Fragment 9

Quelques lois classiques

On renvoie principalement à [Bor95], [Rud75], [Bou86] et [DCD82b]. Dans toute la suite, on note $\mathcal{L}(X)$ la loi de la v.a. X et $*$ la convolution des fonctions ou des lois de probabilité, définies par :

$$(f * g)(x) := \int f(x-y)g(y)\,dy = \int f(z)g(x-z)\,dz = (g * f)(x),$$

et

$$\mathbf{E}_{\mu*\nu}(\varphi) := \int \varphi(x+y)\,d(\mu\otimes\nu)(x,y) = \int \varphi(x+y)\,d(\nu\otimes\mu)(x,y) = \mathbf{E}_{\nu*\mu}(\varphi)\,.$$

Si μ et ν ont pour densités respectives f et g par rapport à la mesure de Lebesgue, alors $\mu * \nu$ a pour densité $f * g$. On rappelle que la convolée $\mu * \nu$ de deux mesures de probabilité représente la loi de la somme de deux v.a. indépendantes X, Y de loi respectives μ et ν :

$$X, Y \text{ indépendantes} \Rightarrow \mathcal{L}(X+Y) = \mathcal{L}(X) * \mathcal{L}(Y).$$

La réciproque est fausse en général. On note μ^{*n} la loi obtenue en convolant n fois μ :

$$\mu^{*n} := \underbrace{\mu * \cdots * \mu}_{n \text{ fois.}}.$$

D'autre part :

$$X, Y \text{ indépendantes} \Leftrightarrow \mathcal{L}((X,Y)) = \mathcal{L}(X) \otimes \mathcal{L}(Y).$$

On note δ_a la masse de Dirac en a, c'est à dire la mesure de probabilité qui affecte la masse 1 aux ensembles mesurables contenant le point a et 0 aux autres. C'est la loi d'une v.a. constante de valeur a. Sa moyenne vaut a et sa variance est nulle. Pour tout μ et ν, on a $\mu * \nu = \nu * \mu$ et $\mu * \delta_0 = \delta_0 * \mu = \mu$ et $\delta_a * \delta_b = \delta_{a+b}$. La masse de Dirac en 0 est donc l'élément neutre de la convolution.

Heuristique 9.0.1 *(Convolution = lissage).* Convoler une fonction f par une fonction g revient à prendre en x une moyenne des valeurs de f pondérées par la translation de g en x. Il est donc naturel que la masse de Dirac en 0 soit l'élément neutre de la convolution. Si g est lisse, on fait donc du lissage de f. Pour ce faire, on prend souvent pour g une fonction $\mathcal{C}^\infty$ paire et à décroissance rapide voire même à support compact de façon à remplacer $f(x)$ par une « g-moyenne » de f au voisinage de x. Penser à la technique de régularisation par convolution utilisée en analyse, en particulier pour montrer la densité de $\mathcal{C}_c^\infty$ dans les espaces $\mathbf{L}^p$ de Lebesgue.

Heuristique 9.0.2 *(Convolution des densités = addition d'un bruit indépendant sur la v.a.).* Soit X et Y des v.a. de densités respectives f et g. Dire que g est *lisse* signifie que la v.a. Y est purement *continue*, c'est à dire que les valeurs prises sont bien réparties et sans sauts. Ainsi, si X est indépendante de Y, la v.a. $X+Y$ aura en général des valeurs mieux réparties que celles de X lorsque g est lisse. La densité de $X+Y$ est la convolée de la densité f de la v.a. X par la densité g de la v.a. Y. Si g est lisse, il en sera de même pour $f * g$. Ainsi, lisser une densité se traduit sur les v.a. par l'addition d'une v.a. indépendante à densité lisse (penser par exemple au lissage d'une v.a. discrète X par addition d'une v.a. indépendante Y de loi uniforme où gaussienne, centrée et de faible variance par rapport à l'écart entre les valeurs discrètes prises pas X).

Remarque 9.0.3 *(Fonctions de répartition et densité).* Considérons une loi de probabilité μ sur $\mathbb{R}$, de fonction de répartition F. En tant que fonction de répartition, la fonction F est croissante, continue à droite, et $F(\mathbb{R}) = [0,1]$. On montre alors qu'elle ne peut avoir qu'un nombre au plus infini dénombrable de points de discontinuité, notés $(s_n)_n \subset \mathbb{R}$, cf. [BL98, prop. III.2.2 page 48]. On montre de plus, [BL98, III, pages 49-50], qu'il existe trois réels positifs ou nuls α, β, γ vérifiant $\alpha+\beta+\gamma = 1$, tels que F possède la décomposition suivante :

$$F = \alpha F_{\text{disc}} + \beta F_{\text{sing}} + \gamma F_{\text{cont}},$$

où les fonctions $F_\bullet$, qui vérifient les propriétés suivantes, n'existent que lorsque le coefficient associé est non nul :

1. F_{disc} est constante par morceaux, et c'est la fonction de répartition de la loi de probabilité discrète $\sum_n \delta_{s_n}$
2. F_{cont} est absolument continue et c'est la fonction de répartition d'une loi de probabilité à densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$
3. F_{sing} est continue et c'est la fonction de répartition d'une mesure de probabilité singulière par rapport à la mesure de Lebesgue sur $\mathbb{R}$

F est continue si et seulement si $\alpha = 0$ et μ admet une densité si et seulement si $\alpha = \beta = 0$. Il ne suffit donc pas que F soit continue pour que μ admette une densité par rapport à la mesure de Lebesgue. En d'autre termes, la composante étrangère à la mesure de Lebegue de μ dans sa décomposition de Lebesgue-Radon-Nikodym (cf. [Rud75, thm 6.9, page 117]) ne se réduit pas aux sauts de la fonction F. Un exemple simple est fourni par les lois fractales de la section 1.5.1 page 28. Bien entendu, si F est C^1, alors μ admet une densité qui n'est rien d'autre que F'.

Remarque 9.0.4 *(Densité au sens de la théorie des distributions).* Une loi de probabilité est, en tant que mesure de Radon, une distribution. Elle s'identifie à sa densité lorsque cette dernière existe. La fonction de répartition F_X d'une v.a.r. X existe toujours. Sa dérivée F'_X au sens des distributions a donc toujours un sens. Lorsque la loi $\mathcal{L}(X)$ admet une densité f, cette dernière n'est rien d'autre que F'_X. Les points de discontinuité de F_X se traduisent sur F'_X par des masses de Dirac au sens des distributions, tandisque les « plats » du graphe de F_X correspondent à des intervalles qui ne sont par portées par $\mathcal{L}(X)$. Il ne suffit pas que F_X soit continue pour que $\mathcal{L}(X)$ admette une densité, i.e. que F'_X soit une fonction en tant que distribution. Les F_X pour lesquelles c'est le cas sont dites « absolument continues » et s'écrivent comme l'intégrale d'une fonction $\mathbf{L}^1$ (qui est alors la densité associée). Une condition suffisante d'absolue continuité est que F_X soit $\mathcal{C}^1$.

Remarque 9.0.5. Les générateurs aléatoires de la bibliothèque Stixbox pour Matlab sont listés dans la table 1.1 page 39. Le tableau 9.1 page 145 recense quelques lois usuelles. Pour les démonstrations et des extensions, on renvoie principalement à [DCD82b] et [Bou86].

