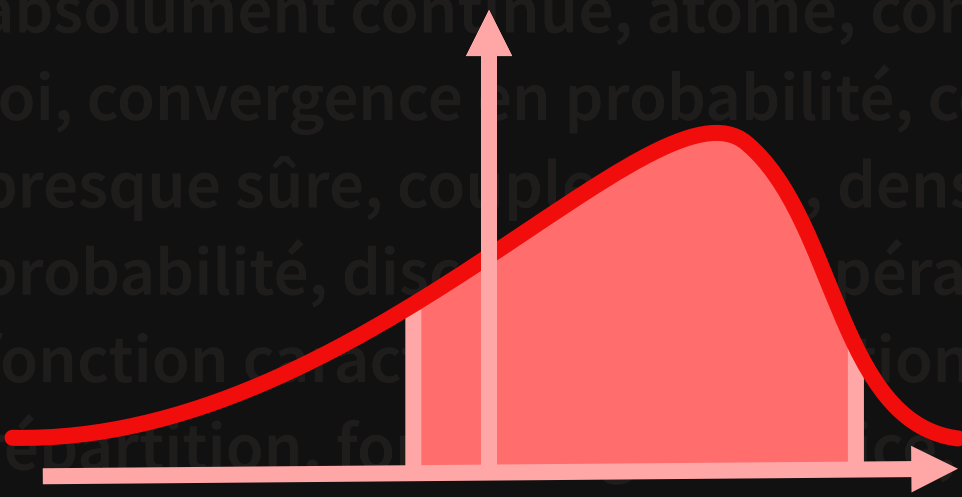
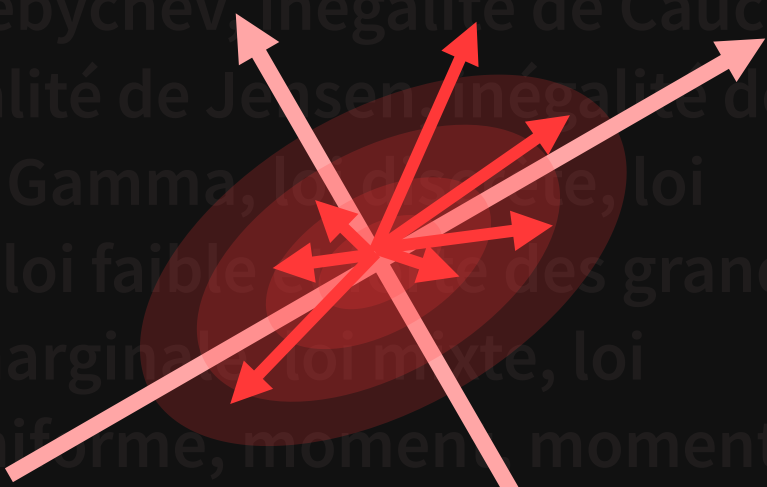
X31M010.

Probabilités continues et convergences.





Probabilités continues et convergences.

Ceci est un de 4 *pdf* de préparation à la L3. Il ne contient que les notions et connaissances qui seront enseignées l'année prochaine d'après le programme. Il n'y a pas de preuves; le but est juste de s'introduire aux concepts, pas de faire un cours entier (et rigoureux).

Les pages sont essentiellement (et en fait, quasiment entièrement) basées sur les polycopiés des années précédentes dans les matières ; il n'y a pas grand-chose ici, qu'on ne trouverait pas dans le cours.

Chaque page-fiche commence par une motivation; une sorte de dialogue entre un personnage D (Deuxième année), et un autre T (Troisième année), que vous pourrez appeler, dans votre tête, comme vous le souhaitez (*e.g.* <u>D</u>amien & <u>T</u>om Worro).

Pour cette fiche en particulier : il est (fortement) recommandé de lire d'abord la dernière page de Probabilités du livret de « Théorie de la mesure ».

<u>Sommaire.</u>

Basé sur les polycopié (de Licence de maths, 3e année) de *<u>Probabilités continues et convergences</u>* de Paul-Éric Chaudru de Raynal (d'après les notes de cours de Nicolas Pétrélis).

→ Fonctions de répartitions.

Contrairement aux v.a. discrètes, on ne peut pas décrire les variables continues par les probabilités de chaque évènement.

■ La motivation.

