13

Probabilités

Convergence de suites de variables aléatoires

Introduction...

Ce chapitre met l'accent sur deux théorèmes fondamentaux en théorie des probabilités et en statistiques : la loi faible des grands nombres et le théorème central limite.
La loi faible des grands nombres est due à Pafnouti Tchebychev (1821-1894, russe) qui la démontra en utilisant une inégalité énoncée par Irénée-Jules Bienaymé (1796-1878, français) et démontrée par Tchebychev lui-même.
Quant au théorème central limite : des cas particuliers ont d'abord été démontrés par Abraham de Moivre (1667-1754, français) et Laplace, mais concernant la version générale, qui l'a démontrée en premier ? En 1920, il semble que Alexandre Liapounov (1857-1918, russe) et Jarl Waldemar Lindeberg (1876-1932, finlandais) en aient tous deux fourni une démonstration différente. En 1922, Paul Lévy (1886-1971, français) démontre le théorème qui porte son nom, dont le TCL est une conséquence immédiate, et permet ainsi à Lindeberg d'en donner une version avec des hypothèses amoindries ! Comme bien souvent, ce théorème est le fruit de collaborations plus ou moins volontaires entre des mathématiciens d'époques et de nationalités variées.

Il est important de noter que l'Histoire se rapproche ! Nous sommes à présent au début du XXème siècle, qui demeurera sans aucun doute le siècle de l'essor des probabilités comme branche légitime des mathématiques.

Pour bien démarrer...