9.1 Fonctions génératrices – la transformée du pauvre

Si X est une variable aléatoire entière (i.e. à valeurs dans $\bar{\mathbb{N}} := \mathbb{N} \cup (+\infty)$), on définit sa fonction génératrice par :

$$\begin{aligned} G_X : [0,1[&\to \mathbb{R} \\ s &\mapsto G_X(s) := \mathbf{E}\big(s^X\big) = \sum_{k=0}^{+\infty} s^k\,\mathbb{P}(X=k). \end{aligned}$$

Elle ne dépend que de la loi de X, qu'elle caractérise. En d'autre termes, si X et Y sont deux v.a. entières, alors $\mathcal{L}(X) = \mathcal{L}(Y)$ si et seulement si $G_X = G_Y$. De plus, $G_{X+Y} = G_X G_Y$ lorsque X et Y sont indépendantes (la réciproque est fausse). On peut définir la fonction génératrice du couple de v.a. entières (X,Y) en $(s_1,s_2) \in [0,1[^2$ par : $G_{(X,Y)}(s_1,s_2) := \mathbf{E}\big(s_1^X s_2^Y\big)$. Les v.a. X et Y sont alors indépendantes si et seulement si $G_{(X,Y)}(s_1,s_2) = G_X(s_1)G_Y(s_2)$ pour tout $(s_1,s_2) \in [0,1[^2$.

La fonction génératrice permet également de calculer les moments puisque $\mathbb{P}(X=n) = \frac{1}{n!}\,G_X^n(0)$ et $G'_X(1) = \mathbf{E}(X)$ et plus généralement : $G_X^{(n)}(1) = \mathbf{E}(X(X-1)\cdots(X-n+1))$.

9.2 Lois discrètes à support sur entiers naturels

On rappelle que si X et Y sont deux v.a. entières (i.e. à valeurs dans $\bar{\mathbb{N}}$) indépendantes alors :

$$\mathbb{P}(X+Y=n) = \sum_{k_1+k_2=n} \mathbb{P}(X=k_1)\mathbb{P}(Y=k_2).$$

En terme de lois, cela donne :

$$(\mu * \nu) = \sum_{k_1+k_2=n} \mu(\{k_1\})\,\nu(\{k_2\})\,\delta_n.$$

9.2.1 Loi uniforme discrète

C'est la loi donnée par : $(\delta_{a_1} + \cdots + \delta_{a_n})/n$, où les a_i sont tous différents. Elle a pour moyenne $(a_1 + \cdots + a_n)/n$ et pour variance $(n\|a\|_2^2/n - (a_1 + \cdots + a_n)^2)/n^2$. On l'utilise pour modéliser un phénomène pouvant prendre n valeurs différentes $a_1, \ldots, a_n$ avec la même probabilité $1/n$. Elle a pour propriété remarquable de maximiser, à n fixé, l'entropie de Shannon $\mathbf{H}(p_1, \ldots, p_n)$ définie pour une loi de probabilité $p_1\delta_{a_1} + \cdots + p_n\delta_{a_n}$ par :

$$\mathbf{H}(p_1, \ldots, p_n) := -\sum_{k=1}^{n} p_i \log p_i.$$

9.2.2 Loi de Bernoulli

C'est la loi $q\delta_a + p\delta_b$ où $p \in [0,1]$, $q := 1-p$ et $a \neq b$. Le plus souvent, on utilise $q\delta_{-1} + p\delta_1$ ou $q\delta_0 + p\delta_1$. La loi $q\delta_0 + p\delta_1$, notée $\mathcal{B}(1,p)$, est de moyenne p et de variance pq. Elle permet de modéliser le résultat d'une expérience à deux issues possibles, comme par exemple un jet de pièce dans un jeu de pile ou face, avec probabilité p de gagner (codé 1) et probabilité $1-p$ de perdre (codé 0). Elle a pour fonction génératrice $q + ps$.

9.2.3 Loi binomiale

Notée $\mathcal{B}(n,p)$, c'est la loi :

$$\sum_{k=0}^{n} C_n^k\, p^k (1-p)^{n-k}\, \delta_k = (q\delta_0 + p\delta_1)^{*n}.$$

Elle permet de modéliser le nombre de gains pour n lancés au jeu de pile ou face avec probabilité de gain de p. Si $X_1, \ldots, X_n$ sont i.i.d. de loi de Bernoulli $q\delta_0 + p\delta_1$, alors $X_1 + \cdots + X_n$ suit la loi $\mathcal{B}(n,p)$. La probabilité de gagner k fois en n lancés vaut $C_n^k p^k (1-p)^{n-k}$. Sa moyenne est np et sa variance npq. Elle a pour fonction génératrice :

$$G_{\mathcal{B}(n,p)}(s) := \mathbf{E}_{\mathcal{B}(n,p)}(s^\bullet) = (q + ps)^n.$$

La loi gaussienne peut être obtenue par passage à la limite, cf. théorème 2.2.2, page 45.

9.2.4 Loi géométrique

La loi géométrique $\mathcal{G}(p)$ de paramètre $p \in [0,1]$, ou loi de Pascal, correspond à la loi du premier succès à un jeu de pile ou face avec probabilité de gain p. Si $(X_n, n \in \mathbb{N}^*)$ est une suite de v.a. i.i.d. de loi de Bernoulli $q\delta_0 + p\delta_1$, la loi géométrique de paramètre p est la loi de la v.a. « premier succès » définie par :

$$T(p) := \inf\{n \in \mathbb{N}^*, X_n = 1\}.$$

Elle est donnée par :

$$\mathcal{G}(p) := pq^{-1} \sum_{k=1}^{+\infty} q^k \delta_k.$$

Sa moyenne vaut $1/p$, sa variance q/p^2 et sa fonction génératrice est $sp/(1-sq)$. On préfère parfois numéroter les lancés de pile ou face à partir de 0, on a alors la loi géométrique de base 0 :

$$p \sum_{k=0}^{+\infty} q^k \delta_k,$$

qui a pour moyenne q/p et pour variance q/p^2. Sous certains aspects, la loi géométrique peut être vue comme l'analogue discrète de la loi exponentielle.

9.2.5 Loi négative-binomiale

C'est la loi du nombre de lancés nécessaires à un jeu de pile ou face de probabilité de gain p pour obtenir exactement m succès. C'est donc la loi de la v.a. $T_m(p)$ où $T_1(p) = T(p)$ (de loi géométrique) et $T_i(p)$ $(i \geqslant 1)$ est défini par récurrence par :

$$T_i(p) := \inf\{n > T_{i-1}, X_n = 1\}.$$

Il est facile de voir que $T_m(p)$ à la même loi que la somme de m v.a. i.i.d. de loi géométrique de paramètre p. On a donc :

$$\mathcal{G}_m(p) = \sum_{k=m}^{+\infty} C_{n-1}^{m-1}\, p^m q^{k-m}\, \delta_k = \mathcal{G}(p)^{*m}.$$

On retrouve bien la loi géométrique pour $m = 1$: $\mathcal{G}(p) = \mathcal{G}_1(p)$. La loi $\mathcal{G}_m(p)$ a pour moyenne m/p, pour variance mq/p^2 et pour fonction génératrice $(ps/(1-sq))^m$. Elle tire son nom du fait que les coefficients de sa fonction génératrice proviennent d'une sorte de « formule du binôme inversé ».

