THÉORÈME DE LA LIMITE CENTRALE (C7, E1)
(09 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

On appelle **théorème de la limite centrale** (tlc), ou parfois **théorème de la limite locale**, une propriété caractérisant la **convergence en loi** d'une **suite** $S = (S_N)_{N \in \mathbf{N}^*}$ liée à des sommes de **variables aléatoires** tq :

(1) $\quad T_N = \Sigma_{n=1}^N X_n , \qquad \forall N \in \mathbf{N}^*,$

vers la **loi normale**, lorsque N tend vers l'infini (cf **problème de la limite centrale**).

(i) Dans le cas équidistribué, soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** et $\xi : \Omega \mapsto \mathcal{X}$ une **variable aléatoire**, où $\mathcal{X}$ est un **espace de BANACH** réel séparable, noté $(E, ||.||)$. On étudie une suite $X = (X_n)_{n \in \mathbf{N}}$ constituée de **copies** indépendantes et de même loi P^ξ que ξ (**suite iid** ou **processus purement aléatoire** équidistribué en **temps** discret).

On dit alors que ξ (ou X) vérifie le **théorème de la limite centrale** ssi il existe un **vecteur aléatoire** gaussien $\gamma \sim \mathcal{N}_{\mathcal{X}}(0 , \Sigma)$ tq :

(2) $\quad \mathcal{L}(N^{-1/2} T_N) \rightarrow_{N \rightarrow +\infty} \mathcal{N}_{\mathcal{X}}(0 , \Sigma).$

où T_N est défini en (1).

Un **vecteur aléatoire gaussien** $\gamma : \Omega \mapsto \mathcal{X}$ est, en effet, défini comme un **élément aléatoire** tq, si $\mathcal{X}'$ désigne le **dual topologique** de $\mathcal{X}$ et $<. , .>$ le **produit scalaire** canonique $\mathcal{X}' \times \mathcal{X} \mapsto \mathcal{X}$, les **formes linéaires** (γ , g), aléatoires et à valeurs dans **R**, sont gaussiennes pour tout $g \in \mathcal{X}'$. On note alors $\gamma \sim \mathcal{N}_{\mathcal{X}}(\mu , \Sigma)$, expression dans laquelle $\mu \in \mathcal{X}$ et Σ est l'**opérateur de covariance** associé à γ.

(iii) Les hypothèses qui permettent cette propriété de convergence sont diverses. Elles portent notamment :

(a) sur la **nature de la suite** à considérer : eg **total** (comme en (1)) ou **moyenne arithmétique**. On doit aussi la modifier pour obtenir la convergence voulue : soit par transformation algébrique (cf **variable centrée**, **variable réduite**, **variable normée**), soit en fonction des propriétés stochastiques de la suite (cf **modification**, **suite équidistribuée**, **suite indépendante**, **suite iid**) ;

(b) sur le **mode de convergence** de la suite $(T_N)_{N \in \mathbf{N}^*}$ ou de ses transformations ;

(c) sur la **loi limite** (ou **loi asymptotique**) spécifiée à l'avance, qui est gaussienne dans les cas les plus courants.

Il existe donc de nombreuses versions de ce théorème.

(iv) On considère ainsi un espace probabilisé $(\Omega, \mathcal{T}, P)$ et une suite $X = (X_n)_{n \in \mathbf{N}^*}$ constituée de **vecteurs aléatoires** $X_n : \Omega \mapsto \mathbf{R}^K$. On distingue ci-après selon que X est une **suite équidistribuée** ou non, et l'on ne considère que des convergences vers une loi normale.

(a) **cas équidistribué** :

(1) **version vectorielle de J.W. LINDEBERG - P.P. LÉVY**. On suppose que X est une **suite iid**, que $X_n \in \mathcal{L}_{\mathbf{R}K}^2 (\Omega, \mathcal{T}, P)$ et l'on note $E\, X_n = \mu$ et $V\, X_n = \Sigma$ (avec $\mathrm{rg}\, \Sigma \leq K$). On définit la suite de vecteurs aléatoires suivante :

(3) $S_N = N^{-1/2} (T_N - N\, \mu)$, où $N\, \mu = E\, T_N$, $\forall\, N \in \mathbf{N}^*$.

