

CONVERGENCE ET APPROXIMATION

CORRECTION

Inégalité de Bienaymé-Tchebychev

Exercice 1 : Application directe

Soit X , une variable aléatoire suivant la loi exponentielle de paramètre $\lambda = 1$.

1. La variable X admet une espérance $E(X) = 1$ et une variance $V(X) = 1$ donc on ne peut pas appliquer l'inégalité de Bienaymé-Tchebychev :

$$\forall \varepsilon > 0, \quad P(|X - 1| > \varepsilon) \leq \frac{1}{\varepsilon^2}.$$

2. Remarquons que $(|X - 1| > 2) = (X - 1 > 2) \cup (X - 1 < -2)$ donne $(|X - 1| > 2) = (X > 3) \cup (X < -1)$. Et donc par incompatibilité :

$$P(|X - 1| > 2) = P(X > 3) + P(X < -1).$$

Or $X \mapsto E(1)$ donc $X(\Omega) = \mathbb{R}_+$ et donc $P(X < -1) = 0$ et de plus, X étant à densité, on a $P(X > 3) = P(X \geq 3)$.

On a donc montré que $P(|X - 1| > 2) = P(X \geq 3)$.

Il s'ensuit, d'après la question précédente en prenant $\varepsilon = 2$:

$$P(|X - 1| > 2) \leq \frac{1}{2^2} \iff P(X \geq 3) \leq \frac{1}{4}.$$

3. On donne $e^{-3} \approx 0,05$.

On a :

$$\begin{aligned} P(X \geq 3) &= 1 - F_X(3) \\ &= 1 - (1 - e^{-3}) \\ &= e^{-3} \\ &\approx 0,05 < 0,25. \end{aligned}$$

Ccl : ce qui confirme bien que $P(X \geq 3) \leq \frac{1}{4}$.

Exercice 2 : Inégalité de Bienaymé-Tchebychev et intervalle de confiance

1. Les variables Y_n admettent toutes une espérance donc par linéarité $T_n = \frac{1}{n} \sum_{k=1}^n Y_k$ admet une espérance.

De plus les variables Y_n admettant toutes une variance, T_n admet elle aussi une variance.

De plus :

$$\begin{aligned}
 E(T_n) &= E\left(\frac{1}{n} \sum_{k=1}^n Y_k\right) \\
 &= \frac{1}{n} \sum_{k=1}^n E(Y_k) \quad \text{par linéarité de l'espérance} \\
 &= \frac{1}{n} \sum_{k=1}^n \theta \quad \text{car les } Y_k \text{ ont toutes la même espérance } \theta \\
 &= \frac{1}{n} \times n\theta \\
 &= \theta
 \end{aligned}$$

$$\begin{aligned}
 V(T_n) &= V\left(\frac{1}{n} \sum_{k=1}^n Y_k\right) \\
 &= \frac{1}{n^2} \sum_{k=1}^n E(Y_k) \quad \text{par prop de la variance puisque les variables } (Y_1, \dots, Y_n) \text{ sont indép.} \\
 &= \frac{1}{n^2} \sum_{k=1}^n \theta^2 \quad \text{car les } Y_k \text{ ont toutes la même variance } \theta^2 \\
 &= \frac{1}{n^2} \times n\theta^2 \\
 &= \frac{\theta^2}{n}
 \end{aligned}$$

On peut donc appliquer l'inégalité de Bienaymé-Tchebychev à T_n :

2. On en déduit que :

$$\forall \varepsilon > 0, \quad P(|T_n - \theta| > \varepsilon) \leq \frac{\theta^2}{\varepsilon^2} = \frac{\theta^2}{n\varepsilon^2}$$

Puis, en remarquant que $P(|T_n - \theta| \leq \varepsilon) = 1 - P(|T_n - \theta| > \varepsilon)$, que :

$$\forall \varepsilon > 0, \quad P(|T_n - \theta| \leq \varepsilon) \geq 1 - \frac{\theta^2}{n\varepsilon^2}$$

Remarquons ensuite que :

$$(|T_n - \theta| \leq \varepsilon) = (-\varepsilon \leq T_n - \theta \leq \varepsilon) = (T_n - \varepsilon \leq \theta \leq T_n + \varepsilon) = (\theta \in [T_n - \varepsilon, T_n + \varepsilon]).$$

On en déduit donc

$$\forall \varepsilon > 0, \quad P(\theta \in [T_n - \varepsilon, T_n + \varepsilon]) \geq 1 - \frac{\theta^2}{n\varepsilon^2}.$$