Remarque 9.2.1 *(Schéma de Bernoulli).* Le schéma de Bernoulli correspond à un jeu de pile ou face infini, avec probabilité $p \in]0,1[$ de succès à chaque lancé (les cas extrêmes sont triviaux). On le modélise par la donnée d'une suite $(X_n, n \in \mathbb{N}^*)$ de v.a. i.i.d. de loi de Bernoulli de paramètre p sur $\{0,1\}$: à chaque lancé, on code 0 un échec (probabilité $q := 1-p$) et 1 un succès (probabilité p). Le théorème de Caratheodory permet de définir une tribu et une loi de probabilité $\mathcal{L}$ sur l'ensemble $\{0,1\}^{\mathbb{N}^*}$ des suites infinies de 0 et de 1 telle que $(X_1, X_2, \ldots)$ soit de loi $\mathcal{L}$. Bien entendu, pour toute suite s dans cet espace, on a $\mathcal{L}(\{s\}) = 0$. En revanche, on a pour tout $n \in \mathbb{N}^*$ et $a \in \{0,1\}^n$, par indépendance des $X_1, \ldots, X_n$: $\mathbb{P}(X_1 = a_1, \ldots, X_n = a_n) = \mathbb{P}(X_1 = a_1) \cdots \mathbb{P}(X_n = a_n)$. Cela correspond à l'événement cylindrique $\{a_1\} \times \cdots \times \{a_n\} \times \{0,1\} \times \{0,1\} \times \cdots$. La plupart des questions que l'on peut se poser sur le jeu de pile ou face reviennent à étudier certaines propriétés de la loi $\mathcal{L}$. Ses marginales sont des lois de Bernoulli $\mathcal{B}(1,p)$, et la loi de la somme de k termes est binomiale $\mathcal{B}(k,p)$. Sous $\mathcal{L}$, la loi du rang du premier 1 dans la suite infinie est géométrique $\mathcal{G}(p)$, et plus généralement la loi du rang du $m^{\text{ième}}$ 1 est négative-binomiale $\mathcal{G}_m(p)$. Les schémas de Bernoulli sont beaucoup moins triviaux qu'on le pense, et fournissent un très bon terrain d'investigation (y compris pour les systèmes dynamiques...). La section 1.5.1 page 28 donne une propriété fractale amusante de certains schémas de Bernoulli.

9.2.6 Loi de Poisson

La loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$ est définie par :

$$e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} \delta_n,$$

C'est la loi du nombre de v.a.r. i.i.d. de loi exponentielle de paramètre λ qu'il faut additionner pour dépasser le seuil 1. Elle a pour moyenne et variance λ et pour fonction génératrice $e^{\lambda(s-1)}$. La loi de Poisson est en quelque sorte l'analogue discret de la loi gaussienne. On a par exemple $\mathcal{P}(\lambda_1) * \mathcal{P}(\lambda_2) = \mathcal{P}(\lambda_1 + \lambda_2)$. On dispose même d'un « théorème central limite poissonien » : $\mathcal{B}(n, \lambda/n)$ converge étroitement vers $\mathcal{P}(\lambda)$ lorsque $n \to +\infty$, cf. [BL98, th. V.5.6 p. 148].

9.2.7 Loi multinomiale

La loi multinomiale $\mathcal{M}(n, k, p_1, \ldots, p_k)$ de taille $n \in \mathbb{N}^*$, de dimension $k \in \mathbb{N}^*$ et de paramètre $(p_1, \ldots, p_k) \in \mathbb{R}_+^k$ avec $p_1 + \cdots + p_k = 1$ est la loi portée par l'ensemble

$$\left\{(n_1, \ldots, n_k) \in \mathbb{N}^k, \text{ t.q. } n_1 + \cdots + n_k = n\right\}$$

et donnée par la formule :

$$\mathcal{M}(n, k, p_1, \ldots, p_k)(\{n_1, \ldots, n_k\}) := \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}.$$

Elle tire son nom de la formule du multinôme :

$$(a_1 + \cdots + a_k)^n = \sum_{\substack{0 \leqslant n_1, \ldots, n_k \leqslant n \\ n_1 + \cdots + n_k = n}} \frac{n!}{n_1! \cdots n_k!} a_1^{n_1} \cdots a_k^{n_k},$$

qui donne :

$$\mathcal{M}(n, k, p_1, \ldots, p_k) = \sum_{\substack{0 \leqslant n_1, \ldots, n_k \leqslant n \\ n_1 + \cdots + n_k = n}} \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k} \delta_{(n_1, \ldots, n_k)} = (p_1 \delta_{e_1} + \cdots + p_k \delta_{e_k})^{*n},$$

où $\{e_1, \ldots, e_k\}$ est la base canonique de $\mathbb{R}^k$. On retrouve la loi binomiale pour $k = 1$:

$$\mathcal{M}(n, 1, p) = \mathcal{B}(n, p).$$

Si $(X_1, \ldots, X_n)$ sont des v.a. i.i.d. de loi $(p_1, \ldots, p_k)$ sur $\{1, \ldots, k\}$, et si $N_i := \#\{j = 1, \ldots, n, \text{ t.q. } X_j = i\}$ alors $(N_1, \ldots, N_k)$ suit la loi $\mathcal{M}(n, k, p_1, \ldots, p_k)$. Remarquons que par définition, $N_1 + \cdots + N_k = n$. La loi multinomiale permet par exemple de modéliser les résultats de n lancés indépendants d'un dé à k faces, N_i représentant alors le nombre de fois où la face numéro i a été obtenue. On a enfin la convergence étroite suivante vers une loi normale :

$$\left(\frac{N_1 - np_1}{\sqrt{np_1}}, \ldots, \frac{N_k - np_k}{\sqrt{np_k}}\right) \xrightarrow[n\to+\infty]{\text{étr.}} \mathcal{N}(0, \Sigma_{p_1,\ldots,p_k}),$$

où

$$\Sigma_{p_1,\ldots,p_k} := \mathbf{I}_k - (\sqrt{p_1}, \ldots, \sqrt{p_k})^\top (\sqrt{p_1}, \ldots, \sqrt{p_k}).$$

La loi normale $\mathcal{N}(0, \Sigma(p_1, \ldots, p_k))$ est la loi de la projection d'un vecteur aléatoire de loi $\mathcal{N}(0, \mathbf{I}_k)$ sur l'hyperplan orthogonal au vecteur $\left(\sqrt{p_1}, \ldots, \sqrt{p_k}\right)^\top$. Sa matrice de covariance est de rang $k - 1$, elle est donc dégénérée (i.e. singulière).

9.2.8 Loi hypergéométrique

La loi hypergéométrique de paramètres $(N, N_1, n) \in \mathbb{N}^* \times \mathbb{N}^* \times \mathbb{N}^*$ avec $N_1 \leqslant N$ et $n \leqslant N$ est la loi sur $\{0, \ldots, n\}$ donnée par :

$$\sum_{k=0}^{n} \frac{C_{N_1}^k C_{N-N_1}^{n-k}}{C_N^n} \delta_k.$$

Si l'on considère une population de N individus dont N_1 possèdent un caractère particulier, par exemple une maladie, alors la loi hypergéométrique de paramètre (N, N_1, n) représente la loi du nombre d'individus possédant ce caractère dans un échantillon de taille n de la population. C'est donc la loi de $X_1 + \cdots + X_n$ où les v.a. X_i sont i.i.d. et ont pour loi $((N - N_1)/N)\delta_0 + (N_1/N)\delta_1$.

