Statistique Mathématique

Chapitre 4: Convergences et théorèmes limites

Mohamed Essaied Hamrita

Mastère de Recherche: Finance & Actuariat
IHEC Sousse

Novembre 2021

1. Inégalités
2. Convergences
3. Théorèmes limites

Définition 1 (Inégalité de Markov)

Si X est une variable aléatoire positive dont l'espérance existe, alors, pour tout $\lambda > 0$,

$$\mathbb{P}(\mathbf{X} \geq \lambda \mathbb{E}(\mathbf{X})) \leq \frac{\mathbf{1}}{\lambda} \text{ et } \mathbb{P}(\mathbf{X} \geq \lambda) \leq \frac{\mathbb{E}(\mathbf{X})}{\lambda}$$

Définition 1 (Inégalité de Markov)

Si X est une variable aléatoire positive dont l'espérance existe, alors, pour tout $\lambda > 0$,

$$\mathbb{P}(\mathbf{X} \geq \lambda \mathbb{E}(\mathbf{X})) \leq \frac{\mathbf{1}}{\lambda} \text{ et } \mathbb{P}(\mathbf{X} \geq \lambda) \leq \frac{\mathbb{E}(\mathbf{X})}{\lambda}$$

Exemple 1

Soit $X \sim B(1, \frac{1}{25})$. En utilisant l'inégalité de Markov, trouver une borne de la probabilité de $X \geq 5$.

Définition 1 (Inégalité de Markov)

Si X est une variable aléatoire positive dont l'espérance existe, alors, pour tout $\lambda > 0$,

$$\mathbb{P}(\mathbf{X} \geq \lambda\mathbb{E}(\mathbf{X})) \leq \frac{\mathbf{1}}{\lambda} \text{ et } \mathbb{P}(\mathbf{X} \geq \lambda) \leq \frac{\mathbb{E}(\mathbf{X})}{\lambda}$$

Exemple 1

Soit $X \sim B(1, \frac{1}{25})$. En utilisant l'inégalité de Markov, trouver une borne de la probabilité de $X \geq 5$.

$$\mathbb{P}(X \geq 5) \leq \frac{\mathbb{E}(X)}{5} = \frac{1/5}{5} = \frac{1}{25}.$$

Proposition 1

Soit X une v.a positive et $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction non négative, alors:

$$\mathbb{P}(g(X) \geq \lambda) \leq \frac{\mathbb{E}(g(X))}{\lambda}$$

En particulier, on a:

$$\mathbb{P}(|\mathbf{X}|^{\mathbf{k}} \geq \lambda) \leq \frac{\mathbb{E}\left(|\mathbf{X}|^{\mathbf{k}}\right)}{\lambda} \text{ et } \mathbb{P}(|\mathbf{X}| \geq \epsilon) \leq \frac{\mathbb{E}\left(|\mathbf{X}|^{\mathbf{k}}\right)}{\epsilon^{\mathbf{k}}};\ \lambda > 0,\ \epsilon > 0,\ k \in \mathbb{N}^*$$

Définition 2 (Inégalité de Bienaymé-Tchebychev)

Si X est une variable aléatoire telle que $\mathbb{V}(X)$ existe, alors, pour tout $\epsilon > 0$,

$$\mathbb{P}\left(|\mathbf{X} - \mathbb{E}(\mathbf{X})| \geq \epsilon\right) \leq \frac{\mathbb{V}(\mathbf{X})}{\epsilon^{\mathbf{2}}}$$

Définition 2 (Inégalité de Bienaymé-Tchebychev)

Si X est une variable aléatoire telle que $\mathbb{V}(X)$ existe, alors, pour tout $\epsilon > 0$,

$$\mathbb{P}\left(|\mathbf{X}-\mathbb{E}(\mathbf{X})| \geq \epsilon\right) \leq \frac{\mathbb{V}(\mathbf{X})}{\epsilon^{\mathbf{2}}}$$

Cas particulier des variables centrées réduites: *si X est une variable aléatoire telle que $\mathbb{V}(X)=\sigma^2$ existe, alors, pour tout $a>1$,*

$$\mathbb{P}\left(\left|\frac{\mathbf{X}-\mathbb{E}(\mathbf{X})}{\sigma}\right| \geq \mathbf{a}\right) \leq \frac{\mathbf{1}}{\mathbf{a}^{\mathbf{2}}}$$

Définition 2 (Inégalité de Bienaymé-Tchebychev)

