Chapitre 22

Convergences et approximations en probabilités

On considère une expérience aléatoire modélisée par un triplet $(\Omega, \mathscr{F}, \mathbb{P})$.

1 Convergence en probabilité

1.1 Inégalité de Markov

Soit X une variable aléatoire définie sur $(\Omega, \mathscr{F}, \mathbb{P})$, admettant une espérance.

On cherche à estimer la probabilité que X se disperse, c'est-à-dire qu'elle prenne des valeurs éloignées de son espérance. Pour cela on a le résultat général suivant.

Théorème 1 Inégalité de Markov

1. Si X est une variable aléatoire **positive** admettant une espérance, alors :

$$\forall a > 0, \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

2. Si X est une variable aléatoire admettant une espérance, alors :

$$\forall a > 0, \quad \mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}\big(|X|\big)}{a}$$

***Exemple* :** Si X est une variable aléatoire admettant une espérance, alors :

$$\lim_{a \to +\infty} \mathbb{P}(X > a) = 0$$

Par convergence monotone, on savait déjà que $\lim\limits_{n \to +\infty} \mathbb{P}(X > n) = 0$. Quelle est la différence entre ces deux résultats ?

1.2 Inégalité de Bienaymé-Tchebychev

Sous l'hypothèse que X a un moment d'ordre 2 on a un résultat plus précis.

Théorème 2 Inégalité de Bienaymé-Tchebychev
Soit X une VARD admettant un moment d'ordre 2. Alors :

$$\forall \varepsilon > 0, \quad \mathbb{P}\big(|X - \mathbb{E}(X)| \geq \varepsilon\big) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}$$

Cette inégalité confirme l'utilisation de l'écart-type comme mesure de dispersion.

1.3 Convergence en probabilité

Soient X une VAR définie sur $(\Omega, \mathscr{F}, \mathbb{P})$ et $(X_n)_{n\in\mathbb{N}}$ une suite de VAR toutes définies sur $(\Omega, \mathscr{F}, \mathbb{P})$.

On souhaite définir une notion de convergence de la suite X_n vers X, pas en tant que suite de nombres mais en tant que suite de variables aléatoires.

Définition 3 Convergence en probabilité
On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en probabilité** vers X lorsque :

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} \mathbb{P}\big(|X_n - X| \geq \varepsilon\big) = 0$$

On le note $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$.

Remarquer que ce ne sont pas les expressions de X_n et X qui interviennent dans cette définition, mais leur loi.

L'inégalité de Markov donne une condition suffisante de convergence en probabilité, dans le cas de variables admettant une espérance :

$$\lim_{n\to+\infty} \mathbb{E}\big(|X_n - X|\big) = 0$$

Si $\forall n \in \mathbb{N}, \mathbb{E}(X_n) = \mathbb{E}(X)$, l'inégalité de Bienaymé-Tchebychev donne une condition suffisante de convergence en probabilité, dans le cas de variables admettant un moment d'ordre 2 :

$$\lim_{n\to+\infty} V(X_n - X) = 0$$

On peut généraliser au cas où $\big(\mathbb{E}(X_n)\big)_{n\in\mathbb{N}}$ converge lorsque $n \to +\infty$, comme nous le verrons en TD.

⚠ Il n'y a pas unicité de la limite.

⚠ Si $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ on ne peut rien dire que $\mathbb{E}(X_n)$.
Considérer X_n telle que $\mathbb{P}\,(X_n = 0) = 1 - \dfrac{1}{n}$ et $\mathbb{P}(X_n = n) = \dfrac{1}{n}$.

***Exemple* :** Si X est une variable aléatoire, on pose $X_n = \dfrac{1}{2^n}\lfloor 2^n X\rfloor$ pour tout $n \in \mathbb{N}$. Montrer que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$.