3. On dit que I_n est un intervalle de confiance pour θ au niveau de confiance 90% lorsqu'on a :

$$P(\theta \in I_n) \geq 1 - 0,1.$$

On choisit $n = 100$. Avec la condition $\theta \in [0, 1/2]$, on a alors

$$\frac{\theta^2}{n\varepsilon^2} \leq \frac{1}{400\varepsilon^2}$$

Ainsi, l'intervalle $I_n = [T_n - \varepsilon, T_n + \varepsilon]$ est un intervalle de confiance de θ au seuil de confiance 90% si

$$\frac{1}{400\varepsilon^2} \leq 0,1 \iff \frac{1}{4 \cdot 10^2 \varepsilon^2} \leq 10^{-1} \iff \varepsilon^2 \geq \frac{1}{4 \cdot 10^2 \cdot 10^{-1}} = \frac{1}{40} \iff \varepsilon \geq \sqrt{\frac{1}{40}}$$

Ccl : Autrement dit $\varepsilon = \sqrt{\frac{1}{40}} \approx \frac{1}{2} \times 0,32$, soit $\varepsilon \approx 0,16$ convient.

Exercice 3 : Inégalité de Bienaymé-Tchebychev et intervalle de confiance

Soit une variable aléatoire X qui suit la loi de Bernoulli de paramètre p , d'espérance p et de variance $p(1-p)$.

Soit (X_1, \dots, X_n) un échantillon de taille n de la variable X .

On a déjà vu que l'on pouvait appliquer l'I-B-T à la variable $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$, qui donne :

$$\forall \varepsilon > 0, \quad P(|\bar{X}_n - p| > \varepsilon) \leq \frac{\frac{p(1-p)}{n}}{\varepsilon^2}.$$

Avec $\varepsilon = 0,1$ on a donc :

$$P(|\bar{X}_n - p| > 0,1) \leq \frac{p(1-p)}{n \cdot 0,1^2} \iff P(|\bar{X}_n - p| \leq 0,1) \geq 1 - \frac{p(1-p)}{n \cdot 10^{-2}} \iff P(p \in [\bar{X}_n - 0,1; \bar{X}_n + 0,1]) \geq 1 - \frac{p(1-p)}{n \cdot 10^{-2}}.$$

Ainsi l'intervalle $I_n = [\bar{X}_n - 0,1; \bar{X}_n + 0,1]$ est un intervalle de confiance de p au seuil 95% dès que

$$\frac{p(1-p)}{n \cdot 10^{-2}} \leq 0,05.$$

Or, on a vu en cours que $\forall p \in [0,1], \quad p(1-p) \leq \frac{1}{4}$ (étude de fonction), donc l'intervalle $I_n = [\bar{X}_n - 0,1; \bar{X}_n + 0,1]$ est un intervalle de confiance de p au seuil 90% dès que :

$$\frac{1}{4 \cdot 10^{-2} n} \leq 5 \cdot 10^{-2} \iff n \geq \frac{1}{4 \cdot 10^{-2} \times 5 \cdot 10^{-2}} = \frac{1}{20 \cdot 10^{-4}} \iff n \geq \frac{1}{20} \cdot 10^4 = 0,05 \times 10\,000 \iff n \geq 500.$$

Ccl : l'intervalle $I_n = [\bar{X}_n - 0,1; \bar{X}_n + 0,1]$ est un intervalle de confiance de p au seuil 95% dès que $n \geq 500$.

Exercice 4 : Loi faible des grands nombres

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant une même loi de Bernoulli de paramètre p .

On pose $Y_n = X_n + X_{n+1}$ et $Z_n = \frac{1}{n} \sum_{i=1}^n Y_i$.

1. Déterminer la loi de Y_n .
2. En déduire l'espérance et la variance de Z_n .
3. Démontrer que :

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|Z_n - 2p| > \varepsilon) = 0.$$

Convergence en loi

Exercice 5 : Convergence en loi de lois de Bernoulli

Cours : On rappelle que si X_n et X sont à support contenus dans \mathbb{Z} (entiers relatifs) alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff \forall k \in X(\Omega), \lim_{n \rightarrow +\infty} P(X_n = k) = P(X = k)$.

On considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires telles que pour tout $n \geq 1$, $X_n \hookrightarrow \mathcal{B}\left(\left(1 - \frac{1}{n}\right)^n\right)$ et X suivant une loi $\mathcal{B}(e^{-1})$.