1. Pour rappel :

	Variables aléatoires discrètes	Variables aléatoires à densité
Donnée de la loi de probabilité	$X(\Omega)$ est fini ou dénombrable et on donne la *fonction de masse* : • $\mathbb{P}\big([X=n]\big)$ pour tout $n \in X(\Omega)$ • $\mathbb{P}\big([X=n]\big)=0$ pour tout n en dehors de $X(\Omega)$ • On a : $\displaystyle\sum_{n\in X(\Omega)} \mathbb{P}\big([X=n]\big)=1$	$X(\Omega)$ est un intervalle (ou union d'intervalles) de $\mathbb{R}$ et on donne une *fonction de densité* : • f_X est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points • f_X est positive sur $\mathbb{R}$ (peut être choisie nulle en dehors de $X(\Omega)$) • $\displaystyle\int_{-\infty}^{+\infty} f_X(t)dt$ est convergente et vaut 1
Calculs de probabilités	$\mathbb{P}\big([a\leqslant X\leqslant b]\big)=\displaystyle\sum_{\substack{n\in X(\Omega)\\ a\leqslant n\leqslant b}} \mathbb{P}\big([X=n]\big)=F_X(b)-F_X(a)+\mathbb{P}\big([X=a]\big)$ $\mathbb{P}\big([X\leqslant b]\big)=\displaystyle\sum_{\substack{n\in X(\Omega)\\ n\leqslant b}} \mathbb{P}\big([X=n]\big)$ $\mathbb{P}\big([X\geqslant a]\big)=\displaystyle\sum_{\substack{n\in X(\Omega)\\ n\geqslant a}} \mathbb{P}\big([X=n]\big)$	$\mathbb{P}\big([a\leqslant X\leqslant b]\big)=\displaystyle\int_a^b f_X(t)dt=F_X(b)-F_X(a)$ $\mathbb{P}\big([X\leqslant b]\big)=\displaystyle\int_{-\infty}^b f_X(t)dt$ $\mathbb{P}\big([X\geqslant a]\big)=\displaystyle\int_a^{+\infty} f_X(t)dt$ $\mathbb{P}\big([X=a]\big)=0$; $\mathbb{P}\big([X<b]\big)=\mathbb{P}\big([X\leqslant b]\big)$; $\mathbb{P}\big([X>a]\big)=\mathbb{P}\big([X\geqslant a]\big)$
Propriétés de la fonction de répartition F_X	• F_X est constante par morceaux • F_X est discontinue en chaque $n\in X(\Omega)$ • le saut de continuité en n est égal à $\mathbb{P}(X=n)$ • $\forall n\in X(\Omega),\ \mathbb{P}\big([X\leqslant n]\big)-\mathbb{P}\big([X\leqslant n-1]\big)=\mathbb{P}(X=n)$	• F_X continue sur $\mathbb{R}$ • F_X est $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points • pour tout x où F_X est $\mathscr{C}^1$: $F_X'(x)=f_X(x)$
Indépendance	Pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$: $\mathbb{P}\big([X=x]\cap[Y=y]\big)=\mathbb{P}\big([X=x]\big)\times\mathbb{P}\big([Y=y]\big)$	Pour tous intervalles I,J de $\mathbb{R}$: $\mathbb{P}\big([X\in I]\cap[Y\in J]\big)=\mathbb{P}\big([X\in I]\big)\times\mathbb{P}\big([Y\in J]\big)$
Espérance (si existence)	• X a une espérance ssi $\displaystyle\sum_{n\in X(\Omega)} \big\lvert n\mathbb{P}\big([X=n]\big)\big\rvert$ est convergente • $\mathbb{E}(X)=\displaystyle\sum_{n\in X(\Omega)} n\mathbb{P}\big([X=n]\big)$	• X a une espérance ssi $\displaystyle\int_{-\infty}^{+\infty} \lvert tf_X(t)\rvert dt$ est convergente • $\mathbb{E}(X)=\displaystyle\int_{-\infty}^{+\infty} tf_X(t)dt$
Théorème de transfert	• $g(X)$ a une espérance ssi $\displaystyle\sum_{n\in X(\Omega)} \lvert g(n)\mathbb{P}(X=n)\rvert$ est convergente • $\mathbb{E}\big(g(X)\big)=\displaystyle\sum_{n\in X(\Omega)} g(n)\mathbb{P}\big([X=n]\big)$ et en particulier : $\mathbb{E}(X^2)=\displaystyle\sum_{n\in X(\Omega)} n^2\mathbb{P}\big([X=n]\big)$	Si g est continue sur $X(\Omega)$: • $g(X)$ a une espérance ssi $\displaystyle\int_{X(\Omega)} \lvert g(t)f_X(t)\rvert dt$ est convergente • $\mathbb{E}\big(g(X)\big)=\displaystyle\int_{X(\Omega)} g(t)f_X(t)dt$ et en particulier : $\mathbb{E}(X^2)=\displaystyle\int_{X(\Omega)} t^2 f_X(t)dt$
Variance (si existence)	$\mathbb{V}(X)=\mathbb{E}\Big(\big(X-\mathbb{E}(X)\big)^2\Big)$	
Formule de Keonig-Huygens	$\mathbb{V}(X)=\mathbb{E}(X^2)-\mathbb{E}(X)^2$	
Propriétés de l'espérance et de la variance	Soient $a,b\in\mathbb{R}$ et $n\in\mathbb{N}^*$. • Linéarité de l'espérance. Si X et Y ont une espérance, alors, $aX+bY$ aussi et : $\mathbb{E}(aX+bY)=a\mathbb{E}(X)+b\mathbb{E}(Y)$. • Croissance de l'espérance. Si X et Y ont une espérance et $\mathbb{P}\big([X\leqslant Y]\big)=1$, alors $\mathbb{E}(X)\leqslant\mathbb{E}(Y)$. • Si X a une variance, alors $aX+b$ aussi et : $\mathbb{V}(aX+b)=a^2\mathbb{V}(X)$. • Si X et Y ont une espérance et sont indépendantes, alors XY a une espérance et $\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$. Si $X_1,\dots,X_n$ ont une espérance et sont indépendantes, alors $\displaystyle\prod_{k=1}^n X_k$ a une espérance et $\mathbb{E}\left(\displaystyle\prod_{k=1}^n X_k\right)=\displaystyle\prod_{k=1}^n \mathbb{E}(X_k)$. • Si X et Y ont une variance et sont indépendantes, alors $X+Y$ a une variance et $\mathbb{V}(X+Y)=\mathbb{V}(X)+\mathbb{V}(Y)$. Si $X_1,\dots,X_n$ ont une variance et sont indépendantes, alors $\displaystyle\sum_{k=1}^n X_k$ a une variance et $\mathbb{V}\left(\displaystyle\sum_{k=1}^n X_k\right)=\displaystyle\sum_{k=1}^n \mathbb{V}(X_k)$.	

On admet que l'on peut étendre les propriétés sur l'espérance et la variance, vues dans le cas discret (et démontrées dans le chapitre 2) au cas des variables aléatoires à densité. Il nous manque des outils pour les démontrer dans ce cas-ci.
On admet également que l'on peut définir l'indépendance (celle dans le cas à densité), l'espérance et la variance pour des variables aléatoires ni discrètes, ni à densité. Puis on admet ensuite que les propriétés énoncées en fin de tableau sont valables dans ce cas de figure.

