CH2. Les théorèmes limites

2.1 Convergence des variables aléatoires.

La notion de convergence d'une v.a. est fondamentale pour les applications statistiques.

Soit une v.a. X définie sur (Ω, A, P).
On définit une suite des Variables Aléatoires Indépendantes définies sur l'espace (Ω, A, P) notées $(X_1, X_2, ..., X_n)$
X_i est la valeur observée à la $i-$ème expérience.

Dans la suite du cours

On considère une suite des v.a. $X_1, \ldots, X_n$ sur $(\Omega,\, \mathcal{A},\, \mathcal{P})$
On notera cette suite des *v.a.* par $(X_n; n \geq 1)$.

2.1 Convergence des variables aléatoires.

1. Convergence Presque Sûre
2. Convergence en Probabilité
3. Convergence en Moyenne quadratique
4. Convergence en Loi

2. Convergence Presque Sûre

La notion "Presque Sûr"

On dit qu'un événement A se réalise "presque sûrement" si $P(A) = 1$

exemple

Si X est une v.a. continue et x est un réel, alors la $P(X \neq x) = 1$ presque sûrement. Puisque $P(X = x) = 0$

1. Convergence Presque Sûre

Définition : On dit que la suite des v.a. $(X_n)_n \geq 1$ converge vers la v.a. X *presque sûrement* ou avec probabilité 1, on note :

$$\lim_{n\to\infty} X_n \underset{ps}{=} X$$

si

$$P\left[\lim_{n\to\infty} X_n = X\right] = 1$$

c.a.d. si

$$P[\lim_{n\to\infty} X_n(\omega) - X(\omega)] = 1 \quad \forall\omega \in \Omega - \Omega_0 \quad ; \quad \Omega_0 \subset \Omega \quad \text{tel que } P(\Omega_0) = 0$$

La convergence presque sûre est la convergence en moyenne sur tout Ω, sauf éventuellement sur un sous-ensemble Ω_0 de mesure nulle.

Remarques :

1. C'est l'extension naturelle des opérations limites, utilisées en calcul standard
2. C.a.d. : pour $n \to \infty$ la différence entre deux v.a. devient très petite
3. C'est une notion très forte parce qu'elle concerne $\forall \omega \in \Omega$
4. Cette définition dit que l'ensemble des points de divergence est de probabilité nulle.

3. Convergence en Probabilité

Définition : La suite des *v.a.* $(X_n)_{n\geq 1}$ converge en *probabilité* vers la v.a. X et l'on note :

$$\lim_{n\to\infty} X_n \underset{p}{=} X$$

si

$$\forall\epsilon > 0 : \lim_{n\to\infty} P\left[|X_n - X| < \epsilon\right] = 1$$

Si la *v.a.* $X = c$ est *dégénérée* on dit que X_n converge en *probabilité* vers la constante C et on écrit :

$$\lim_{n\to\infty} X_n \underset{p}{=} C$$

Exemple :

Soit la fonction de densité $f(x) = \frac{ne^{-nx}}{(1+e^{-nx})^2}$

On va montrer que $\lim\limits_{n\to\infty} X_n \underset{p}{=} 0$

C.a.d. $\forall\epsilon > 0 \lim\limits_{n\to\infty} P[|X_n - X| \leq \epsilon] = 1$

Preuve :

$P[|X_n - 0| \leq \epsilon] = \int\limits_{-\epsilon}^{\epsilon} \frac{ne^{-nx}}{(1+e^{-nx})^2} dx = [\frac{1}{1+e^{-nx}}]_{-\epsilon}^{\epsilon} = \frac{1}{1+e^{-n\epsilon}} - \frac{1}{1+e^{n\epsilon}} \underset{+\infty}{\longrightarrow} 1$

Alors $\lim\limits_{n\to+\infty} P[|X_n| \leq \epsilon] = 1$ et donc $\lim\limits_{n\to\infty} X_n \underset{p}{=} 0$

3.2 Convergence en probabilité et Convergence presque sûre

La convergence presque sûre établit en plus que la relation est valable pour chaque $\omega \in \Omega - \Omega_0$

4. Convergence en Moyenne Quadratique

On fait l'hypothèse $E(X^2) < \infty$

Définition : On dit que la suite $(X_n)_n$ converge en *moyenne quadratique* vers X et l'on note

$$\lim_{n \to \infty} X_n \underset{mq}{=} X$$

si

$$\lim_{n \to \infty} E\left[|X_n - X|^2\right] = 0$$

c'est-à-dire que $\forall \epsilon > 0$ il existe un entier $N(\epsilon) > 0$ t.q. $\forall n \geq N(\epsilon)$ $E\left[|X_n - X|^2\right] = 0$.