1.4 Loi faible des grands nombres pour la loi binomiale

Théorème 4 Loi faible des grands nombres pour la loi binomiale
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que $\forall n \in \mathbb{N}$, $X_n \hookrightarrow \mathscr{B}(n,p)$. Alors :

$$\frac{1}{n}X_n \xrightarrow[n\to+\infty]{\mathbb{P}} p$$

(où p désigne une variable aléatoire presque sûrement constante égale à p).

Ce résultat permet une justification partielle, a posteriori, de la notion de probabilité d'un évènement, introduite intuitivement : « la fréquence de succès sur n essais indépendants converge vers la probabilité d'un succès ».

2 Convergence en loi

Soient X une VAR définie sur $(\Omega, \mathscr{F}, \mathbb{P})$ et $(X_n)_{n\in\mathbb{N}}$ (chacune de ses variables est définies sur un espace probabilisé qui lui est propre).

On souhaite définir une notion de convergence de la suite X_n vers X, plus faible que la notion de convergence en probabilité.

Définition 5 Convergence en loi
On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X lorsque :

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x)$$

en tout point de continuité de F_X. On le note $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$.

Dans le cas où X a pour densité f, on a donc :

$$\forall (a,b) \in \mathbb{R}^2, \quad a < b \Longrightarrow \lim_{n\to+\infty} \mathbb{P}(a < X_n \leq b) = \int_a^b f(t)\,dt$$

Il existe d'autres définitions équivalentes (et très différentes) mais elles ne sont pas au programme d'ECS.

Nous verrons en TD que la convergence en loi est une conséquence de la convergence en probabilité (mais ce n'est pas au programme).

⚠ Il n'y a pas unicité de la limite.

⚠ Si $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$ on ne peut rien dire sur $\mathbb{E}(X_n)$.
Considérer X_n telle que $\mathbb{P}\left(X_n = \frac{1}{n}\right) = 1 - \frac{1}{n}$ et $\mathbb{P}(X_n = n) = \frac{1}{n}$.

Théorème 6 Convergence en loi pour des variables aléatoires à valeurs dans $\mathbb{N}$
Si toutes les variables aléatoires X_n, $\in \mathbb{N}$, et X sont à valeurs dans $\mathbb{N}$ alors $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} x$ si, et seulement si :

$$\forall k \in X(\Omega), \quad \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$$

⚠ En général ceci est faux en général pour des variables discrètes !
Contre-exemple : dans l'exemple précédent, $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} 0$ mais $\lim_{n\to+\infty} \mathbb{P}(X_n = 0) \neq 1$.

2.1 Approximation binomiale-Poisson

Théorème 7 Approximation binomiale-Poisson
Soit X_n une VARD de loi $\mathscr{B}\left(n, p_n\right)$ avec $\lim_{n\to+\infty} np_n = \lambda > 0$. On a alors :

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

ie $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} \mathscr{P}(\lambda)$.

Ce résultat explique pourquoi la loi de Poisson est appelée « loi des évènements rares ».

On l'applique souvent avec $p_n = \dfrac{\lambda}{n}$.

2.2 Théorème central limite

Théorème 8 Théorème central limite pour la loi binomiale (de De Moivre Laplace)
Soit $(X_n)_{n\in\mathbb{N}}$ suite de variables aléatoires telle que $\forall n \in \mathbb{N}$, $X_n \hookrightarrow \mathscr{B}(n,p)$.
On considère les variables aléatoires centrées réduites associées :

$$\forall n \in \mathbb{N}, \quad X_n^\star = \frac{X_n - \mathbb{E}(X_n)}{\sqrt{V(X_n)}} = \frac{X_n - np}{\sqrt{np(1-p)}}$$

Alors :

$$X_n^\star \xrightarrow[n\to+\infty]{\mathscr{L}} \mathscr{N}(0,1)$$

ie :

$$\forall (a,b) \in \mathbb{R}^2, \quad a < b \Longrightarrow \lim_{n\to+\infty} \mathbb{P}\left(a < \frac{X_n - np}{\sqrt{np(1-p)}} \leq b\right) = \int_a^b f(t)\,dt$$

On peut donc approximer la loi $\mathscr{B}(n,p)$ par la loi $\mathscr{N}\left(np, np(1-p)\right)$ pour de grandes valeurs de n.

