

Cours

Charles Vin

Date

1 Introduction

Exemple 1.1. Soit $(X_i)_{i \in \mathbb{N}}$ une suite de v.a. i.i.d., admettant un moment d'ordre 2.

En particulier, les X_i admettent aussi un moment d'ordre 1. Ainsi d'après la loi forte des grands nombres

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{p.s.} E(X_1).$$

Pour n fixé, comment quantifier \bar{X}_n proche de $E(X_1)$. "Avec n fixé très grand, on sais que c'est autour de l'espérance mais comment?" Si je reformule, quelle est la taille typique de la variable $\bar{X}_n - E(X_1)$

- Espérance? : $E(\bar{X}_n - E(X_1)) = E(\bar{X}_n) - E(X_1) = E(X_1) - E(X_1) = 0$
- Variance? : On pose $Y_n = \bar{X}_n - E(X_1)$

$$Var(Y_n) = E((Y_n - E(Y_n))^2) = E(Y_n^2) = Var(\bar{X}_n) = \frac{Var(X_1)}{n}$$

→ Ma variance est petite

Y_n est une variable d'espérance 0 et de variance $\frac{Var(X_1)}{n}$. Par quoi fait-il multiplier Y_n pour que sa variance soit d'ordre 1 (volontairement impréci, je crois que ça veut dire fini et de grandeur humaine).

Si on regarde $\sqrt{n}Y_n$ que voit on?

$$Var(\sqrt{n}Y_n) = n * Var(Y_n) = n * \frac{Var(X_1)}{n} = Var(X_1).$$

La variance ne dépend plus de n ! Conclusion : Pour étudier les fluctuation de \bar{X}_n autour de sa limite $E(X_1)$, il semble raisonnable d'étudier $Y_n = \sqrt{n}(\bar{X}_n - E(X_1))$.

Définition 1.1. Une variable X est dit centré et réduite si

- Elle admet un moment d'ordre 2
- Si $E(X) = 0, Var(X) = 1$

Si X est une variable admettant un moment d'ordre 2, la variable centré réduite issue de X est $Z = \frac{X - E(X)}{\sqrt{Var(X)}}$

La centrée réduite de \bar{X}_n est

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - E(X_1))}{\sqrt{Var(X_1)}}.$$

Preuve importante à faire en exercice

2 Notion de convergence en loi

Définition 2.1 (Convergence en loi). Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire.

On dit que $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X si $\forall t$ où F_X est continue en t

$$F_{X_n}(t) \rightarrow_{n \rightarrow +\infty} F_X(t).$$

On le note

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X.$$

Remarque. — Ca veut dire quoi $\{t \in \mathbb{R} \text{ tq } F_X \text{ est continue en } t\}$?

F_X est continue en $t \Leftrightarrow P(X = t) = 0$

— Si X est à densité, F_X est continue sur \mathbb{R}

— Si X est discrète, F_X est discontinue en chaque point tel que $P(X = k) > 0$

— "Cette notion de convergence est foireuse" Cette notion de convergence n'identifie **que** la loi de X

Si on a $X_n \xrightarrow{\mathcal{L}} X$ et $Y \perp X$ de même loi que X alors $X_n \xrightarrow{\mathcal{L}} Y$

Preuve : Soit $t \in \mathbb{R}$ tel que F_Y est continue en t . Comme $F_X = F_Y$, F_X est aussi continue en t , et on a comme $X_n \xrightarrow{\mathcal{L}} X$

$$F_{X_n}(t) \rightarrow F_X(t) = F_Y(t).$$

donc $X_n \xrightarrow{\mathcal{L}} Y$ □

— La remarque 2 implique : si $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} Y$

$$\lim_{n \rightarrow \infty} X_n + Y_n \neq X + Y.$$

Exemple : $X_n, Y_n, X \sim \mathcal{N}(0, 1)$

$$X_n \xrightarrow{\mathcal{L}} X \text{ car } \forall n, t, F_{X_n}(t) = F_X(t).$$

$$Y_n \xrightarrow{\mathcal{L}} -X.$$

$$\forall t \in \mathbb{R}, F_{Y_n}(t) = F_X(t) = F_{-X}(t).$$

Comme la gaussienne est symétrique $-X$ est une gaussienne également ect

Mais on a **pas** $X_n + Y_n \xrightarrow{\mathcal{L}} X_X = 0$ FAUX car X_1 indép Y_1 , $X_n + Y_n \sim \mathcal{N}(0, 2)$

Proposition 2.1 ((admise)). *Une suite de v.a. $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une variable aléatoire X ssi*

$$\forall f \in C_b^o(\mathbb{R}), E(f(X_n)) \rightarrow E(f(X)).$$

Exercice : En admettant cette proposition, montrer que si $X_n \xrightarrow{\mathcal{L}} X$ et si $h \in C^o(\mathbb{R})$ alors $h(X_n) \xrightarrow{\mathcal{L}} h(X)$

Méthode :

1. Calculer F_X
2. Déterminer $A = \{t \in \mathbb{R}, F_x \text{ est } C^o \text{ en } t\}$
3. Calculer $F_{X_n}(t)$ pour $t \in A$ et montrer $F_{X_n}(t) \rightarrow F_X(t)$

Autre cas :

Si je vous donne $(X_n)_{n \in \mathbb{N}}$ et que je vous demande si $(X_n)_{n \in \mathbb{N}}$ converge en loi. Que faut-il faire?

