Séance 2 (mercredi 9 décembre)

Exercice 1 (Loi de Pareto) Soit Y une v.a. de loi exponentielle de paramètre λ. La variable aléatoire $X = e^Y$ est appelée *variable aléatoire de Pareto*, de loi de Pareto $\mathcal{P}(\lambda, 1)$.

1. On pose $r(x) = \mathbb{P}(X \geq x)$ (c'est la *fonction de survie* de X). Montrer que r est donnée par :
$$r(x) = \begin{cases} x^{-\lambda}, & \text{si } x \geq 1; \\ 1, & \text{si } x < 1. \end{cases}$$
et la densité par :
$$f(x) = \begin{cases} \lambda / x^{\lambda+1}, & \text{si } x \geq 1; \\ 0, & \text{si } x < 1. \end{cases}$$
2. Calculer $\mathbb{E}(X)$ (attention, il y a deux cas à distinguer).
3. Pour tout entier $k \geq 1$ évaluer $\mathbb{E}(X^k)$ en déterminant la fonction de survie de X^k.
4. On appelle *fonction génératrice* de X la fonction suivante : $g(u) = \mathbb{E}(e^{uX})$ pour tout $u \in \mathbb{R}$. Montrer que pour tout $u \in \mathbb{R}$:
$$g(u) = \sum_{k=0}^{+\infty} \mathbb{E}(X^k) \frac{u^k}{k!}.$$
5. En déduire que $g(u) < +\infty$ si et seulement si $u = 0$.

Réponse. Corrigé le 11 décembre. La loi de Pareto a un moment d'ordre k si et seulement si $k < \lambda$. Ainsi si $\lambda \leq 1$, la loi des grands nombres (LGN) ne s'applique pas. Si $1 < \lambda \leq 2$, la LGN fonctionne mais pas le théorème de la limite centrale (TCL). Et si $\lambda > 2$, le TCL et la LGN s'appliquent ! $\square$

Exercice 2 Soient X et Y deux variables aléatoires indépendantes et suivant une même loi normale centrée réduite.

1. Montrer que $C = X/Y$ suit une loi de Cauchy.
2. Montrer que $U = \dfrac{X^2 + Y^2}{2}$ suit une loi exponentielle dont on calculera le paramètre.
3. On considère $V = \dfrac{X^2}{X^2 + Y^2}$. Calculer la loi de (U, V). Montrer que U et V sont indépendants. Donner la loi de V (on calculera sa densité). On appelle cette dernière loi *"loi Arc sinus"*.
4. Soit C une variable suivant une loi de Cauchy de paramètre 1. Montrer que $(1 + C^2)^{-1}$ suit une loi Arc sinus.

Réponse.

