

Remise à niveau - Statistique

Rappels de probabilités

Claire Boyer

1. Définitions de base

2. Modes de convergence

3. Inégalités classiques

4. Théorèmes asymptotiques

5. Opérations sur les limites

Espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$

- ▶ Ω représente l'ensemble de toutes les éventualités possibles, toutes les déterminations du hasard dans l'expérience considérée
- ▶ $\mathcal{A}$ est l'ensemble des événements, qui sont les parties de Ω dont on peut évaluer la probabilité

$(\Omega, \mathcal{A})$ est un espace mesurable

- ▶ Pour $A \in \mathcal{A}$, $\mathbb{P}(A)$ représente la probabilité d'occurrence de l'événement A

Variable aléatoire

- ▶ Soit $(E, \mathcal{E})$ un espace mesurable. Une application mesurable $X : \Omega \rightarrow E$ est appelée variable aléatoire à valeurs dans E

1. Définitions de base

2. Modes de convergence

3. Inégalités classiques

4. Théorèmes asymptotiques

5. Opérations sur les limites

On considère une suite de variables aléatoires (X_n) et on s'intéresse à ses différents modes de convergence :

- ▶ Convergence en loi
- ▶ Convergence en probabilité
- ▶ Convergence presque sûre
- ▶ Convergence en moyenne quadratique

Définition

Soit (X_n) une suite de variables aléatoires. On dit que (X_n) **converge en loi** vers X si, pour toute fonction (uniformément) continue bornée φ ,

$$\mathbb{E}[\varphi(X_n)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[\varphi(X)].$$

Exemple 1 Si $X_n \sim \mathcal{B}\left(p + \frac{1-p}{n}\right)$, alors

$$X_n \xrightarrow{\mathcal{L}} \mathcal{B}(p).$$

Exemple 2 Si $X_n = -X$, avec $X \sim \mathcal{N}(0, 1)$, alors

$$X_n \xrightarrow{\mathcal{L}} X.$$

Proposition

Soit $(X_n)_n$ une suite de variables aléatoires. X_n converge en loi vers une variable aléatoire X

1. ssi en tout point de continuité x de la *fonction de répartition* F_X de X on a

$$F_{X_n}(x) \xrightarrow[n \rightarrow \infty]{} F_X(x)$$

2. si^a les *fonctions densités* de X_n et X , supposées *absolument continues*,

$$f_{X_n}(x) \xrightarrow[n \rightarrow \infty]{} f_X(x)$$

3. ssi pour tout $t \in \mathbb{R}$, les *fonctions caractéristiques* vérifient

$$\Phi_{X_n}(t) \xrightarrow[n \rightarrow \infty]{} \Phi_X(t) \quad \text{avec} \quad \Phi_X(t) := \mathbb{E} [e^{itX}]$$

a. suffisant, e.g. $F_{X_n}(x) = x - \sin(2\pi nx)/2\pi n$ sur $[0, 1]$ converge vers $F_{\mathcal{U}([0,1])}$, mais pas en terme de densité

Définition

On dit que (X_n) converge en probabilité vers X si, pour tout $\varepsilon > 0$,

$$\mathbb{P}(|X_n - X| > \varepsilon) \xrightarrow[n \rightarrow \infty]{} 0.$$

Proposition

Si X_n converge en probabilité vers X , alors X_n converge en loi vers X .

- ▶ Attention : réciproque fausse en général
- ▶ Contre-exemple : $X_n = -X$ avec $X \sim \mathcal{N}(0, 1)$

Proposition

Si X_n converge en loi vers une constante c , alors X_n converge en probabilité vers c .

Définition

Soit (X_n) une suite de variables aléatoires. On dit que X_n converge presque sûrement vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$ si

$$\mathbb{P} \left(\lim_{n \rightarrow +\infty} X_n = X \right) = \mathbb{P} \left(\left\{ \omega \in \Omega, \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega) \right\} \right) = 1$$

A retenir

$$X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X \implies X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} X \implies X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$$

Application au lemme de Borel-Cantelli : Si pour tout $\varepsilon > 0$, $\sum_n \mathbb{P}(|X_n - X| \geq \varepsilon) < +\infty$ alors (X_n) converge presque sûrement vers X .

Exemple : Soit (X_n) telle que pour tout $n \geq 1$, $\mathbb{P}(X_n = 0) = 1 - 1/n^2$ et $\mathbb{P}(X_n = n) = 1/n^2$

Définition

Soit (X_n) une suite de variables aléatoires de carré intégrable. On dit que X_n converge en moyenne quadratique vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{L^2} X$ si

$$\mathbb{E} [|X_n - X|^2] \xrightarrow[n \rightarrow +\infty]{} 0$$

Exemple : Soit (X_n) telle que pour tout $n \geq 0$, $\mathbb{P}(X_n = 0) = 1 - 1/n$ et $\mathbb{P}(X_n = 1) = 1/n$

A retenir

$$X_n \xrightarrow[n \rightarrow +\infty]{L^2} X \implies X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} X$$