- Je calcule $F_{X_n}(t), \forall t \in \mathbb{R}$.
- Je montre $F_{X_n}(t) \rightarrow_{n \rightarrow +\infty} g(t)$ (pas forcément pour tous les t)
- Je dois reconnaître une variable Z tq $F_Z = g$
- Il faudra avoir montré la convergence pour tous les t tel que F_Z est continue en t .

Exemple 2.1. Soit (p_n) une suite qui converge vers $p \in]0, 1[$. X_n v.a. de loi $\text{Ber}(p_n)$

Montrons que si X est une v.a. de loi $\text{Ber}(p)$ on a

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

1. Soit $t \in \mathbb{R}$

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & \text{si } t < 0 \\ -p & \text{si } t \in [0; 1[+ 1 \\ 1 & \text{si } t \geq 1 \end{cases}.$$

2. F_X est continue en t sur $\mathbb{R} \setminus \{0, 1\}$

3. Soit $t \in \mathbb{R} \setminus \{0, 1\}$

$$F_{X_n}(t) = \begin{cases} 0 & \text{si } t < 0 \\ -pn & \text{si } t \in [0; 1[\\ 1 & \text{si } t \geq 1 \end{cases}.$$

Si $t < 0$, alors

$$F_{X_n} = 0 \rightarrow 0 = F_X(t).$$

Si $t \in]0, 1[$

$$F_{X_n}(t) = 1 - p_n \rightarrow 1 - p = F_X(t) \text{ car } p_n \rightarrow p.$$

Si $t > 1$

$$F_{X_n}(t) = 1 \rightarrow 1 = F_X(t).$$

Ainsi, $\forall t \in \mathbb{R} \setminus \{0, 1\}, F_{X_n}(t) \rightarrow F_X(t)$ donc $X_n \rightarrow^{\mathcal{L}} X$

Exemple 2.2. $\lambda_n \rightarrow \lambda, \forall \lambda_n > 0, \lambda > 0$. Soit $(X_n)_{n \in \mathbb{N}}$ suite de v.a. de loi $\mathcal{E}(\lambda_n)$ et X une v.a. de loi $\mathcal{E}(\lambda)$. Montrer que $X_n \rightarrow^{\mathcal{L}} X$

1. Soit $t \in \mathbb{R}$

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ \int_0^t \lambda e^{-\lambda x} dx & \text{si } t \geq 0 \end{cases} = \begin{cases} 0 & \text{si } t \leq 0 \\ (1 - e^{-\lambda t}) & \text{si } t \geq 0 \end{cases}.$$

2. F_X est continue sur \mathbb{R} (X est à densité)

3. Soit $T \in \mathbb{R}$

$$F_{X_n}(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ (1 - e^{-\lambda_n t}) & \text{si } t \geq 0 \end{cases}.$$

Si $t \leq 0, F_{X_n}(t) = 0 \rightarrow 0 = F_X(t)$

Si $t > 0, F_{X_n}(t) = (1 - e^{-\lambda_n t}) \rightarrow (1 - e^{-\lambda t}) = F_X(t)$ car $\lambda_n \rightarrow \lambda$ et l'expo est continue

Ainsi, $\forall t \in \mathbb{R}, F_{X_n}(t) \rightarrow F_X(t)$ donc $X_n \rightarrow^{\mathcal{L}} X$

Nouveau cours du 21/10

3 Le Théorème Central Limite

Théorème 3.1 (Théorème Central Limite). Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variable aléatoire **i.i.d.** admettant un **moment d'ordre 2**.

Soit $m = E(X_i), \sigma = \sqrt{Var(X_i)}$ alors

$$Z_n = \frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{i=1}^n X_i - m \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z \sim \mathcal{N}(0, 1).$$

Cela signifie :

$$\forall t \in \mathbb{R}, P(Z_n \leq t) \xrightarrow[n \rightarrow \infty]{} \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Remarque. $n = 10, X_n$ i.i.d. uniforme sur $[0, 1], Var X_i = 1/12, \sigma = \frac{1}{\sqrt{12}}$

$$\begin{aligned} Z_{10} &= \frac{\sqrt{10}}{1/\sqrt{12}} \left(\frac{1}{10} \sum_{i=1}^{10} X_i - \frac{1}{2} \right) \\ &= \sqrt{120} \left(\frac{1}{10} \sum_{i=1}^{10} X_i - 1/2 \right) \end{aligned}$$

Comment voir la loi de Z_{10} ?

$Z_{10}', \dots, Z_{10}^{(N)}$ copies i.i.d. de Z_{10} .

Glivenko-Conteli nous assure que

$$t \in \mathbb{R} \sup \left| \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{Z_{10}^{(i)} \leq t} - F_{Z_{10}}(t) \right| \xrightarrow[n \rightarrow \infty]{p.s.} 0.$$

C'est bien mais super visuel.