Elle a pour moyenne nN_1/N et pour variance $n(N-n)N_1(N-N_1)(N-1)^{-1}N^{-2}$. Bien sûr, de la même façon que la loi multinomiale généralise la loi binomiale, on pourrait définir une loi hypergéométrique multiple lorsque l'on s'intéresse à un nombre $m \geqslant 1$ de caractères dans la population.

9.3 Transformées de Fourier & de Laplace

Pour les lois continues, la fonction génératrice est remplacée par la transformée de Fourier ou de Laplace, qui sont plus puissantes et gardent leur sens pour les lois discrètes. Soit μ une mesure de probabilité sur $(\Omega, \mathcal{A})$ où Ω est un espace vectoriel topologique et $\mathcal{A}$ sa tribu borélienne. On note Ω' l'espace dual topologique de Ω, constitué des formes linéaires continues sur Ω. La transformée de Fourier de μ est définie par :

$$\begin{aligned} \mathcal{F}ourier(\mu) : \Omega' &\to \mathbb{R} \\ f &\mapsto \mathbf{E}_\mu(\exp(if)) \end{aligned}$$

Cette fonction est bien définie puisque pour tout $f \in \Omega'$, la fonction $\exp(if) : \Omega \to \mathbb{C}$ est continue et bornée. Lorsque Ω est un espace Hilbert $\mathbb{H}$ ($\mathbb{R}^n$ par exemple), on identifie Ω' à $\mathbb{H}$ et on pose :

$$\begin{aligned} \mathcal{F}ourier(\mu) : \mathbb{H} &\to \mathbb{C} \\ s &\mapsto \int_\Omega \exp(i < s, x >)\, d\mu(x). \end{aligned}$$

La transformée de Laplace, quant à elle, est définie lorsque cela a un sens par :

$$\mathcal{L}aplace(\mu)(u) := \int_\Omega \exp\left(- < u, x >\right) d\mu(x).$$

Formellement, il suffit de prendre $s = iu$ dans la définition de la transformée de Fourier pour obtenir la transformée de Laplace, lorsqu'elle existe.

La transformée de Fourier (ou de Laplace) caractérise la loi et l'indépendance. Ainsi, deux v.a. ont la même loi si et seulement si elles ont même transformée et X et Y sont indépendantes si et seulement si la transformée du couple (X, Y) en (s_1, s_2) est égale au produit des transformées respectives en s_1 et s_2 pour tout (s_1, s_2). bien entendu, la transformée en s de la somme de deux v.a. indépendantes est égale au produit des transformées respectives en s (la réciproque est fausse).

La transformée de Fourier d'une loi symétrique (i.e. invariante par la symétrie centrale de centre 0) est réelle.

La transformée de Fourier de δ_a avec $a \in \mathbb{R}^n$ est donnée par $e^{i<s,a>}$, elle permet de calculer les transformées de Fourier des lois discrètes (i.e. à support au plus dénombrable sans point d'accumulation). Ainsi, les lois binomiales $\mathcal{B}(n, p)$ et de Poisson $\mathcal{P}(\lambda)$ ont pour transformée de Fourier respectives :

$$(pe^{is} + q)^n \quad \text{et} \quad e^{\lambda(e^{is}-1)}.$$

Quant à la loi géométrique $\mathcal{G}(p)$, sur $\mathbb{N}^*$, et à la loi négative-binomiale $\mathcal{G}_m(p)$), elles ont pour transformée de Fourier respectives :

$$\frac{pe^{is}}{1 - qe^{is}} \quad \text{et} \quad \frac{p^m e^{mis}}{(1 - qe^{is})^m}.$$

Heuristique 9.3.1 *(Transformation de Fourier = analyse en fréquences).* La transformation de Fourier d'une fonction f exprime les composantes en fréquences (on dit aussi spectre), indexées par s, de son argument f. La fonction $x \mapsto e^{isx}$ représente une fonction périodique élémentaire, dite *harmonique de fréquence* $2\pi/s$. Il faut beaucoup de spectre pour reconstituer une fonction f constante ou plus généralement « peu périodique » à partir des harmoniques. La transformation de Fourier change la convolution en un produit. Pour tronquer le spectre d'un signal f, il suffit de le convoler avec la transformée de Fourier inverse de la fonction indicatrice de troncature.

9.4 Lois continues

9.4.1 Loi uniforme

La loi uniforme $\mathcal{U}(a, b)$ sur l'intervalle compact non vide $[a, b] \subset \mathbb{R}^n$ a pour densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{1}{b-a} \mathrm{I}_{[a,b]}(x).$$

Elle a pour moyenne $(a + b)/2$ et pour variance $(b - a)^2/12$. Sa transformée de Fourier est :

$$e^{is(a+b)/2} \frac{\sin(s(b-a)/2)}{s(b-a)/2}.$$

Tout ceci se généralise sans difficultés à un pavé $[a_1, b_1] \times \cdots \times [a_n, b_n]$ de $\mathbb{R}^n$.

9.4.2 Loi triangulaire

C'est la loi « uniforme » sur le triangle du plan de base $[-a, a]$ et de sommet de coordonnées $(0, 1)$. Elle a pour densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{1}{a}\left(1 - \frac{|x|}{a}\right) \mathrm{I}_{[-a,a]}(x),$$

pour moyenne 0, pour variance $a^2/6$ et pour transformée de Fourier :

$$\frac{2(1 - \cos(as))}{a^2 s^2}.$$

9.4.3 Loi gaussienne

On parle aussi de loi normale, de Gauss, ou encore de Maxwell ! En Amérique du nord, on parle souvent de « bell curve », i.e. de courbe en cloche. Ces multiples appélations témoignent de l'importance de cette loi. On pourrait définir les lois gaussiennes sur des espaces très généraux, mais cela nous mènerait bien trop loin[1]. La loi gaussienne $\mathcal{N}(m, \Gamma)$, sur $\mathbb{R}^n$, de moyenne $m \in \mathbb{R}^n$ et de matrice de covariance $\Sigma \in \mathbb{S}_n^+(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{R})$, a pour densité par rapport à la mesure de Lebesgue $dx_1 \cdots dx_n$ sur $\mathbb{R}^n$:

$$(2\pi)^{-n/2}|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}\langle x - m, \Sigma(x - m)\rangle\right).$$

Elle a pour transformée de Fourier :

$$\exp\left(i\langle s, m\rangle - \frac{1}{2}\langle \Sigma\, s, s\rangle\right).$$

Elle a pour propriété de maximiser l'entropie de Shannon définie sur les lois des vecteurs aléatoires à densité f de covariance fixée Σ par :

$$\mathbf{H}(f) := -\int_{\mathbb{R}^n} f(x) \log f(x)\, dx.$$

9.4.4 Loi exponentielle et loi gamma

La loi gamma $\Gamma(a, \lambda)$ de paramètres $(a, \lambda) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ a pour densité sur $\mathbb{R}$ par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{\lambda^a}{\Gamma(a)}\, x^{a-1}\, e^{-\lambda x}\, \mathrm{I}_{\mathbb{R}^+}(x),$$

où $\Gamma(a)$ est la fonction Gamma d'Euler définie par :

$$\Gamma(a) := \int_0^{+\infty} x^{a-1}\, e^{-x}\, dx, \quad \text{en particulier } \Gamma(n) = (n-1)!.$$

Parfois, on choisit un paramétrage différent de la loi gamma, en posant $b = 1/\lambda$, et l'on note alors $G(a, b) = \Gamma(a, 1/b)$. Pour $a = 1$, on obtient la loi exponentielle :

$$\Gamma(1, \lambda) = G(1, 1/\lambda) = \mathcal{E}(\lambda).$$

[1]Sur un espace préhilbertien $\mathbb{H}$, les choses sont simples : une v.a. à valeurs dans $\mathbb{H}$ est gaussienne si et seulement si pour tout v dans $\mathbb{H}$, la v.a.r. $\langle X, v\rangle_{\mathbb{H}}$ est gaussienne réelle. On retrouve facilement les vecteurs gaussiens usuels en prenant $\mathbb{H} = \mathbb{R}^n$.

