CHAPITRE 16

Convergences et approximations

Ce calcul délicat s'étend aux questions les plus importantes de la vie, qui ne sont en effet, pour la plupart, que des problèmes de probabilité.

PIERRE-SIMON, MARQUIS DE LAPLACE
Mathématicien, physicien français (1749-1827)

L'objectif de ce chapitre est d'approfondir les notions de convergences vues en première année afin de pouvoir énoncer et illustrer deux des grands théorèmes des probabilités : la loi des grands nombres et le théorème limite central.

1 Inégalités de concentration

PROPOSITION **inégalités de Markov et de Bienaymé-Tchebychev**

- *Soit* Z *une variable aléatoire positive admettant une espérance, alors*

$$\forall \lambda \in \mathbb{R}_*^+, \qquad \mathbf{P}\big([Z \geqslant \lambda]\big) \leqslant \frac{\mathbf{E}(Z)}{\lambda}.$$

- *Soit* X *une variable aléatoire réelle admettant une variance, alors*

$$\forall \varepsilon \in \mathbb{R}_*^+, \qquad \mathbf{P}\big(|X-\mathbf{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

Preuve. Pour tout événement A, on définit l'application *indicatrice de* A, notée $\mathbf{1}_A$, par

$$\forall \omega \in \Omega, \qquad \mathbf{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases}$$

L'application $\mathbf{1}_A$ est une variable aléatoire qui suit une loi de Bernoulli de paramètre $p = \mathbf{P}(A)$. Ainsi

$$\mathbf{P}(A) = \mathbf{E}(\mathbf{1}_A).$$

Par conséquent, pour toute variable aléatoire positive Z, la croissance de l'espérance donne

$$Z \geqslant \lambda \mathbf{1}_{[Z\geqslant\lambda]} \quad \text{puis} \quad \mathbf{E}(Z) \geqslant \lambda \mathbf{E}\big(\mathbf{1}_{[Z\geqslant\lambda]}\big) = \lambda \mathbf{P}\big([Z \geqslant \lambda]\big).$$

On en déduit la première inégalité. La seconde inégalité en découle directement en considérant $Z = \big(X - \mathbf{E}(X)\big)^2$ et $\lambda = \varepsilon^2$.

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

■

Exercice 1

✦ **Preuve dans le cas discret et à densité**

1. Soit Z une variable aléatoire discrète positive admettant une espérance. Notons $Z(\Omega) = \{z_i \mid i \in I\}$. Soit $\lambda \in \mathbb{R}_*^+$. On pose $J = \{i \in I \mid z_i \geqslant \lambda\}$.

a) Vérifier que $\mathbf{P}\big([Z \geqslant \lambda]\big) = \sum\limits_{i \in J} \mathbf{P}\big([Z = z_i]\big)$.

b) En déduire l'inégalité de Markov.

2. Faire la preuve dans le cas d'une variable aléatoire à densité.

p. 29

Exemple. Si $\mathbf{E}(X) = 10$ et $\sigma(X) = 0.1$ sont les valeurs exactes de la moyenne et de l'écart-type d'une variable aléatoire X, sans aucune indication particulière sur la loi de X, on sait que X prendra des valeurs entre 9.7 et 10.3 avec une probabilité supérieure à 88%. En effet,

$$\mathbf{P}\big(|X - 10| \geqslant 0,3\big) \leqslant \frac{(0,1)^2}{(0,3)^2} = \frac{1}{9}, \quad \text{puis,} \quad \mathbf{P}\big(|X - 10| < 0,3\big) = 1 - \mathbf{P}\big(|X - 10| \geqslant 0,3\big) \geqslant 1 - \frac{1}{9} \simeq 88\%.$$

Si on connait la loi de X, on a des estimations beaucoup plus précises.

Exercice 2

✦ En moyenne, une personne sur 100 sait placer dans le bon ordre les pays baltes sur une carte. On choisit au hasard n personnes et de manière indépendante, notons Y_n le pourcentage des personnes capables de donner le bon ordre.

1. Donner l'espérance et la variance de Y_n.

2. Par application de l'inégalité de Bienaymé-Tchebychev, donner une valeur de n à partir de laquelle Y_n se trouve dans l'intervalle

$$I =]0,009; 0,011[$$

avec une probabilité supérieure à 0,9.

p. 29

Remarque. Ces inégalités ont peu d'applications pratiques, car la majoration qu'elles fournissent est la plupart du temps excessive, mais elles sont valables quelle que soit la loi de X, pourvu que l'on puisse définir une espérance ou une variance. L'inégalité de Bienaymé-Tchebychev nous permettra toutefois de démontrer la loi faible des grands nombres (voir théorème page 4).

Exercice 3

✦ Soit X une variable aléatoire possédant une espérance de 6 et une variance de 2. Appliquer, lorsque cela est possible, l'inégalité de Bienaymé-Tchebychev pour majorer ou minorer les probabilités des événements suivants. Préciser si le résultat obtenu est intéressant.

1. $[2 \leqslant X \leqslant 10]$;

2. $[5 < X < 7]$;

3. $[X \leqslant 7]$;

4. $[|X - 6| \geqslant 1]$;

5. $[X \geqslant 11]$;

6. $[X \geqslant 4]$.

p. 30

Exercice 4

✦✦ ✎ Soit X une variable à densité dont une densité f est nulle sur $\mathbb{R}^-$. On suppose qu'il existe $\lambda \in \mathbb{R}_*^+$ tel que $\displaystyle\int_0^{+\infty} f(t)\mathrm{e}^{\lambda t}\,\mathrm{d}t$ soit convergente. Montrer que

$$\forall a \in \mathbb{R}_*^+, \quad \forall x \in]0, \lambda[, \qquad \mathbf{P}(X \geqslant a) \leqslant \mathrm{e}^{-ax}\mathbf{E}\left(\mathrm{e}^{xX}\right).$$

p. ??

2 Convergence en probabilité

2.1 Définition et exemples

Définition — Convergence en probabilité

Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ et X une variable aléatoire définie aussi sur cet espace.
*On dit que la suite $(X_n)_n$ **converge en probabilité** vers la variable aléatoire X, noté $X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X$ si :*

$$\forall\,\varepsilon\in\mathbb{R}_*^+,\qquad \mathbf{P}\big(|X_n - X|\geqslant\varepsilon\big)\underset{n\to\infty}{\longrightarrow} 0..$$

Exemple. Soit (X_n) une suite de variables aléatoires mutuellement indépendantes. On suppose de plus que

$$\forall\,n\in\mathbb{N},\quad X_n \hookrightarrow \mathscr{U}([0;1]).$$

Étudions la convergence en probabilité de la variable $Y_n = \max(X_1,X_2,\ldots,X_n)$. Soit $\varepsilon\in\,]0;1[$.

$$\begin{aligned}
\mathbf{P}\big(|Y_n-1|\geqslant\varepsilon\big) = \mathbf{P}(1-Y_n\geqslant\varepsilon) = \mathbf{P}(Y_n\leqslant 1-\varepsilon) &= \mathbf{P}\left(\bigcap_{i=1}^{n}[X_i\leqslant 1-\varepsilon]\right)\\
&= \prod_{i=1}^{n}\mathbf{P}(X_i\leqslant 1-\varepsilon)\ \text{indépendance des variables aléatoires}\\
&= \big(F(1-\varepsilon)\big)^n \qquad \text{où F est la fonction de répartion de la loi } \mathscr{U}([0;1])\\
\mathbf{P}(|Y_n-1|\geqslant\varepsilon) = (1-\varepsilon)^n & \qquad\qquad \text{car } 1-\varepsilon\in\,]0;1[.
\end{aligned}$$

On constate que $$\mathbf{P}(|Y_n-1|\geqslant\varepsilon)\underset{n\to\infty}{\longrightarrow} 0.$$

Notons que si $\varepsilon\geqslant 1$, la conclusion est identique.
On vient de montrer que $(Y_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire presque surement constante à 1.

Exercice 5

✧ ✎ Pour tout $n\in\mathbb{N}^*$, on pose $Z_n = \min(X_1,X_2,\ldots,X_n)$.
Reprendre le calcul précédent pour justifier la convergence en probabilité de $(Z_n)_n$ vers une variable aléatoire que l'on précisera. p. 30

Remarque. Une suite de variables à densité peut donc converger vers une variable discrète. Nous verrons que la réciproque est aussi possible.

2.2 Les théorèmes de convergence en probabilité

Règles de calcul

⚠ **Attention.** Contrairement au cas des suites réelles ou des fonctions numériques, il n'y a pas unicité de la limite (si elle existe). Plus précisément, si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ et X, X' des variables aléatoires définies sur le même espace et telles que

$$X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X \quad\text{et}\quad X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X' \quad\text{alors}\quad \mathbf{P}\big(X\neq X'\big)=0.$$

Pour les détails de la démonstration, on pourra consulter l'exercice **??**, p.**??**.

Proposition — convergence en probabilité et combinaisons linéaires

Soient $(X_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ deux suites de variables aléatoires sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$.

Si $X_n \xrightarrow[n\to+\infty]{\mathbf{P}} X$ *et* $Y_n \xrightarrow[n\to+\infty]{\mathbf{P}} Y$

Alors *pour tout $\lambda\in\mathbb{R}$,* $\lambda X_n \xrightarrow[n\to+\infty]{\mathbf{P}} \lambda X$ *et* $X_n+Y_n \xrightarrow[n\to+\infty]{\mathbf{P}} X+Y.$

Exercice 6

✦ Prouver cet énoncé. *Pour le second point, on pourra utiliser l'encadrement*

$$\left[|X_n - X| < \frac{\varepsilon}{2}\right] \cap \left[|Y_n - Y| < \frac{\varepsilon}{2}\right] \subset \left[\left|X_n + Y_n - (X+Y)\right| < \varepsilon\right].$$

p. 30

Exemple. Reprenons l'exercice 5 avec $(X_n)_n$ une suite de variables aléatoires mutuellement indépendantes suivant toutes une loi $\mathscr{U}([0;1])$. On a

$$\frac{1}{2}\left(\max\left(X_1, X_2, \ldots, X_n\right) + \min\left(X_1, X_2, \ldots, X_n\right)\right) \xrightarrow[n\to+\infty]{\mathbf{P}} \frac{1}{2}(1+0) = \frac{1}{2}.$$

PROPOSITION — **composition par une fonction continue**

Soit $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires sur un espace probabilisé* $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- *La suite* $(X_n)_{n\in\mathbb{N}}$ *converge en probabilité vers* X.
- *La fonction* f *est continue sur* $\mathbb{R}$ *à valeurs réelles.*

Alors

$$f(X_n) \xrightarrow[n\to+\infty]{\mathbf{P}} f(X).$$

Résultat admis.

Remarque. On peut affiner le théorème en ne supposant seulement que f est continue sur un intervalle I tel que pour tout $n \in \mathbb{N}$, $\mathbf{P}(X_n \in I) = 1$.

Exercice 7

✦✦ **Preuve dans deux cas particuliers**

Soient $(X_n)_n$ une suite de variables aléatoires définies sur $(\Omega, \mathscr{A}, \mathbf{P})$ et $f : \mathbb{R} \to \mathbb{R}$.

1. On suppose qu'il existe $\alpha \in \mathbb{R}_*^+$ tel que : $\quad \forall\, x, y \in \mathbb{R}, \quad |f(x) - f(y)| \leqslant \alpha|x - y|$.
Montrer que $f(X_n) \xrightarrow[n\to+\infty]{\mathbf{P}} f(X)$.

p. 30

2. Justifier que si la fonction f est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ avec les variables X_n à valeurs dans $[a;b]$, alors $f(X_n) \xrightarrow[n\to+\infty]{\mathbf{P}} f(X)$.

Loi faible des grands nombres

THÉORÈME — **loi faible des grands nombres**

Soient $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires définies sur un même espace probabilisé* $(\Omega, \mathscr{A}, \mathbf{P})$ *et* X *une variable aléatoire définie aussi sur cet espace.*

Si
- *La variable* X *admet un moment d'ordre* 2.
- *Les variables* $(X_n)_{n\in\mathbb{N}^*}$ *sont mutuellement indépendantes et de même loi que* X.