A la place, on préfère tracer des histogrammes (voir le cours de Mma. Fradon). On regarde si l'histogramme suit la densité d'une $\mathcal{N}(0, 1)$

$$\text{On note } \phi(t) = \int_{-\infty}^t \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

Proposition 3.2. Soit les hypothèses précédentes, on a

$$\forall a, b \in \mathbb{R}, P(a \leq Z_n \leq b) \rightarrow P(a \leq Z \leq b) = \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} = \phi(b) - \phi(a).$$

De même pour

$$P(a < Z_n < b) \rightarrow \phi(b) - \phi(a)$$

$$P(a \leq Z_n < b) \rightarrow \phi(b) - \phi(a)$$

$$P(a < Z_n \leq b) \rightarrow \phi(b) - \phi(a)$$

Preuve : Uniquement pour $a < Z_n \leq b$

Soit $a, b \in \mathbb{R}$

$$P(a < Z_n \leq b) = F_{Z_n}(b) - F_{Z_n}(a) \text{ car } P(Z_n \leq b) - P(Z_n \leq a).$$

Or d'après le théorème central limite

$$F_{Z_n}(b) \rightarrow \phi(b).$$

De même

$$F_{Z_n}(a) \rightarrow \phi(a).$$

Donc :

$$P(a < Z_n \leq b) = F_{Z_n}(b) - F_{Z_n}(a) \rightarrow \phi(b) - \phi(a).$$

Pour tous les autres cas, ce n'est pas immédiat. Cela provient de la continuité de ϕ

□

Définition 3.1 (Intervalle de confiance asymptotique). Soit $(\Omega, \mathcal{F}, (P_\theta)_{\theta \in \Theta})$ un modèle statistique. Soit X_1, \dots, X_n un échantillon d'une loi P_θ .

Un intervalle de confiance asymptotique de niveau $1 - \alpha$ pour une quantité $g(\theta)$ est un intervalle aléatoire $I(X_1, \dots, X_n)$, ne dépendant pas de θ et tel que

$$\forall \theta \in \Theta, \lim_{n \rightarrow \infty} P(g(\theta) \in I(X_1, \dots, X_n)) \geq 1 - \alpha.$$

Exemple 3.1 (Application). Détermination d'un intervalle de confiance asymptotique à l'aide du théorème central limite

Soit $(X_i)_{i \in \mathbb{N}}$ une suite de v.a. i.i.d. d'espérance inconnue et de variance égale à 35.

On cherche à estimer m et à donner un IC. asymptotique.

Soit $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Comme $(X_i)_{i \in \mathbb{N}}$ est une suite de v.a. i.i.d. admettant un moment d'ordre 2, d'après le théorème central limite, on a

$$Z_n = \frac{\sqrt{n}}{\sqrt{35}}(\bar{X}_n - m) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z \sim \mathcal{N}(0, 1).$$

Ainsi

$$\forall a \in \mathbb{R}, P(-a \leq Z_n \leq a) \xrightarrow[n \rightarrow \infty]{} \int_{-a}^a \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \phi(a) - \phi(-a).$$

Si je souhaite un intervalle de confiance asymptotique de niveau 99%.

On cherche a tel que : $\phi(a) - \phi(-a) \geq 0.99$

On cherche a tel que : $P(Z \geq a) = 0.005$ (l'aire restante à droite de la gaussienne) $P(Z \geq a) = 1 - \phi(a) = 0.005$. Ainsi, on cherche $a \in \mathbb{R}$ tel que $\phi(a) = 0.995$. Lecture de la table : si $a = 2.58$, $\phi(a) = 0.995$. Ainsi, $P(-2.58 \leq Z \leq 2.58) \geq 0.99$ On a donc :

$$\lim_{n \rightarrow \infty} P(-2.58 \leq \frac{\sqrt{n}}{\sqrt{35}}(\bar{X}_n - m) \leq 2.58) \geq 0.99.$$

Ré exprimons $-2.58 \leq Z_n \leq 2.58$ en $m \in I(X_1, \dots, X_n)$

$$\begin{aligned} -2.58 &\leq \frac{\sqrt{n}}{\sqrt{35}}(\bar{X}_n - m) \leq 2.58 \\ \frac{2.58 * \sqrt{35}}{\sqrt{n}} &\geq m - \bar{X}_n \geq \frac{-2.58 * \sqrt{35}}{\sqrt{n}} \\ m &\in [\bar{X}_n - \frac{2.58 * \sqrt{35}}{\sqrt{n}}, \bar{X}_n + \frac{2.58 * \sqrt{35}}{\sqrt{n}}] \\ 2.58 * \sqrt{35} &\approx 15.26 \end{aligned}$$

On obtient alors

$$\lim_{n \rightarrow \infty} P(m \in [\bar{X}_n \pm \frac{15.26}{\sqrt{n}}]) \geq 0.99.$$

Un intervalle de confiance asymptotique de niveau 99% est

$$[\bar{X}_n \pm \frac{2.56 * \sqrt{35}}{\sqrt{n}}].$$

Nouveau cours du 28/10

4 Précision du théorème central limite

Le TCL stipule que

$$\forall a \in \mathbb{R}, P(Z_n \leq a) \rightarrow_{n \rightarrow +\infty} P(Z \leq a) = \phi(a).$$

Ou bien :

$$\forall a \in \mathbb{R}, P(-a \leq Z_n \leq a) \rightarrow_{n \rightarrow \infty} \phi(a) - \phi(-a).$$

Malheureusement, à ce stade, nous n'avons aucune idée de la valeur de $|P(Z_n \leq a) - \phi(a)|$ pour une valeur de n fixée.