Si X est une variable aléatoire telle que $\mathbb{V}(X)$ existe, alors, pour tout $\epsilon > 0$,

$$\mathbb{P}\left(|\mathbf{X} - \mathbb{E}(\mathbf{X})| \geq \epsilon\right) \leq \frac{\mathbb{V}(\mathbf{X})}{\epsilon^{\mathbf{2}}}$$

Cas particulier des variables centrées réduites: *si X est une variable aléatoire telle que $\mathbb{V}(X) = \sigma^2$ existe, alors, pour tout $a > 1$,*

$$\mathbb{P}\left(\left|\frac{\mathbf{X} - \mathbb{E}(\mathbf{X})}{\sigma}\right| \geq \mathbf{a}\right) \leq \frac{\mathbf{1}}{\mathbf{a}^{\mathbf{2}}}$$

Inégalité de Jensen: Si g est une fonction réelle convexe sur un intervalle I de $\mathbb{R}$ et si $\mathbb{E}(X)$ et $\mathbb{E}(g(X))$ existent, alors

$$\mathbf{g}(\mathbb{E}(\mathbf{X})) \leq \mathbb{E}(\mathbf{g}(\mathbf{X}))$$

1 Inégalités

2 Convergences
- convergence en probabilité
- Convergences en moyennes d'ordre p
- Convergence en distribution

3 Théorèmes limites

Définition 3

La suite de variables aléatoires (X_n) ***converge en probabilité*** *vers une variable aléatoire X si, pour tout $\epsilon > 0$,*

$$\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{P}\left(|\mathbf{X_n} - \mathbf{X}| < \epsilon\right) = \mathbf{1} \quad \text{ou} \quad \lim_{\mathbf{n}\longrightarrow\infty} \mathbb{P}\left(|\mathbf{X_n} - \mathbf{X}| > \epsilon\right) = \mathbf{0}$$

On note $X_n \xrightarrow[n\mapsto\infty]{prob} X$ ou plim $X_n = X$

Définition 3

La suite de variables aléatoires (X_n) ***converge en probabilité*** *vers une variable aléatoire* X *si, pour tout* $\epsilon > 0$,

$$\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{P}\left(|\mathbf{X_n} - \mathbf{X}| < \epsilon\right) = \mathbf{1} \quad \textit{ou} \quad \lim_{\mathbf{n}\longrightarrow\infty} \mathbb{P}\left(|\mathbf{X_n} - \mathbf{X}| > \epsilon\right) = \mathbf{0}$$

On note $X_n \xrightarrow[n\mapsto\infty]{prob} X$ *ou* $\text{plim } X_n = X$

Exercice 1

Soient $X_i \overset{iid}{\sim} \mathcal{U}[0,1]$, $Y_n = \min\{X_i\}$ *et* $Z_n = \max\{X_i\}$. *Montrer que* Y_n *et* Z_n *convergent en probabilité vers* 0 *et* 1 *respectivement.*

Soit $\epsilon > 0$, $X_i \geq 0$ et $Y_n = \min\{X_i\}$, donc $Y_n \geq 0$.
$\mathbb{P}(|Y_n| \geq \epsilon) = \mathbb{P}(Y_n \geq \epsilon)$.
Si $\epsilon > \mathbf{1}$, alors $\mathbb{P}(Y_n \geq \epsilon) = 0$.
Si $\mathbf{0} < \epsilon \leq \mathbf{1}$, alors

$$\begin{aligned}\mathbb{P}(Y_n \geq \epsilon) =& \mathbb{P}(X_1 \geq \epsilon, X_2 \geq \epsilon, \ldots, X_n \geq \epsilon)\\ =& [\mathbb{P}(X_i \geq \epsilon]^n\\ =& (1-\epsilon)^n \xrightarrow[n\longrightarrow\infty]{} 0.\end{aligned}$$

Donc, pour $\epsilon > 0$, on a $\lim_{n\mapsto\infty} \mathbb{P}(Y_n \geq \epsilon) = 0$. Ainsi Y_n converge en probabilité vers 0. On note $Y_n \xrightarrow{prob} 0$.

Définition 4 (Convergences en moyennes d'ordre p)

La suite de variables aléatoires (X_n) ***converge en moyenne d'ordre*** $\mathbf{p > 0}$ *vers une variable aléatoire X si*

$$\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^{\mathbf{p}}) = \mathbf{0} \text{ et l'on note } X_n \xrightarrow[n\mapsto\infty]{m.p} X.$$

Cas particuliers:

- *En moyenne:* $\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|) = \mathbf{0}$.