1. Il y a deux façons de calculer la loi de Z.
   (a) Méthode 1 : Considérer $\psi(x,y) = (t = y/x, u = x)$ qui est un difféomorphisme de $D = \mathbb{R} \times \mathbb{R}^*$ dans $\Delta = \mathbb{R} \times \mathbb{R}^*$. On a $x = u$ et $y = tu$. Donc $J(\psi^{-1})(u,v) = -u$. Le couple (X,Y) a pour densité $\frac{1}{2\pi}e^{-(x^2+y^2)/2}\mathbf{1}_D(x,y)$. Donc
   $$\mathbb{E}(f(T,U)) = \int_D f(y/x,x)\frac{1}{2\pi}e^{-(x^2+y^2)/2}dxdy = \int_\Delta f(t,u)\frac{1}{2\pi}e^{-(u^2+(tu)^2)/2}|u|dudt.$$
   Donc (T,U) a pour densité $\frac{1}{2\pi}|u|e^{-u^2(1+t^2)/2}\mathbf{1}_\Delta(t,u)$. De là T a pour densité
   $$q(t) = \int_{\mathbb{R}} \frac{1}{2\pi}|u|e^{-u^2(1+t^2)/2}du = \frac{1}{\pi}\int_0^\infty e^{-u^2(1+t^2)/2}udu = \frac{1}{\pi(1+t^2)}.$$
   (b) Méthode 2 : calculer directement
   $$\begin{aligned}\mathbb{E}(f(T)) &= \frac{1}{2\pi}\int\int f(y/x)e^{-(x^2+y^2)/2}dxdy = \frac{1}{2\pi}\int_0^{2\pi}\int_0^\infty f(\tan(\theta))e^{-r^2/2}rd\theta dr \\ &= \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} f(\tan(\theta))d\theta = \frac{1}{\pi}\int_{-\infty}^\infty f(t)\frac{1}{1+t^2}dt.\end{aligned}$$
2. Méthode la plus efficace :
   $$\begin{aligned}\mathbb{E}(G(U,V)) &= \iint_{\mathbb{R}^2} G\left(\frac{x^2+y^2}{2}, \frac{x^2}{x^2+y^2}\right)\frac{1}{2\pi}e^{-(x^2+y^2)/2}dxdy \\ &= 4\iint_D G\left(\frac{x^2+y^2}{2}, \frac{x^2}{x^2+y^2}\right)\frac{1}{2\pi}e^{-(x^2+y^2)/2}dxdy\end{aligned}$$
   où $D =]0,+\infty[^2$. On fait le changement de variables : pour $(x,y) \in D$
   $$u = \frac{x^2+y^2}{2}, \quad v = \frac{x^2}{x^2+y^2},$$
   soit
   $$x = \sqrt{2uv}, \quad y = \sqrt{2u(1-v)}.$$
   Et $(u,v) \in]0,+\infty[\times]0,1[= \Delta$. On calcule la matrice jacobienne de ce chgt de variables :
   $$\begin{pmatrix} \frac{\sqrt{2v}}{2\sqrt{u}} & \frac{\sqrt{2u}}{2\sqrt{v}} \\ \frac{\sqrt{2(1-v)}}{2\sqrt{u}} & -\frac{\sqrt{2u}}{2\sqrt{1-v}} \end{pmatrix}$$
   dont le déterminant (jacobien) est $J(u,v) = \frac{-1}{2\sqrt{v(1-v)}}$. De là
   $$\begin{aligned}&4\iint_D G\left(\frac{x^2+y^2}{2}, \frac{x^2}{x^2+y^2}\right)\frac{1}{2\pi}e^{-(x^2+y^2)/2}dxdy = 4\iint_\Delta G(u,v)\frac{1}{2\pi}e^{-u}|J(u,v)|dudv \\ &= \iint_\Delta G(u,v)\frac{1}{\pi\sqrt{v(1-v)}}e^{-u}dudv = \iint_{\mathbb{R}^2} G(u,v)\frac{e^{-u}}{\pi\sqrt{v(1-v)}}\mathbf{1}_{]0,+\infty[}(u)\mathbf{1}_{]0,1[}(v)dudv.\end{aligned}$$
   Donc
   $$\mathbb{E}(G(U,V)) = \iint_{\mathbb{R}^2} G(u,v)h(u,v)dudv.$$

Le vecteur aléatoire (U, V) admet donc pour densité la fonction h. Celle-ci se décompose en

$$h(u,v) = \phi(u)\psi(v) = e^{-u}\mathbf{1}_{]0,+\infty[}(u)\frac{1}{\pi\sqrt{v(1-v)}}\mathbf{1}_{]0,1[}(v).$$

Donc U et V sont indépendantes et ϕ est (proportionnelle à) la densité de U (loi exponentielle de paramètre 1). ψ est (proportionnelle à) la densité de V.

$\square$

Exercice 3 Soit (X, Y) un couple de variables aléatoires dont la densité jointe est définie par :

$$f(x,y) = \begin{cases} 24xy, & x \geq 0,\ y \geq 0,\ x+y \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

1. Vérifier que $f(x, y)$ est bien une fonction de densité.
2. Déterminer les densités marginales de X et de Y. Les variables aléatoires X et Y sont-elles indépendantes ?

Réponse. L'ensemble

$$D = \{(x,y),\ x \geq 0,\ y \geq 0,\ x+y \leq 1\}$$

est un triangle et donc on ne peut pas écrire $f(x, y) = g(x)h(y)$. Donc les deux composantes ne sont pas indépendantes.