Pour a fixé, quelle est la qualité de cette approximation? Pour mesurer cette différence, je souhaite contrôler

$$\sup_{a \in \mathbb{R}} |F_{Z_n}(a) - \phi(a)|.$$

Note. C'est la distance entre les deux fonctions de répartition dans le pire des cas

Remarque. Cela n'a RIEN à voir avec le théorème de Glivenko-Cantelli

Théorème 4.1 (Inégalité de Berry-Esseen). Soit $(X_i)_{i \in \mathbb{N}}$ une suite de v.a. i.i.d., admettant un moment d'ordre 3.

On note $m = E(X_1)$, $\sigma = \sqrt{\text{Var}(X_1)}$, $\rho = E(|X_1 - m|^3)$, $Z_n = \frac{\sqrt{n}}{\sigma}((\frac{1}{n} \sum_{i=1}^n X_i) - m)$ on a :

$$\forall a \in \mathbb{R}, |F_{Z_n}(a) - \phi(a)| \leq \frac{1}{2} \frac{\rho}{\sigma^3 \sqrt{n}}.$$

Remarque. Pour pouvoir utiliser ce théorème, il faut connaître la valeur exacte de σ et ρ . Calculer $\rho = E(|X - m|^3)$

Exemple 4.1. Si $X \sim \mathcal{B}(p)$, $m = p$

$$\begin{aligned} E(|X - p|^3) &= P(X = 0)p^3 + P(X = 1)|1 - p|^3 \\ &= (1 - p)p^3 + p(1 - p)^3 \end{aligned}$$

Si $X \sim \mathcal{E}(\lambda)$, $m = 1/\lambda$

$$E\left(\left|X - \frac{1}{\lambda}\right|^3\right) = \int_0^{+\infty} \left|x - \frac{1}{\lambda}\right|^3 \lambda e^{-\lambda x} dx.$$

Comment se servir de ce théorème ? Application aux intervalles de confiance.

X_1, \dots, X_n v.a. i.i.d. d'espérance m inconnue, mais admettant un moment d'ordre 3, avec $\sigma = \sqrt{\text{Var}(X_1)}$ et $\rho = E(|X_1 - m|^3)$ connus.
D'après le théorème central limite, appliqué à la suite $(X_i)_{i \in \mathbb{N}}$

$$\forall a \in \mathbb{R}, P(Z_n \leq a) \rightarrow_{n \rightarrow +\infty} \phi(a).$$

Or d'après l'inégalité de Berry-Esseen, on a

$$|P(Z_n \leq a) - \phi(a)| \leq \frac{1}{2} \frac{\rho}{\sigma^3 \sqrt{n}}.$$

Cela donne :

$$\phi(a) - \frac{1}{2} \frac{\rho}{\sigma^3 \sqrt{n}} \leq P(Z_n \leq a) \leq \phi(a) + \frac{1}{2} \frac{\rho}{\sigma^3 \sqrt{n}}.$$

Si on veut a tq $\phi(a) = 0.95 \Leftrightarrow a = 1.65$ alors

$$0.95 - \frac{1}{2} \frac{\rho}{\sigma^3 \sqrt{n}} \leq P(Z_n \leq 1.65) \leq 0.95 + \frac{1}{2} \frac{\rho}{\sigma^3 \sqrt{n}}.$$

On change d'exemple : Si on regarde $P(-a \leq Z_n \leq a) \approx \phi(a) - \phi(-a)$ on regarde en fait $F_{Z_n}(a) - F_{Z_n}(-a) \approx \phi(a) - \phi(-a)$.

Si on choisit a tq $\phi(a) - \phi(-a) = 0.96 \Leftrightarrow \phi(a) = 0.98 \Leftrightarrow a \approx 2.06$

D'après le TCL, on avait

$$P(-2.06 \leq Z_n \leq 2.06) \rightarrow_{n \rightarrow +\infty} 0.96$$

$$P(m \in (\bar{X}_n - \frac{2.06\sigma}{\sqrt{n}}, \bar{X}_n + \frac{2.06\sigma}{\sqrt{n}})).$$

$$\begin{aligned} P(-a \leq Z_n \leq a) &= P(Z_n \leq a) - P(Z_n < -a) \rightarrow \phi(a) - \phi(-a) \\ &\Leftrightarrow |P(-a \leq Z_n \leq a) - (\phi(a) - \phi(-a))| \\ &= |P(Z_n \leq a) - P(Z_n < -a) - (\phi(a) - \phi(-a))| \\ &= |(P(Z_n \leq a) - \phi(a)) - (P(Z_n < -a) - \phi(-a))| \\ &\leq |P(Z_n \leq a) - \phi(a)| + |P(Z_n < -a) - \phi(-a)| \end{aligned}$$

D'après l'inégalité de Berry-Esseen, on a :

$$\begin{aligned} |P(-a \leq Z_n \leq a) - (\phi(a) - \phi(-a))| &\leq \frac{\phi}{2\sigma^3 \sqrt{n}} + \frac{\rho}{2\sigma^3 \sqrt{n}} \\ &\leq \frac{\rho}{\sigma^3 \sqrt{n}} \end{aligned}$$

Pour $a = 2.06$ cela donne

$$0.96 - \frac{\rho}{\sigma^3 \sqrt{n}} \leq P(-2.06 \leq Z_n \leq 2.06) \leq 0.96 + \frac{\rho}{\sigma^3 \sqrt{n}}.$$

Pour $\rho = 10, \sigma = 1, n = 10000$

$$\frac{\rho}{\sigma^3 \sqrt{n}} = \frac{10}{100} = 0.1, P(-2.06 \leq Z_n \leq 2.06) \geq 0.86.$$

L'inégalité de Berry-Esseen permet de passer d'un intervalle de confiance asymptotique de niveau $1 - \alpha$ à un intervalle de confiance **exact** de niveau $1 - \alpha$ avec α' explicite. $\alpha' = \alpha - \frac{\rho}{\sigma^3 \sqrt{n}}$
Vous pouvez comparer avec Tchebychev. Le meilleur dépendra des valeurs de ρ, n, σ . En général, n grand \rightarrow TCL + Berry-Esseen // n petit \rightarrow Tchebychev.