Notons que cette convergence est définie pour tout $\omega \in \Omega$.

Exemple 1

Soit (X_n) une suite des v.a.i.i.d et X une v.a.
On fait l'approximation de (X_n) par X .
Soit $\epsilon_n = X_n - X$ l'erreur commise par l'approximation.

$$\lim_{n \to +\infty} E(X_n - X)^2 = \lim_{n \to +\infty} (\epsilon_n)^2$$

Si $\lim_{n \to +\infty} (\epsilon_n)^2 = 0$
c.a.d. si la variation autour de X tend vers 0, alors $\lim_{n \to \infty} X_n \underset{mq}{=} X$

Exemple 2

Soit $(X_n), n \geq 1, \quad X_n \sim \mathcal{B}(n,p)$
On pose $f_n = \frac{X_n}{n}$, la fréquence d'apparition d'un événement

Pb : Est-ce que f_n converge en $m.q$ vers p (p probabilité d'apparition)
Si oui, on pourra estimer le paramètre p par f_n et on notera $\widehat{p} = \frac{X_n}{n}$

On a : $E[(\dfrac{X_n}{n} - p)^2] = \dfrac{1}{n^2} E[(X_n - np)^2]$

$$= \frac{1}{n^2} E[(X_n - E(X_n))^2] = \frac{1}{n^2} V(X_n) = \frac{1}{n^2} p(1-p) \underset{n \to +\infty}{\longrightarrow} 0$$

Donc $f_n \underset{m.q}{\longrightarrow} p$ quand $n \to \infty$

5. Convergence en Loi

On s'intéresse au voisinage des fonctions de répartition des v.a.

Définition : On dit que la suite $(X_n)_n$ converge en loi vers X et l'on note

$$\lim_{n \to +\infty} X_n \underset{l}{=} X$$

si

$$\lim_{n \to +\infty} F_{X_n}(x) = F_X(x) \quad \forall x \in \mathbb{R} \ \text{ pour le quel la fonction } F_X \ \text{ est continue}$$

Exemple

Soit X une v.a., $X \sim \mathcal{B}(p)$, $P(X=1) = P(X=-1) = \frac{1}{2}$
On pose $X_n = X \quad \forall n \in \mathbb{N}$.
Alors la suite (X_n) converge en loi vers X
mais

si Y est une v.a., $Y \sim \mathcal{B}(p)$, $P(Y=0) = P(Y=1) = \frac{1}{2}$

La suite X_n converge en loi vers Y aussi et généralement vers toute v.a. de Bernoulli $\mathcal{B}(p)$ avec , $p = \frac{1}{2}$

Remarques

– La convergence en loi concerne les lois,c.a.d. leur fonction de repartition.
– Il est possible que X_n et X ne soient pas définies dans le même espace.
– Concrètement, elle nous permet d'approcher la loi de X_n, souvent inconnue ou difficilement manipulable par une loi suivie par X connue.