On établit alors la **convergence légale** suivante :

(2)' $S_N \to^{\mathcal{L}}{}_{N \to +\infty} \mathcal{N}_K(0 , \Sigma)$.

En particulier, lorsque K = 1, on obtient :

(2)'' $(N\, \sigma^2)^{-1/2} . (T_N - N\, \mu) \to^{\mathcal{L}}{}_{N \to +\infty} \mathcal{N}(0, 1)$.

(2) **version avec hypothèse sur la fonction caractéristique**. On suppose encore que X est iid selon P^ξ et que :

(4) $\xi \in \mathcal{L}_{\mathbf{R}K}^2 (\Omega, \mathcal{T}, P)$, $\varphi \in \mathcal{L}_{\mathbf{C}}^1 (\mathbf{R}^K, \mathcal{B}(\mathbf{R}^K), \lambda_K)$,

où **R**K désigne $\mathbf{R}^K$ et φ la **fonction caractéristique** de P^ξ.

Alors, lorsque N parcourt $\mathbf{N}^*$, la suite :

(3)' $S_N' = N^{-1/2} (T_N - E\, T_N) = N^{1/2} (N^{-1}\, T_N - E\, \xi) = N^{1/2} (\overline{X}_N - E\, \xi)$,

dans laquelle $\overline{X}_N$ désigne la **moyenne empirique**, est constituée de vecteurs aléatoires S_N' dont les **lois** resp possèdent des **densités de probabilité** h_N ($N \in \mathbf{N}^*$) absolument continues par rapport à la **mesure de LEBESGUE** λ_K (cf **loi absolument continue**) et qui convergent uniformément sur $\mathbf{R}^K$ vers la densité de la **loi normale** $\mathcal{N}(0, V\, \xi)$ (cf **convergence uniforme**). La **matrice des covariances** $V\, \xi$ de ξ est alors une **matrice régulière** ;

(b) **cas non (nécessairement) équidistribué**. Les théorèmes précédents se généralisent au cas où X n'est pas équidistribuée. On se limite au cas scalaire (K = 1), dans lequel X est constituée de **vars** $X_n : \Omega \mapsto \mathbf{R}$:

(1) **version scalaire de A.M. LIAPOUNOV (première version)**. On suppose que X est une **suite indépendante**, que les variables X_n suivent resp des lois $\mathscr{L}(X_n) = P^{X_n} = X_n(P)$, $\forall\ n \in \mathbf{N}^*$, et que les deux conditions suivantes sont vérifiées :

(5)
$$X_n \in \mathscr{L}_{\mathbf{R}}^3(\Omega, \mathscr{T}, P), \qquad \forall\ n \in \mathbf{N}^*,$$
$$\lim_N (A_N)^{1/3} / (B_N)^{1/2} = 0,$$

avec $A_N = \Sigma_{n=1}^N E\ |X_n - E\ X_n|^3$ et $B_N = \Sigma_{n=1}^N E\ |X_n - E\ X_n|^2$.

Alors la suite, indexée par N, des va $S_N = (V\ T_N)^{-1/2}\ (T_N - E\ T_N)$ tend en loi vers $\mathscr{N}(0, 1)$, ie :

$$(2)''' \quad (N\ \sigma^2)^{-1/2} . (T_N - N\ \mu) \to^{\mathscr{L}} \mathscr{N}(0, 1) \qquad \text{lorsque } N \to +\infty.$$

(2) **version scalaire de A.M. LIAPOUNOV (seconde version)**. On suppose X indépendante avec $X_n \sim \mathscr{L}(X_n) = P^{X_n}$, $\forall\ n \in \mathbf{N}^*$. Alors, pour que :

$$(6) \qquad \mathscr{L}(X_n) \to_{n \to +\infty} \mathscr{N}(0, 1),$$

il faut et il suffit que les deux conditions suivantes soient réalisées :