Si on veut garantir un niveau fixé, on peut anticiper. C'est à dire choisir α tq, par exemple, $1 - \alpha \geq 0.95$. (Quand c'est possible, je ne peux pas toujours anticiper car $\alpha \geq 0$). Reprenons l'exemple :

On sait

$$(\phi(a) - \phi(-a)) - \frac{\rho}{\sigma^3 \sqrt{n}} \leq P(-a \leq Z_n \leq a).$$

On cherche alors a tq

$$\begin{aligned}\phi(a) - \phi(-a) - \frac{\rho}{\sigma^3\sqrt{n}} &= 1 - \alpha \\ \Leftrightarrow \phi(a) - \phi(-a) &= 1 - \alpha + \frac{\rho}{\sigma^3\sqrt{n}} \\ \Leftrightarrow \phi(a) &= 1 - \frac{1}{2}\left(\alpha - \frac{\phi}{\sigma^3\sqrt{n}}\right) = 1 - \frac{\alpha}{2} + \frac{\phi}{\sigma^3\sqrt{n}}\end{aligned}$$

Il est possible, si $1 - \frac{\alpha}{2} + \frac{\phi}{\sigma^3\sqrt{n}} > 1$ qu'il n'y ait pas de solution.

Si $\rho = 10, \sigma = 1, n = 10000, 1 - \alpha = 0.95$:

$$\phi(a) = 1 - 0.025 + 0.05 = 1.025 > 1 \text{ impossible.}$$

Nouveau cours du 18/11

Corollaire (Inégalité de Berry-Esseen). *Sous les hypothèses de l'inégalité de Berry-Esseen, on a :*

$$\forall a \in \mathbb{R}^+, P(-a \leq Z_n \leq a) \geq P(-a \leq Z \leq a) - \frac{\rho}{\sigma^3\sqrt{n}}.$$

Avec $Z \sim \mathcal{N}(0, 1)$

Preuve : La démonstration a été faite le cours précédent. □

Interprétation :

Si a est tel que $P(-a \leq Z \leq a) = 1 - \alpha$ alors $P(-a \leq Z_n \leq a) \geq 1 - \alpha - \frac{\rho}{\sigma^3\sqrt{n}}$.

On en déduit alors le niveau **réel** de l'intervalle de confiance asymptotique.

Lien entre $P(-a \leq Z_n \leq A)$ et $P(m \in)$

$$P(-a \leq Z_n \leq a) = P(m \in [\bar{X}_n - \frac{a\sigma}{\sqrt{n}}; \bar{X}_n + \frac{a\sigma}{\sqrt{n}}]).$$

5 Théorème central limite avec autonormalisation

Le point faible de ce qu'on a vu précédemment est qu'on requiert la connaissance de la variance des variables aléatoires considérées.

Dans de nombreux exemples, la connaissance de l'espérance et de la variance sont équivalentes.

Exemple :

- si $X \sim \mathcal{B}(p)$, $E(X) = p$, $Var(X) = p(1-p)$. Si je connais $p(1-p)$, je connais p .
- Si $X \sim \mathcal{E}(\lambda)$, $\lambda > 0$, $E(X) = \frac{1}{\lambda}$, $Var(X) = \frac{1}{\lambda^2}$. Si $Var(X) = \frac{1}{100}$ alors $\lambda = 10$ et on sait tout ...
- On ne peut pas dire qu'on connaît la variance si on connaît pas l'espérance : sob :

Pour beaucoup d'applications, on ne peut pas connaître la variance des données et le TCL est inapplicable. Le théorème suivant est alors utilisé.

Théorème 5.1 (Théorème central limite avec autonormalisation). *Soit $(X_i)_{i \in \mathbb{N}}$ une suite de variable aléatoire i.i.d. admettant un moment d'ordre 2 (inconnu). De moyenne $m = E(X_1)$*

Soit V_n un estimateur faiblement consistant de $Var(X_1)$. Alors :

$$Z_n = \frac{\sqrt{n}}{\sqrt{V_n}} \left(\frac{1}{n} \sum_{i=1}^n X_i - m \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z \text{ de loi } \mathcal{N}(0, 1).$$

Note. 3 méthodes pour estimateur faiblement consistant

- CVG presque sur
- CVG en moyenne quadratique
- A la main Tchebychev

Remarque. Les avantages :

- On peut faire des intervalles de confiance asymptotique ! Du type

$$\left[\bar{X}_n - \frac{1.96\sqrt{V_n}}{\sqrt{n}}, \bar{X}_n + \frac{1.96\sqrt{V_n}}{\sqrt{n}} \right].$$

Les désavantages :

