SUJETS DE TD

Table des matières

Exercice 1. — *Convergences en probabilité.*

(1) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoire de loi exponentielle de paramètre n. Démontrer, en utilisant la définition, que $X_n \underset{P}{\longrightarrow} 0$

Définition de la convergence en probabilité :

$$\forall\, \varepsilon > 0 \ \lim_{n\to+\infty} P\left(|X_n - X| > \varepsilon\right) = 0$$

Soit $\varepsilon > 0$. Ici, $X = 0$

$$\begin{aligned} P\left(|X_n| > \varepsilon\right) &= P\left(\varepsilon < X_n\right) \\ &= \int_\varepsilon^{+\infty} n e^{-nx} \mathrm{d}\,x \\ &= \left[-e^{-nx}\right]_\varepsilon^{+\infty} \\ &= e^{-n\varepsilon} \end{aligned}$$

comme pour $\varepsilon > 0$, $\lim_{n\to+\infty} e^{-n\varepsilon} = 0$, on a bien $\lim_{n\to+\infty} P\left(|X_n| > \varepsilon\right) = 0$.

Remarque :

$$\int_0^{+\infty} n e^{-nx} \mathrm{d}\,x = 1 \ \forall n > 0$$

Alors que $\lim_{n\to+\infty} ne^{-nx} = 0$ et donc $\int_0^{+\infty} \lim_{n\to+\infty} e^{-n\varepsilon} = 0$. On a un exemple tel que

$$1 = \lim_{n\to+\infty} \int_0^{+\infty} ne^{-nx} \mathrm{d}\,x \neq \int_0^{+\infty} \lim_{n\to+\infty} ne^{-nx} \mathrm{d}\,x = 0$$

(2) Soit $(p_n)_{n\geq 1}$ une suite dans $[0,1]$ qui tend vers 0. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires de Bernoulli de paramètre p_n. En utilisant la définition, montrer que (X_n) converge vers 0 en probabilité.

$$P\left(|X_n| > \varepsilon\right) = P\left(X_n > \varepsilon\right)$$

X_n ne prend que deux valeurs : 0 ou 1. On a $P(X_n = 1) = p_n$ donc, pour $\varepsilon > 0$, $P\left(X_n > \varepsilon\right) = p_n$, ce qui donne

$$\lim_{n\to+\infty} P\left(|X_n| > \varepsilon\right) = 0$$

par hypothèse sur p_n.

(3) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoire indépendantes de loi $\mathcal{N}(0,1)$. Soit

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 e^{-\frac{X_i^2}{3}} \ .$$

Démontrer que S_n converge en probabilité vers une limite que l'on précisera.

Les X_i étant indépendantes, les $X_i^2 e^{-\frac{X_i^2}{3}}$ sont aussi indépendantes. On peut donc appliquer la loi faible des grands nombres :

$$\forall\, \varepsilon > 0 \ \lim_{n\to+\infty} P\left(\left|S_n - \mathrm{E}\left[\mathrm{X}_1^2 \mathrm{e}^{-\frac{\mathrm{X}_1^2}{3}}\right]\right| > \varepsilon\right) = 0$$

On calcule l'espérance voulue

$$\begin{aligned}
\mathrm{E}\left[\mathrm{X}_1^2 \mathrm{e}^{-\frac{\mathrm{X}_1^2}{3}}\right] &= \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{3}} f_{X_1}(x) \mathrm{d}\,x \\
&= \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{3}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}\,x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{5x^2}{6}} \mathrm{d}\,x \\
&= \sqrt{\frac{3}{5}} \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{\frac{3}{5}2\pi}} e^{-\frac{x^2}{2\frac{3}{5}}} \mathrm{d}\,x \\
&= \sqrt{\frac{3}{5}} \frac{3}{5}
\end{aligned}$$

Exercice 2. — *Convergences en loi d'une variable de Bernoulli.*

(1) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoire de loi binomiale de paramètre n et $\frac{\lambda}{n}$. En utilisant les fonctions de répartitions, étudier la convergence en loi de (X_n).

$$\begin{aligned} F_{X_n}(x) &= P(X_n \leq x) = \sum_{k=0}^{x} P(X_n = k) \\ &= \sum_{k=0}^{x} \binom{n}{k} \frac{\lambda}{n}^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

et on s'intéresse à la limite en $+\infty$. On réécrit

$$\begin{aligned} F_{X_n}(x) &= \sum_{k=0}^{x} \frac{\lambda^k}{k!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} \frac{n!}{n^k(n-k)!} \\ &= \sum_{k=0}^{x} \frac{\lambda^k}{k!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} \frac{n(n-1)\dots(n-k+1)}{n^k} \end{aligned}$$