La loi exponentielle est l'analogue continu de la loi géométrique, elle apparaît comme loi du temps inter-sauts du processus de Poisson. La loi exponentielle est également sans mémoire : si $\mathcal{L}(X) = \mathcal{E}(\lambda)$, alors $\mathbb{P}(X \geqslant t+s \mid X \geqslant t) = \exp(-\lambda s)$ ne dépend que de s et pas de t. Cette propriété la caractérise dans l'ensemble des lois continues, et traduit bien le comportement de durées de vie en fiabilité, du temps d'arrivée du prochain client dans une file d'attente, etc. Elle explique également la stationnarité des accroissements du processus de Poisson (renouvellement).

La loi gamma $\Gamma(a,\lambda)$ a pour moyenne a/λ, pour variance a/λ^2 et pour transformée de Fourier $(1 - is\lambda^{-1})^{-a}$. La loi $\Gamma(n,\lambda)$ avec $n \in \mathbb{N}^*$ est parfois appelée loi de Erlang. Ces lois sont très importantes puisqu'elles apparaissent lorsque l'on fait la somme de v.a.r. i.i.d. de loi exponentielle. La famille des lois gamma est stable par convolution et constitue donc en quelque sorte l'analogue continu de la famille des lois négatives-binomiales. La loi du temps de depassement d'un seuil pour un processus de Poisson est une loi gamma, car convolution de lois exponentielles.

9.4.5 Loi double exponentielle

Également appelée loi de Laplace. Elle a pour densité sur $\mathbb{R}$ par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$: $\lambda\exp(-\lambda|x|)/2$, où λ est un paramètre dans $\mathbb{R}_+$. Sa moyenne vaut 0, sa variance $2/\lambda^2$ et sa transformée de Fourier est $1/(1+s^2/\lambda^2)$.

9.4.6 Loi de Dirichlet

La loi de Dirichlet de taille $n \in \mathbb{N}^*$ et de paramètre $\alpha \in (\mathbb{R}_+^*)^n$ est la loi du vecteur aléatoire $(X_1,\ldots,X_n)/(X_1+\cdots+X_n)$ lorsque les variables aléatoires $X_1,\ldots,X_n$ sont i.i.d. et suivent des lois exponentielles de paramètres respectifs $\alpha_1,\ldots,\alpha_n$, cf. [DCD82b, exercice 8.2.15, page 192]. Elle est portée par le simplexe $\mathbb{S}(n,1) := \{y \in \mathbb{R}^n,\ y_i \geqslant 0,\ y_1+\cdots+y_n = 1\}$. La densité de $(Y_1,\cdots,Y_{n-1})$ est donnée par :

$$\frac{\Gamma(\alpha_1+\cdots+\alpha_n)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_n)}\left(\prod_{i=1}^{n-1} y_i^{\alpha_i-1}\right)\left(1-\sum_{i=1}^{n-1} y_i\right)^{\alpha_k-1}.$$

La loi de Dirichlet de taille n et de paramètre $\alpha = (1,\ldots,1)$ n'est rien d'autre que la loi uniforme sur le simplexe $\mathbb{S}(n,1)$, cf. section 1.7 page 34. Elle permet par exemple de choisir « uniformément » une loi de probabilité discrète.

9.4.7 Loi du chi-deux

La loi du $\chi^2(n)$ à n degrès de liberté n'est rien d'autre que la loi de la somme des carrés de n v.a.r. i.i.d. de loi normale centrées réduites sur $\mathbb{R}$. C'est aussi la loi gamma $G(n/2,2)$ de paramètres $(a,b) := (n/2,2)$, i.e. la loi gamma $\Gamma(n/2,1/2)$ de paramètres $(a,\lambda) := (n/2,1/2)$. La loi du $\chi^2(n)$ intervient dans le test du même nom (cf. 3.4.2 page 70).

9.4.8 Loi exponentielle et loi de Weibull

La loi de Weibull $W(\alpha,\lambda)$ de paramètres $(\alpha,\lambda) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ est la loi sur $\mathbb{R}$ dont la densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$ est :

$$\alpha\lambda\, x^{\alpha-1}\, \exp\left(-\lambda x^\alpha\right) \mathrm{I}_{\mathbb{R}_+}(x).$$

Elle est plus communément donnée par sa fonction de répartition :

$$F_{\alpha,\lambda}(u) = \left(1-\exp\left(-\lambda x^\alpha\right)\right) \mathrm{I}_{\mathbb{R}_+}.$$

Sa moyenne vaut $\lambda^{-1/\alpha}\Gamma(1+\alpha^{-1})$ et sa variance $\lambda^{-2/\alpha}\Big(\Gamma(1+2\alpha^{-1})-\big(\Gamma(1+\alpha^{-1})\big)^2\Big)$. Cette loi intervient en fiabilité des systèmes. Notons que l'on retrouve la loi exponentielle de paramètre λ lorsque $\alpha = 1$:

$$W(1,\lambda) = \mathcal{E}(\lambda).$$

Pour $\alpha = 2$, on parle parfois de loi de Rayleigh.

9.4.9 Loi de Cauchy

C'est la loi sur $\mathbb{R}$ de densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{c}{\pi}\,\frac{1}{c^2+x^2},$$

où c est un paramètre positif. Elle n'a pas de moyenne et donc pas de variance. Sa transformée de Fourier est $\exp\left(-c|s|^2\right)$.

9.4.10 Loi Beta

La loi bêta de paramètres $(a,b) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ est la loi sur $\mathbb{R}$ de densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{1}{B(a,b)}\,x^{a-1}(1-x)^{b-1}\,\mathrm{I}_{\{0<x<1\}}(x),$$

où $B(a,b)$ est la fonction bêta d'Euler définie par :

$$B(a,b) := \int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

La loi bêta de paramètres (a,b) a pour moyenne $a/(a+b)$ et pour variance $ab(a+b)^{-2}(a+b+1)^{-1}$.

9.4.11 Loi de Pareto

Si $\mathcal{L}(X) = \mathcal{E}(\lambda)$ et si $r \in \mathbb{R}$, la v.a. $r\exp(X)$ suit par définition une loi de Pareto de paramètres λ et r. Sa densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ vaut :

$$\frac{\lambda\,r^{\lambda}}{x^{1+\lambda}}\,\mathrm{I}_{\{x>r\}}(x).$$

Elle a pour moyenne $\alpha r/(\alpha-1)$ si $\alpha > 1$ et n'a pas moyenne sinon. Si $\alpha > 2$, elle a pour variance $\alpha r^2(\alpha-1)^{-2}(\alpha-2)^{-1}$, et elle n'a pas de moment d'ordre deux sinon.