Les variables X_n et X sont de support $\{0, 1\}$ donc le cours s'applique et nous devons étudier les limites des suites $P(X_n = k)$ pour $k \in \{0, 1\}$.

- $P(X_n = 1) = \left(1 - \frac{1}{n}\right)^n = \exp\left(n \ln\left(1 - \frac{1}{n}\right)\right)$.

Or par équivalent usuel $\ln\left(1 - \frac{1}{n}\right) \underset{+\infty}{\sim} -\frac{1}{n}$ car $\frac{1}{n} \rightarrow 0$ et donc $n \ln\left(1 - \frac{1}{n}\right) \underset{+\infty}{\sim} -1$ ce qui implique que

$$\lim_{n \rightarrow +\infty} n \ln\left(1 - \frac{1}{n}\right) = -1.$$

Par continuité de la fonction exp en -1 on obtient donc la limite (ultra classique) suivante :

$$\lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1} = P(X = 1) \quad \text{car } X \hookrightarrow \mathcal{B}(e^{-1}).$$

Ceci démontre donc que :

$$\lim_{n \rightarrow +\infty} P(X_n = 1) = P(X = 1).$$

- Comme $P(X_n = 0) = 1 - P(X_n = 1)$ on déduit du résultat précédent que $\lim_{n \rightarrow +\infty} P(X_n = 0) = 1 - e^{-1} = 1 - P(X = 1)$ et donc que $\lim_{n \rightarrow +\infty} P(X_n = 0) = P(X = 0)$.

D'où la convergence en loi.

Exercice 6 : Convergence en loi de lois uniformes

Cours : On rappelle que si X_n et X alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff$ en tout point de continuité de F_X : $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$.

On considère une suite $(X_n)_{n \geq 1}$ de variables aléatoires telles que pour tout $n \geq 1$, $X_n \hookrightarrow \mathcal{U}\left(\left[0; 1 - \frac{1}{n}\right]\right)$.

Les variables X_n et X sont à densité donc on applique la définition du cours :

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{1 - \frac{1}{n}} & \text{si } x \in \left[0, 1 - \frac{1}{n}\right] \\ 1 & \text{si } x \geq 1 - \frac{1}{n} \end{cases}$$

En passant à la limite et en remarquant $1 - \frac{1}{n} \rightarrow 1$, on trouve la fonction

$$x \mapsto \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } x \in [0, 1] \\ 1 & \text{si } x \geq 1 \end{cases}$$

qui n'est d'autre que la fonction de répartition F_X de la loi $\mathcal{U}[0, 1]$.

D'où la convergence en loi.

Exercice 7 : Convergence en loi de tirages avec un grand nombre de boules

On considère une urne constituée de $2n$ boules ($n \geq 2$) dont la moitié sont noires et l'autre moitié sont blanches. On pioche dans cette urne deux boules successivement et sans remise, et on note X_n le nombre de boules blanches obtenues.

1. L'urne contenant au moins 2 boules blanches et 2 boules noires, à l'issue des deux tirages on aura tiré ou bien 0, 1 ou 2 boules blanches $\Rightarrow X_n(\Omega) = \{0, 1, 2\}$.

• $(X_n = 0)$ est réalisé si on tire successivement deux boules noires. Les tirages étant fait sans remise on a :

$$P(X_n = 0) = \frac{n}{2n} \times \frac{n-1}{2n-1} \Rightarrow P(X_n = 0) = \frac{n(n-1)}{2n(2n-1)}.$$

• $(X_n = 1)$ est réalisé si on tire une boule blanche et une boule noire. La boule blanche pouvant être tirée au 1er ou au 2nd tirage on a :

$$P(X_n = 1) = \frac{n}{2n} \times \frac{n}{2n-1} + \frac{n}{2n} \times \frac{n}{2n-1} = \frac{2n}{2n} \times \frac{n}{2n-1} \Rightarrow P(X_n = 1) = \frac{n}{2n-1}.$$

• $(X_n = 2)$ est réalisé si on tire successivement deux boules blanches. Les tirages étant fait sans remise on a :

$$P(X_n = 2) = \frac{n}{2n} \times \frac{n-1}{2n-1} \Rightarrow P(X_n = 2) = \frac{n(n-1)}{2n(2n-1)}.$$

2. Les variables X_n et $Z \hookrightarrow \mathcal{B}\left(2, \frac{1}{2}\right)$ étant à support contenus dans \mathbb{N} on peut appliquer la propriété du cours et étudier les 3 limites $\lim_{n \rightarrow +\infty} P(X_n = 0)$, $\lim_{n \rightarrow +\infty} P(X_n = 1)$ et $\lim_{n \rightarrow +\infty} P(X_n = 2)$

• $P(X_n = 0) = \frac{n(n-1)}{2n(2n-1)} \underset{+\infty}{\sim} \frac{n^2}{4n^2} \Rightarrow \lim_{n \rightarrow +\infty} P(X_n = 0) = \frac{1}{4}$.