Dans tout le chapitre, $(\Omega, \mathcal{A}, \mathbb{P})$ désigne un espace probabilisé et X une variable aléatoire définie sur cet espace. Sauf précision, les variables aléatoires étudiées seront quelconques (discrètes, à densité ou ni discrète ni à densité).

I Inégalités de Markov et Bienaymé-Tchebychev

Théorème 1 — Inégalité de Markov

Si X est à valeurs positives et admet une espérance, alors :

$$\forall a \in \mathbb{R}_*^+,\ \mathbb{P}\big([X \geqslant a]\big) \leqslant \frac{\mathbb{E}(X)}{a}$$

Petite remarque

Selon le programme, la connaissance de l'inégalité de Markov n'est pas exigible, mais elle semble parfois nécessaire à l'oral d'HEC... Il faut, dans tous les cas, en connaître des démonstrations.

Interprétation

• Si $a \leqslant \mathbb{E}(X)$, l'inégalité n'apporte rien, car alors $\dfrac{\mathbb{E}(X)}{a} \geqslant 1$...
• Le résultat est donc intéressant si a est grand par rapport à $\mathbb{E}(X)$. Dans ce cas, l'inégalité affirme qu'il est peu probable que X prenne des valeurs trop grandes par rapport à $\mathbb{E}(X)$.
• Cette inégalité a l'avantage d'être valable pour toutes les variables aléatoires positives ; avantage : elle est très générale. Inconvénient : la majoration de $\mathbb{P}\big([X \geqslant a]\big)$ n'est pas souvent très fine.

Démonstration : Soit $a \in \mathbb{R}_*^+$.

- **Si X est discrète.**
  On a :

$$\begin{aligned} a\mathbb{P}\big([X \geqslant a]\big) &= a \sum_{x \in X(\Omega),\ x \geqslant a} \mathbb{P}\big([X = x]\big) \\ &= \sum_{x \in X(\Omega),\ x \geqslant a} a\mathbb{P}\big([X = x]\big) \\ &\leqslant \sum_{x \in X(\Omega),\ x \geqslant a} x\mathbb{P}\big([X = x]\big) && \forall x \geqslant a,\ a\mathbb{P}\big([X = x]\big) \leqslant x\mathbb{P}\big([X = x]\big) \text{ (car une probabilité est positive)} \\ &\leqslant \underbrace{\sum_{x \in X(\Omega)} x\mathbb{P}\big([X = x]\big)}_{=\mathbb{E}(X)} && X \text{ est à valeurs positives} \end{aligned}$$

  D'où le résultat, puisque $a > 0$.

- **Si X est à densité.**
  Notons f une densité de X. On a :

$$\begin{aligned} a\mathbb{P}\big([X \geqslant a]\big) &= a \int_a^{+\infty} f(x)dx \\ &= \int_a^{+\infty} af(x)dx \\ &\leqslant \int_a^{+\infty} xf(x)dx && \forall x \geqslant a,\ af(x) \leqslant xf(x) \text{ (car } f \text{ positive) et par croissance de l'intégrale} \\ &\leqslant \underbrace{\int_0^{+\infty} xf(x)dx}_{=\mathbb{E}(X),\ \text{car } X(\Omega) \subset \mathbb{R}^+} && a > 0 \text{ et pour tout } x \in \mathbb{R}^+,\ xf(x) \geqslant 0 \end{aligned}$$

  D'où le résultat, puisque $a > 0$.

- Dans le cas où X est quelconque, voir QC38.

Exemple 1

Soient $r \in \mathbb{N}^*$ et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ admettant un moment d'ordre r. Démontrons :

$$\forall \alpha \in \mathbb{R}_*^+,\ \mathbb{P}\big([|X| \geqslant \alpha]\big) \leqslant \frac{\mathbb{E}\big(|X|^r\big)}{\alpha^r}$$

☞ Pour info...

On a même le résultat suivant : si f est une fonction strictement croissante et positive sur un intervalle I, si X est une variable aléatoire à valeurs dans I, alors pour tout $b \in I$ tel que $f(b) > 0$, on a : $\mathbb{P}\big([X \geqslant b]\big) \leqslant \dfrac{\mathbb{E}\big(f(X)\big)}{f(b)}$.