Théorème 9 Théorème central limite pour la loi de Poisson
Soit $(X_n)_{n\in\mathbb{N}}$ suite de variables aléatoires telle que $\forall n \in \mathbb{N}$, $X_n \hookrightarrow \mathscr{P}(n\lambda)$.
On considère les variables aléatoires centrées réduites associées :

$$\forall n \in \mathbb{N}, \quad X_n^\star = \frac{X_n - \mathbb{E}(X_n)}{\sqrt{V(X_n)}} = \frac{X_n - n\lambda}{\sqrt{n\lambda}}$$

Alors :

$$X_n^\star \xrightarrow[n\to+\infty]{\mathscr{L}} \mathscr{N}(0,1)$$

ie :

$$\forall (a,b) \in \mathbb{R}^2, \quad a < b \Longrightarrow \lim_{n\to+\infty} \mathbb{P}\left(a < \frac{X_n - n\lambda}{\sqrt{n\lambda}} \leq b\right) = \int_a^b f(t)\,dt$$

On peut donc approximer la loi $\mathscr{P}(n\lambda)$ par la loi $\mathscr{N}\left(n\lambda, n\lambda\right)$ pour de grandes valeurs de n.

3 Exercices

Exercice 1 (Convergence en loi et lois uniformes)

1. Pour $n \geq 1$, on suppose que $X_n \hookrightarrow \left(\left[-\frac{1}{n}, \frac{1}{n}\right]\right)$.
   Montrer que $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} 0$.
2. Pour $n \geq 1$, on suppose que $X_n \hookrightarrow \left(\left\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\right\}\right)$.
   Montrer que $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} \mathscr{U}\big([0,1]\big)$.

Exercice 2 Soit X_n de densité :

$$f_n(t) = \frac{1}{n!} \mathrm{e}^{x/n} \left(\frac{x}{n}\right)^{n-1} \mathbb{1}_{x>0}$$

Montrer que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} 1$.

Exercice 3 (Approximation d'une loi hypergéométrique par une loi binomiale)
On considère une urne composée de N boules dont une proportion p de boules blanches. On effectue n tirages d'une boule sans remise et on note X_N le nombre de boules blanches obtenues.
Montrer que $X_N \xrightarrow[N\to+\infty]{\mathscr{L}} \mathscr{B}(n,p)$.

Exercice 4 (Convergences vers la loi de Poisson)
Soit n un entier de $\mathbb{N}^*$.

1. On note F_n l'ensemble des **permutations** de $[\![1,n]\!]$. On choisit au hasard une de ces permutations, et on définit la variable aléatoire Y_n égale au nombre de points fixes de cette permutation.
   (a) Montrer que le nombre d'éléments de F_n n'admettant aucun point fixe vaut :
   $$d_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$
   (b) Déterminer la loi de Y_n.
   (c) Montrer que la suite $(Y_n)_n$ converge en loi vers une variable aléatoire dont on déterminera la loi.
2. On note E_n l'ensemble des **applications** de $[\![1,n]\!]$ dans $[\![1,n]\!]$. On choisit au hasard une de ces applications, et on définit la variable X_n égale au nombre de points fixes de cette application.
   (a) Déterminer la loi de X_n.
   (b) Montrer que la suite $(X_n)_n$ converge en loi vers une variable aléatoire dont on déterminera la loi.

Exercice 5 Pour tout $n \in \mathbb{N}^*$, on considère la fonction f_n définie par :

$$f_n(x) = n^2 x \mathrm{e}^{-n^2x^2/2} \mathbb{1}_{x\geq 0}$$

1. Montrer que f_n est la densité d'une variable aléatoire.
2. Soit X_n de densité f_n. Montrer que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

Exercice 6 Pour $n \in \mathbb{N}^*$ et $a \in \mathbb{R}$, on considère la fonction f_n définie par :

$$f_n(x) = \frac{an}{\pi(1+n^2x^2)}$$

1. Déterminer a pour que f_n soit une densité de variable aléatoire.
2. Soit X_n de densité f_n. Admet-elle des moments ? Montrer que $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} 0$.