D – Quand on veut parler d'une distribution discrète, on peut donner la probabilité de chaque issue. Mais pour une distribution continue comme une loi uniforme sur [0 ;1], $\mathbb{P}(X=r)$ est nul pour n'importe quel r.
T – Ça ne fait pas de sens de parler d'un évènement X = r ; il est plus adapté de parler d'un évènement de la forme $a \leq X \leq b$, sur un intervalle. Et par additivité, comme $\mathbb{P}(a<X\leq b) = \mathbb{P}(X\leq b) - \mathbb{P}(X\leq a)$, il suffit de donner $\mathbb{P}(X\leq r)$ pour tout r.
D – Et ça marche aussi pour les lois discrètes !
T – Oui ; sauf que pour des lois suffisamment continues, on peut dériver pour avoir une sorte de densité de répartition.

■ La fonction de répartition.

Pour X une v.a. réelle, on pose F_X, sa fonction de répartition, définie comme $F_X(x) = \mathbb{P}(X\leq x)$ pour tout x. Notons que cette définition s'applique aussi au v.a. à valeurs dans N (vu que N⊆R).

■ L'unicité.

Si X et Y sont deux v.a. réelles, $F_X = F_Y$ *ssi* X et Y ont même loi; c'est-à-dire *ssi* $\mathbb{P}_X = \mathbb{P}_Y$. On dit qu'une variable aléatoire est caractérisée par sa fonction de répartition.

■ Les propriétés.

Les propriétés les plus importantes/utiles sont que :
i) F est croissante sur $\mathbb{R}$.
ii) F est continue à droite en tout point de $\mathbb{R}$, et admet une limite à gauche en tout point de $\mathbb{R}$. C'est-à-dire : pour tout $a\in\mathbb{R}$, $\lim_{x\to a, x<a} F_X(x)$ et $\lim_{x\to a,x>a} F_X(x)$ existent, et la deuxième vaut $F_X(a)$. On dénotera par la suite $\lim_{x\to a,x<a} f(x)$ par $f(a^-)$.
iii) $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to+\infty} F_X(x) = 1$.

■ La probabilité d'un intervalle.

Par définition, $\mathbb{P}(a<X\leq b) = F_X(b)-F_X(a)$. Pour avoir $P(a\leq X\leq b)$, il faut remplacer $F_X(a)$ par $F_X(a^-)$. De même, pour avoir $\mathbb{P}(a<X<b)$, il faut remplacer $F_X(b)$ par $F_X(b^-)$. Cela nous dit que $\mathbb{P}(X = a) = \mathbb{P}(a\leq X\leq a) = F_X(a) - F_X(a^-)$.

■ Les discontinuités.

Si en un point, $F_X(x)$ est différent de $F_X(x^-)$, alors x est un atome de X; cela signifie que $\mathbb{P}(X = x)$ n'est pas nul, et vaut $F_X(x)-F_X(x^-)$. On montre qu'il y a au plus un nombre dénombrable d'atomes.

■ Le cas absolument continu.

On dira que X est absolument continue si il existe $f_X:\mathbb{R}\to\mathbb{R}^+$ fonction continue par morceaux, et intégrable sur $\mathbb{R}$, telle que

$$F_X(t) = \int_{-\infty}^{t} f_X(x)\,\mathrm{d}x$$

si c'est le cas, f_X est appelée la densité de probabilité de X. Alors, $\int f_X(x)dx = 1$, F_X est continue (X sans atomes), et si f est continue en x_0, alors F_X est dérivable en x_0, avec $F_X'(x_0) = f_X(x_0)$.

D'ailleurs, pour n'importe quelle $f:\mathbb{R}\to\mathbb{R}^+$ continue par morceaux est intégrable avec $\int f(x)dx = 1$; alors on peut trouver une v.a. qui a pour densité f.

■ L'unicité.

Dans le cas des densités de probabilités, $\mathbb{P}_X = \mathbb{P}_Y$ *ssi* $f_X = f_Y$ sauf (possiblement) sur un ensemble de points isolés.

■ La condition utile.

Si F_X est continue, et continûment dérivable par morceaux, alors X est absolument continue et $f_X(x) = F_X'(x)$ en tout point où F_X est dérivable.

■ Les exemples.