Alors *la suite des variables aléatoires* $\overline{X_n}$, *moyenne arithmétique des* n *variables* $X_1, X_2, \cdots, X_n$, *converge en probabilité vers son espérance mathématique* $\mathbf{E}(X)$. *Autrement dit,*

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{\mathbf{P}} \mathbf{E}(X).$$

Preuve. Calculons l'espérance mathématique et la variance de la moyenne arithmétique $\overline{X_n}$. Par linéarité de l'espérance sachant que les X_i ont même loi et sachant que les variables ont même loi et donc même espérance $\mathbf{E}(X)$, il vient

$$\mathbf{E}(\overline{X_n}) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(X_i) = \frac{n\mathbf{E}(X)}{n} = \mathbf{E}(X).$$

Par indépendance des variables X_i,
$$\mathbf{V}(\overline{X_n}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{V}(X_i) = \frac{\mathbf{V}(X)}{n}.$$

Appliquons l'inégalité de Bienaymé-Tchebychev avec la variable $\overline{X_n}$,

$$\mathbf{P}\big(|\overline{X_n} - \mathbf{E}(X)| \geqslant \varepsilon\big) = \mathbf{P}\big(|\overline{X_n} - \mathbf{E}(\overline{X_n})| \geqslant \varepsilon\big) \leqslant \frac{\mathbf{V}(\overline{X_n})}{\varepsilon^2} = \frac{\mathbf{V}(X)}{n\varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0.$$

D'où le résultat. ■

Remarque. Loi faible des grands nombres dans le cas de la loi binomiale
Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ avec

$$\forall\, n \in \mathbb{N}^*, \qquad Y_n \hookrightarrow \mathscr{B}(n; p).$$

Il existe une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ mutuellement indépendantes de même loi (de Bernoulli de paramètre p) telles que pour tout $n \in \mathbb{N}^*$, Y_n et $\sum_{i=1}^{n} X_i$ soient égaux en loi. D'après la loi faible des grands nombres, on obtient la convergence en probabilité :

$$\frac{Y_n}{n} \xrightarrow[n\to+\infty]{\mathbf{P}} Z, \qquad \text{où Z est une variable aléatoire certaine égale à } p.$$

Application. Considérons une expérience aléatoire, et un événement A de probabilité théorique p associé à cette expérience. Soit $n \in \mathbb{N}^*$. Répétons n fois l'expérience de manière indépendante et désignons par X_n le nombre de succès (c'est-à dire le nombre de fois où A est réalisé). Y_n suit donc une loi binomiale de paramètres n, p.
Posons de plus,

$$F_n(A) = \frac{Y_n}{n}, \text{ la fréquence empirique d'apparition de l'événement A.}$$

Par la loi faible des grands nombres, on en déduit l'énoncé suivant :

COROLLAIRE — **interprétation d'une probabilité**

Lorsque le nombre d'expériences aléatoires augmente indéfiniment, la fréquence d'apparition $F_n(A)$ *d'un événement* A *converge en probabilité vers sa probabilité théorique* p*. Autrement dit*

$$\forall\, \varepsilon \in \mathbb{R}_*^+, \qquad \mathbf{P}\Big(\big|F_n(A) - p\big| > \varepsilon\Big) \underset{n\to+\infty}{\longrightarrow} 0.$$

On a une première formulation mathématique de l'interprétation intuitive d'une probabilité d'un événement.

> La probabilité d'un événement est la fréquence que l'on observerait si on effectuait une infinité de fois l'expérience dans « des conditions parfaitement identiques ».

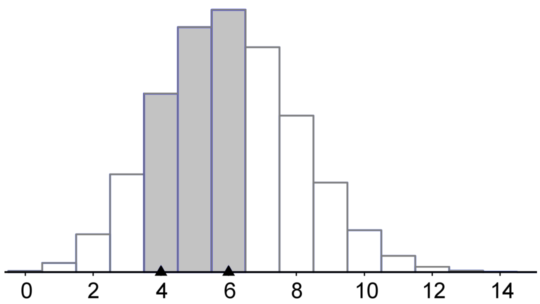
Remarques.
- Il existe de nombreuses versions de la loi des grands nombres. Par exemple, la loi *forte* des grands nombres établit la convergence *presque sûr* de la suite $(X_n)_{n\in\mathbb{N}}$ (voir exercice 43, p.27).
- Nous verrons une nouvelle application avec les méthodes dites de Monte-Carlo (p. 19).

3 Convergence en loi

3.1 Rappels : représentations graphiques des lois

Cas des variables aléatoires discrètes

• Soit X une variable finie avec $X(\Omega) = \{x_1, x_2, \cdots, x_m\}$.
Nous avons vu que l'on peut résumer une loi d'une variable finie par un tableau. Pour chaque indice i, on indique la probabilité $\mathbf{P}(X = x_i)$.
On peut aussi utiliser les diagrammes en bâtons, en abscisse, on place les valeurs x_i. Dans la suite, on s'arrange pour que la hauteur du bâton partant de x_i soit telle que l'aire du rectangle s'identifie à $\mathbf{P}(X = x_i)$.

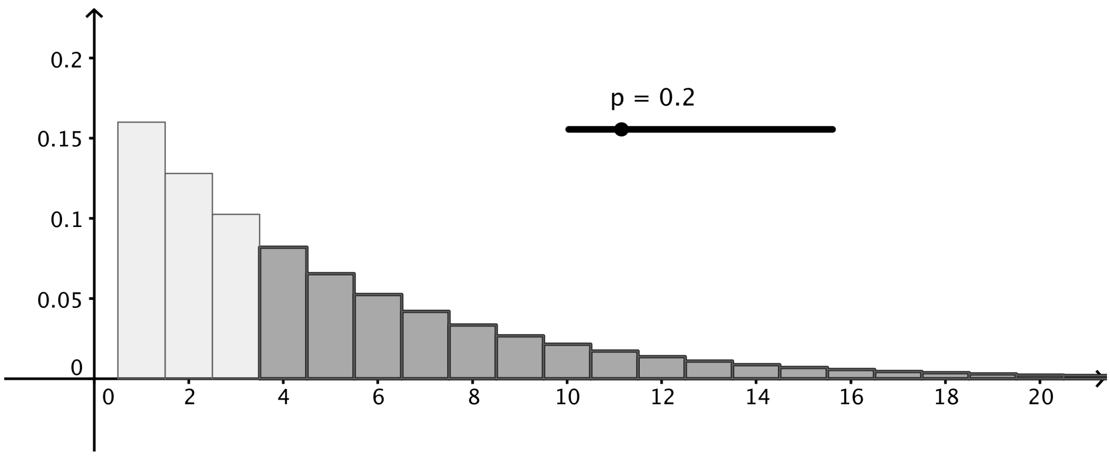


Exemple. Ci-contre, le cas de la loi binomiale de paramètres $n = 20$, $p = 0,3$.
Notons que pour avoir la probabilité $\mathbf{P}\Big([X \in [a;b]]\Big)$, il suffit de sommer les aires des rectangles compris entre les abscisses a et b. En particulier, la somme des aires totales des rectangles est $\mathbf{P}(\Omega) = 1$.
Dans notre exemple, la partie grisée a pour aire

$$\mathbf{P}\Big([X \in [4;6]]\Big).$$

• Lorsque X est une variable aléatoire discrète dénombrable ($X(\Omega) = \{x_1, x_2, \cdots, x_n, \cdots\}$), on ne considère qu'un nombre fini de valeurs $\{x_1, x_2, \cdots, x_n\}$. En géneral, les valeurs où la probabilité n'est pas négligeable.

Donnons l'exemple de la loi géométrique de paramètre $p = 0,2$ où on s'est limité à $[\![0;20]\!]$.



L'aire de la partie la plus grisée correspond à une approximation de $\mathbf{P}([X \geqslant 4])$.

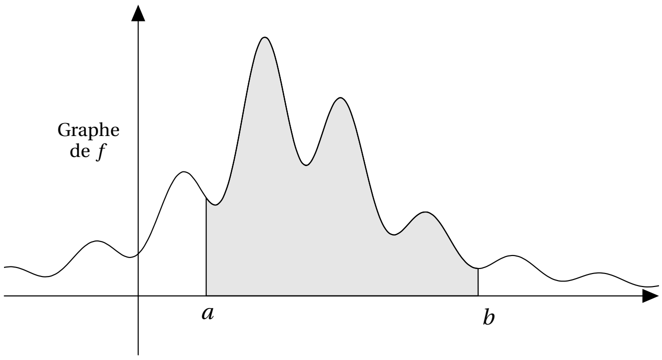
Graphe des densités des variables aléatoires à densité

Soit X une variable aléatoire à densité dont f, est une densité.
L'aire de la partie grisée comprise entre la courbe, l'axe des abscisses et les droites d'équation $x = a$, $x = b$ correspond exactement à la probabilité que X prenne les valeurs entre a et b.

$$\text{Aire} = \mathbf{P}\big([a \leqslant X \leqslant b]\big) = \int_a^b f(t)\,\mathrm{d}t.$$

Lorsque $a \to -\infty$ et $b \to +\infty$, on retrouve bien

$$1 = \mathbf{P}(\Omega) = \int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t.$$



3.2 Définition et exemples

DÉFINITION — convergence en loi

Soient $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires et* X *une variable aléatoire toutes définies sur* $(\Omega, \mathscr{A}, \mathbf{P})$.

Notons
- *pour tout* $n \in \mathbb{N}$, F_n, *la fonction de répartition de la variable* X_n,
- F, *la fonction de répartition de la variable* X.

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ ***converge en loi*** *vers* X, *noté* $X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$, *si en tout point* x *de continuité de* F :

$$F_n(x) \underset{n\to\infty}{\longrightarrow} F(x).$$

⚠ **Attention.** Il n'y a pas unicité de la limite lors d'une convergence en loi. Si X et Y sont deux variables aléatoires de même loi, alors

$$X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X \iff X_n \xrightarrow[n\to+\infty]{\mathscr{L}} Y.$$

Exemples.
- *Exemple 1.*

Pour tout $n \in \mathbb{N}^*$, on vérifie que la fonction f_n définie sur $\mathbb{R}$ par

$$f_n(t) = n^2 t \exp\left(-n^2 t^2/2\right) \mathbf{1}_{\mathbb{R}_+}(x)$$

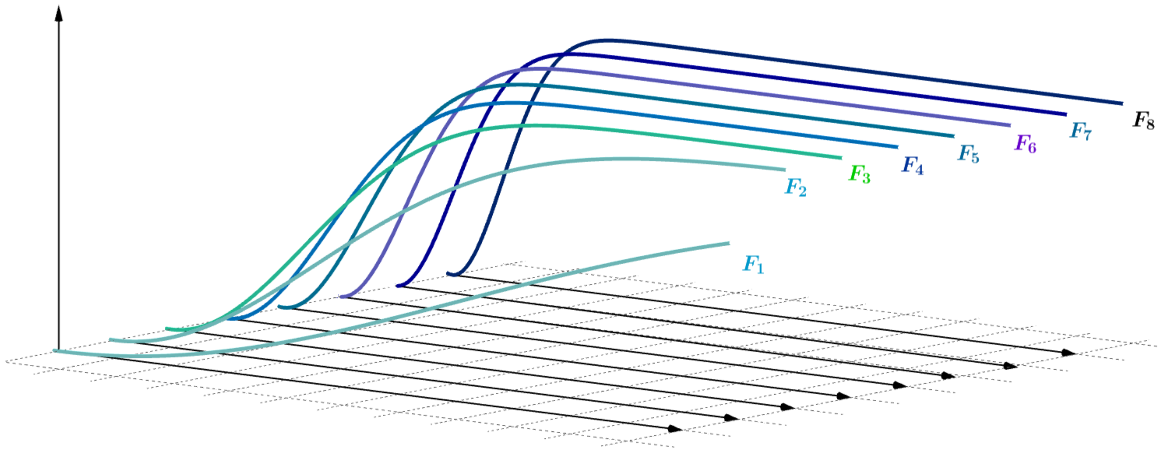
est une densité de probabilité. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires à densité où f_n est une densité de X_n. Dans un premier temps, on montre que la fonction de répartition F_n de X_n est donnée par

$$\forall\, x \in \mathbb{R}^+, \quad F_n(x) = \int_{-\infty}^{x} f_n(t)\,\mathrm{d}t = \int_0^x n^2 t \exp\left(-n^2 t^2/2\right)\mathrm{d}t = \left[-\exp\left(-n^2 t^2/2\right)\right]_0^x = 1 - \exp\left(-n^2 x^2/2\right).$$

Vérifions par le calcul que $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X où X est une variable presque surement constante à 0.
Soit $x \in \mathbb{R}$ **fixé.** Procédons par disjonction des cas :

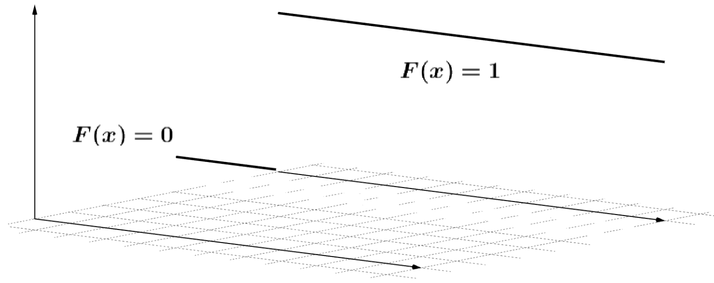
- Si $x \in \mathbb{R}_*^-$, $F_n(x) = 0 \underset{n\to\infty}{\longrightarrow} 0 = F(x)$.
- Si $x \in \mathbb{R}_*^+$, $F_n(x) = 1 - \exp\left(-n^2 x^2/2\right) \underset{n\to\infty}{\longrightarrow} 1 = F(x)$.
- Il n'est pas utile de considérer le cas $x = 0$. C'est un point de discontinuité de F.