• $P(X_n = 1) = \frac{n}{2n-1} \underset{+\infty}{\sim} \frac{n}{2} \Rightarrow \lim_{n \rightarrow +\infty} P(X_n = 1) = \frac{1}{2}$.

• $P(X_n = 2) = P(X_n = 0) \Rightarrow \lim_{n \rightarrow +\infty} P(X_n = 2) = \frac{1}{4}$.

Or la loi de Z est donnée par

• $P(Z = 0) = \binom{2}{0} \left(\frac{1}{2}\right)^0 \left(1 - \frac{1}{2}\right)^{2-0} = 1 \times 1 \times \frac{1}{4} \Rightarrow P(Z = 0) = \frac{1}{4}$

• $P(Z = 1) = \binom{2}{1} \left(\frac{1}{2}\right)^1 \left(1 - \frac{1}{2}\right)^{2-1} = 2 \times \frac{1}{2} \times \frac{1}{2} \Rightarrow P(Z = 1) = \frac{1}{2}$

• $P(Z = 2) = \binom{2}{2} \left(\frac{1}{2}\right)^2 \left(1 - \frac{1}{2}\right)^{2-2} \Rightarrow P(Z = 2) = \frac{1}{4}$.

On retrouve bien les limites trouvées plus haut, d'où la convergence en loi.

3. Interprétation du résultat : deux tirages successifs sans remise dans une urne contenant un grand nombre de boules avec parité de blanches et de noires peut être approché par le tirage successif de 2 boules avec remise dans une urne contenant un nombre de boules avec parité de blanches et de noires.

En bref les tirages sans remise reviennent à des tirages avec remise plus on augmente le nombre de boules dans l'urne.

Exercice 8 : Convergence en loi vers une loi exponentielle

Soit X une variable aléatoire suivant la loi uniforme sur $[0; 1]$ et (X_1, \dots, X_n) un n -échantillon de X .

On considère les variables aléatoires

$$M_n = \max(X_1, \dots, X_n) \text{ et } Y_n = n(1 - M_n).$$

On admet que M_n et Y_n sont des variables à densité.

1. De manière très classique en utilisant le fait que (X_1, \dots, X_n) un n -échantillon de X , on trouve :

$$\forall x \in \mathbb{R}, \quad F_{M_n}(x) = (F_X(x))^n$$

et donc

$$F_{M_n}(x) = \begin{cases} 0 & x < 0 \\ x^n & x \in [0, 1] \\ 1 & x > 1 \end{cases}.$$

Ensuite :

$$\begin{aligned} F_{Y_n}(x) &= P(Y_n \leq x) \\ &= P(n(1 - M_n) \leq x) \\ &= P\left(M_n \geq 1 - \frac{x}{n}\right) \\ &= 1 - F_{M_n}\left(1 - \frac{x}{n}\right) \end{aligned}$$

Remarquons que

$$1 - \frac{x}{n} \in [0, 1] \iff \frac{x}{n} \in [0, 1] \iff x \in [0, n].$$

Ainsi on obtient :

$$F_{Y_n}(x) = \begin{cases} 0 & x < 0 \\ 1 - \left(1 - \frac{x}{n}\right)^n & x \in [0, n] \\ 1 & x > n \end{cases}$$

2. Etudions la limite $\lim_{n \rightarrow +\infty} F_{Y_n}(x)$ pour $x \in \mathbb{R}$.

Soit $x \in \mathbb{R}$.

1er cas : $x < 0$: $F_{Y_n}(x) = 0$ donc $\lim_{n \rightarrow +\infty} F_{Y_n}(x) = 0$.

2ème cas : $x \geq 0$:

Soit $n \in \mathbb{N}$ suffisamment grand de sorte que $x \in [0, n]$ (il suffit de prendre $n = \lfloor x \rfloor + 1$ par exemple).

Alors $F_{Y_n}(x) = 1 - \left(1 - \frac{x}{n}\right)^n = 1 - \exp\left(n \ln\left(1 - \frac{x}{n}\right)\right)$

Or $\ln(1-u) \underset{u \rightarrow 0}{\sim} -u$ donc en posant $u = \frac{x}{n}$ avec $n \rightarrow +\infty$ il vient $\ln\left(1 - \frac{x}{n}\right) \underset{n \rightarrow +\infty}{\sim} -\frac{x}{n}$ et donc $n \ln\left(1 - \frac{x}{n}\right) \underset{n \rightarrow +\infty}{\sim} -x$.