En conséquence de l'inégalité de Markov :

Théorème 2 — **Inégalité de Bienaymé-Tchebychev**

Si X admet une variance, alors :

$$\forall a \in \mathbb{R}_*^+,\ \mathbb{P}\big([|X-\mathbb{E}(X)| \geqslant a]\big) \leqslant \frac{\mathbb{V}(X)}{a^2}$$

☞ Rappel...

X admet une variance si, et seulement si, X admet un moment d'ordre 2.

☞ Pour info...

L'inégalité de BT est un cas *d'inégalité de concentration.*
Elle permet d'estimer la probabilité qu'une variable aléatoire s'écarte de sa moyenne.
Tout comme l'inégalité de Markov, elle est générale mais la majoration fournie n'est pas très fine...
On peut déjà remarquer que si l'on prend $a \leqslant \sigma(X)$, l'inégalité de BT est inutile !

⋆ Démonstration :

⋆

Exemples 2

E1 Lors de l'épreuve de mathématiques organisée par l'EMLyon, la moyenne des notes est égale à 10 et l'écart-type à 4. On modélise par une variable aléatoire X la note d'un candidat choisi au hasard. A l'aide de l'inégalité de Bienaymé-Tchebychev, démontrons que $\mathbb{P}\big([5 \leqslant X \leqslant 15]\big) \geqslant 0,36$.

On suppose maintenant que $X \hookrightarrow \mathcal{N}(10; 4^2)$. A l'aide de la table de la loi $\mathcal{N}(0; 1)$, déterminons une valeur approchée de $\mathbb{P}\big([5 \leqslant X \leqslant 15]\big)$.

E2 Soient $n \in [\![2; +\infty[\![$ et $X_1, X_2, \dots, X_n$ des variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes et suivant toutes la même loi de Bernoulli de paramètre $p \in]0; 1[$. On note $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.

Démontrons : $\forall \varepsilon > 0,\ \mathbb{P}\big([|\overline{X_n} - p| \geqslant \varepsilon]\big) \leqslant \dfrac{1}{4n\varepsilon^2}$.

II Loi faible des grands nombres

Avant d'énoncer ce célèbre théorème, commençons par deux résultats immédiats sur la moyenne empirique...
Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes admettant toutes la même espérance notée m et la même variance notée σ^2. Soit $n\in\mathbb{N}^*$. Notons $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$. Justifions que $\overline{X_n}$ possède une espérance et une variance et déterminons-les.

Vocabulaire
$\overline{X_n}$ est la **moyenne empirique** de $X_1, X_2, \dots, X_n$.

Théorème 3 — Loi faible des grands nombres

Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$. Pour tout $n\in\mathbb{N}^*$, on note $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.

Si $(X_k)_{k\in\mathbb{N}^*}$ est une suite de variables aléatoires **indépendantes**, admettant toutes **la même espérance m** et **la même variance σ^2** (c'est le cas si elles ont toutes la même loi) alors :

$$\forall \varepsilon > 0,\quad \lim_{n\to+\infty} \mathbb{P}\big([|\overline{X_n} - m| \geqslant \varepsilon]\big) = 0$$

Vocabulaire
On dit que la suite $(\overline{X_n})_{n\in\mathbb{N}^*}$ converge en probabilité (vocabulaire HP) vers la variable aléatoire constante égale à m.

Utile ?
Pour tout $\varepsilon > 0$, on a :
$[|\overline{X_n} - m| > \varepsilon] \subset [|\overline{X_n} - m| \geqslant \varepsilon]$.
On a donc également :
$\lim\limits_{n\to+\infty} \mathbb{P}\big([|\overline{X_n} - m| > \varepsilon]\big) = 0$

⋆ Démonstration : Voir QCI39. ⋆

Exemples 3

> **Petite remarque**
> Pour une autre application de la LfGN, voir le TP sur la Méthode de Monté-Carlo.