Représentons les courbes des premières fonctions de répartition.



On constate que « les courbes convergent vers une courbe limite » qui n'est autre que la courbe de la fonction de répartition d'une loi presque surement constante :

$$\forall\, x \in \mathbb{R}, \quad F(x) = \begin{cases} 1 & \text{si } x \geqslant 0 \\ 0 & \text{sinon.} \end{cases}$$



- *Exemple 2.*

Reprenons le cas de $(X_n)_{n\in\mathbb{N}^*}$, une suite de variables aléatoires mutuellement indépendantes telles que pour tout $n \in \mathbb{N}^*$,

$$X_n \hookrightarrow \mathscr{U}([0;1]).$$

Étudions la convergence en loi des variables $Y_n = \max(X_1, X_2, \ldots, X_n)$.
Tout d'abord, pour tout $x \in \mathbb{R}$, on a

$$F_{Y_n}(x) = G(x)^n \quad \text{où} \quad \text{G est la fonction de répartion de } \mathscr{U}([0;1]).$$

Soit $x \in \mathbb{R}$ **fixé.** Procédons par disjonction des cas.

- Lorsque $x \in \mathbb{R}_*^-$, on a directement $F_{Y_n}(x) = 0 \underset{n\to\infty}{\longrightarrow} 0$.
- Pour $x \in [0;1[$, on a $F_{Y_n}(x) = x^n \underset{n\to\infty}{\longrightarrow} 0$.
- Enfin, pour $x \in [1;+\infty[$, $F_{Y_n}(x) = 1 \underset{n\to\infty}{\longrightarrow} 1$.

En résumé, si F désigne la fonction de répartition d'une variable aléatoire presque surement constante, pour tout point x de continuité de F (ici $x \in \mathbb{R} \setminus \{1\}$), on a

$$F_{Y_n}(x) \underset{n\to\infty}{\longrightarrow} F(x).$$

C'est la définition, de la convergence en loi de la suite $(Y_n)_{n\in\mathbb{N}^*}$ vers une variable presque surement constante en 1.

- L'exercice suivant donne un exemple où la suite converge vers une variable non presque surement constante.

Exercice 8

✦ ✎ Reprenons les notations de l'exemple précédent et posons, pour tout $n \in \mathbb{N}^*$,

$$M_n = n(1 - Y_n).$$

p. 31

Étudier la convergence en loi de la suite $(M_n)_{n\in\mathbb{N}^*}$.

PROPOSITION — convergence en loi

La convergence en loi de la suite $(X_n)_{n\in\mathbb{N}}$ *vers une variable* X *impose pour tous points* a, b *de continuité de* F

$$\mathbf{P}\Big(\big[a < X_n \leqslant b\big]\Big) \underset{n\to\infty}{\longrightarrow} \mathbf{P}\Big(\big[a < X \leqslant b\big]\Big).$$

Preuve. Il suffit de rappeler que pour tout $n \in \mathbb{N}$,

$$\mathbf{P}\Big([X_n \in]a;b]\,]\Big) = F_{X_n}(b) - F_{X_n}(a), \quad \mathbf{P}\Big([X \in]a;b]\,]\Big) = F_X(b) - F_X(a)$$

et, en tous points a, b de continuité de F,

$$F_{X_n}(a) \underset{n\to\infty}{\longrightarrow} F_X(a), \quad F_{X_n}(b) \underset{n\to\infty}{\longrightarrow} F_X(b).$$

■

PROPOSITION — convergence en loi dans le cas discret

Soit $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires telles que :*

$$\forall\, n \in \mathbb{N}, \quad X_n(\Omega) \subset \mathbb{N} \quad \textit{et} \quad \forall\, k \in \mathbb{N}, \quad \mathbf{P}\big([X_n = k]\big) \underset{n\to\infty}{\longrightarrow} p_k \in [0;1].$$

Alors $(X_n)_{n\in\mathbb{N}}$ *converge en loi vers une variable aléatoire* X *avec*

$$X(\Omega) \subset \mathbb{N} \quad \textit{et} \quad \forall\, k \in X(\Omega), \qquad \mathbf{P}\big([X = k]\big) = p_k.$$

Preuve. Soit $x \in \mathbb{R}$ fixé.

- Si $x < 0$ alors pour tout $n \in \mathbb{N}$, $F_n(x) = 0$ et $F(x) = 0$. On a bien $F_n(x) \underset{n\to\infty}{\longrightarrow} F(x)$.
- Si $x \geqslant 0$. On a maintenant

$$F_n(x) = \mathbf{P}(X_n \leqslant x) = \mathbf{P}(X_n \leqslant \lfloor x \rfloor) = \sum_{k=0}^{\lfloor x \rfloor} \mathbf{P}(X_n = k).$$

  La somme est finie, le nombre de termes est indépendants de n. On a

$$F_n(x) \underset{n\to\infty}{\longrightarrow} \sum_{k=0}^{\lfloor x \rfloor} p_k = F(x).$$

La définition est vérifiée, $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X.

■

Remarque. Noter que la réciproque est vraie, c'est une conséquence directe de l'énoncé précédent où pour $k \in \mathbb{N}$, on choisit $[a;b] = [k - 1/2; k + 1/2]$ de sorte que

$$\mathbf{P}\Big(\big[a < X_n \leqslant b\big]\Big) = \mathbf{P}\big([X_n = k]\big).$$

Exemple. Si $X_n \hookrightarrow \mathscr{B}(p_n)$ avec $p_n \underset{n\to\infty}{\longrightarrow} p \in]0;1[$. On montre que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire suivant une loi de Bernoulli $\mathscr{B}(p)$. En effet, seules les valeurs $k \in \{0;1\}$ comptent et

$$\mathbf{P}\big([X_n = 1]\big) = p_n \underset{n\to\infty}{\longrightarrow} p \quad \text{et} \quad \mathbf{P}\big([X_n = 0]\big) = 1 - \mathbf{P}\big([X_n = 1]\big) \underset{n\to\infty}{\longrightarrow} 1 - p.$$

Exercice 9

✧ **Exemples** *Les questions 1 et 2 sont indépendantes*

1. Soient $n \in \mathbb{N}^*$ et X_n une variable aléatoire de loi

$$X_n(\Omega) = \{0;1;2\} \quad \text{et} \quad \mathbf{P}\big([X_n = 0]\big) = \frac{n-\alpha}{3n}, \quad \mathbf{P}\big([X_n = 1]\big) = \frac{n+\cos(n)^2}{3n} \quad \mathbf{P}\big([X_n = 2]\big) = \frac{n+\sin(n)^2}{3n}.$$

p. 31

   a) Déterminer la valeur de α.
   b) Vérifier que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une loi usuelle.

2. Soit $X_n \hookrightarrow \mathscr{P}(\lambda_n)$ avec $\lambda_n \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}_*^+$.
   Montrer que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire dont on précisera la loi.

3.3 Les théorèmes de convergence en loi

Règles de calculs sur les limites

⚠ **Attention.** Contrairement à la convergence en probabilité, la convergence en loi de $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ vers X et Y n'implique pas nécessairement la convergence de la suite $(X_n + Y_n)_{n\in\mathbb{N}}$ vers $X + Y$.

Exercice 10

✦ **Contre-exemple**
Soit $X \in \mathscr{B}(1/2)$. Posons pour tout $n \in \mathbb{N}$, $X_n = X$ et $Y_n = X$.
Vérifier que cela fournit bien un contre-exemple.

p. 32

Remarques. *Un peu de hors-programme*

- On montre que la convergence en probabilité implique la convergence en loi :

$$X_n \underset{n\to+\infty}{\overset{\mathbf{P}}{\longrightarrow}} X \quad \Rightarrow \quad X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X \qquad \text{(voir exercice 44, p.27).}$$

Il est à noter que la réciproque est fausse. Il suffit de reprendre le contre-exemple de l'exercice précédent.

Exercice 11

✧ Soient c un réel et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles convergeant en loi vers une variable aléatoire X. Alors la suite de variables aléatoires $(X_n + c)_{n\in\mathbb{N}}$ converge en loi vers $X + c$.

p. 32

- L'exercice s'étend avec le théorème de Slutsky :

Si $X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X$ et si $(Y_n)_{n\in\mathbb{N}}$ converge en probabilité vers une constante c, alors :

- → $(X_n + Y_n)_{n\in\mathbb{N}}$ converge en loi vers $X + c$
- → $(X_n Y_n)_{n\in\mathbb{N}}$ converge en loi vers cX.

La preuve en exercice (page **??**).

Proposition — **composition et convergence en loi**

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- → *La suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X.*
- → *La fonction f est continue sur $\mathbb{R}$ à valeurs réelles.*

Alors

$$f(X_n) \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} f(X).$$

Résultat admis.

Exercice 12

✦ **Cas particulier**

Justifier l'énoncé précédent dans le cas particulier où la fonction f est bijective croissante.

p. 32

Convergence de la loi binomiale vers la loi de Poisson

THÉORÈME — **convergence de la loi binomiale vers la loi de Poisson**

Soient $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires binomiales* $\mathscr{B}(n;p_n)$ *telles que*

$$np_n \underset{n\to+\infty}{\longrightarrow} \lambda \in \mathbb{R}_*^+.$$

Alors la suite $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *converge en loi vers une variable aléatoire suivant une loi de Poisson* $\mathscr{P}(\lambda)$.

$$\mathrm{X}_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} \mathrm{Z} \quad \textit{avec} \quad \mathrm{Z} \hookrightarrow \mathscr{P}(\lambda).$$

Exercice 13

✦✦ Voici la preuve du théorème. Complétez-la.