Exercice 7 (Inégalité de grandes déviations)
Dans cet exercice X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, de densité :

$$f(t) = \frac{1}{(n-1)!} t^{n-1} \mathrm{e}^{-t}$$

ie de loi $\Gamma(1,n) = \gamma(n))$, avec $n \in \mathbb{N}^*$.
On rappelle que $\mathbb{E}(X) = n$ et $V(X) = n$.

1. Pour tout réel $\lambda > 0$, on pose $\psi(\lambda) = \ln\left(\mathbb{E}(\mathrm{e}^{-\lambda(X-\mathbb{E}(X))})\right)$ (transformée de Legendre de X). Calculer $\psi(\lambda)$.
2. Montrer que pour tout $\lambda > 0$, on a $\psi(\lambda) \leq n\dfrac{\lambda^2}{2}$.
3. (a) Montrer que pour tout $x > 0$ et tout $\lambda > 0$, on a

   $$\mathbb{P}(\mathbb{E}(X) - X \geq x) \leq \mathrm{e}^{-\lambda x + \psi(\lambda)}$$

   (b) En déduire que pour tout $x > 0$

   $$\mathbb{P}(\mathbb{E}(X) - X \geq x) \leq \mathrm{e}^{-x^2/(2n)}$$

4. Comparer l'inégalité que l'on vient d'obtenir avec celle obtenue par l'inégalité de Bienaymé-Tchebicheff.

Exercice 8 (Inégalités pour la loi de Poisson)
Dans tout l'exercice, X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ et suivant la loi de Poisson de paramètre $\lambda > 0$.

1. (a) Montrer que $\mathbb{P}(|X-\lambda| \geq \lambda) \leq \dfrac{1}{\lambda}$.

   (b) En déduire l'inégalité $\mathbb{P}(X \geq 2\lambda) \leq \dfrac{1}{\lambda}$.
2. On considère dans toute cette question une variable aléatoire discrète Z définie sur $(\Omega, \mathscr{A}, \mathbb{P})$, d'espérance nulle et de variance σ^2.

   (a) Montrer que

   $$\forall\, a > 0, \forall\, x \geq 0, \mathbb{P}(Z \geq a) \leq \mathbb{P}[(Z+x)^2 \geq (a+x)^2]$$

   (b) Montrer que $\forall\, a > 0, \forall\, x \geq 0, \mathbb{P}(Z \geq a) \leq \dfrac{\sigma^2 + x^2}{(a+x)^2}$.

   (c) Montrer que $\forall\, a > 0, \mathbb{P}(Z \geq a) \leq \dfrac{\sigma^2}{\sigma^2 + a^2}$.

(d) En déduire que $\mathbb{P}(X \geq 2\lambda) \leq \dfrac{1}{\lambda+1}$.

3. Pour tout réel t, on pose $G_X(t) = \sum_{k=0}^{+\infty} \mathbb{P}(X=k)\, t^k$.

   (a) Pour tout réel t, justifier l'existence de $G_X(t)$ et calculer sa valeur.

   (b) Montrer que : $\forall\, t \geq 1, \forall\, a > 0, \mathbb{P}(X \geq a) \leq \dfrac{G_X(t)}{t^a}$.

   (c) En déduire que $P(X \geq 2\lambda) \leq \left(\dfrac{\mathrm{e}}{4}\right)^{\lambda}$.

Exercice 9 (Transformée de Laplace et inégalités exponentielles) Si X est une variable aléatoire réelle, on appelle transformée de Laplace de X (ou fonction génératrice des moments) la fonction $\mathscr{L}_X$ définie par :

$$\forall\, t \in \mathbb{R}, \quad \mathscr{L}_X(t) = \mathbb{E}\left(\mathrm{e}^{tX}\right)$$

1. (a) Montrer que $\mathscr{L}_X$ est toujours définie en 0.

