CH XIV : Convergence et approximation

Dans ce chapitre, on cherche à définir la notion de convergence d'une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$. On étudie ici deux types de convergence :

× la convergence en probabilité,

× la convergence en loi.

On donnera également quelques exemples d'approximation en loi, découlant de la convergence en loi.

Dans la suite, on ne précisera pas sur quel espace probabilisé est défini chaque v.a.r. . Un tel espace probabilisé sera noté, de manière générique, $(\Omega, \mathscr{A}, \mathbb{P})$.

I. Loi faible des grands nombres

I.1. L'inégalité de Markov

Théorème 1 (inégalité de Markov).

Soit X une v.a.r. qui admet une espérance.

On suppose de plus que : $X(\Omega) \subset \mathbb{R}^+$.
(autrement dit, X est à valeurs positives)

$$\boxed{\forall a > 0,\ \mathbb{P}([X \geqslant a]) \ \leqslant\ \frac{\mathbb{E}(X)}{a}}$$

(démonstration demandée dans l'`ESSEC II-2013`*)*

Démonstration.
Soit $a > 0$. Considérons la v.a.r. Y définie par :

$$Y : \Omega \to \mathbb{R}, \quad \omega \mapsto \begin{cases} a & \text{si } X(\omega) \geqslant a \\ 0 & \text{si } X(\omega) < a \end{cases}$$

Listons quelques propriétés de Y :

- $Y(\Omega) = \{0, a\}$ (Y est donc une v.a.r. finie).
- $[Y = a] = [X \geqslant a]$ (puisque : $\forall \omega \in \Omega,\ Y(\omega) = a \Leftrightarrow X(\omega) \geqslant a$).
- $[Y = 0] = [X < a]$ (puisque : $\forall \omega \in \Omega,\ Y(\omega) = 0 \Leftrightarrow X(\omega) < a$).

1) Comme Y est finie, Y admet une espérance. De plus :

$$\mathbb{E}(Y) = 0 \times \underset{\mathbb{P}([X<a])}{\underset{\shortparallel}{\mathbb{P}([Y=0])}} + a \times \underset{\mathbb{P}([X\geqslant a])}{\underset{\shortparallel}{\mathbb{P}([Y=a])}} = a \times \mathbb{P}([X \geqslant a])$$

2) On remarque aussi : $Y \leqslant X$ (*i.e.* : $\forall \omega \in \Omega$, $Y(\omega) \leqslant X(\omega)$).
Démontrons-le. Soit $\omega \in \Omega$.
Deux cas se présentent.

× si $X(\omega) \geqslant a$ alors $Y(\omega) = a \leqslant X(\omega)$.

× si $X(\omega) < a$ alors $Y(\omega) = 0 \leqslant X(\omega)$.

3) Par croissance de l'espérance, on déduit des deux points précédents :

$$\underset{a\,\mathbb{P}([X\geqslant a])}{\underset{\shortparallel}{\mathbb{E}(Y)}} \leqslant \mathbb{E}(X)$$

On conclut en divisant de part et d'autre par $a > 0$. □

Remarque

- Notons que si $a \leqslant \mathbb{E}(X)$, alors l'inégalité est évidente (et donc peu intéressante).
  En effet, dans ce cas, on a directement : $\mathbb{P}([X \geqslant a]) \leqslant 1 \leqslant \dfrac{\mathbb{E}(X)}{a}$.
  Dans le cas où a est « grand » par rapport à $\mathbb{E}(X)$, on obtient un résultat de bon sens : la probabilité que X prenne de grandes valeurs est d'autant plus faible que celles-ci excèdent $\mathbb{E}(X)$.
- Dans le théorème, nous n'avons pas précisé si X est une v.a.r. à densité ou discrète. La démonstration est valide dans les deux cas.
- Mieux : la démonstration est aussi valable lorsque X est un v.a.r. quelconque (ni discrète, ni à densité). Le seul passage subtil de la démonstration est alors la croissance de l'espérance, opérateur seulement défini pour les v.a.r. discrètes et à densité dans le programme d'ECE.
- Cependant, il existe une théorie mathématique qui définit une notion d'espérance valable pour toute v.a.r. (discrète, à densité, ou autres). L'opérateur ainsi défini vérifie les propriétés classiques de l'espérance, notamment la linéarité et la croissance.