1. Pour tout (x, y), $f(x, y) \geq 0$ et on vérifie que $\iint_{\mathbb{R}^2} f(x, y)dxdy = 1$.
2. La densité de X est donné par

$$g(x) = \int_{\mathbb{R}} f(x,y)dy = \begin{cases} 0 & \text{si } x \notin [0,1] \\ \int_0^{1-x} 24xydy = 12x(1-x)^2 & \text{si } x \in [0,1]. \end{cases}$$

Notons que $g \geq 0$ et $\int_0^1 g(x)dx = 1$. C'est bien une densité ! La densité de Y est

$$h(y) = \int_{\mathbb{R}} f(x,y)dx = g(y).$$

Si X et Y étaient indépendantes alors $f(x, y) = g(x)h(y)$, ce qui n'est pas vrai.

$\square$

Les exercices 4, 5 et 6 sont moins importants pour l'agrégation interne. On rappelle que la convergence presque sûre implique la convergence en probabilité qui implique la convergence en loi, tandis que la convergence en moyenne (d'ordre p) implique la convergence en probabilité. Toutes les réciproques sont en général fausses.

Exercice 4 Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que X_n prend les valeurs 0, n et $-n$ avec probabilité $(1 - 1/n^2)$, $1/(2n^2)$ et $1/(2n^2)$. Montrer que cette suite tend vers 0 presque sûrement mais pas en moyenne quadratique.

Réponse. Commençons par calculer

$$\mathbb{E}(|X_n|^p) = n^p \frac{1}{n^2} = n^{p-2}.$$

Donc si $p \geq 2$, la suite X_n ne converge pas en moyenne d'ordre p vers 0. Mais si $p < 2$, alors il y a convergence en moyenne d'ordre p vers 0. En particulier il y a convergence en probabilité vers 0, ce qu'on peut vérifier à la main : pour tout $\varepsilon > 0$

$$\mathbb{P}(|X_n| \geq \varepsilon) = \mathbb{P}(|X_n| = n) = \frac{1}{n^2} \longrightarrow 0.$$

Pour montrer la convergence presque sûre, on utilise le résultat suivant : *s'il existe $r > 0$ tel que la série de terme géneral $\mathbb{E}(|X_n|^r)$ converge, alors la suite (X_n) converge vers 0 presque sûrement.* Ce résultat est lui même une conséquence de : *si pour tout $\varepsilon > 0$ la série de terme géneral $\mathbb{P}(|X_n| > \varepsilon)$ converge, alors la suite (X_n) converge vers 0 presque sûrement.* Ce sont les deux critères les plus rapides pour montrer de la convergencve presque sûre. Ici on peut utiliser $r = 1/2$ pour conclure. □

Exercice 5 Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes telles que X_n prend les valeurs 0 et n avec probabilité $(1 - 1/n)$ et $1/n$.

1. Montrer que cette suite tend vers 0 en probabilité.
2. On pose $Y_n = \dfrac{1}{n} \sum_{k=1}^n X_k$. Montrer que Y_n ne converge pas en probabilité vers zéro.
3. Qu'en déduit-on ?

Réponse.

1. Comme pour $\varepsilon > 0$
$$\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(X_n = n) = 1/n$$
la convergence en probabilité s'en déduit.
2. On remarque que l'événement $\{Y_{2n} \geq 1/2\}$ contient
$$\bigcup_{n<k\leq 2n} \{X_n = k\}.$$
Donc
$$\mathbb{P}(Y_{2n} \geq 1/2) \geq \mathbb{P}\left(\bigcup_{n<k\leq 2n} \{X_n = k\}\right) = 1 - \mathbb{P}\left(\bigcap_{n<k\leq 2n} \{X_n = 0\}\right) = 1 - \prod_{n<k\leq 2n} (1 - (1/k)).$$
Or $1 - 1/k \leq \exp(-1/k)$ et ainsi on peut minorer cette probabilité :
$$\mathbb{P}(Y_{2n} \geq 1/2) \geq 1 - \prod_{n<k\leq 2n} \exp(-1/k) = 1 - \exp(- \sum_{n<k\leq 2n} 1/k) \geq 1 - \exp(-1/2).$$
Ainsi $\mathbb{P}(Y_{2n} \geq 1/2)$ ne tend pas vers zéro lorsque n tend vers $+\infty$. Il n'y a pas convergence en probabilité.
3. Le théorème de Césaro n'est pas vrai pour la convergence en probabilité.