Soit $k\in\mathbb{N}$. Pour $n\geqslant k$, on a $\quad \mathbf{P}\big([\mathrm{X}_n=k]\big)=\binom{n}{k}{p_n}^k(1-p_n)^{n-k}=\binom{n}{k}(1-p_n)^n\left(\dfrac{p_n}{1-p_n}\right)^k.$

1. Justifier les équivalents de chacun des facteurs lorsque n tend vers $+\infty$ (k est fixé).

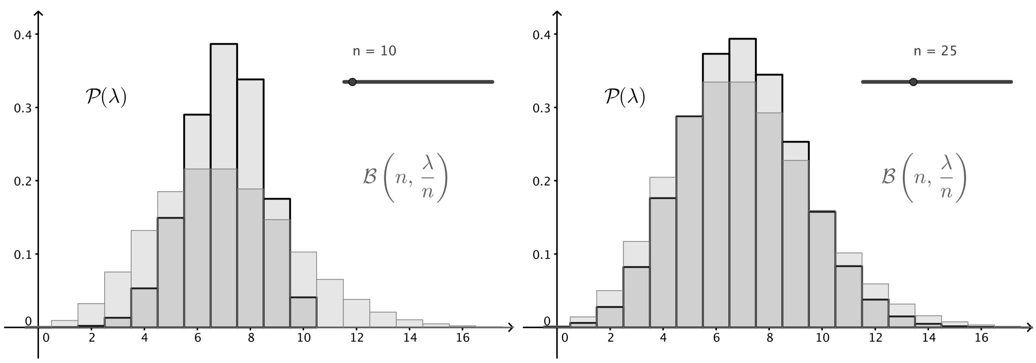
p. 32

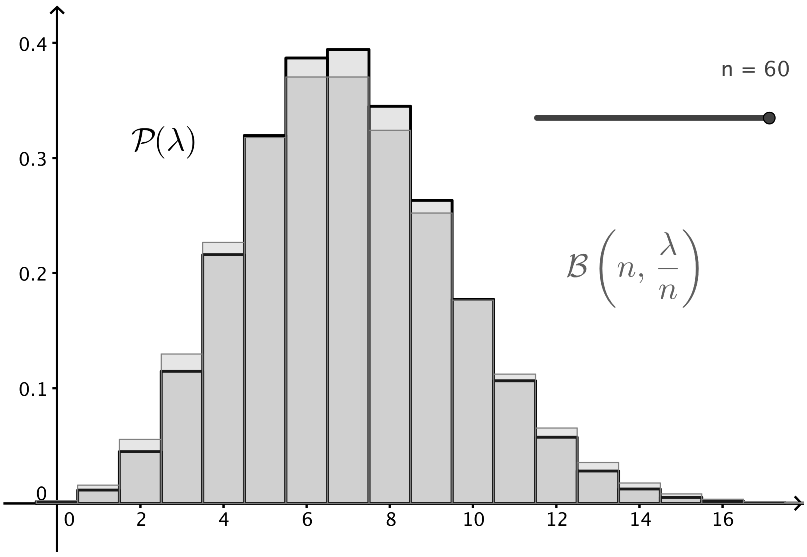
$(a)\quad \left(\dfrac{p_n}{1-p_n}\right)^k \underset{n+\infty}{\sim} \left(\dfrac{\lambda}{n}\right)^k;\quad (b)\quad \ln(1-p_n)\underset{n+\infty}{\sim} -\dfrac{\lambda}{n},\ (1-p_n)^n \underset{n+\infty}{\sim} \mathrm{e}^{-\lambda};\quad (c)\quad \binom{n}{k}\underset{n+\infty}{\sim}\dfrac{n^k}{k!}.$

2. Conclure.

Interprétation graphique

On trace les diagrammes représentant les lois binomiales $\mathscr{B}(n;\lambda/n)$ et de Poisson $\mathscr{P}(\lambda)$. On constate que plus n est grand, plus les diagrammes associés aux lois $\mathscr{P}(\lambda)$ et $\mathscr{B}(n;\lambda/n)$ se confondent.





Application à l'approximation

Considérons $X \hookrightarrow \mathscr{B}(n;p)$, on a pour $k \in [\![0;n]\!]$,

$$\mathbf{P}\big([X=k]\big) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Ce produit peut être assez difficile à évaluer numériquement lorsque n est « très grand » et p « petit ». En effet, on doit faire le produit entre p^k extrêmement petit et $\binom{n}{k}$ potentiellement très grand. Notons de plus, que les coefficients binomiaux ne sont pas si aisés à calculer [1].
L'idée est donc dans les cas limites (n est « très grand » et p « petit ») d'avoir une expression approchée plus simple de la probabilité en posant $\lambda = np$ et

$$\mathbf{P}\big([X=k]\big) \simeq \mathbf{P}\big([Z=k]\big) \quad \text{où} \quad Z \hookrightarrow \mathscr{P}(\lambda).$$

Remarque. Dans la pratique, dès que $n \geqslant 30$, $p \leqslant 0,1$ et $np < 15$, on approche $\mathscr{B}(n,p)$ par $\mathscr{P}(np)$.

Exemple. En moyenne un étudiant d'ECG fait une faute d'orthographe tous les 500 mots.
Lors d'une rédaction de 2000 mots, quelle est la probabilité de faire plus de 5 fautes ?

Modélisons la situation : la probabilité de faire une faute sur un mot est $p = 1/500$. On suppose que les fautes arrivent de manière indépendantes (c'est une hypothèse forte). Lors de la rédaction, on a donc une succession de 2000 expériences de Bernoulli de paramètre p mutuellement indépendantes. Si X est la variable aléatoire qui compte le nombre de fautes alors X suit une loi binomiale de paramètres $n = 2000$ (le nombre de répétitions) et $p = 1/500$ (le paramètre de la loi de Bernoulli). Ainsi, la probabilité recherchée est

$$\mathbf{P}\big([X \geqslant 5]\big) \quad = \quad 1 - \mathbf{P}\big([X<5]\big) = 1 - \sum_{k=0}^{4} \mathbf{P}\big([X=k]\big) = 1 - \sum_{k=0}^{4} \binom{2000}{k} \left(\frac{1}{500}\right)^k \left(\frac{499}{500}\right)^{2000-k}.$$

Cette somme est difficile à évaluer. On peut la calculer avec PYTHON :

1. On pourra penser à la formule $\dfrac{n!}{(n-k)!k!}$. Cette formule est peu pratique à cause de la forte croissance de la factorielle.

Console

```
p=1/500
q=1-p
n=2000
print(1-(q**n+n*p*q**(n-1)+n*(n-1)/2*p**2*q**(n-2)+n*(n-1)*(n-2)/6*p**(3)*q**(n-3)+n*(n-1)
    *(n-2)*(n-3)/24*p**4*q**(n-4)))

>>> # script executed

0.371162999567369
```

Comme $n \gg 1$ et $np = 4$. D'après ce qui précède, on peut approximer par la loi de Poisson de paramètre $\lambda = 4$. Soit $Z \hookrightarrow \mathscr{P}(4)$, alors

$$\mathbf{P}\big([X \geqslant 5]\big) \simeq \mathbf{P}\big([Z \geqslant 5]\big) = 1 - \mathbf{P}\big([Z \leqslant 4]\big) = 1 - e^{-\lambda} \sum_{k=0}^{4} \frac{\lambda^k}{k!}.$$

PYTHON donne

Console

```
lda=4
print(1-np.exp(-lda)*(1+lda+lda**2/2+lda**3/6+lda**4/24))

>>> # script executed
0.3711630648201266
```

L'erreur relative est d'environ 3×10^{-7}%. L'approximation est ici excellente.

Exercice 14. Soit $Z \hookrightarrow \mathscr{P}(\lambda)$.

1. **a)** Justifier que pour tout réel x, $\dfrac{e^x + e^{-x}}{2} = \displaystyle\sum_{k=0}^{+\infty} \frac{x^{2k}}{(2k)!}$.

   b) En déduire $\mathbf{P}\big([Z \text{ est pair}]\big)$.

2. *Application*. Une ligne de transmission entre émetteur et récepteur transporte des données représentées par 10 000 bits (un bit est un élément de $\{0; 1\}$). La probabilité que la transmission d'un bit soit erronée est estimée à 10^{-5} et on admet que les erreurs sont mutuellement indépendantes les unes des autres. On contrôle la qualité de la transmission avec un calcul de parité sur le nombre de « 1 » envoyés :

   - S'il y a un nombre impair d'erreurs, un message d'erreur apparaît.
   - Sinon, c'est-à-dire s'il y a un nombre pair d'erreurs, la transmission est acceptée.

   a) Considérons X la variable aléatoire associant à chaque envoi de données, le nombre d'erreurs lors de la transmission, c'est-à-dire le nombre de bits parmi les 10 000 dont la transmission est erronée. Quelle est la loi de X ?

   b) Calculer la probabilité qu'il n'y ait aucune erreur sachant que la transmission est acceptée.
   On admettra que l'on peut approximer le problème par une loi de Poisson.

Exercice 15

✦ Proposer une méthode pour simuler une loi de Poisson $\mathscr{P}(5)$ uniquement à partir de la commande **rd.random()**.

p. 33

4 Théorème limite central

4.1 Le théorème

THÉORÈME **limite central**

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

Si
- *Les variables $(X_n)_{n\in\mathbb{N}^*}$ sont mutuellement indépendantes.*
- *Les variables $(X_n)_{n\in\mathbb{N}^*}$ ont même loi et admettent une espérance m et une variance $\sigma^2 \neq 0$.*
- *On note $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $\overline{X_n}^* = \frac{\sqrt{n}}{\sigma}(\overline{X_n} - m)$.*

Alors
$$\left(\overline{X_n}^*\right)_{n\in\mathbb{N}^*} \xrightarrow[n\to+\infty]{\mathscr{L}} Z \quad \textit{avec} \quad Z \hookrightarrow \mathscr{N}(0;1).$$

Autrement dit, pour tous $a < b$,
$$\mathbf{P}\left(\left[a \leqslant \overline{X_n}^* \leqslant b\right]\right) \xrightarrow[n\to\infty]{} \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}\,dt,$$

où Φ est la fonction de répartition de la loi normale centrée réduite.

Résultat admis.

L'énoncé du théorème central limite est parfois surprenant et n'a souvent rien à voir avec celui du programme.

Rapport de Jury : Oral HEC 2021

4.2 Cas particuliers

Rappelons que si X est une variable aléatoire admettant une variance σ^2 (et donc une espérance), on définit la variable aléatoire centrée réduite associée à X, notée X^* par

$$\boxed{X^* = \frac{X - \mathbf{E}(X)}{\sigma} \quad \text{avec} \quad \mathbf{E}(X^*) = 0 \quad \text{et} \quad \mathbf{V}(X^*) = 1.}$$

Par exemple, si $X \hookrightarrow \mathscr{B}(n;p)$ alors $X^* = \dfrac{X - np}{\sqrt{np(1-p)}}$.

THÉORÈME **de Moivre-Laplace**

Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires suivant des lois binomiales $\mathscr{B}(n;p)$ avec $p \in]0;1[$, alors

$$X_n^* = \frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to+\infty]{\mathscr{L}} Z \quad \textit{avec} \quad Z \hookrightarrow \mathscr{N}(0;1).$$

Autrement dit, pour tous $a < b$, on a

$$\mathbf{P}\left(\left[a \leqslant X_n^* \leqslant b\right]\right) \xrightarrow[n\to\infty]{} \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}\,dt,$$

où Φ est la fonction de répartition de la loi normale centrée réduite.

Preuve.

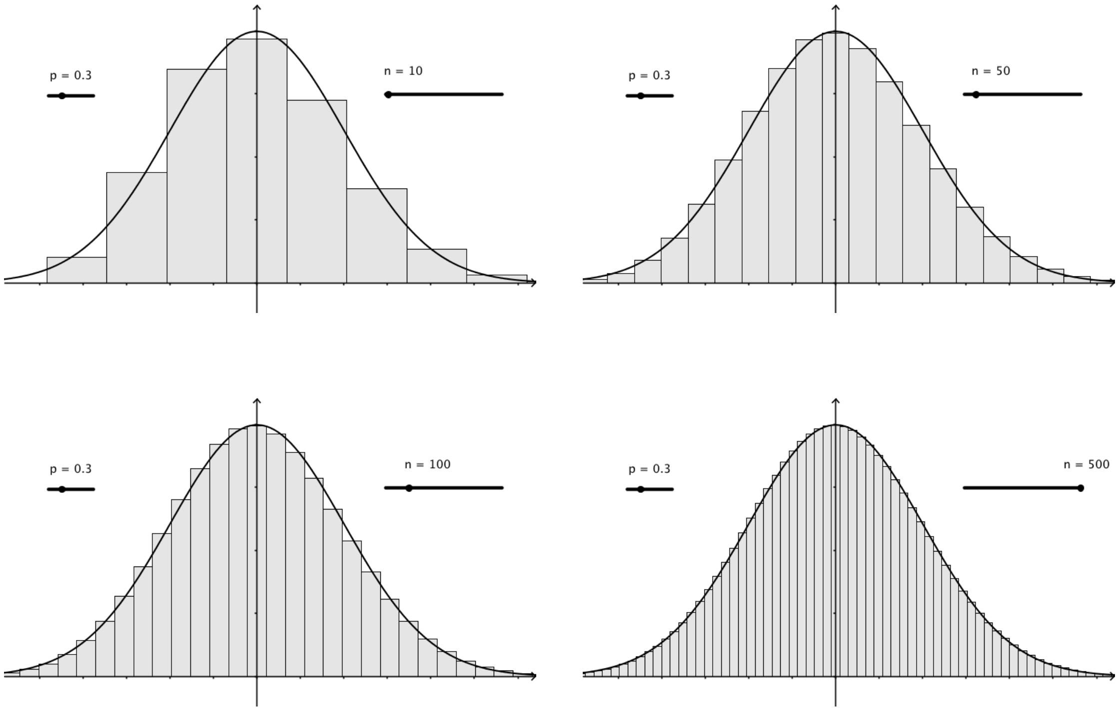
..

