Chapitre 3: Théorème central limite

Prof. Mohamed El Merouani

Université Abdelmalek Essaâdi
Faculté des Sciences de Tétouan
Département de Mathématiques
https://elmerouani.jimdofree.com/calcul-des-probabilités

2020/2021

Plan

Introduction :

- Nous allons étudier la tendance d'une suite de variable aléatoire vers une variables aléatoires Normale et nous allons établir quelques théorèmes fournissant des conditions suffisantes pour que cette tendance ait lieu.
- Tout théorème de ce type porte le nom du théorème central limite.
- Le théorème central limite donne des conditions suffisantes dans lesquelles une somme finie de variables aléatoires réelles convenablement normalisée suit approximativement une loi Normale.

Théorème de Lindeberg-Lévy :

Théorème de Lindeberg-Lévy :

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, équidistribuées, de carrés intégrables. Posons : $\mu = E(X_1)$; $\sigma^2 = Var(X_1)$ $(0 < \sigma < +\infty)$; $S_n = \sum_{k=1}^{n} X_k$; $\overline{X_n} = \frac{S_n}{n}$; $Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}}$. Alors la suite $(Y_n)_{n\geq 1}$ converge en loi vers une variable aléatoire de loi $\mathcal{N}(0,1)$, c'est-à-dire que pour tout réel x, on a : $P(Y_n \leq x) \underset{n\to+\infty}{\longrightarrow} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$

Preuve :

Désignons par φ la fonction caractéristique de $X_1 - \mu$; cette variable aléatoire étant centrée, de carré intégrable, la formule de Taylor jusqu'à l'odre deux de $\varphi(t)$ au voisinage de $t = 0$ fournit

$\varphi(t) = 1 - \frac{t^2}{2}\sigma^2 + o(t^2)$, $(t \to 0)$.

Théorème de Lindeberg-Lévy :

Preuve :

Or $Y_n = \sum_{k=1}^{n} \frac{X_k - \mu}{\sigma\sqrt{n}}$,

d'où $\varphi_{Y_n}(t) = \left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n = \left(1 - \frac{t^2}{2n} + o(\frac{t^2}{n})\right)^n \to e^{-\frac{t^2}{2}}$

ce qui établit le résultat en vertu du théorème de Paul-Lévy. ■

Remarque 1 :

L'importance du théorème de Lindeberg-Lévy réside dans le fait que pour n "grand" les lois de probabilité (en général compliquées) de S_n, $\overline{X_n}$ peuvent être approchées par les lois Normales $\mathcal{N}(n\mu, \sigma\sqrt{n})$, $\mathcal{N}(\mu, \sigma/\sqrt{n})$.

Le théorème de Lindeberg-Lévy comporte un certain nombre de cas particuliers que nous allons examiner.

Cas particulier 1 : Théorème de Moivre-Laplace

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, de loi commune des X_n la loi de Bernoulli de paramètre p. Alors S_n suit la loi binomiale $\mathcal{B}(n,p)$. On a $E(S_n) = np$, $Var(S_n) = npq$ et le théorème précédent affirme :

Proposition 1 :

La suite des variable aleatoire de terme général $Y_n = \frac{S_n - np}{\sqrt{npq}}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0,1)$

Il en résulte que, pour n "grand" la loi binomiale $\mathcal{B}(n,p)$ peut être approchée par la loi Normale $\mathcal{N}(np, \sqrt{npq})$.

Cas particulier 2 :

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, de loi commune des X_n la loi de Poisson de paramètre 1. Alors S_n suit la loi de Poisson de paramètre n ; on a $E(S_n) = n$; $Var(S_n) = n$ et le théorème précédent affirme :

Proposition 2 :

La suite des variables aléatoire de terme général $Y_n = \frac{S_n - n}{\sqrt{n}}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0, 1)$.

Il en résulte que, pour n "grand" la loi de Poisson $\mathcal{P}(n)$ peut être approchée par la loi Normale $\mathcal{N}(n, \sqrt{n})$.

Cas particulier 2 :

Remarque 2 : (Bernstein)

Il résulte de la proposition 2 que lorsque n tend vers l'infini, on a :
$P\left(\frac{S_n-n}{\sqrt{n}} \leq 0\right) \underset{n\to+\infty}{\longrightarrow} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx = \frac{1}{2}$
c'est-à-dire que $P(S_n \leq n) = e^{-n}\sum_{k=0}^{n} \frac{n^k}{k!} \underset{n\to+\infty}{\longrightarrow} \frac{1}{2}$
(On remarquera que ce dernier passage à la limite n'est pas aisé à établir sans faire appel au théorème central limite)

Remarque 3 :

La proposition 2 peut être étendue au cas d'une famille $(X_\lambda)_{\lambda>0}$ de variables aléatoires, où le terme général X_λ suit la loi de Poisson $\mathcal{P}(\lambda)$ de paramètre λ. On peut montrer directement que lorsque λ tend vers l'infini, $\frac{X_\lambda-\lambda}{\sqrt{\lambda}}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0,1)$, de sorte que, pour λ "grand", la loi de Poisson $\mathcal{P}(\lambda)$ peut être approchée par la loi Normale $\mathcal{N}(\lambda, \sqrt{\lambda})$.