   (b) Si X est positive, montrer que $\mathscr{L}_X$ est définie au moins sur $\mathbb{R}^-$.

   (c) Calculer $\mathscr{L}_X$ dans les cas suivants : $X \hookrightarrow \mathscr{U}\left(\llbracket 1,n \rrbracket\right)$, $X \hookrightarrow \mathscr{B}(n,p)$, $X \hookrightarrow \mathscr{G}(p)$, $X \hookrightarrow \mathscr{P}(\lambda)$, $X \hookrightarrow \mathscr{U}\left([a,b]\right)$, $X \hookrightarrow \mathscr{E}\left(\lambda\right)$ et $X \hookrightarrow \mathscr{N}(0,1)$. On précisera à chaque fois son ensemble de définition

2. (a) Montrer que si X est une variable aléatoire réelle et $t > 0$ tel que e^{tX} admet une espérance, alors :
   $$\forall\, a \in \mathbb{R}, \quad \mathbb{P}(X \geq a) \leq \mathrm{e}^{-ta}\mathbb{E}\left(\mathrm{e}^{tX}\right)$$

   (b) Dans le cas où $X \hookrightarrow \mathscr{N}(0,1)$ en déduire :
   $$\forall\, a > 0, \quad \mathbb{P}(X \geq a) \leq \mathrm{e}^{-a^2/2}$$

Exercice 10 (Une condition suffisante de convergence en probabilité)
Soit $(X_n)_{n\in\mathbb{N}}$ suite de variables aléatoires admettant un moment d'ordre 2 et telle que :

$$\lim_{n\to+\infty} \mathbb{E}(X_n) = a \quad \text{et} \quad \lim_{n\to+\infty} V(X_n) = 0$$

Montrer que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$.

Exercice 11 (Convergence en probabilité et fonctions continues)
Si $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ et $f : \mathbb{R} \longrightarrow \mathbb{R}$ est continue sur $\mathbb{R}$, alors $f(X_n) \xrightarrow[n\to+\infty]{\mathbb{P}} f(X)$.

Exercice 12 (La convergence en probabilité donne la convergence en loi) On suppose que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$.

1. Vérifier que si Y et Z sont deux variables aléatoires réelles, si $c \in \mathbb{R}$ et $\varepsilon > 0$ alors :
   $$\mathbb{P}(Y \leq c) \leq \mathbb{P}(Z \leq c + \varepsilon) + \mathbb{P}(Z - Y > \varepsilon)$$

2. Soient $\varepsilon > 0$ et $a \in \mathbb{R}$. Etablir que, pour tout $n \in \mathbb{N}$:
   $$\mathbb{P}(X \leq a - \varepsilon) - \mathbb{P}\left(|X_n - X| > \varepsilon\right) \leq \mathbb{P}(X_n \leq a) \leq \mathbb{P}(X \leq a + \varepsilon) + \mathbb{P}\left(|X_n - X| > \varepsilon\right)$$

3. Conclure que $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$.

Exercice 13 On considère une suite de variables aléatoires $(X_n)_{n\geq 1}$, définies sur $(\Omega, \mathscr{A}, \mathbb{P})$, où, pour tout entier naturel $n \geq 1$, X_n suit la loi normale d'espérance nulle et d'écart-type $1/n$.
On définit, pour tout entier naturel n non nul, l'application Y_n par :

$$\forall \omega \in \Omega, Y_n(\omega) = \int_{-1}^{1} \left| X_n(\omega) - t \right| dt$$

On admet que, pour tout entier naturel non nul, Y_n est une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbb{P})$.
On note F_{Y_n} la fonction de répartition de Y_n et Φ celle de la loi normale centrée réduite.

1. Exprimer, pour tout réel y, $F_{Y_n}(y)$ en fonction de $\Phi(y)$ et de n.
2. Montrer que la suite $(Y_n)_{n\geq 1}$ converge en loi.