On a

$$\lim_{n\to+\infty} \left(1 - \frac{\lambda}{n}\right)^k = 1$$

$$\lim_{n\to+\infty} \frac{n!}{n^k(n-k)!} = 1$$

De plus

$$\begin{aligned} \left(1 - \frac{\lambda}{n}\right)^n &= e^{n\ln\left(1-\frac{\lambda}{n}\right)} \\ &= e^{-\frac{\lambda}{N}\ln(1+N)} \quad \text{avec } N = -\frac{\lambda}{n} \end{aligned}$$

or

$$\lim_{N\to 0} \frac{\ln(1+N)}{N} = 1$$

donc

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

Finalement

$$\lim_{n\to+\infty} F_{X_n}(x) = \sum_{k=0}^{x} \frac{\lambda^k}{k!} e^{-\lambda}$$

et on reconnait donc la distribution d'une loi de Poisson de paramètre λ.

(2) Application pratique : on lance une pièce truquée 100 fois. On suppose que la probabilité d'apparition de "pile" est $p = 0,05$. Quelle est la probabilité pour avoir 2 "piles" sans, et avec approximation ?

Il s'agit bien de 100 épreuve de Bernoulli de paramètre 0.05. Donc, sans approximation :

$$P(X = 2) = \binom{100}{2} 0.05^2 0.95^{100-2} \simeq 0.081$$

Avec approximation, on pose $\frac{\lambda}{n} = 0.05$ et $n = 100$, ce qui donne $\lambda = 5$.

$$P(X = 2) \simeq \frac{5^2}{2!}e^{-5} \simeq 0.084$$

Exercice 3. — *Convergences.*

Soit la suite de variables aléatoires X_n définie par

$$\begin{cases} P(X_n = 0) & = 1 - \frac{1}{n} \\ P(X_n = n) & = \frac{1}{n} \end{cases}$$

(1) Montrer que $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers $X = 0$

Remarque :

$$P(X_n = m) = 0 \ \forall m \neq n \text{ et } m \neq 0$$

Soit $\varepsilon > 0$

$$P(|X_n| > \varepsilon) = P(X_n = n) = \frac{1}{n}$$

donc

$$\lim_{n\to+\infty} P(|X_n| > \varepsilon) = 0$$

(2) En utilisant le théorème de Levy, montrer que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers $X = 0$

$$\varphi_{X_n}(t) = \mathrm{E}\left[\mathrm{e}^{\mathrm{itX_n}}\right] = 1 - \frac{1}{n} + \frac{1}{n}e^{itn}$$

donc

$$\lim_{n\to+\infty} \varphi_{X_n}(t) = 1 = \mathrm{E}\left[\mathrm{e}^{\mathrm{it0}}\right]$$

en appliquant le théorème de Levy, on a $X_n \to 0$ en loi.

(3) Montrer que $(X_n)_{n\in\mathbb{N}}$ ne converge pas en moyenne quadratique vers $X = 0$

$$\mathrm{E}\left[\mathrm{X_n^2}\right] = n^2\frac{1}{n} = n$$

et

$$\lim_{n\to+\infty} n = +\infty$$

donc X_n ne converge pas en moyenne quadratique

Exercice 4. — *Convergence en loi.*

Soit Y une variable aléatoire de loi uniforme sur $[-1, 1]$.

(1) Déterminer la fonction caractéristique de Y.

$$\begin{aligned}\varphi_Y(t) &= \int_{-1}^{1} \frac{1}{2} e^{itx} \mathrm{d}\, x \\ &= \frac{1}{2}\left[\frac{1}{it}e^{itx}\right]_{-1}^{1} \\ &= \frac{e^{it}-e^{-it}}{2it} \\ &= \frac{\sin(t)}{t} \\ &= \operatorname{sinc}(t)\end{aligned}$$

où $\operatorname{sinc}(t) = \frac{\sin(t)}{t}$

On considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes de loi

$$P\left(X_n = \frac{1}{2^n}\right) = \frac{1}{2} \text{ et } P\left(X_n = -\frac{1}{2^n}\right) = \frac{1}{2}$$

et on pose

$$S_n = \sum_{k=1}^{n} X_k$$

(2) Déterminer la fonction caractéristique des X_k

$$\varphi_{X_k}(t) = \frac{e^{\frac{it}{2^n}} + e^{\frac{-it}{2^n}}}{2} = \cos\left(\frac{t}{2^n}\right)$$

(3) En déduire celle de S_n

Les variables étants indépendantes, on a

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \varphi_{X_k}(t) = \prod_{k=1}^{n} \cos\left(\frac{t}{2^k}\right)$$