9.4.12 Loi log-normale

C'est la loi de l'exponentielle d'une variable aléaroite de loi gaussienne $\mathcal{N}(m,\sigma^2)$. Elle a pour densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$(2\pi\sigma^2)^{-1/2}\,x^{-1}\,\exp\left(-\frac{(\log(x)-m)^2}{2\sigma^2}\right)\mathrm{I}_{\{x>0\}}(x),$$

pour moyenne $\exp\left(m+\sigma^2/2\right)$ et pour variance $\exp\left(2m+\sigma^2\right)\left(e^{\sigma^2}-1\right)$.

9.4.13 Loi de Ficher-Snedecor

La loi de Fisher-Snedecor $\mathcal{F}(n,m)$ à n et m degrés de liberté est la loi de

$$\frac{X/n}{Y/m},$$

où X et Y sont indépendantes et ont pour lois respectives $\chi^2(n)$ et $\chi^2(m)$. Elle a pour densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{\Gamma((n+m)/2)n^{n/2}m^{n/2}}{\Gamma(n/2)\Gamma(m/2)}\frac{x^{n/2-1}}{(m+nx)^{(n+m)/2}}.$$

Pour $m > 4$, sa moyenne vaut $m/(m-2)$ et sa variance $2m^2(m+n-2)n^{-1}(m-4)^{-1}(m-2)^{-2}$.

9.4.14 Loi de Student

La loi de Student $\mathcal{T}_n$ à n degrés de liberté est la loi de

$$\frac{\sqrt{n}\,X}{\sqrt{Y}},$$

où X et et Y sont indépendantes et ont pour lois respectives $\mathcal{N}(0,1)$ et $\chi^2(n)$. Elle a pour densité par rapport à la mesure de Lebesgue dx sur $\mathbb{R}$:

$$\frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}.$$

Elle est centrée et sa variance vaut $n/(n-2)$ pour $n > 2$.

9.4.15 Loi de Kolmogorov-Smirnov

C'est la loi sur $\mathbb{R}$ de fonction de répartition :

$$\mathbf{F}(u) := \begin{cases} 0 & \text{si } u \leqslant 0, \\ 1+2\displaystyle\sum_{k=1}^{+\infty}(-1)^k e^{-2k^2u^2} & \text{si } u > 0. \end{cases}$$

Elle intervient dans le test d'adéquation de Kolmogorov-Smirnov (cf. 3.4.3 page 73). C'est aussi la loi sur supremum de la valeur absolue d'un pont brownien sur $[0,1]$ (cf. section 2.4 page 49).

9.4.16 Statistique d'ordre, quantiles, médiane

Soit μ une loi de probabilité sur $\mathbb{R}$, de fonction de répartition F. On note $F(x^-)$ sa limite à gauche en x. On dit qu'un réel m est une *médiane* pour μ lorsque $F(m^-) \leqslant 1/2 \leqslant F(m)$. Plus généralement, pour tout $\alpha \in]0,1[$, on dit qu'un réel q_α est un *quantile* d'ordre α pour μ lorsque $F(q_\alpha^-) \leqslant \alpha \leqslant F(q_\alpha)$. Pour tout ordre α, un quantile d'ordre α existe toujours, car la fonction de répartition F est croissante, et a pour limite 0 en $-\infty$ et 1 en $+\infty$. La croissance de F entraîne également que l'ensemble des quantiles d'ordre α est un intervalle. Si q est un quantile, l'intervalle des quantiles du même ordre est réduit à $\{q\}$ si et seulement si $\mu(\{q_\alpha\}) = 0$, et alors F est continue en α (on rappelle qu'une fonction de répartition est toujours continue à droite). Souvent, on appelle encore quantile d'ordre α le milieu de l'intervalle des quantiles d'ordre α, ce qui permet de parler *du* quantile d'ordre α, même quand il y en a plusieurs. Enfin, lorsque la loi μ possède une densité, sa fonction de répartition est continue et les intervalles quantiles sont tous des singletons.

On parle de *déciles* pour les quantiles d'ordre $1/10$, de *quartiles* pour ceux d'ordre $1/4$, etc. On parle aussi parfois de *percentiles* pour les quantiles. Une médiane est un quantile d'ordre $1/2$. La fonction Matlab `median` permet d'obtenir la médiane empirique d'un échantillon (i.e. la médiane de la fonction de répartition empirique), alors que la fonction Stixbox `quantile` permet d'obtenir un quantile empirique d'ordre quelconque d'un échantillon.

Si $(x_1, \dots, x_n)$ est un vecteur de $\mathbb{R}^n$, on note $(x_{(1)}, \dots, x_{(n)})$ le vecteur de $\mathbb{R}^n$ obtenu en réordonnant de façon croissante les composantes de x. Si $X := (X_1, \dots, X_n)$ est un vecteur aléatoire de $\mathbb{R}^n$, on appelle statistique d'ordre de X le vecteur aléatoire $\tilde{X} := (X_{(1)}, \dots, X_{(n)})$ obtenu en réordonnant les coordonnées de X à ω fixé. La fonction Matlab `sort` permet de trier de façon croissante un vecteur.

Si $X := (X_1, \dots, X_n)$ est un échantillon d'une loi μ sur $\mathbb{R}$, alors la médiane empirique n'est rien d'autre que $X_{((n+1)/2)}$ si n est impair et $(X_{((n/2)} + X_{((1+n/2))})/2$ si n est pair. De manière générale, si α n'est pas un multiple de $1/n$, le quantile d'ordre α empirique vaut $(X_{([n\alpha])} + X_{([n\alpha]+1)})/2$. Enfin, on peut montrer que si μ a une densité f par rapport à la mesure de Lebesgue, alors la loi de $X_{(i)}$ a pour densité :

$$f_i(x) = \frac{n!}{(i-1)!(n-i)!} F^{i-1}(x)(1-F(x))^{n-i} f(x),$$

où F est la fonction de répartition de μ. On retrouve bien f pour $n = 1$. Si $x := (x_1, \dots, x_n) \in \mathbb{R}^n$, on appelle *rang* de x_i dans la suite x la quantité

$$r_i(x) := \sum_{j=1}^{n} \mathrm{I}_{x_j \leqslant x_i},$$

qui représente le nombre de composantes de x qui sont inférieures ou égales à x_i. Il peut arriver que $r_i = r_j$, mais on a toujours $x_{r_i(x)} = x_{(i)}$. Lorsque les composantes de x sont toutes différentes, l'application $r : i \in \{1, \dots, n\} \mapsto r_i(x)\{1, \dots, n\}$ est une permutation. Le vecteur $r(x) := (r_1(x), \dots, r_n(x))$ est appelé *vecteur rang* de x. On définit naturellement le vecteur aléatoire rang $R(X)$ d'un vecteur aléatoire $X := (X_1, \dots, X_n)$ par $R(X)(\omega) := r(X(\omega))$. On montre, cf. [DCD82b, thm 4.4.29 page 107] que pour un échantillon $X := (X_1, \dots, X_n)$ d'une loi μ sur $\mathbb{R}$, continue :

- le vecteur rang $R(X)$ est presque-sûrement une permutation
- le vecteur rang $R(X)$ suit la loi uniforme sur l'ensemble des permutations de $\{1, \dots, n\}$
- la statistique d'ordre $\tilde{X}$ de X et son rang $R(X)$ sont indépendants
- la statistique d'ordre $\tilde{X}$ a pour loi la loi trace de $n!\mu \otimes \cdots \otimes \mu$ sur $\{x \in \mathbb{R}^n,\, x_1 \leqslant \cdots \leqslant x_n\}$.