E1 On considère une variable aléatoire X admettant une espérance et une variance. On suppose l'existence d'une fonction **Python** nommée **simule_X** permettant de simuler une réalisation de X.
A quoi peut-on s'attendre lors de l'exécution du programme suivant ?

```python
L=[simule_X() for k in range(10000)]
print(sum(L)/len(L))
```

E2 Dans le même contexte, à quoi peut-on s'attendre lors de l'exécution du programme suivant ?

```python
c=0
for k in range(10000):
    if simule_X()>0:
        c=c+1
print(c/10000)
```

III Convergence en loi

III.1 Définition et premiers exemples

Définition 1 — Convergence en loi

Soient X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$. On note F_X la fonction de répartition de X et, pour tout $n \in \mathbb{N}^*$, on note F_{X_n} la fonction de répartition de X_n.
On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ **converge en loi vers X** lorsque, **pour tout x où F_X est continue**, on a :

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

Notation

On note $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$.
Attention : ~~$X_n(\omega) \to X(\omega)$~~. Il n'y a d'ailleurs pas unicité de X : toute variable aléatoire ayant la même loi que X convient. En revanche, il y a unicité de la fonction limite de $(F_{X_n})_{n\in\mathbb{N}^*}$. Et comme la fonction de répartition caractérise la loi, c'est parfait : il y a unicité de la *loi limite*.

Important !

On examine, **à x fixé**, la limite de $F_{X_n}(x)$ quand $n \to +\infty$ (on parle de convergence simple de la suite de fonctions $(F_{X_n})_{n\in\mathbb{N}^*}$: vocabulaire HP).

♣ Méthode 1 ♣ Pour établir que $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$:

- On commence par rappeler (ou déterminer) la fonction de répartition de X et celles de X_n, pour tout $n \in \mathbb{N}^*$.
- On **fixe x** dans $\mathbb{R}$ **sauf en les points de discontinuité de F_X** (si X est à densité, F_X est continue, donc x sera dans $\mathbb{R}$ entier).
- On examine alors $\lim\limits_{n\to+\infty} F_{X_n}(x)$, en observant bien les disjonctions de cas sur $F_X(x)$...

♣ Méthode !

Il se peut que l'énoncé soit formulé ainsi "Montrer que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X dont on précisera la loi." ; la méthode ne change alors pas beaucoup...

Exemples 4

E1 Pour tout $n \in \mathbb{N}^*$, on considère une variable aléatoire X_n suivant la loi exponentielle de paramètre $1 + \dfrac{1}{n}$.
Montrons que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X suivant la loi exponentielle de paramètre 1.

ici

E2 Pour tout $n \in \mathbb{N}^*$, on considère une variable aléatoire X_n suivant la loi uniforme sur $\left[\dfrac{-1}{n} ; \dfrac{1}{n}\right]$. Montrons que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X suivant la loi certaine égale à 0.

ici

$\boxed{\text{E3}}$ Pour tout $n \in \mathbb{N}^*$, on considère une variable aléatoire X_n telle que $\mathbb{P}\big([X_n = 0]\big) = 1 - \dfrac{1}{n}$ et $\mathbb{P}\big([X_n = n]\big) = \dfrac{1}{n}$. Montrons que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X dont on précisera la loi.

ici

$\boxed{\text{E4}}$ Pour tout $n \in \mathbb{N}^*$, on considère une variable aléatoire X_n suivant la loi exponentielle de paramètre $\dfrac{1}{n}$. Montrons que $(X_n)_{n\in\mathbb{N}^*}$ ne converge pas en loi.

Petite remarque

Les exemples ci-dessus mettent en évidence qu'il n'y a pas toujours convergence en loi ; et que, si convergence, les cas "discret" → "discret", "à densité" → "à densité" et "à densité" → "discret" peuvent se produire. Le cas "discret" → "à densité" est également possible...

Théorème 4 — **Cas "discret" → "discret" dans $\mathbb{N}$.**

Soient X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si $X(\Omega) \subset \mathbb{N}$ **et, pour tout** $n \in \mathbb{N}^*$**,** $X_n(\Omega) \subset \mathbb{N}$, alors :

$$\left(X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X\right) \iff \left(\forall k \in \mathbb{N},\ \lim_{n\to+\infty} \mathbb{P}\big([X_n = k]\big) = \mathbb{P}\big([X = k]\big)\right)$$

Important !

Toutes les variables aléatoires en jeu doivent être discrètes !

Petite remarque

C'est donc en particulier le cas si les variables aléatoires sont à valeurs dans un sous-ensemble fini de $\mathbb{N}$ (le résultat est en fait valable si les variables aléatoires sont discrètes, non nécessairement à valeurs dans $\mathbb{N}$...).

★ Démonstration : Supposons que X ainsi que $X_1, X_2, \ldots$ sont à valeurs dans $\mathbb{N}$. Raisonnons ensuite par double-implication.