- Pas d'inégalité de Berry-Esseen, donc pas d'IC, ni de correction

Exemple 5.1 (Application aux sondages). On effectue un sondage dans la population, deux choix sont possibles : A et B. Les votants sont modélisés par des variables aléatoires i.i.d. de loi de Bernoulli de paramètre p inconnu

$$X_i = 1 \Leftrightarrow \text{le } i\text{ème votant vote A}$$

$$X_i = 0 \Leftrightarrow \text{le } i\text{ème votant vote B}$$

But : Intervalle de confiance asymptotique pour p de niveau 0.96

1. Choix de l'estimateur : On relie p aux propriétés de X_i . Comme $E(X_i) = p$. On propose comme estimateur de p : $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Exo : montrer que c'est un estimateur sans biais et fortement consistant de p .
2. On cherche un intervalle de confiance asymptotique de niveau 0.96 :
 - Les variables $(X_i)_{i \in \mathbb{N}}$ sont i.i.d. et admettent un moment d'ordre 2 (leurs variances valent $p(1-p)$)
 - Estimateur fortement consistant? $V_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ OU plus simple $V_n = \bar{X}_n(1 - \bar{X}_n)$
Comme :
 - $\bar{X}_n \xrightarrow{p.s} p$
 - $x \mapsto x(1-x)$ continue

Alors

$$\bar{X}_n(1 - \bar{X}_n).$$

est aussi un estimateur fortement consistant pour $Var(X_i)$ il est donc faiblement consistant. Ainsi d'après le TCL avec autonormalisation appliqué à la suite de variable i.i.d. $(X_i)_{i \in \mathbb{N}}$ admettant un moment d'ordre 2 et à un estimateur faiblement consistant de $V'_n = Var(X_1)$ on a :

$$\forall a \in \mathbb{R}^+, P(-a \leq \frac{\sqrt{n}}{\sqrt{V_n}} (\frac{1}{n} \sum_{i=1}^n X_i - p) \leq a) \rightarrow_{n \rightarrow +\infty} P(-a \leq Z \leq a) (Z \sim \mathcal{N}(0, 1)).$$

On cherche a dans la table tel que

$$P(-a \leq Z \leq a) = 0.96.$$

C'est à dire

$$P(Z \leq a) = 0.98 \Leftrightarrow a = 2.06.$$

Ainsi

$$P(-2.06 \leq \frac{\sqrt{n}}{\sqrt{V_n}} (\frac{1}{n} \sum_{i=1}^n X_i - p) \leq 2.06) \rightarrow 0.96.$$

Or

$$\begin{aligned} -2.06 &\leq \frac{\sqrt{n}}{\sqrt{V_n}} (\frac{1}{n} \sum_{i=1}^n X_i - p) \leq 2.06 \\ \Leftrightarrow \frac{-2.06\sqrt{V_n}}{\sqrt{n}} &\leq \bar{X}_n - p \leq \frac{2.06\sqrt{V_n}}{\sqrt{n}} \\ \Leftrightarrow \bar{X}_n - \frac{2.06\sqrt{V_n}}{\sqrt{n}} &\leq p \leq \bar{X}_n + \frac{2.06\sqrt{V_n}}{\sqrt{n}} \end{aligned}$$

Ainsi

$$P(p \in [\bar{X}_n - \frac{2.06\sqrt{V_n}}{\sqrt{n}}, \bar{X}_n + \frac{2.06\sqrt{V_n}}{\sqrt{n}}]) \rightarrow 0.96.$$

$[\bar{X}_n - \frac{2.06\sqrt{V_n}}{\sqrt{n}}, \bar{X}_n + \frac{2.06\sqrt{V_n}}{\sqrt{n}}]$ est un intervalle de confiance asymptotique de niveau 0.96 pour p

Nouveau cours du 07/12

On a vu le théorème central limite avec autonormalisation. Malheureusement, ce résultat n'est qu'asymptotique.

Que faire si on veut quelque chose de précis?

6 Le cas des variables gaussienne

6.1 Rappel : variance connue

X_1, \dots, X_n va. iid. de loi $\mathcal{N}(m, \sigma^2)$ avec m inconnu et σ^2 connue. On a vu précédemment que

$$\forall n \in \mathbb{N}^*, Z_n = \frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{i=1}^n X_i - m \right) \sim \mathcal{N}(0, 1).$$

On obtient alors facilement des intervalles de confiance exacts et optimaux.

Que faire si la variance est inconnu ?

6.2 Cas de la variance inconnue

X_1, \dots, X_n va. iid. de la $\mathcal{N}(m, \sigma^2)$ avec m et σ^2 inconnues. Dans toute cette section :

$$V_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

On ne va considérer que cet estimateur. Tout ce qu'on va raconter dans la suite est faux pour un autre choix d'estimateur de la variance. Pourquoi ce modèle est-il important ?