■

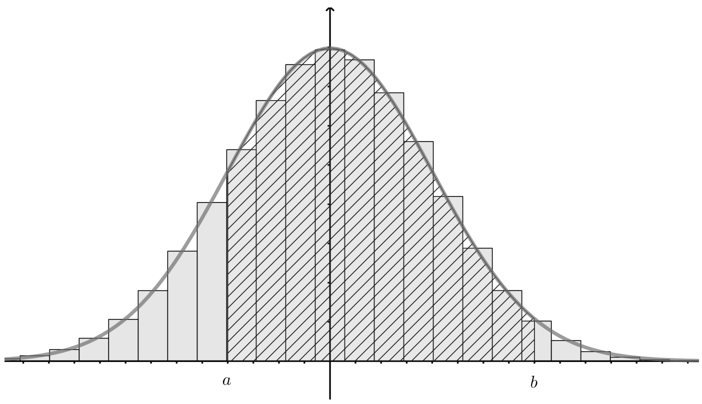
Remarque. Dans la pratique, dès que $n \geqslant 30$, $np \geqslant 5$ et $nq \geqslant 5$, on approche $\mathscr{B}(n,p)$ par $\mathscr{N}(np, npq)$.

Interprétation graphique

On trace les histogrammes associés aux lois de ${X_n}^*$ pour différentes valeurs de n (10, 50, 100 et 500). On superpose la courbe représentative de la densité de loi normale centrée réduite.



On constate que plus n est grand, plus les histogrammes épousent la forme de la courbe.



Interprétons.
Soit $X_n \hookrightarrow \mathscr{B}(n;p)$. De nouveau, plaçons l'histogramme associé à la loi ${X_n}^*$. L'aire hachurée est l'aire sous la courbe représentative de la densité. Elle vaut donc

$$\int_a^b f(t)\,\mathrm{d}t = \mathbf{P}\big([a \leqslant Z \leqslant b]\big),$$

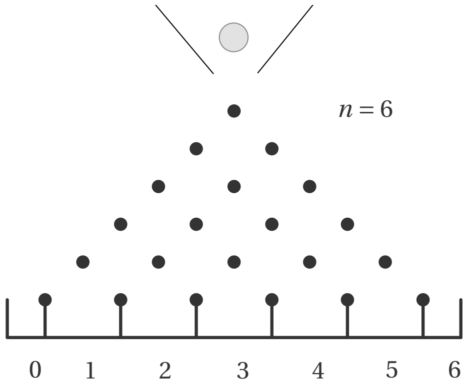
où $\quad f : t \mapsto \dfrac{1}{\sqrt{2\pi}} \exp(-t^2) \quad$ et $\quad Z \hookrightarrow \mathcal{N}(0;1)$.

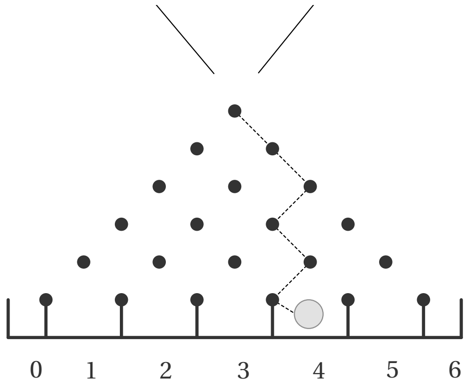
Cette aire s'identifie approximativement à l'aire des rectangles compris entre les droites $x = a$ et $x = b$. Or, cette dernière est par construction $\mathbf{P}\big([a \leqslant {X_n}^* \leqslant b]\big)$. On en déduit l'approximation :

$$\mathbf{P}\Big(\big[a \leqslant X_n^* \leqslant b\big]\Big) \simeq \mathbf{P}\Big(\big[a \leqslant Z \leqslant b\big]\Big).$$

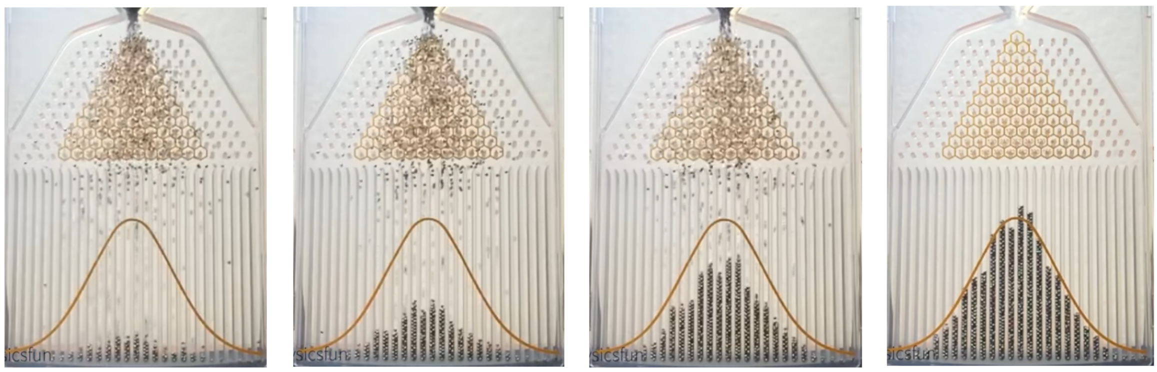
Exemple. *La planche de Galton.*
Des clous sont plantés dans une planche de manière régulière. On place la planche à la verticale de sorte qu'une bille

lachée au sommet percute les clous sous l'effet de la gravité.





On modélise la situation en supposant que lors du choc avec le clou, la bille a une probabilité de $1/2$ d'aller à gauche et une probabilité de $1/2$ d'aller à droite. On suppose que les chocs sont mutuellement indépendants. Soit n le nombre de lignes de clous. Si on numérote les cases de 0 à n, notons X la variable aléatoire, qui a une bille associe la case de chute.
Alors X compte le nombre de fois où la bille va à droite, X suit une loi binomiale $\mathscr{B}(n;1/2)$.
On peut illustrer physiquement la situation :



Lien vers la vidéo
Planche de Galton

Attention. Il y a deux théorèmes de convergence impliquant des lois binomiales. Précisons la différence :

- Dans le cas de convergence vers une loi de Poisson : $n \to +\infty$ mais $np_n \underset{n\to\infty}{\longrightarrow} \lambda > 0$, sous-entendu, $p_n \underset{n\to\infty}{\longrightarrow} 0$.
- Dans le cas de convergence vers une loi normale : $n \to +\infty$ mais p correspond à une probabilité fixée strictement positive.

THÉORÈME **convergence des lois de Poisson**

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telle que $X_n \hookrightarrow \mathscr{P}(n\lambda)$ pour tout $n \in \mathbb{N}^*$.
Alors la suite des variables aléatoires centrées réduites $({X_n}^*)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire suivant une loi normale centrée réduite.

$$ {X_n}^* = \frac{X_n - n\lambda}{\sqrt{n\lambda}} \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} Z \quad \textit{avec} \quad Z \hookrightarrow \mathcal{N}(0;1). $$

Remarque. Dans la pratique, dès que $\lambda \geqslant 18$, on approche la loi $\mathscr{P}(\lambda)$ par la loi normale $\mathcal{N}(\lambda,\lambda)$.

Le déterminisme laplacien **Le saviez-vous**

Pierre-Simon de Laplace (1749-1827) est un mathématicien, astronome et physicien français. Il est l'un des principaux scientifiques de la période napoléonienne; il a apporté des contributions fondamentales dans divers domaines allant de la mécanique céleste à la théorie des probabilités (dont le théorème précédent). Philosophiquement, il fut l'un des membres les plus influents du déterminisme.

« Une intelligence qui, pour un instant donné, connaitrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux. L'esprit humain offre, dans la perfection qu'il a su donner à l'astronomie, une faible esquisse de cette intelligence. »

C'est justement la faiblesse de l'esprit humain qui rend l'utilisation des probabilités inévitables. Si tout phénomène a une cause, on parle de hasard lorsqu'on n'est pas en mesure d'expliciter les causes et les enchaînements précisément. **Le hasard n'est alors « qu'une mesure de notre ignorance »**.

Tous les événements, ceux même qui par leur petitesse, semblent ne pas tenir aux grandes lois de la nature, en sont une suite aussi nécessaire que les révolutions du soleil. Dans l'ignorance des liens qui les unissent au système entier de l'univers, on les a fait dépendre des causes finales, ou du hasard, suivant qu'ils arrivaient et se succédaient avec régularité, ou sans ordre apparent; mais ces causes imaginaires ont été successivement reculées avec les bornes de nos connaissances, et disparaissent entièrement devant la saine philosophie, qui ne voit en elles, que l'expression de l'ignorance où nous sommes de véritables causes.

Essai philosophique sur les probabilités, LAPLACE (1814).

4.3 Applications à l'approximation

Méthode

Comment approximer une probabilité à l'aide d'une loi normale?

On lance une pièce équilibrée 10000 fois et on souhaite calculer la probabilité que le nombre de « PILE » soit compris dans l'intervalle $[4900;5100]$.
On suppose les lancers mutuellement indépendants.
Ainsi si X est la variable aléatoire qui compte le nombre de « PILE », X suit une loi binomiale de paramètres $n=10000$, $p=\frac{1}{2}$. L'espérance de X est $np=5000$, l'écart type est $\sqrt{np(1-p)}=50$. Évaluons $\mathbf{P}\big([4900\leqslant X\leqslant 5100]\big)$.

- La première étape consiste à renormaliser en introduisant X^* :

$$\begin{aligned}\mathbf{P}\big([4900\leqslant X\leqslant 5100]\big) &= \mathbf{P}\big([np-100\leqslant X\leqslant np+100]\big) = \mathbf{P}\big([-100\leqslant X-np\leqslant 100]\big)\\ &= \mathbf{P}\left(\left[-2\leqslant \frac{X-np}{\sqrt{np(1-p)}}\leqslant 2\right]\right) = \mathbf{P}\big([-2\leqslant X^*\leqslant 2]\big).\end{aligned}$$

- Puis, on applique le théorème précédent.

$$\mathbf{P}\left(\left[-2\leqslant X^*\leqslant 2\right]\right)\simeq \mathbf{P}\big([-2\leqslant Z\leqslant 2]\big)\quad\text{avec}\quad Z\hookrightarrow\mathcal{N}(0;1).$$

Notons Φ la fonction de répartition de la loi normale centrée réduite.

$$\mathbf{P}\big([4900\leqslant X\leqslant 5000]\big)\simeq \Phi(2)-\Phi(-2)=2\Phi(2)-1.$$

À l'aide de Python ou de table de la loi normale, on sait que $\Phi(2)\simeq 0,9772$.

Conclusion : la probabilité recherchée vaut environ $0,9544$.

Exemple. Surbooking.
Afin d'augmenter le nombre de personnes transportées, une compagnie aérienne vend plus de billets qu'elle n'a de places en pariant sur les absences de certains de ses passagers.
Considérons un vol d'un appareil contenant $\ell=400$ places. On suppose que $q=8\%$ des passagers ne pourront être à l'heure pour l'embarquement. La compagnie vend $n=420$ billets. On cherche à calculer le risque de surbooking, c'est-à-dire la probabilité qu'il y ait strictement plus de passagers présents à l'embarquement que de places disponibles.

On pose X le nombre de personnes présentes à l'embarquement. Si on suppose la présence des passagers indépendante les unes des autres (c'est une hypothèse forte, il peut avoir des familles, un problème d'accès à l'aéroport), alors X suit une loi binomiale de paramètres n,q avec $p=1-q=92\%$. Précisons que $\mathbf{E}(X)=nq=386.4$ et l'écart-type est $\sigma_X=\sqrt{npq}\simeq 5.56$.
Il y a surbooking si l'événement $[X>\ell]$ est réalisé.