Définition 4 (Convergences en moyennes d'ordre p)

La suite de variables aléatoires (X_n) ***converge en moyenne d'ordre*** $\mathbf{p > 0}$ *vers une variable aléatoire* X *si*

$$\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^{\mathbf{p}}) = \mathbf{0} \text{ et l'on note } X_n \xrightarrow[n\mapsto\infty]{m.p} X.$$

Cas particuliers:

- *En moyenne:* $\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|) = \mathbf{0}$.
- *En moyenne quadratique:* $\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^{\mathbf{2}}) = \mathbf{0}$

Définition 4 (Convergences en moyennes d'ordre p)

La suite de variables aléatoires (X_n) ***converge en moyenne d'ordre $\mathbf{p > 0}$*** *vers une variable aléatoire X si*

$$\lim_{\mathbf{n} \longrightarrow \infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^{\mathbf{p}}) = \mathbf{0} \text{ et l'on note } X_n \xrightarrow[n \mapsto \infty]{m.p} X.$$

Cas particuliers:

- *En moyenne:* $\lim_{\mathbf{n} \longrightarrow \infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|) = \mathbf{0}$.
- *En moyenne quadratique:* $\lim_{\mathbf{n} \longrightarrow \infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^2) = \mathbf{0}$

Définition 4 (Convergences en moyennes d'ordre p)

La suite de variables aléatoires (X_n) ***converge en moyenne d'ordre*** $\mathbf{p > 0}$ *vers une variable aléatoire X si*

$$\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^{\mathbf{p}}) = \mathbf{0} \text{ et l'on note } X_n \xrightarrow[n\mapsto\infty]{m.p} X.$$

Cas particuliers:

- *En moyenne:* $\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|) = \mathbf{0}$.
- *En moyenne quadratique:* $\lim_{\mathbf{n}\longrightarrow\infty} \mathbb{E}(|\mathbf{X_n} - \mathbf{X}|^{\mathbf{2}}) = \mathbf{0}$

(X_n) converge en moyenne quadratique vers une variable aléatoire X si:

$$\mathbb{E}(X_n) \longrightarrow X \text{ et } \mathbb{V}(X_n) \longrightarrow 0$$

Exercice 2

Soit $X_n \overset{iid}{\sim} \mathcal{U}[0, \frac{1}{n}]$. *Montrer que* X_n *converge en moyenne d'ordre* $p \geq 1$ *vers* 0.

Exercice 2

Soit $X_n \overset{iid}{\sim} \mathcal{U}[0, \frac{1}{n}]$. *Montrer que* X_n *converge en moyenne d'ordre* $p \geq 1$ *vers* 0.

La densité de X_n est définie par

$$f(x_n) = \begin{cases} n & \text{si } 0 \leq x_n \leq \frac{1}{n} \\ 0 & \text{sinon} \end{cases}$$

On a donc,

$$\mathbb{E}(|X_n|^p) = \int_0^{\frac{1}{n}} n x^p dx = \frac{1}{(p+1)n^p} \xrightarrow[n \longrightarrow \infty]{} 0$$

Proposition 2

$$Si\ X_n \xrightarrow[n\longrightarrow\infty]{m.p} X,\ alors\ X_n \xrightarrow[n\longrightarrow\infty]{prob} X$$

Proposition 2

Si $X_n \xrightarrow[n \longrightarrow \infty]{m.p} X$, *alors* $X_n \xrightarrow[n \longrightarrow \infty]{prob} X$

Exercice 3

Soit $(X_n)_{n \in \mathbb{N}^*}$ *tel que:*

$$X_n = \begin{cases} n^2 \textit{ avec une probabilité } \frac{1}{n} \\ 0 \textit{ avec une probabilité } 1 - \frac{1}{n} \end{cases}$$

Étudier la convergence en probabilité, en moyenne puis en moyenne quadratique de la suite X_n vers 0.

En probabilité: $\forall \epsilon > 0,\ \mathbb{P}(X_n \geq \epsilon) = \mathbb{P}(X_n = n^2) = \frac{1}{n} \longrightarrow 0.$

Donc $X_n \xrightarrow[n\longrightarrow\infty]{prob} 0.$

En moyenne: $\mathbb{E}(|X_n|) = n^2 \times \frac{1}{n} + 0 \times \left(1 - \frac{1}{n}\right) = n \longrightarrow +\infty.$

Donc, X_n ne converge pas en moyenne vers 0.

En moyenne quadratique:

$\mathbb{E}(X_n^2) = n^4 \times \frac{1}{n} + 0^2 \times \left(1 - \frac{1}{n}\right) = n^3 \longrightarrow +\infty.$

Donc, X_n ne converge pas en moyenne quadratique vers 0.