□

Exercice 6 Soient X une variable aléatoire à valeurs dans $[0, +\infty[$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de même loi que X.

1. Supposons que $\mathbb{P}(X > x) = o(1/x)$ lorsque x tend vers $+\infty$. Montrer que $Z_n = \frac{1}{n}\max(X_1, \ldots, X_n)$ tend en loi vers 0.
2. Si $\mathbb{P}(X > x) \sim \alpha/x^\lambda$ lorsque x tend vers $+\infty$, montrer que
$$Z_n = \frac{1}{n^{1/\lambda}}\max(X_1, \ldots, X_n)$$
tend en loi vers Y de loi de Fréchet, dont la fonction de répartition est $F_Y(x) = e^{-\alpha x^{-\lambda}}$, $x > 0$.

Réponse.

1. Remarquons d'abord que $Z_n \geq 0$ presque sûrement (comme maximum de v.a. positives). Pour tout $x > 0$
$$\begin{aligned}\mathbb{P}(Z_n \leq x) &= \mathbb{P}(\max(X_1, \ldots, X_n) \leq nx) = \mathbb{P}(\forall k,\ X_k \leq nx) = (\mathbb{P}(X \leq nx))^n \\ &= (1 - \mathbb{P}(X > nx))^n = (1 - o(1/nx))^n.\end{aligned}$$
Donc $\ln(\mathbb{P}(Z_n \leq x)) \leq n\ln(1 - o(1/nx))$ tend vers zéro quand n tend vers $+\infty$. Ainsi $\mathbb{P}(Z_n \leq x)$ tend vers 1, et ceci quel que soit $x > 0$. Donc la fonction de répartition de Z_n tend vers celle d'une v.a. nulle presque sûrement.
À titre d'exemple, on peut considérer X de loi exponentielle ou la valeur absolue de X de loi normale.
2. Le même calcul montre que
$$\mathbb{P}(Z_n \leq x) = (\mathbb{P}(X \leq n^{1/\lambda}x))^n = \left(1 - \frac{\alpha}{nx^\lambda} + o(1/n)\right)^n$$
qui tend vers $e^{-\alpha x^\lambda}$ lorsque n tend vers $+\infty$.
Comme exemple, on peut considérer la loi de Pareto (exercice 1).

□

Exercice 7 Soit f_n la fonction nulle en dehors de l'intervalle $[-1, 1]$ et definie pour tout $x \in [-1, 1]$ par
$$f_n(x) = \sin^2(2\pi nx).$$

1. Montrer que f_n est la densité d'une loi de probabilité $\mathbb{P}_n$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
2. Calculer son espérance et sa variance et montrer que sa fonction caractéristique est :
$$\phi_n(t) = \frac{\sin(t)}{t}\left(1 - \frac{t^2}{t^2 - 16\pi^2n^2}\right)$$
pour $t \neq \pm 4\pi n$.
3. Montrer que si (X_n) est une suite de v.a. de loi $\mathbb{P}_n$, la suite X_n converge en loi vers la loi uniforme sur $[-1, 1]$. A-t-on convergence de la suite (f_n) ? De la suite des variances de $\mathbb{P}_n$?

Réponse.

1. Pour tout $x \in \mathbb{R}$, $f_n(x) \geq 0$ et
$$\int_{\mathbb{R}} f_n(x)dx = \int_{-1}^{1} \sin^2(2\pi nx)dx = \int_{-1}^{1} \frac{1-\cos(4\pi nx)}{2}dx = 1.$$
C'est une densité.

2. Sa moyenne est nulle :
$$\int_{-1}^{1} x f_n(x)dx = 0.$$
Pour sa variance, par intégration par parties :
$$\int_{-1}^{1} x^2 f_n(x)dx = \int_{-1}^{1} x^2 \frac{1-\cos(4\pi nx)}{2}dx = \frac{1}{3} - \frac{1}{2}\int_{-1}^{1} x^2\cos(4\pi nx)dx = \frac{1}{3} - \frac{1}{8\pi^2 n^2}.$$
Fonction caractéristique :
$$\mathbb{E}e^{itX} = \int_{-1}^{1} e^{itx}\sin^2(2\pi nx)dx = \int_{-1}^{1} e^{itx}\frac{1-\cos(4\pi nx)}{2}dx = \frac{\sin(t)}{t} - \frac{1}{2}e^{itx}\cos(4\pi nx)dx.$$
La formule d'Euler permet de conclure.