$\boxed{\Longrightarrow}$ Supposons $\left(X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X\right)$. Soit $k \in \mathbb{N}$.
Puisque X et $X_1, X_2, \ldots$ sont à valeurs entières et que $k \in \mathbb{N}$, on a :

$$\begin{aligned}\mathbb{P}\big([X = k]\big) &= \mathbb{P}\big([X \leqslant k]\big) - \mathbb{P}\big([X \leqslant k-1]\big)\\ &= F_X(k) - F_X(k-1)\end{aligned}$$

et :

$$\begin{aligned}\mathbb{P}\big([X_n = k]\big) &= \mathbb{P}\big([X_n \leqslant k]\big) - \mathbb{P}\big([X_n \leqslant k-1]\big)\\ &= F_{X_n}(k) - F_{X_n}(k-1)\end{aligned}$$

Important !

On part de :

$$[X \leqslant k] = [X = k] \cup [X < k]$$

Ensuite, puisque X est à valeurs entières et que $k \in \mathbb{N}$, on a $[X < k] = [X \leqslant k-1]$... Et on mentionne l'incompatibilité de $[X = k]$ et $[X \leqslant k-1]$...

On a envie de passer à la limite quand $n \to +\infty$, mais on ne peut pas ! En effet, puisque $X(\Omega) \subset \mathbb{N}$, la fonction de répartition de X n'est a priori pas continue en k et $k-1$... Contournons le problème en remarquant que, puisque $X_1, X_2, \ldots$ sont à valeurs entières, on a :

$$\forall n \in \mathbb{N}^*,\ \mathbb{P}\big([X_n \leqslant k]\big) = \mathbb{P}\left(\left[X_n \leqslant k + \frac{1}{2}\right]\right) \quad ; \quad \mathbb{P}\big([X_n \leqslant k-1]\big) = \mathbb{P}\left(\left[X_n \leqslant k - 1 + \frac{1}{2}\right]\right)$$

Ainsi, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned}\mathbb{P}\big([X_n = k]\big) &= \mathbb{P}\left(\left[X_n \leqslant k + \frac{1}{2}\right]\right) - \mathbb{P}\left(\left[X_n \leqslant k - \frac{1}{2}\right]\right)\\ &= F_{X_n}\left(k + \frac{1}{2}\right) - F_{X_n}\left(k + -\frac{1}{2}\right)\end{aligned}$$

Or $k + \dfrac{1}{2}$ et $k - \dfrac{1}{2}$ n'appartiennent pas à $X(\Omega)$, donc F_X est continue en $k + \dfrac{1}{2}$ et $k - \dfrac{1}{2}$. Ainsi, puisque $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X, on a :

$$\lim_{n\to+\infty} F_{X_n}\left(k + \frac{1}{2}\right) = F_X\left(k + \frac{1}{2}\right) \quad ; \quad \lim_{n\to+\infty} F_{X_n}\left(k - \frac{1}{2}\right) = F_X\left(k - \frac{1}{2}\right)$$

D'où :

$$\begin{aligned}\lim_{n\to+\infty} \mathbb{P}\big([X_n = k]\big) &= F_X\left(k + \frac{1}{2}\right) - F_X\left(k - \frac{1}{2}\right)\\ &= \mathbb{P}\left(\left[X \leqslant k + \frac{1}{2}\right]\right) - \mathbb{P}\left(\left[X \leqslant k - \frac{1}{2}\right]\right)\\ &= \mathbb{P}\big([X \leqslant k]\big) - \mathbb{P}\big([X \leqslant k-1]\big) \qquad X \text{ est à valeurs entières}\\ &= \mathbb{P}\big([X = k]\big)\end{aligned}$$

Conclusion : $\left(X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X\right) \Longrightarrow \left(\forall k \in \mathbb{N},\ \lim\limits_{n\to+\infty} \mathbb{P}\big([X_n = k]\big) = \mathbb{P}\big([X = k]\big)\right)$.

