirard