Convergence des variables aléatoires

Cours de É. Bouchet – ECS1

19 mai 2021

Table des matières

1 Inégalités probabilistes

1.1 Inégalité de Markov

Proposition (Inégalité de Markov).

Soit X une variable aléatoire discrète positive admettant une espérance et $\lambda \in \mathbb{R}_+^*$:

$$P(X \geqslant \lambda) \leqslant \frac{E(X)}{\lambda}.$$

Remarque. Le résultat énoncé est celui du programme, cependant la condition « variable discrète » n'est pas nécessaire pour que l'inégalité soit vérifiée.

Démonstration. (démonstration à connaître) Comme X est une variable aléatoire discrète, on peut supposer que $X(\Omega) = \{x_i | i \in I\}$ où I est un sous-ensemble de $\mathbb{Z}$. Alors :

$$P(X \geqslant \lambda) = \sum_{x_i \geqslant \lambda} P(X = x_i).$$

Par ailleurs, par définition de l'espérance et positivité des x_i,

$$E(X) = \sum_{i \in I} x_i P(X = x_i) \geqslant \sum_{x_i \geqslant \lambda} x_i P(X = x_i) \geqslant \sum_{x_i \geqslant \lambda} \lambda P(X = x_i) = \lambda \sum_{x_i \geqslant \lambda} P(X = x_i).$$

Donc $E(X) \geqslant \lambda P(X \geqslant \lambda)$, et en divisant par $\lambda \in \mathbb{R}_+^*$ on obtient le résultat souhaité. □

1.2 Inégalité de Bienaymé-Tchebychev

Proposition (Inégalité de Bienaymé-Tchebychev).

Soit X une variable aléatoire discrète admettant un moment d'ordre 2 et $\varepsilon \in \mathbb{R}_+^*$:

$$P(|X - E(X)| \geqslant \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2}.$$

Démonstration. La variable $(X - E(X))^2$ est une variable aléatoire discrète positive admettant une espérance (car X admet une variance), on peut donc lui appliquer l'inégalité de Markov : pour $\lambda = \varepsilon^2 \in \mathbb{R}_+^*$,

$$P\left((X - E(X))^2 \geqslant \varepsilon^2\right) \leqslant \frac{E\left((X - E(X))^2\right)}{\varepsilon^2} = \frac{V(X)}{\varepsilon^2}.$$

Or, $(X - E(X))^2 \geqslant \varepsilon^2 \iff |X - E(X)| \geqslant \varepsilon$ car $\varepsilon > 0$ et par stricte croissance de la fonction racine carrée de $\mathbb{R}_+$ dans $\mathbb{R}_+$. Donc

$$P\left((X - E(X))^2 \geqslant \varepsilon^2\right) = P(|X - E(X)| \geqslant \varepsilon)$$

d'où le résultat. □

2 Convergence en probabilité

2.1 Convergence en probabilité

Définition (Convergence en probabilité)**.**

Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$. On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en probabilité** vers X lorsque pour tout $\varepsilon > 0$,

$$\lim_{n\to+\infty} P\left(|X_n - X| > \varepsilon\right) = 0.$$

On note $X_n \xrightarrow{P} X$.

Remarque. $X_n - X \xrightarrow{P} 0$ si et seulement si $X_n \xrightarrow{P} X$

Remarque. Il suffit de montrer que $\lim_{n\to+\infty} P\left(|X_n - X| \geqslant \varepsilon\right) = 0$ car $\forall\varepsilon > 0$,

$$0 \leqslant P\left(|X_n - X| > \varepsilon\right) \leqslant P\left(|X_n - X| \geqslant \varepsilon\right).$$

Remarque. Il suffit de montrer que $\forall\varepsilon \in]0,1]$, $\lim_{n\to+\infty} P\left(|X_n - X| > \varepsilon\right) = 0$. En effet, si $\varepsilon > 1$,

$$0 \leqslant P\left(|X_n - X| > \varepsilon\right) \leqslant P\left(|X_n - X| > 1\right).$$