- ▶ **Achtung :** réciproque fausse en général
- ▶ **Contre-exemple :** Soit (X_n) telle que pour tout $n \geq 1$,

$$\mathbb{P}(X_n = 0) = \frac{n-1}{n} \quad \mathbb{P}(X_n = n) = \frac{1}{n}$$

1. Définitions de base
2. Modes de convergence
3. Inégalités classiques
4. Théorèmes asymptotiques
5. Opérations sur les limites

Inégalité de Markov. Pour $c, p > 0$ et X admettant un moment d'ordre p :

$$\mathbb{P}(|X| \geq c) \leq \frac{\mathbb{E}[|X|^p]}{c^p}$$

Inégalité de Bienaymé-Tchebychev. Si X admet une variance finie, alors pour $c > 0$:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq c) \leq \frac{\text{Var}(X)}{c^2}$$

Exemple : lancer de pièces. On considère des v.a. i.i.d. $X_n \sim \mathcal{B}(p)$. Alors pour $\varepsilon > 0$,

$$\mathbb{P}(|\bar{X}_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}$$

Inégalité de Cauchy-Schwarz. Soient X et Y deux v.a. admettant un moment d'ordre 2. Alors

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[|X|^2] \mathbb{E}[|Y|^2]}$$

Inégalité de Holder. Soient X et Y deux v.a. et p, q tels que $\frac{1}{p} + \frac{1}{q} = 1$. Si X et Y admettent resp. un moment d'ordre p et q , alors

$$\mathbb{E}[|XY|] \leq (\mathbb{E}[|X|^p])^{1/p} (\mathbb{E}[|Y|^q])^{1/q}$$

1. Définitions de base
2. Modes de convergence
3. Inégalités classiques
4. Théorèmes asymptotiques
5. Opérations sur les limites

Loi faible des grands nombres. Si (X_i) sont i.i.d. de moyenne m , alors

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} m.$$

Loi forte des grands nombres. Si (X_i) sont i.i.d. de moyenne m , alors

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} m.$$

Théorème Central Limite. Si (X_i) sont i.i.d. de moyenne m et variance σ^2 , alors

$$\sqrt{n}(\overline{X}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

ce qui peut aussi s'écrire $\frac{\sqrt{n}}{\sigma}(\overline{X}_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$

1. Définitions de base
2. Modes de convergence
3. Inégalités classiques
4. Théorèmes asymptotiques
5. Opérations sur les limites

Théorème (Théorème de continuité)

Soit (X_n) une suite de v.a. convergeant (soit *p.s.* ou en $\mathbb{P}$ ou en $\mathcal{L}$) vers X une v.a.. Soit g une fonction dont on appelle D_g l'ensemble de points de discontinuité.

Si $\mathbb{P}(X \in D_g) = 0$, alors la suite $g(X_n)$ hérite du mode de convergence (*p.s.* / $\mathbb{P}$ / $\mathcal{L}$) de X_n .

- **Exemple : Lancer de pièces.** Si $X_i \sim \mathcal{B}(p)$, et $p \in]0, 1[$,

$$\frac{1}{\sqrt{\bar{X}_n(1 - \bar{X}_n)}} \xrightarrow[n \rightarrow +\infty]{p.s.} \frac{1}{\sqrt{p(1 - p)}}$$

- N'est pas vrai en général pour la convergence en moyenne quadratique
- **Contre-exemple :** Soit (X_n) telle que pour tout $n \geq 1$, $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n^3}$ et $\mathbb{P}(X_n = n) = \frac{1}{n^3}$ et la fonction $g : x \mapsto x^2$

Théorème (Théorème de Slutsky)

Si $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ et $Y_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} c$, alors

$$X_n + Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X + c, \quad X_n Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} cX.$$

► **Exemple : lancer de pièces.** On a

$$\sqrt{n} \frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1 - \bar{X}_n)}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

- Généralement faux quand (Y_n) converge en loi
- **Contre-exemple :** Soient $X \sim \mathcal{N}(0, 1)$, et (X_n) et (Y_n) deux suites de v.a. définies par $X_n = -X$ et $Y_n = X$ pour tout n

Corollaire

Soient (v_n) une suite déterministe tendant vers $+\infty$, (X_n) une suite de v.a., X une v.a., et a une constante telles que

$$v_n(X_n - a) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$$

alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} a$

► (TCL) En particulier si $\sqrt{n}(X_n - a) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ alors $X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} a$

Théorème

Soient (v_n) une suite déterministe tendant vers $+\infty$, (X_n) une suite de v.a., X une v.a., et a une constante telles que

$$\sqrt{v_n}(X_n - a) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$$

Soit g une fonction dérivable en a , alors

$$\sqrt{v_n}(g(X_n) - g(a)) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} g'(a)X$$

► En particulier, si $\sqrt{v_n}(X_n - a) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2)$ alors

$$\sqrt{v_n}(g(X_n) - g(a)) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, (g'(a))^2 \sigma^2)$$