- Pragmatiquement, parce que c'est le seule modèle sur lequel je sais faire tant de chose
- Les gaussiennes sont partout. On modélise souvent de l'incertitude (erreur de mesure) par des gaussiennes. Les mesures expérimentales montrent souvent des fluctuation gaussiennes car on mesure des moyennes

Théorème 6.1 (de Cochran). Si X_1, \dots, X_n sont des va. iid. de loi $\mathcal{N}(m, \sigma^2)$ alors

1. \bar{X}_n et V_n sont **indépendante** (contre instinctif car il y a \bar{X}_n dans V_n , miracle gaussien)
2. $\frac{(n-1)V_n}{\sigma^2}$ suit la loi $\mathcal{X}^2(n-1)$

Cette loi a pour densité

$$f_{n-1}(x) = \frac{1}{2^{n-1/2} \Gamma(\frac{n-1}{2})} x^{\frac{n-1}{2}-1} e^{-x} \mathbb{1}_{x>0}.$$

où

$$\Gamma\left(\frac{n-1}{2}\right) = \int_0^\infty x^{\frac{n-1}{2}-1} e^{-x} dx.$$

est la fonction gamma d'Euler

6.3 Loi du khi deux

Quelle est cette loi $\mathcal{X}^2(n-1)$?

On dit que $\mathcal{X}^2(n-1)$ est la loi du chi-deux à (n-1) degrés de liberté.

En général, la loi $\mathcal{X}^2(h)$, $h \in \mathbb{N}$ (h degrés de liberté) est la loi de $\sum_{i=1}^h Y_i^2$ où $Y_i \sim \mathcal{N}(0, 1)$ iid.

$$f_h(x) = \frac{1}{2^{h/2} \Gamma(\frac{h}{2})} x^{h/2-1} e^{-x} \mathbb{1}_{x>0}.$$

Si $Z \sim \mathcal{X}^2(h)$ alors

$$\begin{aligned} E(Z) &= E\left(\sum_{i=1}^h Y_i^2\right) = hE(Y_n^2) = h \\ \text{Var}(Z) &= h\text{Var}(Y_i) \end{aligned}$$

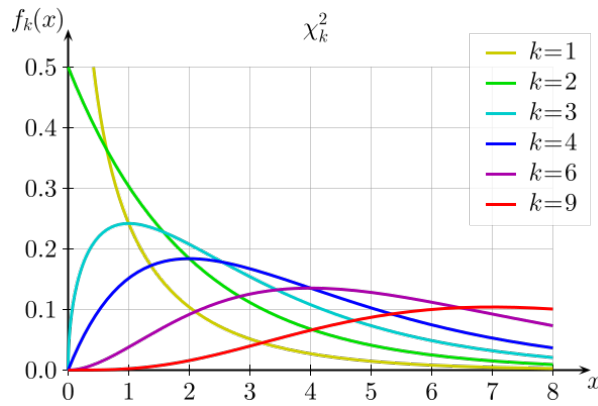


Figure 1 – Loi du khi deux

6.3.1 Intervalle de confiance à 0.95

Comme c'est pas symétrique, il faut trouver a et b tel que $P(Z \leq a) = 0.025$ et $P(Z \leq b) = 0.975$. Utilisons le théorème de Cochran pour trouver un intervalle de confiance de niveau 0.95 pour σ^2

$$\frac{(n-1)V_n}{\sigma^2} \sim \chi^2(n-1).$$

Je trouve a et b tq si $Z \sim \chi^2(n-1)$

$$P(a \leq Z \leq b) = 0.95.$$

On les trouve dans la table en lisant $P(Z \leq a) = 0.025$ et $P(Z \leq b) = 0.975$. Ainsi

$$P(a \leq \frac{(n-1)V_n}{\sigma^2} \leq b) = 0.95 \Leftrightarrow P(\frac{(n-1)V_n}{b} \leq \sigma^2 \leq \frac{(n-1)V_n}{a}) = 0.95$$

$[\frac{(n-1)V_n}{b}, \frac{(n-1)V_n}{a}]$ est un intervalle de confiance de niveau 0.95 pour σ^2 .

6.4 Théorème de Student

Théorème 6.2 (de Student). X_1, \dots, X_n va. iid. de loi $\mathcal{N}(m, \sigma^2)$

$$\frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - m) \sim t(n-1).$$

Loi de student à $(n-1)$ degrés de liberté.

$$t_{n-1}(x) = \frac{1}{\sqrt{(n-1)\pi}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \frac{1}{(1 + \frac{x^2}{n-1})^{n/2}}.$$

Preuve : $T = \frac{Z}{\sqrt{V/h}}$ student si $Z \sim \mathcal{N}(0, 1)$, $V \sim \chi^2(h)$.

On veut regarder

$$\frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - m) = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) * \frac{\sigma}{\sqrt{V_n}}$$

Trouvons la loi de $\frac{\sqrt{V_n}}{\sigma}$

$$\frac{\sqrt{V_n}}{\sigma} = \sqrt{\frac{V_n}{\sigma^2}} = \sqrt{\frac{V_n(n-1)}{\sigma^2}} \underset{(\sim \chi^2(n-1))}{*} \frac{1}{\sqrt{n-1}}.$$

$\frac{\sigma}{\sqrt{V_n}}$ s'écrit comme $\frac{1}{\sqrt{\chi^2(n-1)/(n-1)}}$

Ainsi $\frac{\sqrt{n}}{\sqrt{V_n}(\bar{X}_n - m)}$ est de la forme $\frac{\mathcal{N}(0,1)}{\sqrt{\chi^2(n-1)/(n-1)}}$ et \perp par Cochran!