Exemple 1. Si pour tout $n \in \mathbb{N}^*$, X_n suit une loi exponentielle de paramètre n, la suite des (X_n) converge en probabilités vers la variable aléatoire constante égale à 0.
En effet, si on fixe $\varepsilon > 0$,

$$P\left(|X_n - X| > \varepsilon\right) = P\left(X_n > \varepsilon\right) = 1 - P\left(X_n \leqslant \varepsilon\right).$$

La fonction de répartition de la loi exponentielle donne donc :

$$P\left(|X_n - X| > \varepsilon\right) = 1 - 1 + e^{-n\varepsilon} = e^{-n\varepsilon} \xrightarrow[n\to\infty]{} 0.$$

D'où le résultat (on aurait aussi pu vouloir utiliser l'inégalité de Markov, mais le programme de première année ne le permet pas car c'est une variable à densité).

Remarque. La convergence en probabilités n'entraîne pas la convergence des espérances.

Exemple 2. Pour tout $n \in \mathbb{N}^*$, on définit une variable aléatoire X_n telle que $X_n = n$ avec probabilité $\frac{1}{n}$ et $X_n = 0$ avec probabilité $1 - \frac{1}{n}$. Alors $(X_n)_{n\geqslant 0}$ converge en probabilités vers la variable aléatoire constante égale à 0 mais $E(X_n)$ ne converge pas vers $E(0)$.
Soit $\varepsilon \in]0,1[$ (on peut se ramener à étudier cet intervalle, car ce sont les petits ε qui posent problème) :

$$P\left(|X_n - X| > \varepsilon\right) = P\left(X_n > \varepsilon\right) \overset{\varepsilon>0}{=} P\left(X_n = n\right) = \frac{1}{n}.$$

Comme $\lim_{n\to\infty} \frac{1}{n} = 0$, on en déduit la convergence en probabilités annoncée. Il ne reste plus qu'à calculer l'espérance : pour tout $n \in \mathbb{N}^*$,

$$E(X_n) = 0P(X_n = 0) + nP(X_n = n) = 0 + \frac{n}{n} = 1.$$

Or 1 ne converge pas vers 0. D'où le résultat.

2.2 Loi faible des grands nombres

Théorème (Loi faible des grands nombres)**.**

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que $\forall n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}(n,p)$, alors $\dfrac{1}{n}X_n \xrightarrow{P} p$.

Démonstration. (démonstration à connaître) On utilise l'inégalité de Bienaymé-Tchebychev : les $\dfrac{1}{n}X_n$ sont des variables aléatoires discrètes, qui admettent une variance. Notamment, pour tout entier n,

$$E\left(\frac{X_n}{n}\right) = \frac{1}{n}E(X_n) = \frac{np}{n} = p \quad \text{et} \quad V\left(\frac{X_n}{n}\right) = \frac{1}{n^2}V(X_n) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

On a donc, pour $n \in \mathbb{N}$ et tout $\varepsilon > 0$,

$$0 \leqslant P\left(\left|\frac{X_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{p(1-p)}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0.$$

D'où le résultat par théorème d'encadrement. □

3 Convergence en loi

3.1 Convergence en loi

Définition (Convergence en loi)**.**

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles. On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers une variable aléatoire réelle X lorsque pour tout $x \in \mathbb{R}$ où F_X est continue,

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x).$$

On note alors $X_n \xrightarrow{\mathcal{L}} X$.

Remarque. La nature des variables X_n et de X peut être différente (discrètes ou à densité).