Voici quelques lois absolument continues :
- pour une loi uniforme de paramètre [a,b] (on note $X\sim\mathcal{U}([a,b])$) : $f_X(x) = 1/(b-a)\, \mathbb{1}_{[a,b]}(x)$
- la loi exponentielle de paramètre $\lambda(>0)$ (notée $X\sim\mathcal{Exp}(\lambda)$) :
$$f_X(x) = \lambda e^{-\lambda x}\, \mathbb{1}_{]0,+\infty[}(x)$$
- la loi Gamma de paramètre a,b ($X\sim\gamma(a,b)$) :
$$f_X(x) = b^a x^{a-1}/\Gamma(a)\, e^{-bx}\, \mathbb{1}_{]0,+\infty[}(x)$$
- la loi normale de paramètre m et $\sigma^2(>0)$ ($X\sim\mathcal{N}(m,\sigma^2)$) :
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

En particulier, pour m = 0 et $\sigma^2 = 1$, la loi est dite normale centrée (m=0) et réduite (σ^2=1).

■ L'espérance de g(X).

Si $g:\mathbb{R}\to\mathbb{R}^+$ est une fonction continue par morceaux, $\mathbb{E}[g(X)]$ est un réel ou $+\infty$ donné par $\int_{\mathbb{R}} g(x)\, f_X(x)\, dx$. Plus généralement, pour g à valeurs dans R entier, g(X) est dite intégrable *ssi* $\mathbb{E}[|g(X)|]<+\infty$, ce qu'on note $g(X)\in L^1(\Omega,\mathcal{A},\mathbb{P})$ et dans ce cas, la définition plus haut fonctionne toujours. Cas particulier : X admet un moment d'ordre k quand $X^k\in L^1(\Omega,\mathcal{A},\mathbb{P})$, et alors il est donné par $\mathbb{E}[X^k]$.

■ La méthode de la fonction muette.

Elle sert à déterminer qu'une v.a. est absolument continue, et sa densité. Si il existe une fonction $f:\mathbb{R}\to\mathbb{R}^+$ continue par morceaux telle que $\mathbb{E}[h(X)] = \int h(x)f(x)\, dx$ pour toute fonction $h:\mathbb{R}\to\mathbb{R}$ continue par morceaux et bornée, alors X est absolument continue, de densité f.

→ Vecteur aléatoire.

Souvent les objets étudiés ont plusieurs paramètres ; donc les variables étudiées sont en fait des vecteurs !

■ La motivation.

D – Comment ferait-on pour modéliser un lancer de fléchettes ? C'est 2-dimensionnel, non ?
T – On peut modéliser ça par une paire de variables aléatoires (à vals dans $\mathbb{R}$) sur le même espace, correspondant à abscisse et ordonnée ; ou bien comme une seule variable aléatoire à valeurs vectorielles, dans $\mathbb{R}^2$.
D – Y a-t-il une différence importante ?
T – Les deux donnent les mêmes résultats ; après, pour un changement de variable par exemple, le point de vue vectoriel est plus utile. Dans notre cas, pour passer de coordonnées cartésiennes en coordonnées polaires.

■ Les couples de v.a.

Pour X, Y deux v.a. (à valeurs dans E et F resp.) sur le même espace de probabilité $(\Omega,\mathbb{P})$ discret, le couple (X,Y) à valeurs dans E×F est un exemple simple de vecteur aléatoire, où $(X,Y)(\omega) = (X(\omega),Y(\omega))$. Et vice-versa : si une v.a. discrète Z est à valeurs dans E×F, on peut définir ses lois marginales X et Y telles que $Z(\omega) = (X(\omega),Y(\omega))$ pour tout $\omega\in\Omega$. En particulier, $p_X(x) = \Sigma_{y\in F}\, p_{(X,Y)}(x,y)$ pour tout $x\in E$, et vice-versa.

■ L'indépendance.

Deux variables X et Y sont indépendantes *ssi* $p_X(x)\cdot p_Y(y) = p_{(X,Y)}(x,y)$, et ce pour tout $(x,y)\in E\times F$. Si $X_1, \dots X_n, Y_1, \dots Y_m$ sont des v.a indépendantes, alors $(X_1, \dots X_n)$ et $(Y_1, \dots Y_m)$ aussi; si X et Y sont indépendantes, alors f(X) et g(Y) aussi. Une autre condition pour l'équivalence : si $\forall f:E\to\mathbb{R}^+$ et $\forall g:F\to\mathbb{R}^+$ on a $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\cdot\mathbb{E}[g(Y)]$; on peut même remplacer "$\forall f:E\to\mathbb{R}^+$ et $\forall g:F\to\mathbb{R}^+$" par "$\forall f:E\to\mathbb{R}$ et $\forall g:F\to\mathbb{R}$, avec $f(X),G(Y)\in L^1(\Omega,\mathbb{P})$".

■ Le produit de convolution.

Quand X et Y sont à valeurs dans $\mathbb{N}$, on peut poser le produit de convolution de p_X et p_Y, par

$$\begin{aligned}(p_X \star p_Y)(n) &= \sum_{k=0}^{n} p_X(k)\cdot p_Y(n-k)\\ &= \sum_{k=0}^{n} \mathbb{P}(X=k)\cdot\mathbb{P}(Y=n-k)\end{aligned}$$

Il donne $p_{X+Y}(n)$, c'est-à-dire $\mathbb{P}(X+Y=n)$.

■ Les vecteurs aléatoires abs. cont.

Pour X, Y deux v.a. réelles, le couple (X,Y) est dit absolument continu si il existe une fonction $f:\mathbb{R}^2\to\mathbb{R}$ continue par morceaux, telle que $\mathbb{P}(X\in I, Y\in J) = \int_{I\times J} f(x,y)\,dx\,dy$ (et telle que $\int_{\mathbb{R}\times\mathbb{R}} f(x,y)\,dx\,dy = 1$). Dans ce cas, X et Y sont absolument continues et pour tout $(x,y)\in\mathbb{R}^2$, on a $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dy$, et $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dx$.

■ L'espérance de h(X,Y).

Soit (X,Y) un couple de variables aléatoires absolument continues, de densité $f_{X,Y}$, et $h:\mathbb{R}^2\to\mathbb{R}^+$ continue par morceaux; on appelle espérance de h(X,Y) (notée $\mathbb{E}[h(X,Y)]$) la quantité $\int_{\mathbb{R}^2} h(x,y)\, f_{X,Y}(x,y)\,dx\,dy$. Dans le cas de $h:\mathbb{R}^2\to\mathbb{R}$ continue par morceaux, si $\mathbb{E}[|h(X,Y)|] < +\infty$, alors on peut reprendre la définition de $\mathbb{E}[h(X,Y)]$ donnée plus haut.

■ L'indépendance (2).

Dans le cas continu, pour deux variables aléatoires absolument continues X et Y, X et Y sont indépendantes si pour tous I, J intervalles de $\mathbb{R}$, on a $\mathbb{P}(X\in I, Y\in J) = \mathbb{P}(X\in I)\cdot\mathbb{P}(Y\in J)$. Autre résultat : X et Y sont indépendantes *ssi* (X,Y) est absolument continue et si $f_{X,Y} = f_X\times f_Y$. Et si X et Y sont bien indépendantes, alors pour toutes fonctions h,g continues par morceaux et bornées de $\mathbb{R}$ dans $\mathbb{R}$, $\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$.

→ Lois mixtes.

Toutes les lois ne sont pas forcément discrètes ou absolument continues ; le cas général est en fait la loi mixte.

■ La motivation.

T – Te rappelles-tu de la définition générale (sans dérivées) d'une fonction $\phi:\mathbb{R}\to\mathbb{R}$ convexe ?
D – Euh … pour tous a, b et t entre 0 et 1, on a *[écrit au tableau]* $\phi(a\cdot t+b\cdot(1-t)) \leq t\cdot\phi(a) + (1-t)\cdot\phi(b)$.
T – On généralise sans problème : pour α, β … ω positifs de somme 1, on a *[au tableau]* $\phi(\alpha x_1 + \ldots + \omega x_n) \leq \alpha\cdot\phi(x_1) + \ldots + \omega\cdot\phi(x_n)$, quelque soient les $x_1, \ldots x_n$ réels.
D – Qu'on montre par récurrence ?
T – Bien sûr. Bref, ça se généralise aux intégrales et au v.a. ; on l'appelle alors l'inégalité de Jensen : $\phi(\mathbb{E}[X])\leq\mathbb{E}[\phi(X)]$.

■ Les lois mixtes.

Une loi mixte est une loi qu'on peut écrire comme

$$\mathbb{P}(X \in J) = \sum_{i\in I} p_i \mathbf{1}_J(x_i) + \int_J f(x)\,\mathrm{d}x$$

avec $f:\mathbb{R}\to\mathbb{R}^+$ une fonction continue par morceaux, et $\{x_i\}_{i\in I}$ et $(p_i)_{i\in I} \in[0,1]^I$ un ensemble de points distincts et une suite de coefficients de R. On note alors $\mathbb{P}_X = \Sigma_{i\in I}\, p_i\delta_x^i + f(x)dx$ la loi de X. En particulier, les v.a. discrètes et les v.a. absolument continues sont mixtes (f = 0 et I = Ø respectivement). En général, on peut écrire la loi de X comme une combinaison convexe d'une loi discrète et d'une loi absolument continue (où le coefficient λ utilisé est $\lambda = \Sigma_{i\in I}\, p_i$). On peut là aussi définir, quand g≥0, une espérance de g(X) par

$$\mathbb{E}[g(X)] = \sum_{i\in I} p_i g_i(x) + \int_{\mathbb{R}} g(x)f(x)dx$$

Et réutiliser la définition pour g quelconque si $\mathbb{E}[|g|]<+\infty$, ce qu'on note encore $g\in L^1(\Omega, \mathcal{A}, \mathbb{P})$.

■ Les propriétés.

Donc pour X, Y deux v.a. de loi mixte, alors :
i) si X≥0 et Y≥0, et Y≥X, alors $\mathbb{E}[Y]\geq\mathbb{E}[X]$. Plus généralement, si X, $Y\in L^1(\Omega, \mathcal{A}, \mathbb{P})$ et Y≥X, alors $\mathbb{E}[Y]\geq\mathbb{E}[X]$.
ii) si X, Y≥0 ou X, $Y\in L^1(\Omega, \mathcal{A}, \mathbb{P})$, alors $\mathbb{E}[\mu X+Y] = \mu\mathbb{E}[X]+\mathbb{E}[Y]$ pour tout μ.
iii) si $X\in L^1(\Omega, \mathcal{A}, \mathbb{P})$, alors $|\mathbb{E}[X]|\leq\mathbb{E}[|X|]$.

■ Les définitions supplémentaires.

Si $X\in L^2(\Omega,\mathcal{A},\mathbb{P})$, alors on définit alors $\mathrm{Var}(X) = \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ la variance de X. On dit que X et Y deux v.a. sont indépendantes si pour tous intervalles I, J de $\mathbb{R}$, on a $\mathbb{P}(X\in I, Y\in J) = \mathbb{P}(X\in I)\cdot\mathbb{P}(Y\in J)$; et alors pour toutes fonctions h,g de $\mathbb{R}$ dans $\mathbb{R}$ continues par morceaux avec h(X) et g(Y) dans $L^1(\Omega,\mathcal{A},\mathbb{P})$, alors $h(X)g(Y)\in L^1(\Omega,\mathcal{A},\mathbb{P})$ et $\mathbb{E}[h(X)g(Y)] = \mathbb{E}[h(X)]\times\mathbb{E}[g(Y)]$.

■ Les inégalités courantes.

- Pour X une v.a. à loi mixte, on a, pour tout t>0, l'inégalité de Markov :

  $$\mathbb{P}(|X|\geq t) \leq \mathbb{E}[|X|]/t$$

- Et quand $X\in L^2(\Omega,\mathcal{A},\mathbb{P})$, alors on a l'inégalité de Bienaymé-Tchebychev :

  $$\mathbb{P}(|X-\mathbb{E}[X]|>t) \leq \mathrm{Var}(X)/t^2$$

- Et voici aussi deux inégalités qui viennent de l'intégration : l'inégalité de Cauchy-Schwarz dit que, pour tout $X,Y\in L^2(\Omega,\mathcal{A},\mathbb{P})$ à loi mixte, alors $X\cdot Y\in L^2(\Omega,\mathcal{A},\mathbb{P})$, et

  $$\mathbb{E}[|XY|]\leq\mathbb{E}[X]^{½}\cdot\mathbb{E}[Y]^{½}$$

- On a aussi l'inégalité de Jensen : pour $X\in L^1(\Omega,\mathcal{A},\mathbb{P})$, à valeurs dans un intervalle $I\subseteq\mathbb{R}$, et $f:I\to\mathbb{R}$ une fonction convexe avec $f(X)\in L^1(\Omega,\mathcal{A},\mathbb{P})$, alors :

  $$f(\mathbb{E}[X])\leq\mathbb{E}[f(X)]$$

→ Fonctions génératrices et caractéristiques.

Ces manières de décrire la loi de probabilité d'une v.a. permettent de trouver certaines sommes pénibles, plus vite !

■ La motivation.

D – Pour retenir une loi de Poisson, on peut se rappeler du développement de exp(λ), et on multiplie tout par exp(–λ).
T – Oui ; et on peut même généraliser cette méthode. On peut décrire toute loi discrète (à valeurs dans ℕ) comme une unique fonction, dont les coefficients du développement limité en 0 donne les $\mathbb{P}(X = k)$. Donc pour la loi de Poisson, on peut montrer par le calcul que g(s) = exp(λs)·exp(-λ) = exp(λ(s-1)).
D – Pour une loi binomiale, ce serait donc $g(s) = ((1-p)+s\cdot p)^n$; et ensuite ?
T – Et grâce à ça, on montre que la somme de deux lois de Poisson est une autre loi de Poisson !

■ Les fonctions génératrices.

Pour X une variable aléatoire discrète à valeurs dans ℕ, on peut définir une application g_X dite fonction génératrice, par

$$g_X(s) = \sum_{n\in\mathbb{N}} \mathbb{P}(X = n)s^n = \mathbb{E}[s^X]$$

On vérifie que $g_X(1) = 1$; donc la série qui définit g_X converge sur [-1;1] (et plus généralement, sur un intervalle]-R;R[pour un R≥1).

■ L'unicité.

On a que X = Y (en loi) ⇔ $\mathbb{P}_X = \mathbb{P}_Y$ ⇔ $g_X = g_Y$ sur]-1,1[. En effet, on peut obtenir les $\mathbb{P}(X = n)$ en dérivant g_X n fois en 0, en calculant $g_X^{(n)}(0)/n!$ (c'est le cas de toutes les séries ; on trouve les coefficients de cette manière).

■ Les moments.

On utilisera la notation $f(a^-)$ pour désigner la limite :

$$\lim_{x \xrightarrow{<} a} f(x)$$

Les moments factoriels de X sont définis comme $\mathbb{E}[X(X-1)\ldots(X-r)]$; on peut les obtenir, quand ils existent, comme $g_X^{(r)}(1^-)$. En particulier, l'espérance $\mathbb{E}[X]$ est donnée (quand elle existe), par $g_X'(1^-)$, et la variance (pour rappel, $\mathbb{E}[X^2]-\mathbb{E}[X^2]$) par $G''(1^-)+G'(1^-) - (G'(1^-))^2$.

■ La somme de v.a. indépendantes.

Si X et Y sont deux v.a. discrètes à valeurs dans ℕ indépendantes, alors $g_{X+Y}(s) = g_X(s) \times g_Y(s)$ sur tout [-1;1].

■ Les vecteurs aléatoires.

On peut utiliser les fonctions génératrices pour les vecteurs aléatoires; dans ce cas, on a plusieurs variables. En détail, si $X = (X_1, \ldots, X_n)$ est un vecteur aléatoire à valeurs dans $\mathbb{N}^n$, alors :

$$g_X(s_1, \cdots, s_n) = \mathbb{E}[s_1^{X_1} \times \cdots \times s_n^{X_n}]$$

Et derechef, si X_1, … X_n sont indépendantes, alors on a l'égalité $g_X(s_1, \ldots, s_n) = g_{X_1}(s_1)\times\ldots\times g_{X_n}(s_n)$ pour tout $s_1, \ldots s_n$.

■ Les fonctions caractéristiques.

Pour X une v.a. à loi mixte, on appelle fonction caractéristique de X, notée ϕ_X, la fonction qu'on définit comme $\phi_X(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)]+i\mathbb{E}[\sin(tX)]$. On a que $\phi_X(0) = 1$, et pour tout t réel, $|\phi_X(t)|\leq 1$.
On a, derechef, l'unicité : $\mathbb{P}_X = \mathbb{P}_Y$ *ssi* $\phi_X = \phi_Y$. Et si X et Y sont indépendantes, alors $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$ pour tout t.