Cas particulier 3 :

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes, de loi commune des X_n la loi exponentielle $\mathcal{E}xp(\lambda)$, de paramètre $\lambda > 0$. Alors S_n suit la loi gamma $\Gamma(n, \lambda)$; on a $E(S_n) = \frac{n}{\lambda}$; $Var(S_n) = \frac{n}{\lambda^2}$, et le théorème de Lindeberg-Lévy affirme :

Proposition 3 :

La suite des variables aléatoires de terme général $Y_n = \frac{S_n - \frac{n}{\lambda}}{\sqrt{n}/\lambda}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0, 1)$.

Il en résulte que, pour n "grand", la loi gamma $\Gamma(n, \lambda)$ peut être approchée par la loi Normale $\mathcal{N}(\frac{n}{\lambda}, \frac{\sqrt{n}}{\lambda})$.

Cas particulier 3 :

Remarque 4 :

La proposition 3 peut être étendue au cas d'une famille $(X_p)_{p>0}$ de variables aléatoires, où le terme général X_p suit la loi gamma $\Gamma(p, \lambda)$, avec $\lambda > 0$ fixé. On peut montrer directement que, lorsque p tend vers l'infini, λ restant fixé, la variable aléatoire $\frac{X_p - \frac{p}{\lambda}}{\sqrt{p}/\lambda}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0, 1)$, de sorte que, pour p "grand", la loi gamma $\Gamma(p, \lambda)$ peut être approchée par la loi Normale $\mathcal{N}(\frac{p}{\lambda}, \frac{\sqrt{p}}{\lambda})$.

Le Théorème central limite et la formule de Stirling :

Soit $(X_p)_{p>0}$ une famille de variables aléatoires dont le terme général X_p suit la loi gamma $\Gamma(p+1,1)$. Lorsque p tend vers l'infini, $\frac{X_p-(p+1)}{\sqrt{p+1}}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0,1)$; il est clair que $\frac{X_p-p}{\sqrt{p}}$ a le même comportement asymptotique lorsque p tend vers l'infini. Nous allons montrer que cette dernière propriété est équivalente à la formule de Stirling.

Théorème :

Soit $(X_p)_{p>0}$ une famille de variables aléatoires dont le terme général X_p suit la loi gamma $\Gamma(p+1,1)$. Alors chacune des deux propriétés suivantes peut être déduite de l'autre :

1. $\frac{X_p-p}{\sqrt{p}}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0,1)$, lorsque p tend vers l'infini (théorème central limite) ;
2. $\Gamma(p+1) \sim p^p e^{-p}\sqrt{2\pi p}$, lorsque p tend vers l'infini (formule de Stirling).

Le Théorème central limite et la formule de Stirling :

Preuve :
Désignons les densités de X_p et de $\frac{X_p - p}{\sqrt{p}}$ respectivement par f_p et g_p ;
on a : $f_p(x) = \frac{1}{\Gamma(p+1)} e^{-x} x^p \mathbb{I}_{[0,+\infty[}(x)$; $g_p(x) = \sqrt{p} f_p(p + x\sqrt{p})$.
Pour tout réel x, on peut choisir $p > 0$ assez grand pour que
$p + x\sqrt{p} > 0$; on a alors : $g_p(x) = \sqrt{p} \frac{1}{\Gamma(p+1)} e^{-(p+x\sqrt{p})} (p + x\sqrt{p})^p$

$$= \frac{p^p e^{-p} \sqrt{2\pi p}}{\Gamma(p+1)} \frac{1}{\sqrt{2\pi}} r_p(x)$$

où $r_p(x) = e^{-x\sqrt{p}} \left(1 + \frac{x}{\sqrt{p}}\right)^p$;

de là : $\int_{-\infty}^{x} g_p(u) du = \frac{p^p e^{-p} \sqrt{2\pi p}}{\Gamma(p+1)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} r_p(u) du. \qquad (1)$

Un calcul élémentaire montre que $r_p(u) \underset{p \to +\infty}{\longrightarrow} e^{-\frac{u^2}{2}}$;
d'où, en appliquant le théorème de convergence dominée, on a pour tout réel x

$$\int_{-\infty}^{x} r_p(u) du \underset{p \to +\infty}{\longrightarrow} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du, \qquad (2)$$

Le théorème résulte alors de (1) et (2). ■

Exercices :

Exercice 1 : (Remarque 3)

Soit $(X_\lambda)_{\lambda>0}$ une famille de variables aléatoires, où le terme général X_λ suit la loi de Poisson $\mathcal{P}(\lambda)$. Montrer que $\frac{X_\lambda-\lambda}{\sqrt{\lambda}}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0,1)$, lorsque λ tend vers l'infini.

Exercice 2 : (Remarque 4)

Soit $(X_p)_{p>0}$ une famille de variables aléatoires, où le terme général X_p suit la loi gamma $\Gamma(p,\lambda)$, $(\lambda>0)$. Montrer que $\frac{X_p-\frac{p}{\lambda}}{\sqrt{p}/\lambda}$ converge en loi vers une variable aléatoire de loi Normale $\mathcal{N}(0,1)$, lorsque p tend vers l'infini.

Exercices :

Exercice 3 : (Remarque 2)

Appliquer le théorème limite central à une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes de même loi de Poisson de paramètre 1 pour trouver la limite de la suite $u_n = e^{-n}\sum_{0\leq k\leq n}\frac{n^k}{k!}$, $n\in\mathbb{N}$.

Exercice 4 :

Soit X_n la somme de n variables aléatoires indépendantes uniformement réparties sur le segment $]0,1]$. Trouver la loi limite de la variable aléatoire X_n réduite. De quel théorème retrouve-t-on ici un cas particulier ?