9.5 Convolutions & convergences

On a les propriétés suivantes :

1. $\mathrm{Bern}(p, \{0,1\})^{*n} = \mathcal{B}(n, p)$.
2. $\mathcal{G}(p)^{*m} = \mathcal{G}_m(p)$.
3. $\mathcal{G}_{m_1}(p) * \cdots * \mathcal{G}_{m_n}(p) = \mathcal{G}_{m_1 + \cdots + m_n}(p)$.
4. $\mathcal{E}(\lambda)^{*n} = \Gamma(n, \lambda)$.
5. $\mathcal{N}(m_1, \Sigma_1) * \cdots * \mathcal{N}(m_n, \Sigma_n) = \mathcal{N}(m_1 + \cdots + m_n, \Sigma_1 + \cdots + \Sigma_n)$.
6. $\mathcal{P}(\lambda_1) * \cdots * \mathcal{P}(\lambda_n) = \mathcal{P}(\lambda_1 + \cdots + \lambda_n)$.
7. Si $\mathcal{L}(X) = \mathcal{N}(0, \mathbf{I}_n)$ alors $\mathcal{L}(X_1^2 + \cdots + X_n^2) = G(n/2, 2) = \chi^2(n)$.

8. Si $\mathcal{L}((X,Y)) = \Gamma(a,1) \otimes \Gamma(b,1)$ alors $\mathcal{L}(X/(X+Y)) = \mathrm{Beta}(a,b)$.
9. Si $\mathcal{L}(X) = \Gamma(1,\lambda) = \mathcal{E}(\lambda)$ alors $\mathcal{L}(X^{-\alpha}) = W(\alpha,\lambda)$.
10. Si $\mathcal{L}((X,Y)) = \chi^2(n) \otimes \chi^2(m)$, alors $\mathcal{L}((mX)/(nY)) = \mathcal{F}(n,m)$.
11. Si $\mathcal{L}((X,Y)) = \mathcal{N}(0,1) \otimes \chi^2(n)$, alors $\mathcal{L}(\sqrt{n}X/\sqrt{Y}) = \mathcal{T}(n)$.
12. Si $\mathcal{L}((X_n, n \in \mathbb{N}^*)) = \mathcal{E}(\lambda)^{\mathbb{N}^*}$ et
$$N_t := \sum_{n=1}^{+\infty} \mathrm{I}_{X_1+\cdots+X_n \leqslant t} = \inf\{n \in \mathbb{N},\, X_1 + \cdots + X_n > t\},$$
alors pour tout $t \in \mathbb{R}_+$, $\mathcal{L}(N_t) = \mathcal{P}(\lambda t)$ (processus de Poisson simple d'intensité λ).
13. $\mathcal{B}(n,\lambda/n) \xrightarrow[n\to+\infty]{\text{étr.}} \mathcal{P}(\lambda)$

1. $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$
   (a) ssi $(\forall\varepsilon > 0)$ $\mathbb{P}(\{|X_n - X| \geqslant \varepsilon\}) \xrightarrow[n\to+\infty]{} 0$
   (b) ssi $\mathbf{E}(|X_n - X| \wedge 1) \xrightarrow[n\to+\infty]{} 0$
   (c) ssi $\mathbf{E}(|X_n - X|/(1 + |X_n - X|)) \xrightarrow[n\to+\infty]{} 0$
   (d) ssi de toute sous suite $(X_{n_k}, k \in \mathbb{N}^*)$, on peut extraire une sous-suite qui converge p.s. vers X
2. $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} \mu$ ssi $\mathcal{L}(X_n) \xrightarrow[n\to+\infty]{\text{étr.}} \mu$.
3. $\mu_n \xrightarrow[n\to+\infty]{\text{étr.}} \mu$ (avec μ_n et μ probas)
   (a) ssi $(\forall f \in \mathcal{C}_b)$ $\mathbf{E}_{\mu_n}(f) \xrightarrow[n\to+\infty]{} \mathbf{E}_\mu(f)$
   (b) ssi $(\forall s \in \mathbb{R})$ $\mathbf{E}_{\mu_n}(e^{is\bullet}) \xrightarrow[n\to+\infty]{} \mathbf{E}_\mu(e^{is\bullet})$
   (c) ssi $(\forall t \in \mathrm{Cont}(\mathrm{F}_\mu))$ $F_{\mu_n}(t) \xrightarrow[n\to+\infty]{} F_\mu(t)$
4. La famille de v.a. $(X_i, i \in I) \subset \mathbf{L}^1$ est uniformément intégrable (UI)
   (a) ssi $\sup_{i\in I} \mathbf{E}(|X_i|\mathrm{I}_{\{|X_i|>c\}}) \xrightarrow[c\to+\infty]{} 0$
   (b) ssi[2] $\sup_{i\in I} \mathbf{E}(|X_i|) < +\infty$ et $(\forall\varepsilon > 0)(\exists\delta > 0)(\forall A)(\mathbb{P}(A) \leqslant \delta \Rightarrow \sup_{i\in I} \mathbf{E}(|X_i|\mathrm{I}_A) \leqslant \varepsilon)$
   (c) ssi[3] $\exists\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ telle que $\sup_{i\in I} \mathbf{E}(\Psi(|X_i|)) < +\infty$ et $\Psi(x)/x \xrightarrow[x\to+\infty]{} +\infty$.
5. Une famille bornée dans $\mathbf{L}^1$ est UI.
6. Soit $(X_n, n \in \mathbb{N}^*) \subset \mathbf{L}^1$, il y a équivalence entre :
   (a) $X_n \xrightarrow[n\to+\infty]{\mathbf{L}^1} X$ avec $X \in \mathbf{L}^1$
   (b) $(X_n, n \in \mathbb{N}^*)$ est UI et $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$
7. Si $X_n \xrightarrow[n\to+\infty]{\mathbf{L}^p} X$ avec $p \geqslant 1$, alors de toute sous-suite $(X_{n_k}, k \in \mathbb{N}^*)$, on peut extraire une sous-suite qui converge p.s. vers X
8. Si $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} \mu$ alors[4] $s \mapsto \mathbf{E}(e^{isX_n})$ converge vers $s \mapsto \mathbf{E}_\mu(e^{is\bullet})$ uniformément sur tout compact
9. Soit $(\mu_n, n \in \mathbb{N}^*)$ des probas. Si $s \mapsto \mathbf{E}_{\mu_n}(e^{is\bullet})$ converge simplement vers $s \mapsto G(s)$ continue en 0 alors[5] G est la transformée de Fourier d'une loi μ et $\mu_n \xrightarrow[n\to+\infty]{\text{étr.}} \mu$.

[2] Critère « epsilon-delta » ou encore d'équi-intégrabilité.
[3] Critère de Lavallée-Poussin. cf. [Bor95, p. 10]
[4] Théorème de continuité de Paul Lévy, cf. [Bor95, p. 32].
[5] Théorème de continuité de Paul Lévy, cf. [Bor95, p. 32].