Relations entre les différentes convergences

Schéma d'implication :

$$\begin{array}{lcccc} \text{conv. presque sûre} & \searrow & & & \\ & & \text{conv. en proba.} & \longrightarrow & \text{conv. en loi} \\ \text{moy. quadratique.} & \nearrow & & & \end{array}$$

2.2 Quelques inégalités en probabilités

Théorème 1. *Soit X une v.a. sur $(\Omega, \mathcal{A}, P)$ et $g(X)$ une fonction réelle. Alors $g(X)$ est une v.a. telle que :*

$$P\ [g(X) \geq c] \leq \frac{E\ [g(X)]}{c} \quad , \ \forall c > 0$$

R.q : :
On a : $P\ [|g(X)| \geq c] = P\ [|g(X)|^r \geq c^r] \qquad (1)$

Cas particuliers :

1. **Inégalité de Markov :**
   Posons $g(X) = |X - E(X)|^r$.
   Alors :

$$P\ \left[|X - E(X)|^r \geq c^r\right] \leq \frac{E\ \left[|X - E(X)|^r\right]}{c^r} \quad ;\ \forall c > 0$$

2. Inégalité de Chebychev Si $c = \delta\ \sqrt{V(X)}$, alors :

$$P\ \left[|X - E(X)| \geq \delta\ \sqrt{V(X)}\right] \leq \frac{1}{\delta^2} \quad ;\ \forall c > 0$$

3. Si $g(X) = |X - E(X)|^2$, on utilise la r.q. (1) et on obtient :

$$P\left[|X - E(X)|^2 \geq c^2\right] \leq \frac{E\left[|X - E(X)|^2\right]}{c^2} = \frac{V(X)}{c^2} \quad ; \ \forall c > 0$$

On utilise très souvent cette dernière inégalité.

Exemple

On lance une pièce plusieurs fois.
Soit $f_n = \frac{X_n}{n}$ la fréquence d'apparition de piles.
On cherche n t.q. $P(|f_n - \frac{1}{2}| \geq 10^{-2}) \leq 0.01$.

Rép : On applique B-T avec : $\quad P(|X_n - E(X_n)| \geq c) \leq \frac{\sigma^2}{c^2}$

$E(f_n) = E(X_n) = p = \frac{1}{2}; \quad c^2 = 10^{-4}; V(f_n) = \frac{V(X_n)}{n^2} = \frac{np(1-p)}{n^2} = \frac{1}{4n}$
On obtient :

$$P(|f_n - \tfrac{1}{2}| \geq 10^{-2}) \leq \frac{1}{4n}\frac{1}{10^{-4}} = 0.01$$

On doit faire $n \geq 250.000$ lancers.

Remarques

Ces inégalités :

1. Elles permettent d'obtenir des bornes supérieures pour la probabilité des certains événements
2. Elles sont valables quelle que soit la loi suivie.
3. Elles nous permettent d'obtenir des bornes supérieures pour la probabilité des certains événements.
4. On les utilise quand on ne connaît pas la loi de X, Il suffit de connaître la $V(X)$ et l' $E(X)$ ou les estimer.

2.3 Lois des grands nombres

Question :

$$Y_n = \left(\frac{1}{n} \sum_{i=1}^{k} X_i - E\left(X_i\right) \right)$$

converge vers une limite ?

1. si : $Y_n \underset{ps}{\longrightarrow} 0$,
   alors X_n **obéit à la loi forte des grands nombres**
2. si : $Y_n \underset{p}{\longrightarrow} 0$,
   alors X_n **obéit à la loi faible des grands nombres**

2.3.1 Loi forte des grands nombres

le théorème de Khinchine

Théorème 2. *Soit la suite* $(X_n)_n$ *de v.a.i.i.d.*

Si $E(X_k) = \mu \ \forall, k = 1, 2, \ldots$ *, alors la suite* $(X_n)_n$ *obéit à la* **loi forte des grands nombres**.
c.a.d.

$$\frac{1}{n} \sum_{i=1}^{n} X_i \underset{ps}{\longrightarrow} E\left(X_i\right)$$

Remarque :

On suppose que σ^2 *existe.*

2.3.2 Loi faible des grands nombres

Théorème 3. *Soit* $(X_n)_n$ *une suite de v.a.i.i.d.*

On pose : $E(X_i) = \mu < \infty$ *et* $\sigma^2 = V(X_i) < \infty$.