Étape I. On introduit la variable X^* :

$$\mathbf{P}\big([X>\ell]\big)=\mathbf{P}\big([X-\mathbf{E}(X)>\ell-np]\big)=\mathbf{P}\left(\left[\frac{X-\mathbf{E}(X)}{\sigma_X}>\frac{\ell-np}{\sqrt{npq}}\right]\right)\quad\text{avec}\quad \frac{\ell-np}{\sqrt{npq}}\simeq 2,45.$$

Étape II. On applique le théorème limite central ($n>>1$). Si $Z\hookrightarrow\mathcal{N}(0;1)$, on a

$$\mathbf{P}\big([X>\ell]\big)=\mathbf{P}\left(\left[X^*>\frac{\ell-np}{\sqrt{npq}}\right]\right)\simeq \mathbf{P}\big([Z>2,45]\big)=1-\mathbf{P}\big([Z\leqslant 2,45]\big)=1-\Phi(2,45).$$

En utilisant les tables ou sa calculatrice, il vient $\mathbf{P}\big([X>\ell]\big)\simeq 0.007$. Il y a moins de 1% de chance de rencontrer du surbooking sur ce vol.

5 Applications avec Python

5.1 Simulation d'une loi normale par la méthode des 12 uniformes

D'après le théorème limite central, si $(X_n)_n$ est une suite de variables aléatoires mutuellement indépendantes suivant une loi uniforme sur $[0;1]$, la variable aléatoire $\overline{X_n}^*$ centrée réduite associée à la moyenne arithmétique des variables $X_1, \ldots, X_n$ converge en loi vers la loi normale $\mathcal{N}(0,1)$.

Exercice 16

✦✦

1. En se limitant à douze variables $X_1, \ldots, X_{12}$ suivant des lois uniformes, écrire une fonction Python qui renvoie une simulation de la loi normale centrée réduite.
2. En déduire une seconde fonction qui prend en arguments μ, σ, m et renvoient m simulations d'une loi $\mathcal{N}\left(\mu, \sigma^2\right)$. p. 34
3. Tester votre programme en superposant sur une même figure l'histogramme de 2000 simulations de loi $\mathcal{N}(1,4)$ et la densité de cette loi.

5.2 La méthode de Monte-Carlo pour l'approximation d'intégrales

- **Principe**

  Considérons :

— $g : [0,1] \to [a;b]$ une fonction continue dont on souhaite calculer l'intégrale $\displaystyle\int_0^1 g(t)\,\mathrm{d}t$.

— $(U_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0,1]$.

— Pour tout $i \in \mathbb{N}$, on pose $g(U_i)$. D'après les lemme de l'égalité en loi et le lemme des coalitions, les variables sont indépendantes et de même loi.

Par le théorème de transfert, les variables $X_i = f(U_i)$ admettent toutes une même espérance donnée par

$$\mathbf{E}(X_i) = \int_0^1 g(t)\,\mathrm{d}t.$$

De même, ces variables admettent une variance et la loi faible des grands nombres donne

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{\mathbf{P}} \mathbf{E}(X_1) = \int_0^1 g(t)\,\mathrm{d}t.$$

Dès lors, pour calculer une valeur approchée de l'intégrale, on peut simuler un grand nombre de fois les variables U_i, calculer les images $f(U_i)$ et en faire leur moyenne arithmétique.

Exercice 17

✦ **La théorie : estimation de la probabilité de l'erreur**

Soit X, une variable aléatoire à valeurs dans $[a;b]$.

1. Pour quelle valeur de m, $\mathbf{E}\left((X-m)^2\right)$ atteint son minimum ? Avec $m = \frac{a+b}{2}$, déduire :

$$\mathbf{V}(X) \leqslant \frac{(b-a)^2}{4}.$$

p. 34

2. En reprenant les notations du début, justifier que

$$\forall \varepsilon \in \mathbb{R}_*^+, \quad \mathbf{P}\left(\left|\int_0^1 g(t)\mathrm{d}t - \frac{1}{n}\sum_{k=1}^n g(U_k)\right| > \varepsilon\right) \leqslant \frac{(b-a)^2}{4n\varepsilon^2} \qquad (\bullet)$$

✦ La pratique

Exercice 18

On pose
$$I = \int_0^1 \frac{4}{1+t^2}\,\mathrm{d}t.$$

p. 34

1. Donner la valeur explicite de I. En déduire, une fonction Python qui prend en argument n qui permet d'approcher I.
2. Déterminer n afin d'obtenir une valeur à 10^{-3} près de I avec une probabilité d'au moins 95%. Commenter.

Exercices

Révisions

Exercice 19. ✦ Soit X une variable aléatoire admettant une variance. Montrer que $x\mathbf{P}(X > x) \underset{x\to+\infty}{\longrightarrow} 0$.

≫ Solution p. 34

Exercice 20. ✦ **Estimation**

Une urne contient une proportion p de boules blanches. On souhaite obtenir expérimentalement une approximation de p. Pour cela, on effectue $n \in \mathbb{N}^*$ tirages avec remise et on note X_n le nombres de boules blanches obtenues au cours de ces n tirages. On suppose les tirages mutuellement indépendants.

1. Donner la loi de X_n. Préciser l'espérance et la variance.
2. Justifier que pour tout réel $\varepsilon > 0$, $\quad \mathbf{P}\left(\left|\dfrac{X_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \dfrac{1}{4n\varepsilon^2}$.
3. Combien de tirages faut-il effectuer pour pouvoir affirmer, avec un risque inférieur à 5%, que la fréquence d'obtention de boules blanches au cours des n tirages diffère de p d'au plus 1% ?

≫ Solution p. 34

Exercice 21. ✦✦ **Les souris mutantes**

Un laboratoire élève des souris dont 1/4 sont mutantes. La durée de vie d'une souris mutante est une variable aléatoire dont la moyenne est de 3 ans avec un écart-type de 9 mois, mais elle ne vit jamais plus de 4 ans. La durée de vie d'une souris normale a une moyenne d'un an, avec un écart-type de 6 mois. On ne prend en compte que les souris dont la durée de vie est strictement positive.

Une souris est vivante au bout de deux ans. On note α la probabilité qu'elle soit mutante.

On considère l'événement M : « La souris est une souris mutante »et on note X la variable aléatoire égale à la durée de vie de la souris.

1. Exprimer $\dfrac{\mathbf{P}(M \cap [X \geqslant 2])}{\mathbf{P}(\overline{M} \cap [X \geqslant 2])}$ en fonction de α.
2. En utilisant l'inégalité de Bienaymé-Tchebychev, donner une minoration de α.

≫ Solution p. 35

Exercice 22. ✦✦ ✎ **Inégalité de Chernov**

1. Soit $t \in \mathbb{R}_+^*$. Soit X une variable aléatoire discrète telle que e^{tX} admette une espérance. Montrer, à l'aide de l'inégalité de Markov, que pour tout $a \in \mathbb{R}$,
$$\mathbf{P}(X \geqslant a) \leqslant \frac{\mathbf{E}\left(e^{tX}\right)}{e^{ta}}.$$
2. Soient $n \in \mathbb{N}^*$ et $p \in]0,1[$. On suppose que $X \hookrightarrow \mathscr{B}(n,p)$.
   a) Montrer que pour tout $t \in \mathbb{R}_+^*$, e^{tX} admet une espérance et que :
   $$\mathbf{E}\left(e^{t\mathbf{X}}\right) = \left(1 - p + pe^t\right)^n.$$
   b) Étudier les variations de la fonction $f : t \mapsto (1-p)e^{-\frac{t}{2}} + pe^{\frac{t}{2}}$ sur $\mathbb{R}_+^*$. En déduire que f admet un minimum sur $\mathbb{R}_+^*$, égal à $2\sqrt{p(1-p)}$.
   c) À l'aide de la question 1, montrer que $\mathbf{P}\left(X \geqslant \frac{n}{2}\right) \leqslant 2^n\left(p(1-p)\right)^{\frac{n}{2}}$.

≫ Solution p. 35

Exercice 23. ✦✦ ✎ **Comparaison entre la médiane et l'espérance**

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On suppose que X admet une variance.

1. Soient $\alpha \in \mathbb{R}_*^+$, $\beta \in \mathbb{R}^+$. Démontrer que $\quad \mathbf{P}\left([X \geqslant \mathbf{E}(X) + \alpha]\right) \leqslant \dfrac{\mathbf{E}\left(\left(X - \mathbf{E}(X) + \beta\right)^2\right)}{(\alpha+\beta)^2}$.
2. Avec $\beta = \mathbf{V}(X)/\alpha$, en déduire que $\quad \mathbf{P}\left([X \geqslant \mathbf{E}(X) + \alpha]\right) \leqslant \dfrac{\mathbf{V}(X)}{\mathbf{V}(X) + \alpha^2}$.

3. On suppose dans cette question que X est une variable aléatoire à densité avec une densité strictement positive.

a) Justifier qu'il existe un unique réel m tel que $\mathbf{P}([X \leqslant m]) = \frac{1}{2}$.
Un tel réel m est la médiane de la variable X.

b) À l'aide de la question 2 pour un réel α bien choisi, justifier que $\mathbf{E}(X) + \sigma(X) \geqslant m$ où $\sigma(X)$ désigne l'écart-type de la variable X.

c) En considérant aussi la variable $-X$, conclure en montrant que $|m - \mathbf{E}(X)| \leqslant \sigma(X)$.

≫ Solution p. 36

Exercice 24. ✦✦✦ **Inégalité de Cantelli**
Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, centrée et admettant une variance σ^2.

1. Montrer que pour toute variable aléatoire Y définie sur $(\Omega, \mathscr{A}, \mathbf{P})$ et admettant un moment d'ordre 2 , on a :

$$\mathbf{E}(Y) \leqslant \sqrt{\mathbf{E}(Y^2)\mathbf{P}(Y>0)}$$

2. En déduire que :

$$\forall \varepsilon > 0, \qquad \mathbf{P}(X > \varepsilon) \leqslant \frac{\sigma^2}{\sigma^2 + \varepsilon^2}.$$

≫ Solution p. 37

Convergences en probabilité et en loi

Exercice 25. ✦✦ Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la même loi de Bernoulli de paramètre $p \in]0,1[$. Pour tout $n \in \mathbb{N}^*$, on pose :

$$Y_n = X_n + X_{n+1} \quad \text{et} \quad T_n = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

1. Calculer l'espérance et la variance de T_n.

2. Peut-on appliquer la loi des grands nombres pour étudier la convergence en probabilité de la suite $(T_n)_{n\in\mathbb{N}^*}$?

3. Justifier que $(T_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire presque sûrement constante.

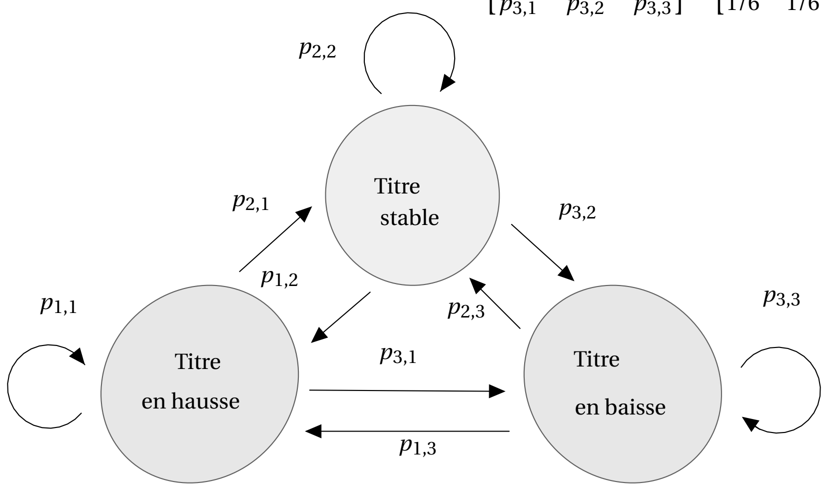
≫ Solution p. 37

Exercice 26. ✦ ✎ **Chaîne de markov : évolution d'un titre boursier**
Dans une bourse de valeurs, un titre peut monter, descendre ou rester stable. On modélise l'évolution du titre.