□

6.4.1 Loi de student

La loi de Student à k degrés de liberté, noté $t(h)$ est la loi de

$$T = \frac{Z}{\sqrt{V/k}} \text{ où } Z \sim \mathcal{N}(0, 1), V \sim \mathcal{X}^2(h), Z \perp V.$$

Elle a pour densité

$$t_h(x) = \frac{1}{\sqrt{h\pi}} \frac{\Gamma(\frac{h+1}{2})}{\Gamma(\frac{h}{2})} \frac{1}{(1 + \frac{t^2}{h})^{\frac{h+1}{2}}}.$$

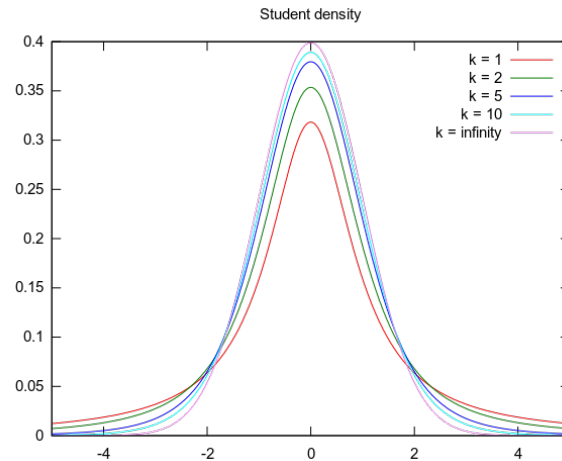


Figure 2 – Loi de student

6.4.2 Intervalle de confiance

IC pour m exact de niveau 0.95 à n fixé. $\frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - m) \sim t(n-1)$ (thm de student).

On cherche $a_t \in \mathbb{R}$ tq si $Z \sim t(n-1)$ on a

$$P(-a_t \leq Z \leq a_t) = 0.95.$$

Par symétrie de la loi de Student. Cela revient à trouver a_t tq

$$P(Z \leq a_t) = 0.975.$$

On trouve a_t dans la table de la loi $t(n-1)$ et on a

$$P(-a_t \leq \frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - m) \leq a_t) = 0.95.$$

On réécrit et on a

$$P(m \in [\bar{X}_n \pm \frac{\sqrt{V_n}}{\sqrt{n}a_t}]).$$

Remarque. Attention : a_t est un quantile de la loi de Student à $n-1$ degrés de liberté. Ce n'est pas le "a" de la loi $\mathcal{N}(0, 1)$.

Plus n est grand, plus a_t se rapproche du quantile de la gaussienne $\mathcal{N}(0, 1)$

7 Complément sur le cours

Théorème 7.1 (de Slutsky). Si $Z_n \xrightarrow{\mathcal{L}} Z$ et $Y_n \xrightarrow{P} y \in \mathbb{R}(\text{constante})$ alors

$$Y_n Z_n \xrightarrow{\mathcal{L}} yZ.$$

Remarque. En général $Y_n = V_n$ estimateur faiblement consistant de la variance

TCL + Slutsky \Rightarrow TCL autonormalisé :

$(X_i)_{i \in \mathbb{N}}$ une suite de va. iid. et admettant un moment d'ordre 2, V_n estimateur faiblement consistant de $\sigma^2 = \text{Var}(X_1)$

$$\frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - m) = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) * \frac{\sigma}{\sqrt{V_n}}$$

Comme

$$Z_n = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, 1)$$

$$Y_n = \frac{\sigma}{\sqrt{V_n}} \xrightarrow{\mathbb{P}} 1$$

Alors

$$Z_n Y_n = \frac{\sqrt{n}}{\sqrt{V_n}}(\bar{X}_n - m) \xrightarrow{\mathcal{L}} 1 \cdot Z = Z \text{ (d'après Slutsky) .}$$

Théorème 7.2 (Méthode Delta). Soit $(T_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et $\theta \in \mathbb{R}$ tels que

$$\sqrt{n}(T_n - \theta) \xrightarrow{\mathcal{L}_{n \rightarrow \infty}} \mathcal{N}(0, \sigma^2).$$

alors si g est une fonction dérivable telle que $g'(\theta) \neq 0$, on a

$$\sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{\mathcal{L}_{n \rightarrow \infty}} \mathcal{N}(0, \sigma^2 g'(\theta)^2).$$

Ce théorème est très utile pour obtenir des intervalles de confiance asymptotique!

8 Récap du cours

Plan du chapitre 2: TCL:

Prérequis: Convergence en \mathbb{R}
Fonction de répartition.

Ex: imaginons des variables X_i : $\mathbb{E}(X_i) = m$, $\text{Var}(X_i) = \sigma^2$, $\mathbb{E}(X_i^3) = w$
imaginons que: $(\bar{X}_n - m) \frac{\cos w^2}{\sigma} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z \sim t(14)$

Slutsky: V_n estimateur faiblement consistant de σ^2
 $\frac{W_n}{\sigma}$ de w
 $\frac{1}{n} \sum_{i=1}^n X_i^3$

Alors: $(\bar{X}_n - m) \frac{\cos W_n^2}{V_n} \xrightarrow{\mathcal{L}} t(14)$

X_i iid $\mathbb{E}(X_i) = 0$ $\text{Var}(X_i) = \mathbb{E}(X_i^2)$
Je veux un ICA pour la variance des X_i ?
Estimateur: $\frac{1}{n} \sum_{i=1}^n X_i^2$ $Y_i = X_i^2$ moyenne $Y_i \rightarrow$ Variance des X_i