(7)
$$|X_n| \le a,\ \forall\ n \in \mathbf{N}^* \text{ (suite bornée)},$$
$$\Sigma_{n \in \mathbf{N}^*} \sigma_n^2 = +\infty \qquad (\text{où } \sigma_n^2 = V\ X_n,\ \forall\ n \in \mathbf{N}^*).$$

(3) **version anonyme**. On suppose X indépendante, centrée ($E\ X = 0$) et de carré intégrable (ie $X_n \in \mathscr{L}_{\mathbf{R}}^2(\Omega, \mathscr{T}, P)$, $\forall\ n \in \mathbf{N}^*$), et que les deux conditions suivantes sont vérifiées :

(8)
$$\Sigma_{n \in \mathbf{N}^*} E\ X_n^2 = +\infty$$
$$\forall\ \varepsilon > 0,\ \lim_N \{C_N(\varepsilon) / D_N\} = 0,$$

où $C_N(\varepsilon) = \Sigma_{n=1}^N \int \mathbf{1}(B_{Nn}(\varepsilon))\ X_n^2\ dP$, $D_N = \Sigma_{n=1}^N X_n^2$, avec $B_{Nn}(\varepsilon) = [X_n > \varepsilon\ \sigma(T_N)]$, $T_N = \Sigma_{n=1}^N X_n$ et $\{\sigma(T_N)\}^2 = V\ T_N$.

Alors, la suite des **lp** des variables $S_N = T_N / \sigma(T_N)$ tend vers $\mathscr{N}(0, 1)$, ie :

(2)'''' $\mathscr{L}\{T_N / \sigma (T_N)\} \rightarrow_{n \rightarrow +\infty} \mathscr{N}(0, 1)$.

La première condition (8) exprime qu'aucune variable X_n n'exerce d'effet prépondérant « au second ordre » (ie en variance) dans la somme T_N .

(4) **version triangulaire**. Ce théorème est utile dans l'étude du **modèle de régression multiple** (« *forme triangulaire* » du tlc). On note $X = (X_{Nn})_{(N,n)}$ une **suite triangulaire**, ie une suite de va $X_{Nn} : \Omega \mapsto \mathbf{R}$ tq $n \in N_N^* = \{1 ,..., N\}$ et $N \in \mathbf{N}^*$. On suppose que X est centrée ($E X_{Nn} = 0$, pour tout (N, n)), et que :

(9) $\Sigma_{n=1}^N \ V X_{Nn} = 1, \qquad \forall N \in \mathbf{N}^*$.

Alors, les deux propriétés suivantes :

$$(10) \quad \begin{cases} \Sigma_{n=1}^N \ X_{Nn} \rightarrow^{\mathscr{L}}_{N \rightarrow +\infty} \mathscr{N}(0, 1), \\ \max_{n=1}^N \ V X_{Nn} \rightarrow_{N \rightarrow +\infty} 0, \end{cases}$$

sont équivalentes à la propriété conjointe suivante, où $P^{XNn} = \mathscr{L}(X_{Nn})$ désigne la loi de X_{Nn} :

(11) $\forall \varepsilon > 0, \qquad \Sigma_{n=1}^N \int \mathbf{1}([|x| \geq \varepsilon]) \ x^2 \ dP^{XNn}(x) \rightarrow_{N \rightarrow +\infty} 0$.

ie, de façon équivalente :

(11)' $\forall \varepsilon > 0, \qquad \Sigma_{n=1}^N \int_{[|x| \geq \varepsilon]} x^2 \ d\mathscr{L}(X_{Nn}) (x) \rightarrow_{N \rightarrow +\infty} 0$,

où XNn désigne (par commodité) X_{Nn} .

(v) Le théorème de la limite centrale usuel fait ainsi jouer une place importante à la **loi normale**, considérée comme **loi limite** de suites de variables aléatoires. Il permet, aussi, de définir cette loi comme **loi asymptotique** d'une suite de sommes de va, ces sommes étant convenablement « normalisées » afin d'assurer un mode de convergence donné.

