Chapitre 11 - Convergence de variables aléatoires

Réviser : si besoin les DL, les équivalents classiques.

Toutes les variables aléatoires considérées dans ce chapitre seront toujours définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

I. Inégalités de Markov et de Bienaymé-Tchebychev (rappel)

I.1) Inégalité de Markov

Théorème I.1
Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire réelle.
Si $X(\Omega) \subset \mathbb{R}_+$ (c'est-à-dire si $X \geqslant 0$), et si X admet une espérance, alors

$$\forall a > 0, \quad P(X \geqslant a) \leqslant \frac{E(X)}{a}$$

I.2) Inégalité de Bienaymé-Tchebychev

Théorème I.2
Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire réelle.
On suppose que la variable X admet un moment d'ordre 2.

$$\forall \varepsilon > 0, \quad P(|X - E(X)| \geqslant \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2}$$

II. Convergence en probabilité

Définition II.1
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles.
Soit X une variable aléatoire réelle.
On dit que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ **converge en probabilité vers la variable aléatoire** X lorsque :

$$\forall \epsilon > 0, \quad \lim_{n\to+\infty} P\Big(|X_n - X| \geqslant \varepsilon\Big) = 0$$

On note alors $X_n \xrightarrow{P} X$.

Méthode :
Dans beaucoup de cas, pour montrer qu'une suite de variables aléatoires converge en probabilité, on écrit l'inégalité de Bienaymé Tchebychev avec X_n où la suite de variable $(X_n)_{n\in\mathbb{N}^*}$ est composée de variables ayant toutes la même espérance m.

$$\forall \varepsilon > 0, \quad P(|X_n - m| \geqslant \varepsilon) \leqslant \frac{V(X_n)}{\varepsilon^2}$$

Il restera ensuite à prouver que la variance de X_n tend vers 0 lorsque n tend vers $+\infty$, puis à appliquer le théorème des encadrements pour conclure.

Exemple

1. On suppose que, pour tout entier naturel n non nul, X_n est une variable aléatoire qui suit une loi $\mathcal{N}\left(0, \frac{1}{n}\right)$.
   Montrer que la suite $(X_n)_{n\geq 1}$ converge en probabilité vers la variable aléatoire certaine nulle.
2. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. vérifiant:

$$\forall n \in \mathbb{N}^*, \quad X_n(\Omega) = \{\frac{1}{n}, n\}, \qquad P(X_n = \frac{1}{n}) = \frac{n}{n+1} \quad \text{et} \quad P(X_n = n) = \frac{1}{n+1}$$

   Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire certaine nulle.

Théorème II.1
Composition par une fonction continue
Soit f une fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles. Soit X une variable aléatoire réelle.

Si la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X, et si la fonction f est **continue** sur $\mathbb{R}$, alors la suite $(f(X_n))_{n\in\mathbb{N}}$ converge en probabilité vers $f(X)$.

Autrement dit, si la fonction f est **continue** sur $\mathbb{R}$ et si $X_n \xrightarrow{P} X$, alors $f(X_n) \xrightarrow{P} f(X)$.

Théorème II.2
Loi faible des grands nombres
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles.
Notons

$$\forall n \in \mathbb{N}^*, \quad \overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{X_1 + \cdots + X_n}{n}$$

La variable $\overline{X_n}$ s'appelle la moyenne empirique de X_1, ..., X_n. Si :

- les variables X_n sont **indépendantes**,
- elles admettent toutes la **même espérance** notée m et la **même variance** notée σ^2,

alors la suite $(\overline{X_n})_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire certaine égale à $m = E(X_1)$

$$\overline{X_n} \xrightarrow{P} m$$

Autrement dit $\forall \varepsilon > 0, \lim_{n\to+\infty} \left(P\left(|\overline{X_n} - m| \geqslant \varepsilon\right)\right) = 0.$

Démonstration à connaître

Exercice 1
Soit $p \in]0, 1[$. Soit S_n une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n, p)$.
Montrer que $\left(\frac{1}{n}S_n\right)_{n\geqslant 1}$ converge en probabilité vers la variable aléatoire certaine p.

Interprétation : Ce résultat prouve que, lors d'une succession d'épreuves de Bernoulli indépendantes, pour lesquelles la probabilité de succès est égale à p, la fréquence des succès $\overline{X_n} = \frac{S_n}{n}$ converge en probabilité vers la probabilité théorique p du succès, ce qui justifie la notion de probabilité telle qu'elle a été définie historiquement à partir de la fréquence statistique.

Exercice 2
Soit une suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a.r. indépendantes, suivant toutes une loi de Poisson de paramètre n.

On note, pour tout entier naturel n non nul, $S_n = \dfrac{X_n}{n}$.
Montrer que la suite $(S_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire certaine égale à 1.

Théorème II.3
Convergence en probabilité et somme (Admis)
Si $X_n \xrightarrow{P} X$ et $Y_n \xrightarrow{P} Y$ alors $X_n + Y_n \xrightarrow{P} X + Y$

III. Convergence en loi

III.1) Définition

Définition III.1
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles.
Soit X une variable aléatoire réelle .
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi vers** la variable X lorsque la fonction de répartition de X_n converge vers la fonction de répartition de X en tout point de continuité de cette dernière, autrement dit, si en tout point de continuité x de F_X :

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

On note $X_n \xrightarrow{\mathscr{L}} X$.

La convergence en loi désigne "la convergence des fonctions de répartitions". Il s'agira donc de faire un calcul de limite, ce qui dépend entièrement du cours d'analyse (et éventuellement équivalents, DL...).

Remarque
$X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{\mathscr{L}} X$
Attention résultat hors-programme. Réciproque fausse.

Définition III.2
Cas des variables discrètes à valeurs dans $\mathbb{N}$

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles.
Soit X une variable aléatoire réelle.
On suppose que $\forall n \in \mathbb{N}$, $X_n(\Omega) \subset \mathbb{N}$ et que $X(\Omega) \subset \mathbb{N}$.
La suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X si et seulement si

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

Remarque
Encore vrai pour des variables à valeurs dans $\mathbb{Z}$.

Théorème III.1
Composition par une fonction continue
Soit f une fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$.
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles. Soit X une une variable aléatoire réelle.
Si la suite de variables aléatoires $(X_n)_{n\geqslant 1}$ **converge en loi** vers X, et si la fonction f est **continue** sur $\mathbb{R}$, alors la suite de variables aléatoires $\big(f(X_n)\big)_{n\geqslant 1}$ converge en loi vers la variable aléatoire $f(X)$.

$$f(X_n) \xrightarrow{\mathscr{L}} f(X)$$

Exercice 3
Soit X une variable aléatoire à densité.
On note, pour tout entier naturel n non nul, $X_n = e^{-\frac{1}{n}}\, X$.
Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X

Exercice 4
On considère une suite $(X_n)_{\mathbb{N}^*}$ de variables aléatoire réelles discrètes.
On suppose que $\forall n \in \mathbb{N}^* \quad X_n \hookrightarrow \mathscr{P}\left(\dfrac{1}{n}\right)$.
Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers la variable certaine égale à 0.

Exercice 5
Soit $\lambda > 0$ et $N = \lfloor \lambda + 1 \rfloor$.
Soit $n \geqslant N$. On suppose que X_n suit une loi géométrique de paramètre $\left(\frac{\lambda}{n}\right)$.

1. Déterminer la fonction de répartition d'une loi géométrique de paramètre $p \in]0,1[$.
2. Montrer que, pour tout réel $x > 0$, $\lfloor nx \rfloor \underset{n\to+\infty}{\sim} nx$.
3. Déterminer la fonction de répartition de $\frac{X_n}{n}$, puis montrer que $\left(\frac{X_n}{n}\right)$ converge en loi vers une variable aléatoire X telle que $X \hookrightarrow \mathcal{E}(\lambda)$.

Exercice 6
On admet que la fonction $F : x \longmapsto e^{-e^{-x}}$ est une fonction de répartition d'une variable aléatoire à densité (loi de Gumbel)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes identiquement distribuées de loi exponentielle de paramètre 1. On note pour tout entier n, $M_n = \max(X_1, \cdots, X_n)$.

Montrer que la suite de variables aléatoires $\big(M_n - \ln(n)\big)$ converge en loi vers une variable aléatoire à densité qui suit la loi de Gumbel.

III.2) Convergence de la loi binomiale vers une loi de Poisson

Théorème III.2
Soit λ un réel strictement positif.
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.
Si pour tout entier naturel n non nul,

$$X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$$

alors la suite $(X_n)_{n\in\mathbb{N}^*}$ converge **en loi** vers une variable aléatoire Y qui suit une loi de Poisson $\mathcal{P}(\lambda)$.

$$X_n \xrightarrow{\mathscr{L}} Y \quad \text{où } Y \hookrightarrow \mathcal{P}(\lambda)$$

Preuve guidée

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires, telle que X_n suit la loi binomiale de paramètre $\left(n, \frac{\lambda}{n}\right)$.

1. Montrer que $\forall k \in \mathbb{N}$, $\displaystyle\binom{n}{k} \underset{n\to+\infty}{\sim} \frac{n^k}{k!}$.

2. En déduire que $\forall k \in \mathbb{N}$, $\lim_{n\to+\infty} P(X_n = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}$.

IV. Théorème limite central

IV.1) Le théorème

Théorème IV.1
Théorème limite central (Admis, preuve très difficile !)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.
Soit m un réel et σ un réel strictement positif.
On suppose que :

- les variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ sont indépendantes,
- les variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ sont identiquement distribuées (elles suivent la même loi),
- toutes ces variables aléatoires admettent une même espérance m et une même variance notée σ^2.

On note : $\forall n \in \mathbb{N}^*$, $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$. La variable aléatoire centrée réduite associée à $\overline{X_n}$ est :

$$\overline{X}_n^* = \frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{V\left(\overline{X}_n\right)}} = \sqrt{n}\left(\frac{\overline{X}_n - m}{\sigma}\right)$$

Alors

la suite $(\overline{X_n}^{\,*})_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire qui suit une loi normale centrée réduite.

$$\overline{X_n}^{\,*} \xrightarrow{\mathscr{L}} N \quad \text{où } N \hookrightarrow \mathcal{N}(0,1)$$

On a donc, pour tout (a,b) tel que $-\infty \leq a \leq b \leq +\infty$,

$$\lim_{n\to+\infty} P(a \leq \overline{X_n}^{\,*} \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)\, dt$$

IV.2) Convergence de la loi binômiale vers la loi normale (Théorème de Moivre-Laplace)

Théorème IV.2
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.
On suppose que, pour tout entier naturel n non nul, X_n suit une loi binomiale $\mathcal{B}(n,p)$.
La suite de variables aléatoires $\left(\dfrac{X_n - np}{\sqrt{np(1-p)}}\right)$ converge en loi vers une variable aléatoire X qui suit une loi normale centrée réduite.

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathscr{L}} N \quad \text{où} \quad N \hookrightarrow \mathcal{N}(0,1)$$

On a donc

$$\forall x \in \mathbb{R}, \quad \lim_{n\to+\infty}\left(P\left(\frac{X_n - np}{\sqrt{np(1-p)}} \leqslant x\right)\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

Remarque
Approximation
On peut approcher une loi binomiale de paramètre (n,p) dès que $n \geq 20$ et p voisin de 0.5 par une loi normale de paramètre $(np, np(1-p))$.

IV.3) Convergence de la loi de Poisson vers la loi normale

Théorème IV.3
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires. Soit λ un réel strictement positif. On suppose que, pour tout entier naturel n non nul, X_n une variable aléatoire de Poisson de paramètre $(n\lambda)$.
La suite de variables aléatoires $\left(\dfrac{X_n - n\lambda}{\sqrt{n\lambda}}\right)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X qui suit une loi normale centrée réduite.

$$\frac{X_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{\mathscr{L}} N \quad \text{où } N \hookrightarrow \mathcal{N}(0,1)$$

D'où

$$\forall x \in \mathbb{R}, \quad \lim_{n\to+\infty}\left(P\left(\frac{X_n - n\lambda}{\sqrt{n\lambda}} \leqslant x\right)\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

Remarque
Approximation
Dès que $\lambda > 10$, on pourra approcher une loi de Poisson de paramètre λ par une loi normale de paramètres (λ, λ).