Nous posons $S_n^2 = \sum_{k=1}^{n} V(X_k)$

Si $\lim_{n\to\infty} \frac{S_n^2}{n^2} = 0$ *alors la suite* $(X_n)_n$ *obéit à la* **loi faible des grands nombres**.

i.e :

$$\frac{1}{n}\sum_{i=1}^{k} X_i \xrightarrow[p]{} E(X_i)$$

Comparaison de la loi Forte et de la loi Faible

Soit : $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

1. **Loi Faible** : $\forall \varepsilon > 0 \quad : \lim_{n\to\infty} P\left(\left\{\left|\overline{X} - \mu\right| < \varepsilon\right\}\right) = 1$
   - Pour toute valeur grande de n la valeur de $\overline{X}$ est probablement très voisine de μ
   - des larges écarts peuvent se produire pour une infinité d'événements dont **la probabilité est très faible**.
2. **Loi Forte** : $P\left[\lim_{n\to\infty} \overline{X} = \mu\right] = 1$
   - elle exclut cette situation et assure que $\overline{X}$ reste avec probabilité 1 proche de μ sauf pour un ensemble d'événements **de mesure nulle.**

Problème à résoudre

Peut-on mesurer l'erreur $\epsilon = \left| \frac{1}{n} \sum_{i=1}^{k} X_i - E(X) \right|$?

1. **Les inégalités B.T.,Markov** donnent des bornes en probabilité de l'écart entre $\overline{X}$ et $\mu = E(X_i)$
2. **Les lois des grands nombres** disent qu'une évaluation est possible car

$\overline{X} \underset{n \to \infty}{\longrightarrow} E(X_i)$, mais on ne connaît pas l'erreur de l'estimation $\epsilon = \left| \overline{X} - \mu \right|$

C'est le Théorème appelé, **Théorème-Central -Limite** qui donnera les fluctuations de cette erreur autour de zéro.

2.4 Théorème Central Limite (T.C.L.)

Théorème 4. 2.4.1 Th. central limite. *Soit la suite* $(X_n)_n$ *de v.a.i.i.d. telles que :*

$$|E(X_k)| = |\mu| < \infty \text{ et } V(X_k) = \sigma^2 < \infty \ ; \forall\, k = 1, 2, \ldots \ .$$

On pose :

$$S_n = \sum_{k=1}^{n} X_k \quad \text{et} \quad Z_n = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

(la v.a. centrée réduite de S_n *).*

Alors :

$$\lim_{n \to \infty} Z_n \underset{l}{=} Z \sim \mathcal{N}(0,1)$$

Remarques

Il y a deux voies d'application :

1. En statistique Mathématique on calcule l'erreur commise $\epsilon = \left|\overline{X} - \mu\right|$.
   C'est pour ça le T.C.L. est dit aussi "Lois des erreurs"
2. Sa grande utilité pratique.
   Il permet de calculer de manière simple, les valeurs approchées des probabilités du type : $P(a \leq S_n \leq b)$, sans connaître la loi de X_i, il suffit d'avoir des Variables Aléatoires Indépendantes Identiquement , Distribuées.

Application :

Comportement de $\overline{X} = \frac{1}{n} \sum_{k=1}^{n} X_k$; quand $n \to \infty$

On pose : $S_n = \frac{1}{n} \sum_{k=1}^{n} X_k$

On a : $E(S_n) = \mu$; $Var(S_n) = \sigma^2_{S_n} = \frac{\sigma^2}{n}$

En appliquant le **TCL** :

$$\frac{\overline{X} - \mu}{\sigma}\sqrt{n} \sim \mathcal{N}\,(0, 1)$$

ou bien

$$\lim_{n \to \infty} S_n \underset{l}{=} S \sim \mathcal{N}\,(\mu, \frac{\sigma}{\sqrt{n}})$$

Exemple : Loi de Moivre

On suppose que $(X_i) \sim B(p)$:
$E(X_k) = p, \quad V(X_k) = \sigma^2 = p(1-p) \quad , i = 1, 2, ..., n$

$S_n = \sum_{k=1}^{n} X_k$

Si on pose $X = S_n, \quad X \sim B(n, p)$
$E(X) = np; V(X) = np(1-p)$

Par le **TCL** : $\frac{X-E(X)}{\sigma_X} \sim \mathcal{N}(0,1)$ donc $\frac{X-np}{\sqrt{p(1-p)n}} \sim \mathcal{N}(0,1)$