L'interprétation, concrète et classique, de ce théorème est illustrée par l'étude de nombreux phénomènes. Un **phénomène** donné est assimilable à un **système** S évoluant au cours du temps. Ce système est souvent lui-même représentable à l'aide d'un **processus stochastique** X. Si ce processus est la résultante (a) d'effets nombreux (N >> 0), (b) chacun exerçant une action additive (cf aussi **additivité**), (c) mais dont aucun ne possède de **variabilité** (ou fluctuation) importante relativement aux autres, alors la résultante « moyenne » de ces effets est distribuée de façon gaussienne.

Les extensions du tlc concernent :

(a) le cas de **vars** à variance infinie (ie hors de $\mathscr{L}_{\mathscr{X}}^2$) ;

(b) des **suites aléatoires** ou, plus généralement, des **familles aléatoires**, de types différents : eg fonctionnelles, etc (au lieu des suites scalaires ou vectorielles précédentes) ;

(c) des **formes d'association** (addition, multiplication, exponentiation, transformation logarithmique, etc) de ces suites ou familles ;

(d) la façon de « résumer » ces associations de façon synthétique (par le choix d'une **caractéristique de centralité**) ;

(e) le **mode de convergence stochastique** de la suite ;

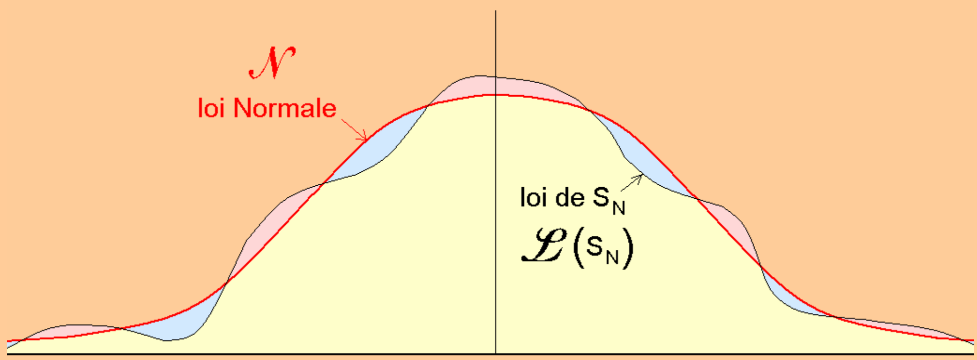
(f) le **type de loi limite** résultant de la convergence de la caractéristique choisie.

On doit donc souvent chercher des hypothèses ou des transformations (normalisation, etc) tq la convergence soit assurée.

(vi) L'expression de « *théorème central limite* » est une traduction incorrecte (et dénuée de sens) de l'expression anglaise « *central limit theorem* ». En effet, l'association sujet-complément devrait être « ((central) limit) theorem » et non « central (limit theorem) », ni « (limit central) theorem ».

A contrario, deux raisons « techniques » justifient l'appellation de « **théorème de la limite centrale** » (cf schéma ci-après) :

illustration de la propriété de "centralité" de la loi gaussienne
en termes de densités



(a) d'une part, ce théorème étudie le comportement « stochastique » limite d'une suite de « grandeurs » S_N (total T_N ou **moyenne empirique** $N^{-1}\ T_N$), que ces grandeurs soient « brutes », ou centrées, ou centrées et réduites. Ces grandeurs concernent les N premiers termes d'une suite X. Ce sont donc bien des **statistiques** ayant la nature d'une « **valeur centrale** » qui sont utilisées ;

(b) d'autre part, lorsque N augmente, la variabilité de ces mêmes grandeurs diminue a priori : en effet, une propriété tq (2) signifie que la stabilité du « **système** »

$X = (X_n)_n$ augmente. Ainsi, dans le cas d'une suite normalisée tq $S_N = (V\ T_N)^{-1} (T_N - E\ T_N)$, lorsque N augmente, la suite des lois $\mathscr{L}(S_N)$ ($N \in \mathbf{N}^*$) « varie » de moins en moins pr à la **loi asymptotique** (en général gaussienne). C'est donc la « fluctuation » (ou la « **distance** », ou l'« **écart** ») entre ces lois qui diminue (zone en gris sur le schéma précédent) : ces mêmes lois finissent par converger asymptotiquement vers une loi fixe, laquelle joue donc un rôle de **loi « centrale »**.