→ Convergences.

Qu'entend-on par le fait qu'une suite de v.a. converge ? Comme pour les fonctions, il y a plusieurs définitions différentes.

La motivation.

D – Il est bien connu que si on lance plusieurs fois une pièce équilibrée, on se rapprochera de plus en plus d'une proportion 50-50 des piles et faces. Comment rédiger et justifier ça, mathématiquement ?

T – Ce qu'on doit dire, c'est que si X_n correspond à « pile au lancer n », alors $(X_1+\ldots+X_n)/n$ converge, presque sûrement, vers ½.

D – Mais si on ne fait pas une infinité de lancers ?

T – On peut être plus précis, et approximer la loi de $S_n = X_1 + \ldots + X_n$ grâce au théorème central limite. En particulier, $\sqrt{n}\cdot(S_n/n - ½)$ converge en loi, vers une v.a. de loi normale $\mathcal{N}(0, ¼)$. Et on peut approximer, par exemple, $\mathbb{P}(0{,}4 \leq S_n/n \leq 0{,}6)$.

Le cadre.

On étudie une suite $(X_n)_n$ de v.a. réelles à loi mixte, définies sur le même espace de probabilité $(\Omega,\mathcal{A},\mathbb{P})$, et X une autre v.a. réelle à loi mixte sur le même espace. On définit plusieurs "convergence de $(X_n)_n$ vers X", comme pour les suites de fonctions.

La convergence en probabilité.

On dit que $(X_n)_n$ converge en probabilité vers X si

$$\forall \varepsilon>0,\ P(|X_n-X|>\varepsilon) \to 0$$

Il est même possible de métriser cette convergence; dans le sens où on peut trouver une distance d sur l'ensemble des fonctions $(\Omega,\mathcal{A},\mathbb{P})\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$, telle que "$X_n$ converge en probabilité vers X" s'écrive "$d(X_n,X) \to 0$".

La convergence presque sûre.

On dit que $(X_n)_n$ converge $\mathbb{P}$-presque sûrement vers X si

$$\mathbb{P}\left(\left\{\omega \in \Omega : X_n(\omega) \xrightarrow[n\to+\infty]{X} (\omega)\right\}\right) = 1$$

La convergence en loi.

On dit que $(X_n)_n$ converge en loi vers X si pour tout $t\in\mathbb{R}$ point de continuité de F_X, on a

$$F_{X_n}(t) \to F_X(t)$$

La loi faible des grands nombres.

Pour une suite $(X_i)_{i\geq1}$ telle que pour tout i, $X_i\in L^2(\Omega,\mathcal{A},\mathbb{P})$, $\mathbb{E}[X_i] = \mathbb{E}[X_1]$ et $\mathrm{Var}(X_i) = \mathrm{Var}(X_1)$; et pour tout $i\neq j$, $\mathrm{Cov}(X_i,X_j) = 0$ (vrai si les X_i sont indépendants), alors

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{\text{Proba}} \mathbb{E}[X_1]$$

La loi forte des grands nombres.

Soit $(X_i)_{i\geq1}$, une suite de v.a. i.i.d. dans $L^2(\Omega,\mathcal{A},\mathbb{P})$, alors :

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{\mathbb{P}\text{-p.s.}} \mathbb{E}[X_1]$$

Les implications.

La convergence $\mathbb{P}$-p.s. implique la convergence en probabilité, qui implique la convergence en loi.

Le théorème de Paul Lévy.

Soit $(X_n)_{n\geq1}$ suite de v.a. alors $\phi_{X_n}(t) \to \phi_X(t)$ pour tout $t\in\mathbb{R}$ est équivalent à ce que X_n converge en loi vers X.

Le théorème central limite.

Soit $(X_i)_{i\geq1}$ une suite de v.a. i.i.d. à loi mixte dans $L^2(\Omega,\mathcal{A},\mathbb{P})$, alors avec $\sigma^2 = \mathrm{Var}(X_1)$, on a :

$$n^{\frac{1}{2}}\left(n^{-1}\sum_{i=1}^{n} X_i - \mathbb{E}[X_1]\right) \xrightarrow[n\to+\infty]{\text{Loi}} \mathcal{N}(0,\sigma^2)$$