3. La fonction ϕ_n converge simplement vers
$$\phi(t) = \frac{\sin(t)}{t} = \int_{-1}^{1} e^{itx}\frac{1}{2}dx.$$
ϕ est la fonction caractéristique de la loi uniforme sur $[-1, 1]$. La convergence simple des fonctions caractéristique entraîne la convergence en loi souhaitée. Et la variance converge vers la variance de la loi uniforme sur $[-1, 1]$ (qui vaut $1/3$).
En revanche la suite de fonctions f_n ne converge pas simplement. D'un point de vue analyse, on a une suite de fonctions qui ne converge pas (au sens de la convergence simple), mais qui converge au sens des distributions.

$\square$

Exercice 8 (Théorème de Weierstrass)

Le théorème de Bernoulli fournit une démonstration remarquable, due à Bernstein, du théorème d'approximation de Weierstrass, qui affirme qu'une fonction *continue* sur un intervalle *borné* peut être approchée par des polynômes, et ceci uniformément sur cet intervalle.

Soit $(X_n)_{n\geq 1}$ une suite de v.a. i.i.d. de loi commune de Bernoulli de paramètre $p \in]0, 1[$. On pose $Y_n = (1/n)\sum_{k=1}^{n} X_k$.

1. Rappeler l'énoncé du théorème de Bernoulli.
2. Soit $h : [0, 1] \to \mathbb{R}$ une fonction continue. Montrer que $\mathbb{E}(h(Y_n)) \to h(p)$ $(n \to +\infty)$ et ceci uniformément en $p \in]0, 1[$.

Réponse.

1. C'est la loi des grands nombres. Presque sûrement Y_n converge vers $\mathbb{E}(X_1) = p$ quand n tend vers $+\infty$.

2. Comme h est continue sur un compact elle est bornée. Donc comme $h(Y_n)$ converge presque sûrement vers p, le théorème de convergence dominée nous dit que $\mathbb{E}(h(Y_n)) \to h(p)$ $(n \to +\infty)$. Or

$$\mathbb{E}(h(Y_n)) = \sum_{k=0}^{n} h\left(\frac{k}{n}\right) \left(\begin{array}{c} n \\ k \end{array}\right) p^k (1-p)^{n-k} = H_n(p),$$

car $\sum_{k=1}^{n} X_k$ suit une loi binomiale (n,p). Et H_n est un polynôme en p ! Cette suite de polynômes converge donc simplement sur $]0,1[$ vers $h(p)$.

Pour obtenir la convergence uniforme, notons que

$$|H_n(p) - H(p)| \leq \sum_{j=0}^{n} |h(j/n) - h(p)| \mathbb{P}(Y_n = j/n).$$

Comme h est continue sur $[0,1]$, elle est bornée et uniformément continue. Donc il existe M tel que $|h(x)| \leq M$ pour tout x et pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $x \in [0,1]$, $|x-p| \leq \delta$ implique que $|h(x) - h(p)| \leq \varepsilon$. Donc

$$\begin{aligned} |H_n(p) - H(p)| &\leq \varepsilon \sum_{0 \leq j \leq n,\ |j/n-p| \leq \delta} \mathbb{P}(Y_n = j/n) + 2M \sum_{0 \leq j \leq n,\ |j/n-p| > \delta} \mathbb{P}(Y_n = j/n) \\ &\leq \varepsilon + 2M\mathbb{E}(|Y_n - p| \geq \delta) \leq \varepsilon + 2M \frac{\mathbb{E}(|Y_n|^2)}{\delta^2} = \varepsilon + 2M \frac{np(1-p)}{n^2\delta^2}. \end{aligned}$$

Pour la dernière inégalité, on a utilisé l'inégalité de Bienaymé-Tchebichev.

□

Notons que d'autres resultats d'analyse s'obtiennent facilement avec le TCL, comme :

— Bernstein :

$$\lim_{n\to+\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

— Formule de Stirling.