

MODÉLISATION STATISTIQUE

Licence 3 MIA SHS – Université de Bordeaux

– Chapitre I –

Notes de cours

Suites de variables aléatoires et théorème central limite

Considérons un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ formé d'un ensemble Ω , d'une tribu \mathcal{A} sur Ω et d'une mesure \mathbb{P} sur \mathcal{A} telle que $\mathbb{P}(\Omega) = 1$.

1 Quelques inégalités célèbres

1.1 Inégalité de Markov

Théorème 1.1 *Soit X une variable aléatoire réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, positive ou nulle et d'espérance finie. Alors, pour tout $\varepsilon > 0$, on a*

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon}.$$

Démonstration.

Corollaire 1.1 *Soit X une variable aléatoire réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, et ϕ une fonction croissante et positive ou nulle sur un intervalle D_ϕ tel que $\mathbb{P}(X \in D_\phi) = 1$. Alors, pour tout $\varepsilon \in D_\phi$ tel que $\phi(\varepsilon) > 0$, on a*

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[\phi(X)]}{\phi(\varepsilon)}.$$

Démonstration.

1.2 Inégalité de Bienaymé-Tchebychev

Théorème 1.2 Soit X une variable aléatoire réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, d'espérance μ et de variance finie. Alors, pour tout $\varepsilon > 0$,

$$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{\mathbb{E}[(X - \mu)^2]}{\varepsilon^2}.$$

Démonstration.

2 Convergence en probabilité

Définition 2.1 Soit (X_n, X) une suite de variables aléatoires réelles définies sur le même espace $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que (X_n) converge en probabilité vers X si, pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0$$

et l'on note

$$X_n \xrightarrow{\mathbb{P}} X.$$

2.1 Loi faible des grands nombres

Théorème 2.1 Soit (X_n) une suite de variables aléatoires réelles définies sur le même espace $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes, d'espérance μ et de variance σ^2 finie. Alors,

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{\mathbb{P}} \mu.$$

Démonstration.

2.2 Théorème de Bernoulli

Théorème 2.2 *Soit (X_n) une suite de variables aléatoires réelles, indépendantes, associées à la répétition d'un même évènement aléatoire dont le succès est noté S . Soit $p = \mathbb{P}(S)$ et, pour tout $n \in \mathbb{N}^*$, $X_n \sim \mathcal{B}(p)$. Alors,*

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{\mathbb{P}} p.$$

Démonstration.

3 Convergence en loi

Définition 3.1 *Soit (X_n, X) une suite de variables aléatoires réelles définies sur le même espace $(\Omega, \mathcal{A}, \mathbb{P})$. On note F_n la fonction de répartition associée à X_n pour tout $n \in \mathbb{N}^*$ et F la fonction de répartition associée à X . On dit que (X_n) converge en loi vers X si, en tout point x de continuité de F ,*

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

et l'on note

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

Exemple. Si (X_n) est une suite de variables aléatoires réelles telles que, pour tout $n \in \mathbb{N}^*$, X_n suit la loi binomiale $\mathcal{B}(n, \lambda/n)$ avec $\lambda > 0$. Alors,

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

où X suit la loi de Poisson $\mathcal{P}(\lambda)$.

En effet,

On en déduit la convergence simple des fonctions de répartition et, par la Définition 3.1, la convergence en loi de X_n vers X .

Proposition 3.1 Soit (X_n, X) une suite de variables aléatoires réelles définies sur le même espace $(\Omega, \mathcal{A}, \mathbb{P})$. Alors,

$$X_n \xrightarrow{\mathbb{P}} X \quad \implies \quad X_n \xrightarrow{\mathcal{L}} X$$

mais la réciproque est généralement fausse.

Démonstration. Admis.

La réciproque de la Proposition 3.1 est vraie lorsque X est une constante. On a alors

$$X_n \xrightarrow{\mathcal{L}} c \quad \implies \quad X_n \xrightarrow{\mathbb{P}} c.$$

Il faut bien noter qu'il s'agit ici d'un cas très particulier où la loi limite de X_n décrit une masse de Dirac centrée en c .

3.1 Théorème de Slutsky

Théorème 3.1 Soit (X_n, X) et (Y_n) deux suites de variables aléatoires réelles telles que

$$X_n \xrightarrow{\mathcal{L}} X \quad \text{et} \quad Y_n \xrightarrow{\mathbb{P}} c$$

avec $c \in \mathbb{R}$. Alors, on a les convergences

$$X_n + Y_n \xrightarrow{\mathcal{L}} X + c \quad \text{et} \quad X_n Y_n \xrightarrow{\mathcal{L}} cX.$$

Démonstration. Admis.