Théorème 2.

Soit $m \in \mathbb{N}$.

Soit X une v.a.r. possédant un moment d'ordre m.

$$\boxed{\forall a > 0,\ \mathbb{P}\big(\,[\ |X| \geqslant a\ ]\,\big) \ \leqslant\ \dfrac{\mathbb{E}\big(|X|^m\big)}{a^m}}$$

Démonstration.
On utilise l'inégalité de Markov à la v.a.r. $|X|^m$ et au réel $a^m > 0$:

$$\begin{array}{ccc} \mathbb{P}([|X|^m \geqslant a^m]) & \leqslant & \dfrac{\mathbb{E}(|X|^m)}{a^m} \\ \shortparallel & & \\ \mathbb{P}([|X| \geqslant a]) & & \end{array}$$

□

I.2. L'inégalité de Bienaymé-Tchebychev

Théorème 3 (Inégalité de Bienaymé-Tchebychev)**.**

Soit X une v.a.r. qui admet une variance.

Alors :

$$\boxed{\forall \varepsilon > 0,\ \mathbb{P}\big(\,[\ |X - \mathbb{E}(X)| \geqslant \varepsilon\ ]\,\big) \ \leqslant\ \dfrac{\mathbb{V}(X)}{\varepsilon^2}}$$

Démonstration.
Soit $\varepsilon > 0$. Considérons la v.a.r. $Y = \big(X - \mathbb{E}(X)\big)^2$. On est dans le cadre d'application de l'inégalité de Markov :

× $Y = \big(X - \mathbb{E}(X)\big)^2$ est à valeurs positives.

× la v.a.r. Y admet une espérance car X admet une variance.

En appliquant l'inégalité de Markov à la v.a.r. Y, avec $a = \varepsilon^2$:

$$\begin{array}{ccccc} \mathbb{P}\big(\,[\ Y \geqslant \varepsilon^2\ ]\,\big) & = & \mathbb{P}\Big(\Big[\ \big(X - \mathbb{E}(X)\big)^2 \geqslant \varepsilon^2\ \Big]\Big) & \leqslant & \dfrac{\mathbb{E}(Y)}{\varepsilon^2} \\ & & \shortparallel & & \shortparallel \\ & & \mathbb{P}\Big(\Big[\ \big(X - \mathbb{E}(X)\big)^2 \geqslant \varepsilon^2\ \Big]\Big) & & \dfrac{\mathbb{E}\big((X - \mathbb{E}(X))^2\big)}{\varepsilon^2} = \dfrac{\mathbb{V}(X)}{\varepsilon^2} \end{array}$$

□

Remarque

- Ce résultat est une inégalité de concentration. Grâce à ce type d'inégalités, on peut mesurer / contrôler la probabilité qu'un phénomène aléatoire puisse s'écarter (on préférera le terme dévier) de la moyenne *i.e.* d'un comportement standard.
- Évidemment, il est rare de constater de grandes déviations. La question que se pose une compagnie d'assurances est de savoir si elle est prête à parier sur le fait qu'une grande déviation (*i.e.* un événement qui s'écarte fortement de la norme) ne se produira pas. Il est alors primordial dans ce cas d'obtenir des inégalités de concentration avec un majorant le plus précis possible.

- L'inégalité de Bienaymé-Tchebychev est peu précise. L'une des raisons est qu'elle s'applique à toute v.a.r. . On peut évidemment obtenir de meilleures majorations en tirant parti des propriétés (*i.e.* de la loi) des v.a.r. étudiées.

I.3. La loi faible des grands nombres

I.3.a) Énoncé et convergence en probabilité

Théorème 4 (Loi faible des grands nombres)**.**

Soit $(X_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires.*

On suppose que les v.a.r. X_i *:*

× *sont indépendantes,*

× *admettent toutes la même espérance* m*,*

× *admettent toutes la même variance* σ^2.

On pose, pour tout $n \in \mathbb{N}^*$, $\overline{X_n} = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$ *(moyenne empirique).*

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

Démonstration.