(4) En utilisant la formule

$$\sin(t/2^{n-1}) = 2\sin\left(\frac{t}{2}\right)\cos\left(\frac{t}{2}\right)$$

vérifier que

$$\varphi_{S_n}(t) = \frac{\sin(t)}{t}\frac{\frac{t}{2^n}}{\sin\left(\frac{t}{2^n}\right)}$$

On a

$$\cos\left(\frac{t}{2}\right) = \frac{1}{2}\frac{\sin(t)}{\sin\left(\frac{t}{2}\right)}$$

et donc

$$\cos\left(\frac{t}{2^k}\right) = \frac{1}{2}\frac{\sin(\frac{t}{2^{k-1}})}{\sin\left(\frac{t}{2^k}\right)}$$

on obtient

$$\begin{aligned}\varphi_{S_n}(t) &= \prod_{k=1}^{n}\cos\left(\frac{t}{2^k}\right)\\ &= \frac{1}{2^n}\frac{\sin(t)}{\sin\left(\frac{t}{2^n}\right)}\\ &= \frac{\sin(t)}{t}\frac{\frac{t}{2^n}}{\sin\left(\frac{t}{2^n}\right)}\end{aligned}$$

(5) En déduire que S_n converge en loi vers une variable qu'on déterminera

On a

$$\lim_{n\to+\infty}\frac{\frac{t}{2^n}}{\sin\left(\frac{t}{2^n}\right)} = 1$$

donc

$$\lim_{n\to+\infty}\varphi_{S_n}(t) = \frac{\sin(t)}{t}$$

En appliquant le théorème de Levy, on a donc S_n converge en loi vers une loi uniforme sur $[-1, 1]$

Exercice 5. — *Réservation.*

Il arrive assez souvent que le nombre de réservations pour une liaison aérienne soit supérieur au nombre de passagers se présentant effectivement le jour du vol. Cela est dû a des empêchements imprévisibles de certains passagers. Pour compenser ce phénomène, une compagnie aérienne exploitant un avion de 300 places décide de faire de la surréservation (surbooking) en prenant pour chaque vol un nombre $n > 300$ de réservations. S'il se présente plus de 300 passagers a l'embarquement, les 300 premiers arrivés prennent leur vol et les autres sont dédommagés financièrement.

(1) On considère que les passagers sont mutuellement indépendants et que la probabilité de désistement de chacun d'eux est 10%. On note n le nombre de réservations prises par la compagnie pour un vol donné et S_n le nombre (aléatoire) de passagers se présentant à l'embarquement pour ce vol. Donner la loi de S_n, son espérance et sa variance.

On pose X_i la variable aléatoire qui vaut 1 si le passager i se présente à l'embarquement et 0 sinon. On a donc $X_i \sim \mathcal{B}(p)$ avec $p = 90\%$. Le nombre S_n (aléatoire) de passagers se présentant à l'embarquement pour ce vol s'écrit donc

$$S_n = \sum_{i=1}^{n} X_i$$

et donc $S_n \sim \mathcal{B}(n, p)$. On a directement

$$\mathrm{E}\left[\mathrm{S_n}\right] = np = n\cdot 0.9 \quad \text{et} \quad \mathrm{Var}\left[\mathrm{S_n}\right] = np(1-p) = n\cdot 0.09$$

(2) Le directeur commercial de la compagnie aimerait connaitre la valeur maximale de n telle que $P(S_n \leq 300) \geq 0.99$. En utilisant le théorème de la limite centrale, proposez une solution approchée de ce problème sachant que pour une variable aléatoire normale centrée réduite Y, on a $P(Y \leq 2.4) \geq 0.99$.

On applique le théorème de la limite centrale (version de Moivre-Laplace)

$$\left(\frac{S_n - np}{\sqrt{np(1-p)}}\right) \longrightarrow \mathcal{N}(0,1)$$

ie

$$P\left(\frac{S_n - np}{\sqrt{np(1-p)}} < x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}\, t$$

$S_n \leq 300$ donne $\frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{300-np}{\sqrt{np(1-p)}}$. De plus on sait que

$$P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq 2.4\right) \geq 0.99$$

On cherche donc n tel que $\frac{300-np}{\sqrt{np(1-p)}} \leq 2.4$ ce qui donne

$$300 - np \leq 2.4\sqrt{n}\sqrt{p(1-p)}$$

avec $N = \sqrt{n}$ et $p = 0.9$

$$0 \leq 0.9N^2 + N0.72 - 300$$

ce qui donne $n = 319$.