$\boxed{\Longleftarrow}$ Supposons "$\forall k \in \mathbb{N},\ \lim\limits_{n\to+\infty} \mathbb{P}\big([X_n = k]\big) = \mathbb{P}\big([X = k]\big)$". Montrons que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X.
Soit $x \in \mathbb{R}$ tel que F_X est continue en x. On a, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned}F_{X_n}(x) &= \mathbb{P}\big([X_n \leqslant x]\big)\\ &= \mathbb{P}\big([X_n \leqslant \lfloor x \rfloor]\big) \qquad X_n \text{ est à valeurs entières}\\ &= \sum_{k=0}^{\lfloor x \rfloor} \mathbb{P}\big([X_n = k]\big)\end{aligned}$$

puisque $X_n(\Omega) \subset \mathbb{N}$, on a $[X_n \leqslant \lfloor x \rfloor] = \bigcup\limits_{k=0}^{\lfloor x \rfloor} [X_n = k]$; puis par incompatibilité des évènements de la famille $\big([X_n = k]\big)_{k\in[\![0;\lfloor x \rfloor]\!]}$

Or : $\forall k \in \mathbb{N},\ \lim\limits_{n\to+\infty} \mathbb{P}\big([X_n = k]\big) = \mathbb{P}\big([X=k]\big)$. D'où, par somme :

$$\begin{aligned}\lim_{n\to+\infty} F_{X_n}(x) &= \sum_{k=0}^{\lfloor x\rfloor} \mathbb{P}\big([X=k]\big) && \text{arguments similaires à ce qui précède}\\ &= \mathbb{P}\big([X \leqslant \lfloor x\rfloor]\big) && X \text{ est à valeurs entières}\\ &= \mathbb{P}\big([X\leqslant x]\big)\\ &= F_X(x)\end{aligned}$$

D'où la convergence en loi de $(X_n)_{n\in\mathbb{N}^*}$ vers X.

★Subtil...★
En fait, nous avons établi $\lim\limits_{n\to+\infty} F_{X_n}(x) = F_X(x)$ pour tous les réels x... même ceux en lesquels F_X n'est pas continue.

Exemple 5

Voir QCl40.

III.2 Théorème central limite

Théorème 5 — **Théorème central limite**

Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires **indépendantes**, **de même loi**, admettant la même espérance m et la même **variance σ^2 non nulle**.
On pose, pour tout $n\in\mathbb{N}^*$:

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k \quad ; \quad \overline{X_n}^* = \frac{\overline{X_n} - \mathbb{E}(\overline{X_n})}{\sigma(\overline{X_n})} = \sqrt{n}\,\frac{\overline{X_n} - m}{\sigma}$$

Dans ce cas, la suite $\left(\overline{X_n}^*\right)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi $\mathcal{N}(0;1)$.
Autrement dit :

$$\forall x\in\mathbb{R},\ \lim_{n\to+\infty} \mathbb{P}\big([\overline{X_n}^* \leqslant x]\big) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}}\, dt$$

Ou encore :

$$\forall (a,b)\in\mathbb{R}^2,\ \left(a<b \Longrightarrow \lim_{n\to+\infty} \mathbb{P}\left([a\leqslant \overline{X_n}^* \leqslant b]\right) = \int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}}\, dt\right)$$

Lorsque n est *suffisamment grand*, on pourra considérer que : $\mathbb{P}\big([a\leqslant \overline{X_n}^* \leqslant b]\big) \simeq \mathbb{P}\big([a\leqslant Z\leqslant b]\big)$ où $Z \hookrightarrow \mathcal{N}(0;1)$.

Un peu d'histoire
Dans sa toute première version, les variables aléatoires $X_1, X_2, \ldots$ suivaient une loi de Bernoulli de paramètre p. Sous cette condition, on parle alors de théorème de Moivre-Laplace.
Abraham de Moivre (1667-1754, français) l'a démontré en 1733 dans le cas où $p = \frac{1}{2}$, puis Laplace dans le cas général au début du XIXème siècle.

Autrement dit :
Quand on centre et qu'on réduit la moyenne empirique d'une suite de VA iid de variance non nulle, la VA obtenue converge en loi vers une VA suivant la loi $\mathcal{N}(0;1)$.

★ Démonstration : Il en existe différentes démonstrations... Dont une très courte, mais hors de portée avec nos outils. D'autres plus longues... On en trouvera une dans la dernière partie du sujet ESSEC II E 2022. ★

IV Approximations de variables aléatoires

Théorème 6 — **Approximation d'une loi binomiale par une loi de Poisson**

Soient $\lambda\in\mathbb{R}_*^+$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si : $\forall n\in\mathbb{N}^*,\ X_n \hookrightarrow \mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$, alors : la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi de Poisson de paramètre λ.

En pratique...
Lorsque $0 < np < 10$ (c'est à dire que si n est grand, p doit être petit) **on approche la loi $\mathcal{B}(n;p)$ par la loi $\mathcal{P}(np)$**.
On voit parfois les conditions : $n\geqslant 30$ et $p\leqslant 0,1$.