1 Inégalités

2 Convergences
- convergence en probabilité
- Convergences en moyennes d'ordre p
- Convergence en distribution

3 Théorèmes limites

Définition 5 (Convergence en loi)

Une suite (X_n) ***converge en loi*** *vers une v.a* X *si* $\lim\limits_{\mathbf{n}\mapsto\infty}\mathbf{F_n(x)}=\mathbf{F(x)}$ *en tout point* x *où* F *est continue. On note* $X_n \xrightarrow[n\longrightarrow\infty]{loi} X$.

Définition 5 (Convergence en loi)

Une suite (X_n) ***converge en loi*** *vers une v.a* X *si* $\lim\limits_{\mathbf{n}\mapsto\infty} \mathbf{F_n(x)} = \mathbf{F(x)}$ *en tout point* x *où* F *est continue. On note* $X_n \xrightarrow[n\longrightarrow\infty]{loi} X$.

Exercice 4

Soit X_n *une suite de variables aléatoires telle que:*

$$F_n(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & si\ x > 0 \\ 0 & sinon \end{cases}$$

Montrer que $X_n \xrightarrow[n\longrightarrow\infty]{loi} \mathcal{E}(1)$.

Soit $F()$ la fonction de répartition de la $v.a$ X. On a, pour $x \leq 0, F_n(x) = F(x) = 0$ et pour $x > 0$, on a:

$$\begin{aligned}
\lim_{n \mapsto +\infty} F_n(x) &= \lim_{n \mapsto +\infty} \left(1 - \left(1 - \frac{1}{n}\right)^{nx}\right) \\
&= 1 - \lim_{n \mapsto +\infty} \left(1 - \frac{1}{n}\right)^{nx} \\
&= 1 - e^{-x} = F(x)
\end{aligned}$$

Ainsi, $X_n \xrightarrow[n \longrightarrow \infty]{loi} \mathcal{E}(1)$.

1. Inégalités
2. Convergences
3. **Théorèmes limites**

Proposition 3 (Loi faible des grands nombres)

Soit X une v.a admettant une variance σ^2. Soit (X_n) une suite de variables ***indépendantes et identiquement distribués*** *selon la loi de X.*

On pose $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Alors, $X_n \xrightarrow[n\longrightarrow\infty]{p} X$

La démonstration se fait à l'aide de linégalité de Bienaymé-Tchebechev.

$\mathbb{E}\left(\overline{X}_n\right) = \mathbb{E}(X_i)$ et $\mathbb{V}\left(\overline{X}_n\right) = \frac{\sigma^2}{n}$. L'inégalité de Bienaymé-Tchebychev donne:

$$\mathbb{P}\left(\left|\overline{X}_n - \mathbb{E}(X)\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} \xrightarrow[n\longrightarrow\infty]{} 0$$

Proposition 4 (Théorème centrale limite (TCL))

Soit $(X_n), n \in \mathbb{N}^*$ *une suite de variables aléatoires indépendantes de même loi, d'espérance m et de variance* σ^2 *finies. Posons:*

$$\forall n \in \mathbb{N}^* \,, \quad \overline{X}_n = \frac{X_1 + \cdots + X_n}{n} \quad et \quad Z_n = \frac{\sqrt{n}}{\sigma}\left(\overline{X}_n - m\right) \,.$$

La loi de Z_n *converge vers la loi normale* $\mathcal{N}(0,1)$*, c'est-à-dire que pour tout* $a < b$ *:* $\lim\limits_{n\to\infty} \mathbb{P}[a < Z_n < b] \; = \; \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$ *.*

Exercice 5

Une usine fabrique des pièces dont une proportion inconnue p est défectueuse, et on souhaite trouver une valeur approchée de p. On effectue un prélèvement de n pièces. On suppose que le prélèvement se fait sur une population très grande, et donc qu'il peut s'apparenter à une suite de n tirages indépendants avec remise. On note X_n la variable aléatoire égale au nombre de pièces défectueuses et on souhaite quantifier le fait que $\dfrac{X_n}{n}$ approche p.

1) *Quelle est la loi de X_n ? Sa moyenne ? Sa variance ?*
2) *Démontrer que, pour tout $\epsilon > 0$, $\mathbb{P}(|\frac{X_n}{n} - p| \geq \epsilon) \leq \dfrac{1}{4n\epsilon^2}$.*
3) *En déduire une condition sur n pour que X_n/n soit une valeur approchée de p à $10-2$ près avec une probabilité supérieure ou égale à 95%.*