- Si un jour n, le titre monte, le jour suivant, il montera avec la probabilité 2/3, restera stable avec la probabilité 1/6, et baissera avec la probabilité 1/6.
- Si un jour n, le titre est stable, le jour $n+1$, il montera avec la probabilité 1/6, restera stable avec la probabilité 2/3, et baissera avec la probabilité 1/6.
- Si un jour n, le titre baisse, le jour $n+1$, il montera avec la probabilité 1/6, restera stable la probabilité 1/6, et baissera avec la probabilité 2/3.

Le premier jour, le titre est stable.

Les probabilités sont spécifiées par une matrice dite de transition : $M = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{bmatrix} = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/6 & 2/3 & 1/6 \\ 1/6 & 1/6 & 2/3 \end{bmatrix}$.



On souhaite connaître l'évolution de ce titre. Pour cela, on introduit pour tout $n \in \mathbb{N}^*$, la variable aléatoire aléatoire X_n définie par

$$X_n = \begin{cases} 1 & \text{si le titre donné monte le jour } n \\ 0 & \text{si le titre est stable le jour } n \\ -1 & \text{si le titre donné baisse le jour } n. \end{cases} \quad \text{et} \quad U_n = \begin{bmatrix} \mathbf{P}(X_n = 1) \\ \mathbf{P}(X_n = 0) \\ \mathbf{P}(X_n = -1) \end{bmatrix}.$$

1. **a)** Vérifier que $U_{n+1} = MU_n$.

b) En déduire U_n en fonction de M et U_1.

2. Donner la loi de X_n.

3. Justifier que $(X_n)_n$ converge en loi vers une variable aléatoire X.

4. Comparer MU et U où $U = \begin{bmatrix} \mathbf{P}(X = 1) \\ \mathbf{P}(X = 0) \\ \mathbf{P}(X = -1) \end{bmatrix}$. Commenter.

≫ Solution p. 38

Exercice 27. ✧ Pour $n \in \mathbb{N}^*$, on considère une variable X_n dont la loi est donnée par :

$$X_n(\Omega) = \{0, n\}, \quad \mathbf{P}(X_n = 0) = 1 - \frac{1}{n} \quad \text{et} \quad \mathbf{P}(X_n = n) = \frac{1}{n}.$$

Étudier la convergence en loi de la suite $(X_n)_{n \in \mathbb{N}^*}$ et la convergence de la suite numérique $\big(\mathbf{E}(X_n)\big)_{n \in \mathbb{N}^*}$. Commenter.

≫ Solution p. 39

Exercice 28. ✦ Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires telles que pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{N}(1; 1/n)$.

1. Étudier la convergence en loi de la suite $(X_n)_{n \in \mathbb{N}^*}$.

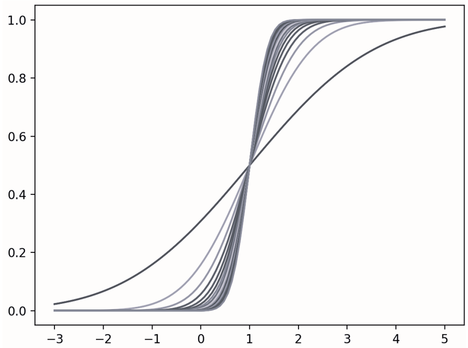
2. Expliquer et commenter le programme Python suivant.
La commande **sp.ndtr(x)** *renvoie* $\Phi(x)$ *où* Φ *est la fonction de répartition de la loi normale centrée réduite.*

≫ Solution p. 39

Editeur

```
import scipy.special as sp
import numpy as np

def Fnormal(x,k):
    return sp.ndtr(k**(1/2)*(x-1)/2)
x=np.linspace(-3,5,100)
for k in range(1,50,3) :
    y=Fnormal(x, k)
    plt.plot(x, y)
plt.show()
```



Exercice 29. ✦✦ ✎ Pour tout $n \in \mathbb{N}$, on considère X_n une variable aléatoire suivant une loi normale $\mathcal{N}(0, {\sigma_n}^2)$ et X qui suit une loi normale $\mathcal{N}(0, \sigma^2)$ avec σ_n, σ strictement positifs. On suppose de plus que les variables $(X_n)_{n \in \mathbb{N}}$ sont mutuellement indépendantes. Montrer l'équivalence entre :

- *i)* La suite de variable aléatoire $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X.
- *ii)* La suite de réels $(\sigma_n)_{n \in \mathbb{N}}$ converge vers σ.

≫ Solution p. 39

Exercice 30. ✦ Soit $(U_n)_{n \in \mathbb{N}^*}$ une suite de variables indépendantes et de même loi uniforme sur $[0, 1]$. Soit $Y \hookrightarrow \mathcal{E}(1)$. On pose $Z_n = \min(U_1, \ldots, U_n)$.

1. Montrer que la suite (nZ_n) converge en loi vers Y.

2. Soit $X \hookrightarrow \mathcal{E}(1)$. Déterminer la loi de $Z = e^{-X}$.

3. On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoire indépendantes suivant toutes la même loi exponentielle de paramètre $\lambda > 0$.
Déterminer la limite en loi de la suite (nT_n) où $T_n = \min\left(e^{-\lambda X_1}, \ldots, e^{-\lambda X_n}\right)$.

≫ Solution p. ??

Exercice 31. ✦✦ **Convergence en loi avec des lois de Cauchy**
Pour tout $n \in \mathbb{N}^*$, on pose

$$\forall\, t \in \mathbb{R}, \quad f_n(t) = \frac{n}{\pi\left(1 + n^2 t^2\right)}.$$

1. Justifier que f_n est une densité de probabilité. Soit X_n une variable aléatoire dont f_n est une densité.
2. Peut-on appliquer l'inégalité de Markov à X_n ?
3. Donner la fonction de répartition de X_n. En déduire la convergence en loi de la suite de variable aléatoire $(X_n)_n$.

≫ Solution p. 39

Exercice 32. ✦✦ ✎ **Variante de la loi faible des grands nombres**
Soit $\left(p_n\right)_{n\in\mathbb{N}^*}$ une suite de réels appartenant à $[0,1]$ et $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes. On suppose que pour tout $k\in\mathbb{N}^*, X_k$ suit une loi de Bernoulli de paramètre p_k. On pose pour tout $n\in\mathbb{N}^*$,

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{et} \quad m_n = \frac{1}{n}\sum_{k=1}^{n} p_k.$$

1. a) Montrer que pour tout $k\in\mathbb{N}^*$: $\mathbf{V}\left(X_k\right)\leqslant \frac{1}{4}$. En déduire, pour tout $n\in\mathbb{N}^*$, une majoration de $\mathbf{V}(Y_n)$. On admet que la variance d'une somme de variables de Bernoulli indépendantes est la somme des variances.
   b) En déduire, à l'aide de l'inégalité de Bienaymé-Tchebychev, que, pour tout $\varepsilon > 0$,

$$\mathbf{P}\left(|Y_n - m_n| < \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 1.$$

2. On suppose que la suite $(m_n)_{n\in\mathbb{N}^*}$ converge vers m.
   a) Soit $\varepsilon > 0$. On suppose $|m_n - m| < \frac{\varepsilon}{2}$. Comparer les événements $\left[|Y_n - m_n| < \frac{\varepsilon}{2}\right]$ et $[|Y_n - m| < \varepsilon]$. En déduire que

$$\mathbf{P}\left(|Y_n - m_n| < \frac{\varepsilon}{2}\right) \leqslant \mathbf{P}\left(|Y_n - m| < \varepsilon\right).$$

   b) En déduire la convergence en probabilité de la suite $(Y_n)_{n\in\mathbb{N}^*}$.

≫ Solution p. 40

Exercice 33. ✦✦ ✎ **Convergence de loi discrète vers une loi à densité**
Pour tout $n\in\mathbb{N}^*$, on considère la variable aléatoire discrète $X_n \hookrightarrow \mathscr{U}([\![1;n]\!])$. On pose $Y_n = X_n/n$.
Justifier que la suite $(Y_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire que l'on précisera.
Indication. On pourra utiliser l'expression de la fonction de répartition de X_n,

$$\forall\, x\in[0;n], \qquad F_n(x) = \frac{\lfloor x\rfloor}{n}.$$

≫ Solution p. 41

Exercice 34. ✦ Soit $\lambda\in\mathbb{N}^*$. Pour tout entier $n > \lambda$, on considère une variable aléatoire X_n suivant une loi géométrique de paramètre $\frac{\lambda}{n}$. Étudier la convergence en loi de la suite (Y_n) définie par

$$Y_n = \frac{X_n}{n}.$$

≫ Solution p. ??

Exercice 35. ✦✦ ✎ **Autour des lois de Cauchy**

- *La fonction arctangente*

1. Rappeler la définition de la fonction arctangente. Donner son graphe avec l'équation de la tangente en 0.
2. Vérifier que pour tout $x\in\mathbb{R}_*^+$, $\arctan(x)+\arctan(1/x) = \pi/2$. Que dire de cette expression si $x\in\mathbb{R}_*^-$?
3. Justifier le développement suivant lorsque $x\to+\infty$: $\qquad \arctan(x) = \dfrac{\pi}{2} - \dfrac{1}{x} + o_{+\infty}\left(\dfrac{1}{x}\right).$

- *Loi de Cauchy*

4. Soit $a\in\mathbb{R}_*^+$. On définit sur $\mathbb{R}$ la fonction f_a par : $f_a(x) = \dfrac{a}{\pi\left(a^2+x^2\right)}$. Montrer que f_a est une densité de probabilité.

   Dans la suite, X est une variable aléatoire réelle sur $(\Omega, \mathscr{A}, \mathbf{P})$ admettant f_a pour densité. On dit alors que X suit une loi de Cauchy de paramètre a et on écrit $X \hookrightarrow \mathscr{C}(a)$.
5. Donner la fonction de répartition de X. Est-ce que X possède une espérance ?
6. Soit $\lambda\in\mathbb{R}_*^+$. Reconnaître la loi de λX lorsque $X\hookrightarrow\mathscr{C}(a)$. Que dire si $\lambda\in\mathbb{R}_*^-$?

- *Maximum et exemple de convergence en loi*
  Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes suivant toutes une loi de Cauchy de paramètre 1. Pour tout $n\in\mathbb{N}^*$, on définit les variables aléatoires :

$$M_n = \max(X_1, X_2, \dots, X_n) \quad \text{et} \quad N_n = n{M_n}^{-1}.$$

7. Pour tout $n\in\mathbb{N}^*$, préciser $\mathbf{P}(N_n \leqslant 0)$. Vérifier ensuite que pour tout $t\in\mathbb{R}_*^+$

$$\mathbf{P}\big([N_n \leqslant t]\cap[M_n \geqslant 0]\big) = 1 - \frac{1}{\pi^n}\left(\frac{\pi}{2}+\arctan\left(\frac{n}{t}\right)\right)^n.$$

8. Conclure en montrant que la suite $(X_n)_n$ converge en loi vers une loi exponentielle dont on précisera le paramètre.

9. Utiliser la question 6 pour reprendre la question précédente en supposant maintenant que les variables $(X_i)_{i\in\mathbb{N}^*}$ suivent une loi de Cauchy $\mathscr{C}(a)$ avec $a\in\mathbb{R}_*^+$.