Par continuité de la fonction \exp sur \mathbb{R} (et donc en x) on obtient donc $\exp\left(n \ln\left(1 - \frac{x}{n}\right)\right) \xrightarrow{n \rightarrow +\infty} \exp(-x)$.

Ainsi $\lim_{n \rightarrow +\infty} F_{Y_n}(x) = 1 - e^{-x}$.

Conclusion : En notant Y une variable suivant la loi exponentielle de paramètre on vient de montrer que $\lim_{n \rightarrow +\infty} F_{Y_n}(x) = F_Y(x)$ et donc que la suite $(Y_n)_n$ converge en loi vers la variable Y .

Exercice 9 : Convergence en loi de lois normales Corrigé en cours.

Exercice 10 : Convergence de lois binomiales vers une loi de Poisson Corrigé en cours.

Exercice 11 : Extrait de concours : convergence vers la loi de Grumbel

On considère une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$, mutuellement indépendantes, de même densité f , où f est définie, pour tout t de \mathbb{R} , par :

$$f(t) = \frac{e^{-t}}{(1 + e^{-t})^2}$$

On pose, pour tout n de \mathbb{N}^* : $T_n = \max(X_1, \dots, X_n)$ et $U_n = T_n - \ln(n)$.

1. (a) De manière ultra classique on a : $F_{T_n}(x) = (F_X(x))^n$ où X est une variable aléatoire admettant f pour d.d.p.

Or $F_X(x) = \int_{-\infty}^x f(t)dt$ et on vérifie que la fonction $t \mapsto \frac{1}{1+e^{-t}}$ est une primitive de f .

D'où, pour $x \in \mathbb{R}$ et $B < x$ on a :

$$\begin{aligned} \int_B^x f(t) &= [g(t)]_B^x \\ &= \frac{1}{1+e^{-x}} - \frac{1}{1+e^{-B}} \\ &\xrightarrow{B \rightarrow -\infty} \frac{1}{1+e^{-x}} \quad \text{puisque } \frac{1}{1+e^{-B}} \xrightarrow{B \rightarrow -\infty} 0. \end{aligned}$$

Ccl : $F_{T_n}(x) = \left(\frac{1}{1+e^{-x}} \right)^n$.

(b)

$$\begin{aligned} P(U_n \leq x) &= P(T_n - \ln(n) \leq x) \\ &= P(T_n \leq x + \ln(n)) \\ &= F_{T_n}(x + \ln(n)) \\ &= \left(\frac{1}{1+e^{-(x+\ln(n))}} \right)^n \\ &= \left(\frac{1}{1+e^{-x-\ln(n)}} \right)^n \\ &= \left(\frac{1}{1+e^{-x} \times e^{-\ln(n)}} \right)^n \\ &= \left(\frac{1}{1+e^{-x} \times \frac{1}{e^{\ln(n)}}} \right)^n \\ &= \left(\frac{1}{1+e^{-x} \times \frac{1}{n}} \right)^n \\ &= \left(\frac{1}{1+\frac{e^{-x}}{n}} \right)^n \\ &= \left(1 + \frac{e^{-x}}{n} \right)^{-n} \end{aligned}$$

On a donc trouvé ici la fonction de répartition de U_n .

2. En utilisant à nouveau l'équivalent usuel $\ln(1+u) \underset{u \rightarrow 0}{\sim} u$ on trouve (et c'est classique) :

$$F_{U_n}(x) = \left(1 + \frac{e^{-x}}{n} \right)^{-n} = \exp \left(-n \ln \left(1 + \frac{e^{-x}}{n} \right) \right) \xrightarrow{n \rightarrow +\infty} e^{-e^{-x}}$$

Ceci démontre que la suite $(U_n)_n$ converge en loi vers une variable aléatoire U admettant pour fonction de répartition $F : x \mapsto e^{-e^{-x}}$.

La fonction F étant de classe \mathcal{C}^1 sur \mathbb{R} elle y est en particulier continue et la variable U est donc à densité.

Une densité de U est donnée par $g = F'$ soit $g : t \mapsto te^{-t}e^{-e^{-t}}$.

On retrouve la d.d.p. de la loi de Grumbel.

Application du théorème limite central

Exercice 12 : Corrigé en cours.