Exemple 3. Pour tout $n \in \mathbb{N}^*$, on définit X_n une variable aléatoire de loi uniforme sur $\left[-\frac{1}{n}, \frac{1}{n}\right]$. Étudier la convergence en loi de la suite $(X_n)_{n\in\mathbb{N}^*}$.
On commence par déterminer la fonction de répartition de X_n, en utilisant la fonction de répartition de la loi uniforme :

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leqslant -\frac{1}{n} \\ \frac{nx+1}{2} & \text{si } x \in \left[-\frac{1}{n}, \frac{1}{n}\right] \\ 1 & \text{si } x \geqslant \frac{1}{n} \end{cases}$$

- Soit $x > 0$. Pour tout $n \geqslant \frac{1}{x}$, on a $x \geqslant \frac{1}{n}$ et donc $F_{X_n}(x) = 1$. Donc $\lim\limits_{n\to+\infty} F_{X_n}(x) = 1$.
- Soit $x < 0$. Pour tout $n \geqslant -\frac{1}{x}$, $x \leqslant -\frac{1}{n}$ et donc $F_{X_n}(x) = 0$. Donc $\lim\limits_{n\to+\infty} F_{X_n}(x) = 0$.

On reconnaît après passage à la limite la fonction de répartition de la variable certaine égale à 0 (qui n'est pas continue en $x = 0$, ce qui justifie de n'avoir pas étudié les limites en ce point). Donc la suite (X_n) converge en loi vers la variable certaine égale à 0.

Proposition (Cas où X_n et X sont à valeurs dans $\mathbb{N}$)**.**

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle. On suppose que $X_n(\Omega)$ et $X(\Omega)$ sont inclus dans $\mathbb{N}$. Alors $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si pour tout $i \in \mathbb{N}$,

$$\lim_{n\to+\infty} P(X_n = i) = P(X = i).$$

Démonstration.

— Supposons que $\forall i \in \mathbb{N}$, $\lim\limits_{n\to+\infty} P(X_n = i) = P(X = i)$. Soit $x \in \mathbb{R}$, par propriété de la fonction de répartition, pour tout $k \in \mathbb{N}$,

$$F_{X_k}(x) = \sum_{i=0}^{\lfloor x \rfloor} P(X_k = i),$$

qui converge par linéarité des limites vers $\sum\limits_{i=0}^{\lfloor x \rfloor} P(X = i) = F_X(x)$. Donc $X_n \xrightarrow{\mathcal{L}} X$.

— Supposons que $X_n \xrightarrow{\mathcal{L}} X$. Alors pour tout $x \in \mathbb{R}\setminus\mathbb{N}$ (qui correspondent aux points de continuité), $\lim F_{X_k}(x) = F_X(x)$. Or on sait que par propriété de la fonction de répartition, pour tout k, et pour tout $i \geqslant 0$,

$$P(X_k = i) = F_{X_k}(i) - F_{X_k}(i-1),$$

sauf que ça n'aboutit pas ici : on ne peut pas passer à la limite, parce que i et $i+1$ ne sont pas des points de continuité. On peut par contre modifier légèrement l'expression pour contourner ce problème : comme la variable aléatoire est à valeurs entières, pour tout k, et pour tout $i \in \mathbb{N}$,

$$P(X_k = i) = F_{X_k}\left(i + \frac{1}{2}\right) - F_{X_k}\left(i - \frac{1}{2}\right),$$

et on est alors bien en des points de continuité : on peut passer à la limite, et

$$\lim_{k\to+\infty} P(X_k = i) = F_X\left(i + \frac{1}{2}\right) - F_X\left(i - \frac{1}{2}\right) = P(X = i).$$

□

3.2 Approximation poissonnienne d'une loi binomiale

Proposition (Approximation poissonnienne d'une loi binomiale)**.**

Soit $(p_n) \in]0,1[^{\mathbb{N}}$ une suite de réels telle que (np_n) converge vers un réel strictement positif λ. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour tout $n \in \mathbb{N}$, $X_n \hookrightarrow \mathcal{B}(n, p_n)$. Alors $X_n \xrightarrow{\mathcal{L}} Y$, où $Y \hookrightarrow \mathcal{P}(\lambda)$.