- Rappelons tout d'abord l'inégalité de Bienaymé-Tchebychev.
  Soit Y une v.a.r. discrète admettant une variance $\mathbb{V}(Y)$.

$$\forall \lambda > 0, \quad \mathbb{P}(|Y - \mathbb{E}(Y)| \geqslant \lambda) \ \leqslant \ \frac{\mathbb{V}(Y)}{\lambda^2}$$

- On applique l'inégalité de Bienaymé-Tchebychev à $Y = \overline{X_n} = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$.

  La v.a.r. $\overline{X_n}$ admet une variance car elle est la CL de v.a.r. qui admettent une variance. De plus :

$$\begin{aligned} \mathbb{E}(\overline{X_n}) &= \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(X_k) \\ &= \frac{1}{n}\sum_{k=1}^{n} m = \frac{1}{\cancel{n}} \cancel{n} \times m = m \end{aligned}$$

Par ailleurs, par indépendance des v.a.r. X_i, on a :

$$\begin{aligned} \mathbb{V}(\overline{X_n}) &= \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\,\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}(X_k) \\ &= \frac{1}{n^2}\sum_{k=1}^{n} \sigma^2 = \frac{1}{n^{\cancel{2}}}\, \cancel{n} \times \sigma^2 = \frac{1}{n}\,\sigma^2 \end{aligned}$$

Soit $\varepsilon > 0$. D'après l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}\left(\left[\left|\overline{X_n} - m\right| \geqslant \varepsilon\right]\right) \ \leqslant \ \frac{\sigma^2}{\varepsilon^2}\,\frac{1}{n} \underset{n\to+\infty}{\longrightarrow} 0$$

□

Remarque

- Il est fréquent de considérer (notamment pour les simulations informatiques) une suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a.r. :
  × indépendantes,
  × **de même loi**,
  × admettant une espérance et une variance.

  Ces hypothèses sont plus strictes que celles énoncés par la LfGN.
  On peut donc bien évidemment utiliser la LfGN dans ce cadre.
- On dit qu'une suite de v.a.r. $(X_n)_{n\in\mathbb{N}^*}$ **converge en probabilté** vers une v.a.r. X si :

$$\forall \varepsilon > 0, \ \lim_{n\to+\infty} \mathbb{P}\left(\left[\ \left|X_n - X\right| \geqslant \varepsilon \ \right] \right) = 0$$

  Dans ce cas, on note : $X_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} X$.
- Via cette définition, on comprend que la LfGN établit un résultat de convergence d'une suite de v.a.r. : la suite de v.a.r. $\left(\overline{X_n}\right)_{n\in\mathbb{N}^*}$ **converge en probabilité** vers la v.a.r. certaine égale à m.
- Il existe une loi forte des grands nombres. Sous des hypothèses plus exigeantes, on établit que la suite de v.a.r. $\left(\overline{X_n}\right)_{n\in\mathbb{N}^*}$ converge vers la v.a.r. certaine égale à m. On retrouve le résultat précédent, à ceci près que le mode de convergence (appelé convergence presque sûre) est un mode plus exigeant que la convergence en probabilité (la convergence presque sûre implique la convergence en probabilité).

I.3.b) Intuition derrière cet énoncé

Exemple *(classique)*

- On considère un dé à 6 faces (truqué ou non).
  - Expérience : on effectue un lancer de ce dé et on note X la v.a.r. égale au résultat obtenu lors de ce lancer.
  - But : on souhaite, à l'aide d'expériences, savoir si le dé peut être considéré comme équilibré. Autrement dit, on souhaite connaître la loi de X. Plus particulièrement, on souhaiterait connaître $\mathbb{P}\big([X=6]\big)$.
- Pour obtenir une approximation de cette probabilité $\mathbb{P}\big([X=6]\big)$, il est naturel de procéder comme suit :
  × on effectue un grand nombre N de tirages.
  × on compte le nombre d'apparitions de la face 6 au cours de ce N tirages.