10. Soit $(X_n, n \in \mathbb{N}^*)$ v.a.r. indép. et $S_n := X_1 + \cdots + X_n$. Il y a équivalence entre[6] :
    (a) $(S_n, n \in \mathbb{N}^*)$ converge p.s.
    (b) $(S_n, n \in \mathbb{N}^*)$ converge en $\mathbb{P}$
    (c) $(S_n, n \in \mathbb{N}^*)$ converge en $\mathcal{L}$

    De plus[7], si $c > 0$ et $Y_n := X_n \mathrm{I}_{\{|X_n| \leqslant c\}}$ alors $(S_n, n \in \mathbb{N}^*)$ converge p.s. ssi on a à la fois :
    (a) $\sum_{n=1}^{+\infty} \mathbb{P}(\{X_n \neq Y_n\}) < +\infty$
    (b) $\sum_{n=1}^{+\infty} \mathbf{E}(Y_n) < +\infty$
    (c) $\sum_{n=1}^{+\infty} \mathbf{Var}(Y_n) < +\infty$
11. Si $X_n \xrightarrow[n\to+\infty]{\text{p.s}} X$ et $Y_n \xrightarrow[n\to+\infty]{\text{p.s}} Y$ et $(a,b) \in \mathbb{R}^2$ alors $aX_n + bY_n \xrightarrow[n\to+\infty]{\text{p.s}} aX + bY$ et $X_nY_n \xrightarrow[n\to+\infty]{\text{p.s}} XY$
12. L'espace $\mathbf{L}^0$ des fonctions mesurables muni de la topologie de la CV en $\mathbb{P}$ associée à la distance $d(X,Y) := \mathbf{E}(|X-Y|/(1+|X-Y|))$ est un espace de Banach[8].
13. Si $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ et $Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} Y$ alors $X_nY_n \xrightarrow[n\to+\infty]{\mathbb{P}} XY$
14. Si $X_n \xrightarrow[n\to+\infty]{\mathbf{L}^1} X$ et $Y_n \xrightarrow[n\to+\infty]{\mathbb{P}} Y$ et $(Y_n, n \in \mathbb{N}^*)$ unif. bornée alors $X_nY_n \xrightarrow[n\to+\infty]{\mathbf{L}^1} XY$
15. Si $(X_n, Y_n) \xrightarrow[n\to+\infty]{\mathcal{L}} (X,Y)$ alors $X_n + Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} X+Y$ et $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ et $Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} Y$
16. Si $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ et $Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} Y$ et $(X_n, Y_n, n \in \mathbb{N}^*)$ indép. alors X et Y indép. et $X_n + Y_n \xrightarrow[n\to+\infty]{\mathcal{L}} X+Y$

$$
\begin{array}{ccccc}
 & & \text{CV } \mathbf{L}^{p>1} & & \\
 & & \Downarrow & & \\
 & & \text{CV } \mathbf{L}^1 & & \\
 & & \Downarrow & & \\
\text{CV p.s.} & \Rightarrow & \text{CV en } \mathbb{P} & \Rightarrow & \text{CV en } \mathcal{L}
\end{array}
$$

$$\text{CV } \mathbf{L}^1 \underset{\text{sous suites}}{\Longrightarrow} \text{CV p.s.}$$

$$\text{CV } \mathbb{P} \underset{\text{sous suites}}{\Longrightarrow} \text{CV p.s.}$$

$$\text{CV } \mathbb{P} \underset{\text{si UI}}{\Longrightarrow} \text{CV } \mathbf{L}^1$$

FIG. 9.1 – Relations entre les différents types de convergences de v.a.

[6]Théorème de Paul Lévy, cf. [Bor95, p. 83]
[7]Théorème des trois séries de Kolmogorov, cf. [Bor95, p. 85]
[8]Donc un espace vectoriel normé, d'où l'additivité des suites convergentes en $\mathbb{P}$.

Loi	Moyenne	Variance	Transformée de Fourier en $s \in \mathbb{C}$
δ_a	a	0	e^{ias}
Bern. $q\delta_0 + p\delta_1$	p	pq	$1 + pe^{is}$
Bern. $q\delta_{-1} + p\delta_1$	$q - p$	pq	$qe^{-is} + pe^{is}$
Unif. sur $\{a_1, \ldots, a_n\}$	$(a_1 + \cdots + a_n)/n$	$(\sum_k a_k^2)/n - (\sum_k a_k)^2/n^2$	$\sum_k e^{ia_k s}/n$
$\mathcal{B}(n,p)$	np	npq	$(1 + pe^{is})^n$
Géom(p) (sur $\mathbb{N}$)	q/p	q/p^2	
$\mathcal{G}(p)$ (sur $\mathbb{N}^*$)	$1/p$	q/p^2	$pe^{is}\left(1 - qe^{is}\right)^{-1}$
$\mathcal{G}_m(p)$	m/p	q/p^2	$p^m e^{mis}\left(1 - qe^{is}\right)^{-m}$
$\mathcal{P}(\lambda)$	λ	λ	$\exp\left(\lambda(e^{is} - 1)\right)$
Unif. $\mathcal{U}(a,b)$	$(a+b)/2$	$(b-a)^2/12$	$2s^{-1}(b-a)^{-1}\exp\left(is(a+b)/2\right)\sin(s(b-a)/2)$
$\mathcal{E}(\lambda)$	$1/\lambda$	$1/\lambda^2$	$\left(1 - is\lambda^{-1}\right)^{-1}$
$\Gamma(a,\lambda)$	a/λ	a/λ^2	$\left(1 - is\lambda^{-1}\right)^{-a}$
$G(a,b)$	ab	ab^2	$(1 - ibs)^{-a}$
$\mathcal{N}(m,\Sigma)$	m	Σ	$\exp\left(i < m, s > - < \Sigma s, s > /2\right)$
$\chi^2(n)$	n	$2n$	$(1 - 2is)^{-n/2}$
$W(\alpha,\lambda)$	$\lambda^{-1/\alpha}\Gamma(1+\alpha-1)$	$\lambda^{-2/\alpha}\left(\Gamma(1+2\alpha^{-1}) - \left(\Gamma(1+\alpha^{-1})\right)^2\right)$	
$\mathcal{F}(n,m)$	$m/(m-2)$ si $m > 4$	$2m^2(m+n-2)n^{-1}(m-4)^{-1}(m-2)^{-2}$	
$\mathcal{T}(n)$	0	$n/(n-2)$ si $n > 2$	

TAB. 9.1 – Tableau récapitulatif de quelques lois usuelles

Loi μ	Fonction de répartition $F_\mu(t) := \mu([t,+\infty[)$
Discrète $\sum_{n\geqslant 0} p_n\delta_{a_n}$ $(a_n < a_{n+1})$	$\sum_{n\geqslant 0} p_n \mathrm{I}_{[a_n,a_{n+1}[}(t)$
Unif. $\mathcal{U}(a,b)$	$(b-a)^{-1}(t-a)\mathrm{I}_{[a,b]}(t) + \mathrm{I}_{]b,+\infty[}(t)$
$\mathcal{E}(\lambda)$	$(1-\exp(-\lambda t))\mathrm{I}_{\mathbb{R}_+}(t)$
$W(\alpha,\lambda)$	$(1-\exp(-\lambda t^\alpha))\mathrm{I}_{\mathbb{R}_+}(t)$
$\mathcal{N}(0,1)$	$\Phi(t) \sim_{+\infty} 1 - t^{-1}\exp\left(-t^2/2\right)$
Kolmogorov-Smirnov	$\left(1+2\sum_{k=1}^{+\infty}(-1)^k e^{-2k^2u^2}\right)\mathrm{I}_{\mathbb{R}_+^*}(t)$

TAB. 9.2 – Quelques fonctions de répartition