Démonstration. (démonstration à connaître) Pour tout $k \in \mathbb{N}$ et pour tout $n \geqslant k$ (car $X_n(\Omega) = [\![0, n]\!]$),

$$P(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}.$$

Or $\binom{n}{k} \sim \frac{n^k}{k!}$ à k fixé, donc $\binom{n}{k} p_n^k \sim \frac{(np_n)^k}{k!}$ va converger vers $\frac{\lambda^k}{k!}$ lorsque n tend vers l'infini. De plus,

$$(1-p_n)^{n-k} = \exp\left((n-k)\ln(1-p_n)\right),$$

et $\ln(1-p_n) \sim -p_n$, car $-p_n \sim -\dfrac{\lambda}{n}$ converge vers 0. Donc $(n-k)\ln(1-p_n) \sim -np_n \sim -\lambda$, et par continuité de l'exponentielle en $-\lambda$,

$$\lim_{n\to+\infty} (1-p_n)^{n-k} = \exp(-\lambda).$$

Par produit, on obtient : $\lim\limits_{n\to+\infty} P(X_n = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}$, ce qui termine la preuve. □

3.3 Théorème limite central

Théorème (Théorème limite central : approximation normale d'une loi binomiale).

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles discrètes telle que pour tout $n \in \mathbb{N}$, $X_n \hookrightarrow \mathcal{B}(n,p)$. Alors

$$X_n^* = \frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} Y,$$

où $Y \hookrightarrow \mathcal{N}(0,1)$.

Théorème (Théorème limite central : approximation normale d'une loi de Poisson).

Soient $\lambda > 0$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles discrètes telle que pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{P}(n\lambda)$. Alors

$$X_n^* = \frac{X_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{\mathcal{L}} Y,$$

où $Y \hookrightarrow \mathcal{N}(0,1)$.

Exemple 4. Étudier la convergence en loi de la suite de variables $\left(\frac{X_n - \frac{n}{3}}{\sqrt{\frac{2n}{9}}}\right)_{n\in\mathbb{N}^*}$ où X_n suit une loi binomiale de paramètre $\left(n, \frac{1}{3}\right)$. En déduire que $\lim\limits_{n\to+\infty} \dfrac{1}{3^n} \sum\limits_{k=0}^{\lfloor n/3\rfloor} \dbinom{n}{k} 2^{n-k} = \dfrac{1}{2}$.

Soit $n \in \mathbb{N}^*$. La variable aléatoire étudiée est la variable centrée réduite X_n^* associée à la variable X_n, car $pn = \frac{n}{3}$, et $p(1-p)n = \frac{n}{3}\frac{2}{3} = \frac{2n}{9}$. Par le théorème limite central, la suite de variables converge donc en loi vers une variable aléatoire Y de loi normale centrée réduite. F_Y est continue sur $\mathbb{R}$, donc pour tout $x \in \mathbb{R}$, $F_{X_n^*}(x)$ converge vers $F_Y(x)$. $\frac{1}{2}$ nous évoque $F_Y(0)$, ce que incite à étudier le cas particulier $x = 0$:

$$\lim_{n\to+\infty} F_{X_n^*}(0) = F_Y(0) = \frac{1}{2}.$$

Par ailleurs,

$$F_{X_n^*}(0) = P\left(\frac{X_n - \dfrac{n}{3}}{\sqrt{\dfrac{2n}{9}}} \leqslant 0\right) = P\left(X_n - \frac{n}{3} \leqslant 0\right) = P\left(X_n \leqslant \frac{n}{3}\right) = P\left(X_n \leqslant \left\lfloor \frac{n}{3} \right\rfloor\right),$$

où la dernière égalité est vraie car X_n est à valeurs entières. Or on connaît la loi de X_n :

$$P\left(X_n \leqslant \frac{n}{3}\right) = \sum_{k=0}^{\lfloor n/3\rfloor} \binom{n}{k}\left(\frac{1}{3}\right)^k \left(1-\frac{1}{3}\right)^{n-k} = \frac{1}{3^n}\sum_{k=0}^{\lfloor n/3\rfloor} \binom{n}{k} 2^{n-k}.$$

En remplaçant dans les expressions précédentes, on obtient donc bien la limite demandée.