Exercice 6. — *Erreurs d'un système de communication numérique.*

Afin de tester les performances d'un système de communications numériques, il est usuel de simuler le fonctionnement de ce système sur un ordinateur. Un des problèmes consiste alors à estimer la probabilité d'erreur p associé à ce système. On estime généralement cette probabilité comme suit

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} X_k$$

où X_k est une variable aléatoire binaire telle que

$$X_k = 1 \text{ s'il y a une erreur pour le } k^{me} \text{ symbole}$$
$$X_k = 0 \text{ s'il n'y a pas d'erreur pour le } k^{me} \text{ symbole}$$

avec $P(X_k = 1) = p$.

(1) Déterminez la moyenne et la variance de $\hat{p}$.

On a

$$\mathrm{E}\left[\hat{\mathrm{p}}\right] = \mathrm{E}\left[\frac{1}{\mathrm{N}} \sum_{\mathrm{k}=1}^{\mathrm{N}} \mathrm{X_k}\right] = \frac{1}{N} \sum_{k=1}^{N} \mathrm{E}\left[\mathrm{X_k}\right] = \frac{1}{N} Np = p$$

et

$$\mathrm{Var}\,[\hat{\mathrm{p}}] = \mathrm{Var}\left[\frac{1}{\mathrm{N}}\sum_{\mathrm{k}=1}^{\mathrm{N}} \mathrm{X_k}\right] \frac{1}{N^2}\sum_{k=1}^{N} \mathrm{Var}\,[\mathrm{X_k}] = \frac{1}{N^2} Np(1-p) = \frac{p(1-p)}{N}$$

(2) En appliquant le théorème de la limite centrale, déterminez la loi de $\hat{p}$ lorsque $N \to +\infty$

Le théorème de la limite centrale version de Moivre-Laplace (appliqué à des variables de Bernoulli) donne

$$\frac{N\hat{p} - Np}{\sqrt{Np(1-p)}} \to \mathcal{N}(0,1)$$

On a donc

$$\hat{p} \to \mathcal{N}\left(p, \frac{p(1-p)}{N}\right)$$

On cherche le nombre de points N nécessaire pour que $\hat{p}$ soit une approximation de p avec une précision relative inférieur à 10%. Pour cela, on se fixe un degré de confiance $\alpha = 95\%$, qui indique la probabilité d'avoir cette précision. Autrement dit

$$P\left(\left|\frac{\hat{p}-p}{p}\right| < \varepsilon\right) = \alpha$$

(3) Déterminez N pour que l'égalité précédente soit vérifiée. On utilisera le résultat suivant : si $Z \sim \mathcal{N}(0,1)$, alors $P(Z > 1.96) = 0.025$.

On a

$$P\left(\left|\frac{\hat{p}-p}{p}\right| < \varepsilon\right) = P\left(-\varepsilon < \frac{\hat{p}-p}{p} < \varepsilon\right)$$

$$P\left(-\varepsilon\sqrt{\frac{pN}{(1-p)}} < \frac{N\hat{p} - Np}{\sqrt{Np(1-p)}} < \varepsilon\sqrt{\frac{pN}{(1-p)}}\right)$$

On pose $Z = \frac{N\hat{p}-Np}{\sqrt{Np(1-p)}}$ de sorte que $Z \sim \mathcal{N}(0,1)$ et $a = \varepsilon\sqrt{\frac{pN}{(1-p)}}$. Ainsi

$$\begin{aligned} P\left(\left|\frac{\hat{p}-p}{p}\right| < \varepsilon\right) &= P(-a < Z < a) \\ &= 1 - 2P(Z > a) \end{aligned}$$

En choisissant $a = 1.96$ on a bien

$$P\left(\left|\frac{N\hat{p} - Np}{\sqrt{Np(1-p)}}\right| < \varepsilon\sqrt{\frac{pN}{(1-p)}}\right) = 2 \cdot (1 - 0.025) = 0.95$$

on choisit donc N tel que

$$\varepsilon\sqrt{\frac{pN}{(1-p)}} = 1.96$$

ce qui donne

$$N = \left(\frac{1.96}{\varepsilon}\right)^2 \frac{1-p}{p}$$

Remarque, en pratique, on utilise plutôt la règle suivante. On sait qu'avec la loi des grands nombres $\hat{p} \simeq p$, ce qui donne, lorsque p est très petit

$$N\hat{p} \simeq Np = \left(\frac{1.96}{\varepsilon}\right)^2 (1-p) \simeq \left(\frac{1.96}{\varepsilon}\right)^2$$

Il suffit donc d'observer $\left(\frac{1.96}{\varepsilon}\right)^2$ **erreurs** (et non pas faire $\left(\frac{1.96}{\varepsilon}\right)^2$ mesures !) pour avoir une estimation de p avec un intervalle de confiance à 95%. Pour $\varepsilon = 20\%$, cela donne une observation d'une centaine d'erreur.