⋙ *Pour des versions similaires, voir EMLyon 2017, EDHEC 2019.*

≫ Solution p. 41

Exercice 36. ✦✦ **Convergence en loi et fonctions génératrices**
Soient une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ et X, définies sur le même espace probabilisé, à valeurs dans $\{x_0, \dots, x_m\}$. On définit les fonctions G_n sur $\mathbb{R}$ par

$$\forall t\in\mathbb{R}, \qquad G_{X_n}(t) = \sum_{k=0}^{m} \mathbf{P}(X_n = x_k)\cdot t^k \quad \text{et} \quad G_X(x) = \sum_{k=0}^{m} \mathbf{P}(X = x_k)\cdot t^k.$$

1. Vérifier que, si $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X, alors :

$$\forall t\in\mathbb{R}, \qquad G_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} G_X(t) \qquad (\bullet)$$

2. L'objectif de la question suivante est d'établir la réciproque. On suppose donc la propriété $(\bullet)$ vérifiée. On pose

$$A = \begin{bmatrix} 1 & x_0 & {x_0}^2 & \dots & {x_0}^m \\ 1 & x_1 & {x_1}^2 & \dots & {x_1}^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & {x_m}^2 & \dots & {x_m}^m \end{bmatrix} \in \mathscr{M}_{m+1}(\mathbb{R}).$$

a) 🔍 Justifier que les colonnes de A forment une famille libre. En déduire l'inversibilité de A.

b) 🔍 Soit $k\in[\![0;m]\!]$. En déduire que la suite $\big(\mathbf{P}(X_n = x_k)\big)_{n\in\mathbb{N}}$ converge. Notons ℓ_k, la limite.

c) Montrer que les réels $(\ell_k)_{k\in[\![0;m]\!]}$ sont les coefficients d'une loi de probabilité.
C'est-à-dire que les réels ℓ_k sont compris dans $[0;1]$ et leur somme vaut 1.

d) En déduire que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi.

3. • *Application*
Soient $m\in\mathbb{N}^*$ fixé et $(p_n)_{n\in\mathbb{N}}$ une suite de réels de $[0;1]$. On suppose que pour tout $n\in\mathbb{N}$, $X_n \hookrightarrow \mathscr{B}(m;p_n)$.

a) Expliciter $G_{X_n}(t)$ pour tout $t\in\mathbb{R}$.

b) En utilisant l'équivalence prouvée aux questions 1 et 2, montrer que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ converge en loi si et seulement si la suite de réels $(p_n)_{n\in\mathbb{N}}$ converge.

≫ Solution p. 43

Approximations

Exercice 37. ✦✦ **Aiguilles de Buffon**
On dispose d'un sol recouvert de parquet, dont les lattes ont une largeur de 20 cm. On dispose aussi de grandes aiguilles, dont la longueur est égale à 20 cm. On admet que si on laisse tomber une aiguille par terre, la probabilité que l'aiguille soit à cheval sur deux lattes est égale à $\frac{2}{\pi}$.

1. Proposer une expérience, à base de lancers d'aiguille, permettant de donner une approximation de $\frac{2}{\pi}$ (dont on déduira une approximation de π).

2. De façon théorique, et en utilisant la valeur de π, combien de lancers faudrait-il réaliser, approximativement, pour obtenir une approximation de $\frac{2}{\pi}$ à moins de 0,01 valable avec une probabilité de 95% ?
Données : $\Phi(1{,}645) \simeq 0{,}95$, $\Phi(1{,}960) \simeq 0{,}975$.

≫ Solution p. 44

Exercice 38. ✦✦✦ **Approximation de π via la méthode de Monte Carlo** *d'après oraux ESCP 2014*

Soit $(U_n)_{n\in\mathbb{N}^*}$ et $(V_n)_{n\in\mathbb{N}^*}$ deux suites de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ et toutes de même loi uniforme sur $[0,1]$. On suppose que toutes les variables U_n et V_n (pour $n \in \mathbb{N}^*$) sont indépendantes.

1. Pour tout réel $x \in]0,1]$, calculer l'intégrale :

$$J(x) = \int_0^x \frac{1}{\sqrt{t(x-t)}}\,dt.$$

On pourra justifier et utiliser le changement de variable (à x fixé) :

$$\varphi :]-\pi/2, \pi/2[\to \mathbb{R}, \qquad \theta \mapsto t = \frac{x}{2}\sin\theta + \frac{x}{2}.$$

2. **a)** Déterminer la loi de U_n^2.

b) Justifier que la variable $U_n^2 + V_n^2$ possède une densité h, que l'on exprimera sous forme d'une intégrale.

c) Déterminer $h(x)$ pour $x \in [0,1]$.

3. On pose :

$$\forall n \in \mathbb{N}^*, \quad X_n = \begin{cases} 1 & \text{si } {U_n}^2 + {V_n}^2 \leqslant 1 \\ 0 & \text{sinon.} \end{cases}$$

Déterminer la loi de X_n.

4. **a)** Prouver que la suite $(Z_n)_{n\in\mathbb{N}^*}$ définie par $\forall n \in \mathbb{N}^*, Z_n = \frac{4}{n}\sum_{k=1}^{n} X_k$, converge en probabilité vers la constante π. C'est-à-dire, pour tout $\varepsilon \in \mathbb{R}_*^+$,

$$\mathbf{P}(|Z_n - \pi| \geqslant \varepsilon) \underset{n\to\infty}{\longrightarrow} 0.$$

b) Soit $\alpha \in]0,1[$ et $\delta > 0$.
Montrer qu'il existe un entier n_0, qu'on exprimera en fonction de α et δ, tel que

$$\forall n \geqslant n_0, \quad \mathbf{P}(|Z_n - \pi| > \delta) \leqslant \alpha$$

c) En déduire un programme Python qui donne une approximation de π.

≫ Solution p. **??**

Les inclassables

Exercice 39. ✦✦ **Application de la formule de Stirling** *D'après EDHEC 2007*

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathscr{A}, P)$, mutuellement indépendantes, et qui suivent toutes la loi exponentielle de paramètre 1. On pose $S_n = \sum_{k=1}^{n} X_k$.

1. Rappeler quelle est la loi suivie par S_n. Donner l'espérance et la variance de S_n.

2. À l'aide du théorème central limite, établir que $\mathbf{P}(S_n \leqslant n) \underset{n\to\infty}{\longrightarrow} \frac{1}{2}$.

3. En déduire la valeur de

$$\lim_{n\to+\infty} \int_0^n \frac{t^{n-1}}{(n-1)!} e^{-t}\,dt.$$

4. **a)** Utiliser le résultat précédent pour montrer que $\displaystyle\int_0^1 z^{n-1} e^{-nz}\,dz \underset{n\to+\infty}{\sim} \frac{n!}{2n^{n+1}}$.

b) On admet que $n! \underset{+\infty}{\sim} \sqrt{2\pi n}\,n^n e^{-n}$. En déduire un nouvel équivalent de $\displaystyle\int_0^1 z^{n-1} e^{-nz}\,dz$.

≫ Solution p. 44

Exercice 40. ✦✦

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes telles que : $X_n \hookrightarrow \mathscr{B}\left(n, \frac{1}{4}\right)$.

1. Calculer $\mathbf{P}\left(X_n \leqslant \frac{n}{4}\right)$.

2. En utilisant le théorème de la limite centrée déterminer un équivalent de $u_n = \displaystyle\sum_{k=0}^{\lfloor n/4 \rfloor} \binom{n}{k} 3^{n-k}$.

≫ Solution p. 45

Exercice 41. ✦✦✦ ✎

1. En appliquant le théorème limite central à une suite de variables aléatoires indépendantes (X_n) suivant toutes une loi de Poisson $\mathscr{P}(1)$, démontrer que
$$e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}\underset{n\to\infty}{\longrightarrow}\frac{1}{2}.$$
2. Écrire la formule de Taylor à la fonction exponentielle entre 0 et n, en déduire
$$\int_0^n e^{-t}t^n\,dt\underset{n+\infty}{\sim}\frac{n!}{2}.$$

≫ Solution p. ??

Exercice 42. **Amélioration de la méthode de Monte-Carlo, Réduction de la variance par méthode des variables antithétiques.**
Soit f, une fonction continue dont on souhaite approcher
$$I=\int_0^1 f(t)\,dt.$$
La méthode de Monte-Carlo part de l'égalité $I=\mathbf{E}\big(f(U)\big)$ où $U\hookrightarrow\mathscr{U}([0;1])$.
Dans la suite, on pose g définie sur $[0;1]$ par $g(t)=(f(t)+f(1-t))/2$.

1. Vérifier que $I=\mathbf{E}\big(g(U)\big)$.
2. Comparer les variances de $f(U)$ et $g(U)$.
3. Reprendre l'exemple précédent avec g. Comparer les deux méthodes.

Compléments théoriques sur les différentes convergences

Exercice 43. ✦✦✦ **Convergence presque sûr**
Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire. Toutes les variables sont définies sur un même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$. On dit que la suite (X_n) converge presque sûrement vers X si :
$$\mathbf{P}\left(\left\{\omega\in\Omega:\lim_{n\to\infty}X_n(\omega)=X(\omega)\right\}\right)=1.$$
Justifier que si la suite $(X_n)_n$ converge presque sûrement vers X, alors elle converge aussi en probabilité vers X.

Exercice 44. ✦✦✦ **La convergence en probabilité implique la convergence en loi** *d'après oraux HEC*
Soient $(X_n)_{n\in\mathbb{N}^*}$ et X des variables aléatoires définies sur un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ telles que $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers X. On note F_n la fonction de répartition de X_n et F la fonction de répartition de X.
Soit x un point de continuité de F, et soit $\delta>0$ fixé.

1. Montrer qu'on peut choisir ε tel que $F(x-\varepsilon)>F(x)-\delta$ et $F(x+\varepsilon)<F(x)+\delta$.
2. Montrer que
$$[X_n\leqslant x]\subset[X\leqslant x+\varepsilon]\cup[|X_n-X|>\varepsilon].$$
Montrer de même que
$$[X\leqslant x-\varepsilon]\subset[X_n\leqslant x]\cup[|X_n-X|>\varepsilon].$$
3. En déduire que
$$F_n(x)\leqslant F(x)+\delta+P(|X_n-X|>\varepsilon)\ \text{ et }F(x)-\delta\leqslant F_n(x)+\mathbf{P}(|X_n-X|>\varepsilon).$$
4. Conclure en prouvant que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X.

Exercice 45. ✦✦ ✎ *D'après Oraux HEC BL 2021*
Soient I un intervalle non trivial de $\mathbb{R}$ et f une fonction définie sur I.
On dit que la fonction f vérifie la propriété $\mathscr{L}_k$ sur I s'il existe un réel $k\in\mathbb{R}^{+^*}$ tel que
$$\forall(x,y)\in I^2,\quad|f(x)-f(y)|\leqslant k|x-y|.$$

1. a) Montrer que les fonctions sinus et valeur absolue vérifient la propriété $\mathscr{L}_1$ sur $\mathbb{R}$.
   b) Montrer que l'on ne peut pas trouver de réel $k\in\mathbb{R}^{+*}$ tel que la fonction racine carrée vérifie la propriété $\mathscr{L}_k$ sur $[0,1]$.
   c) Montrer que s'il existe un réel $k\in\mathbb{R}^{+*}$ tel que f vérifie la propriété $\mathscr{L}_k$ sur I, alors f est continue sur I.
2. Soient un réel $k\in]0,1$, f une fonction définie sur $\mathbb{R}$ et vérifiant la propriété $\mathscr{L}_k$ sur $\mathbb{R}$ et (u_n) la suite définie par la donnée de $u_0\in\mathbb{R}$ et par la relation
$$\forall n\in\mathbb{N},\quad u_{n+1}=f(u_n).$$

a) Montrer que
$$\forall n \in \mathbb{N}, \quad |u_{n+1} - u_n| \leqslant k^n |u_1 - u_0|.$$

b) En déduire que la suite (u_n) converge vers une limite notée ℓ et vérifiant $f(\ell) = \ell$.

3. Soient un réel $k \in]0,1[$, f une fonction définie sur $\mathbb{R}$ et vérifiant la propriété $\mathscr{L}_k$ sur $\mathbb{R}$ et $(\mathrm{T}_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires à densité définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et telles que
$$\forall n \in \mathbb{N}, \qquad \mathrm{T}_{n+1} = f(\mathrm{T}_n).$$

Soit ℓ la limite trouvée à la question 2.

a) Soit $\varepsilon \in \mathbb{R}^+ - *$. Pour tout $n \in \mathbb{N}$, on pose $\mathrm{A}_n = \left[k^n |\mathrm{T}_0 - \ell| \geqslant \varepsilon\right]$. Montrer que
$$\lim_{n\to+\infty} \mathbf{P}(\mathrm{A}_n) = 0.$$

b) Montrer que
$$\forall \varepsilon \in \mathbb{R}_*^+, \quad \lim_{n\to+\infty} \mathbf{P}(|\mathrm{T}_n - \ell| \geqslant \varepsilon) = 0.$$

c) Justifier que $(\mathrm{T}_n)_n$ converge en loi. Reconnaître la loi limite.