  On observe alors la fréquence d'apparition de cette face :

  $$\begin{array}{c}\text{fréquence d'apparition}\\ \textbf{observée}\text{ de la face 6}\end{array} = \frac{\text{nombre de 6 obtenus en } N \text{ tirages}}{N}$$

  Il est naturel de penser que cette fréquence observée est une bonne approximation de $\mathbb{P}\big([X=6]\big)$. La LfGN n'est rien d'autre qu'une validation mathématique de ce procédé. Faisons le lien plus formellement.
- Pour tout $i \in \mathbb{N}^*$, on note X_i la v.a.r. qui donne le résultat du $i^{\text{ème}}$ lancer. On note par ailleurs Z_i la v.a.r. définie par :

  $$Z_i : \omega \mapsto \begin{cases} 1 & \text{si } X_i(\omega) = 6 \\ 0 & \text{sinon} \end{cases}$$

  Autrement dit : $Z_i = \mathbb{1}_{[X_i=6]}$. On a alors :

  $$Z_i \hookrightarrow \mathcal{B}(p) \qquad \text{où} \qquad p = \mathbb{P}\big([X_i=6]\big)$$

  En particulier : $\mathbb{E}(Z_i) = \mathbb{E}\big(\mathbb{1}_{[X_i=6]}\big) = \mathbb{P}\big([X_i=6]\big)$.
- La suite $(Z_n)_{n\in\mathbb{N}^*}$ est une suite de variables :
  × indépendantes (le résultat d'un lancer ne dépend pas des précédents),
  × de même espérance p,
  × de même variance $\sigma^2 = p\,(1-p)$.

  On en déduit, par la loi faible des grands nombres que, pour tout $\varepsilon > 0$:

  $$\lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{Z_1+Z_2+\cdots+Z_n}{n} - p\right| \geqslant \varepsilon\right) = 0$$

  ce que l'on préférera noter : $\overline{Z_n} \xrightarrow[n\to+\infty]{\mathbb{P}} p$.
- La v.a.r. $\overline{Z_n}$ donne la fréquence d'apparition de la face 6 au cours des n premiers lancers. La loi faible des grands nombres affirme que, plus n est grand, plus la fréquence d'apparition du 6 au cours des n lancers est proche de la fréquence théorique ($= \mathbb{P}\big([X=6]\big)$) avec une forte probabilité.

I.3.c) Lien avec la simulation informatique

Exemple *(classique)*

Dans les énoncés de concours, la LfGN apparaît souvent dans les questions informatiques. Considérons un énoncé consistant à l'étude d'une v.a.r. X qui admet une variance.
On trouvera fréquemment les questions suivantes.

1. Écrire une fonction d'entête `function x = simuX()` permettant de simuler la v.a.r. X.

   Cette fonction peut éventuellement prendre des paramètres d'entrée. Généralement, cette simulation est obtenue :
   × en écrivant une fonction permettant de simuler l'expérience aléatoire considérée. La valeur simulée de X est calculée lors de cette expérience (cas classique lorsqu'on travaille avec des v.a.r. discrètes).
   × car X s'écrit à l'aide d'autres v.a.r. qu'on peut facilement simuler (X est une transformée de v.a.r. qui suivent des lois usuelles par exemple).

2. Écrire un programme (généralement on demande plutôt de compléter un programme ou d'expliquer ce que fait un programme) permettant d'obtenir une valeur approchée de l'espérance de la v.a.r. X.

Démonstration.

- L'idée naturelle est la suivante :
  - × on simule un grand nombre N de fois la v.a.r. X. *(généralement, on prend $N = 10^4$ ou 10^5 ou 10^6)*
  - × on détermine la moyenne arithmétique des résultats obtenus.

```
function E = approxEsp(N)
    S = 0
    for i = 1:N
        S = S + simuX()
    end
    E = S / N
function
```

Autre possibilité : on peut créer un tableau contenant les N simulations de la v.a.r. X puis déterminer la moyenne arithmétique de ces observations à l'aide des opérations prédéfinies en **Scilab**.

```
function E = approxEsp(N)
    Obs = zeros(1, N)
    for i = 1:N
        Obs(i) = simuX()
    end
    E = mean(Obs)
function
```

- Mathématiquement parlant, cela consiste à considérer un N-échantillon $(X_1, \dots, X_N)$ de la v.a.r. X. Autrement dit, les v.a.r. $X_1, \dots, X_N$ sont :
  - × indépendantes,
  - × de même loi que X.