Attention : si $Y_n \xrightarrow{\mathcal{L}} Y$ où Y n'est pas une constante alors, dans le cas général, $X_n + Y_n$ ne converge pas en loi vers $X + Y$.

3.2 Théorème central limite

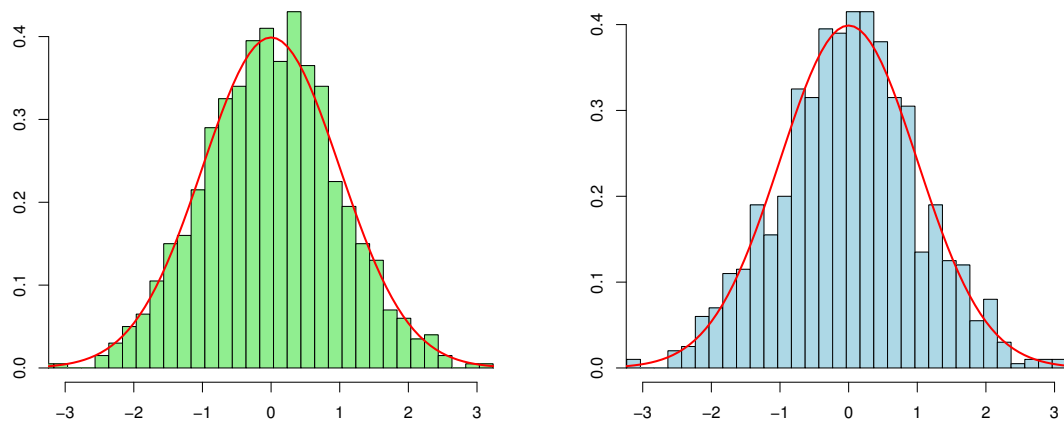
Théorème 3.2 Soit (X_n) une suite de variables aléatoires réelles, indépendantes et identiquement distribuées, d'espérance μ et de variance finie $\sigma^2 > 0$. Soit encore la somme $S_n = X_1 + \dots + X_n$. Alors, en notant $\bar{X}_n = \frac{S_n}{n}$,

$$Z_n = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \xrightarrow{\mathcal{L}} Z$$

où Z suit la loi normale $\mathcal{N}(0, 1)$.

Remarque : On écrit aussi le théorème avec $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$.

Exemples. Les deux graphes ci-dessous représentent l’histogramme associé à $N = 1000$ réalisations de Z_n pour $n = 500$, lorsque la suite aléatoire (X_n) est indépendante et identiquement distribuée, de loi $\mathcal{P}(2)$ puis de loi $\mathcal{B}(0.3)$. Nous y avons superposé la densité de la loi normale $\mathcal{N}(0, 1)$. Ils illustrent assez nettement le théorème central limite.



La connaissance de la loi des X_n n’est en outre pas nécessaire (mais la connaissance de μ et de σ est nécessaire). De plus, cela nous montre qu’une suite de variables aléatoires discrètes peut converger vers une variable aléatoire continue.

Considérons maintenant à titre d’exemple une suite de variables aléatoires (X_n) , indépendantes, et de loi de Bernoulli $\mathcal{B}(p)$. La variable aléatoire définie, pour tout $n \in \mathbb{N}^*$, par

$$S_n = X_1 + \dots + X_n$$

suit une loi binomiale $\mathcal{B}(n, p)$. On a de plus $\mathbb{E}[X_n] = p$, $\mathbb{V}(X_n) = p(1-p)$. Par application du théorème central limite, on a la convergence en loi

$$\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} Z$$

où Z suit la loi normale $\mathcal{N}(0, 1)$. La loi faible des grands nombres nous donne également

$$\hat{p}_n = \bar{X}_n = \frac{S_n}{n} \xrightarrow{\mathbb{P}} p$$

qui est une réécriture du théorème de Bernoulli. Ainsi, par le théorème de Slutsky, nous avons un théorème central limite associé à l’estimateur \hat{p}_n de p , qui s’exprime par

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \xrightarrow{\mathcal{L}} Z$$

et qui décrit la variabilité de l’estimateur \hat{p}_n autour de sa valeur moyenne par l’intermédiaire d’une loi gaussienne.

En effet,