★ Démonstration : Cas particulier de QCl40. ★

Important !
Les deux théorèmes qui suivent ne sont pas explicitement au programme... Il est donc d'autant plus important de savoir les redémontrer.

Théorème 7 — **Approximation d'une loi binomiale par une loi normale (HP ?)**

Soient $p\in]0;1[$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que : $\forall n\in\mathbb{N}^*,\ X_n\hookrightarrow\mathcal{B}(n;p)$.
On pose, pour tout $n\in\mathbb{N}^*$, $X_n^* = \dfrac{X_n - np}{\sqrt{np(1-p)}}$.
Dans ce cas, la suite $(X_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi $\mathcal{N}(0;1)$.

★ Démonstration :

En pratique...
Lorsque $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$ (c'est à dire que si n est grand, p doit être ni trop petit ni trop grand) **on approche la loi $\mathscr{B}(n;p)$ par la loi $\mathscr{N}(np;np(1-p))$**.
On voit parfois les conditions : $n \geqslant 30$ et $p \leqslant 0,1$.
Notons que l'approximation se fait à espérance et variance constantes.

Conséquence : si, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathscr{B}(n;p)$, alors on peut considérer que pour n suffisamment grand, X_n suit approximativement la loi $\mathscr{N}(np;np(1-p))$.
En effet :

avec les notations du théorème 7, la suite $(X_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire Z suivant la loi $\mathscr{N}(0;1)$.
On peut donc considérer que pour n grand, la variable aléatoire X_n se comporte comme la variable aléatoire $\sqrt{np(1-p)}Z+np$.
Or $Z \hookrightarrow \mathscr{N}(0;1)$, donc $\sqrt{np(1-p)}Z+np \hookrightarrow \mathscr{N}(np;np(1-p))$. D'où le résultat.

☞ Rappel...
Soient $a, b \in \mathbb{R}$ avec $a \neq 0$.
Si $X \hookrightarrow \mathscr{N}(\mu;\sigma^2)$, alors $(aX+b) \hookrightarrow \mathscr{N}(a\mu+b;a^2\sigma^2)$

Exemple 6

On effectue 10000 lancers, supposés indépendants, d'une même pièce donnant PILE avec la probabilité $\dfrac{1}{2}$. On note X la variable aléatoire égale au nombre de PILE obtenus. Déterminons une valeur approchée de $\mathbb{P}\big([X \in [\![4900;5100]\!]]\big)$ et de $\mathbb{P}\big([X=5000]\big)$.

Théorème 8 — **Approximation d'une loi de Poisson par une loi normale (HP ?)**

Soient $\alpha > 0$ et $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que : $\forall n \in \mathbb{N}^*,\ X_n \hookrightarrow \mathscr{P}(n\alpha)$.
On pose, pour tout $n \in \mathbb{N}^*$, $X_n^* = \dfrac{X_n - n\alpha}{\sqrt{n\alpha}}$.
Dans ce cas, la suite $(X_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi $\mathscr{N}(0;1)$.

⋆ Démonstration :

⋆

Conséquence : si, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathscr{P}(n\alpha)$, alors on peut considérer que pour n suffisamment grand, X_n suit approximativement la loi $\mathscr{N}(n\alpha; n\alpha)$.
En effet :

avec les notations du théorème 8, la suite $(X_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire Z suivant la loi $\mathscr{N}(0;1)$.
On peut donc considérer que pour n grand, la variable aléatoire X_n se comporte comme la variable aléatoire $\sqrt{n\alpha}Z + n\alpha$.
Or $Z \hookrightarrow \mathscr{N}(0;1)$, donc $\sqrt{n\alpha}Z + n\alpha \hookrightarrow \mathscr{N}(n\alpha; n\alpha)$. D'où le résultat.

En pratique...
Lorsque $\lambda \geqslant 15$ **on approche la loi $\mathscr{P}(\lambda)$ par la loi $\mathscr{N}(\lambda, \lambda)$**.
Là encore, l'approximation se fait à espérance et variance constantes.

☞ Rappel...
Soient $a, b \in \mathbb{R}$ avec $a \neq 0$.
Si $X \hookrightarrow \mathscr{N}(\mu; \sigma^2)$, alors
$(aX + b) \hookrightarrow \mathscr{N}(a\mu + b; a^2\sigma^2)$