  Comme X admet une variance, on se trouve dans un cadre (plus strict) permettant d'appliquer la LfGN. Simuler ce N-échantillon, c'est en obtenir un N-uplet d'observation $(x_1, \dots, x_N)$ (c'est ce que contient le tableau `Obs` précédent). La LfGN permet d'affirmer :

$$\frac{1}{N} \sum_{i=1}^{N} x_i \ \simeq \ \mathbb{E}(X)$$

□

3. Écrire un programme (généralement on demande plutôt de compléter un programme ou d'expliquer ce que fait un programme) permettant d'obtenir une valeur approchée de $\mathbb{P}\big([X \geqslant 3]\big)$.

Démonstration.

- L'idée naturelle est la suivante :
  - × on simule un grand nombre N de fois la v.a.r. X. *(généralement, on prend $N = 10^4$ ou 10^5 ou 10^6)*
  - × on compte le nombre de fois où le résultat obtenu est plus grand que 3.
  - × on divise le nombre obtenu par le nombre N d'observations.

```
function p = approxProba(N)
    S = 0
    for i = 1:N
        if simuX() >= 3 then
            S = S + simuX()
        end
    end
    p = S / N
function
```

Autre possibilité : on peut créer un tableau contenant les N simulations de la v.a.r. X puis déterminer la fréquence d'apparitions de réels plus grands que 6 à l'aide d'opérations prédéfinies en **Scilab**.

```
function p = approxProba(N)
    Obs = zeros(1, N)
    for i = 1:N
        Obs(i) = simuX()
    end
    p = sum(Obs >= 3) / N
function
```

Rappelons que l'instruction `Obs >= 3` permet d'obtenir une matrice de même taille $(1 \times N)$ que `Obs`. Elle est obtenue en testant, pour chaque coefficient de `Obs` (c'est une opération terme à terme), si celui-ci est plus grand que 3. Si c'est le cas, on obtient le booléen `True` (on peut considérer qu'il s'agit de la valeur 1). Dans le cas contraire, on obtient le booléen `False` (on considère qu'il s'agit de la valeur 0).

- Mathématiquement parlant, cela consiste à considérer un N-échantillon $(Y_1, \ldots, Y_N)$ de la v.a.r. $Y = \mathbb{1}_{[X \geqslant 3]}$.
  Autrement dit, les v.a.r. $Y_1, \ldots, Y_N$ sont :
  - × indépendantes,
  - × de même loi que $\mathbb{1}_{[X \geqslant 3]}$.

  Comme $\mathbb{1}_{X \geqslant 3}$ admet une variance, on se trouve dans un cadre (plus strict) permettant d'appliquer la LfGN. Simuler ce N-échantillon, c'est en obtenir un N-uplet d'observation $(y_1, \ldots, y_N)$ (c'est ce que contient le tableau obtenu par l'instruction `Obs >=3`).
  La LfGN permet d'affirmer :

$$\frac{1}{N} \sum_{i=1}^{N} y_i \simeq \mathbb{E}(Y) = \mathbb{E}\big(\mathbb{1}_{[X \geqslant 3]}\big) = \mathbb{P}\big([X \geqslant 3]\big) \qquad \square$$

4. Écrire un programme (généralement on demande plutôt de commenter des graphiques) permettant d'obtenir une représentation de la loi de la v.a.r. X.

Démonstration.
En tout premier lieu, il convient de s'interroger sur ce qu'on entend par « représentation de la loi d'une v.a.r. ». Deux cas se présentent.

1) Si X est une v.a.r. discrète
Dans ce cas, la loi de X est la donnée de :
- × $X(\Omega)$,
- × $\mathbb{P}\big([X = x]\big)$ pour tout $x \in X(\Omega)$.

Il faut donc trouver une méthode permettant de déterminer informatiquement $X(\Omega)$ et de déterminer une valeur approchée des valeurs de $\mathbb{P}\big([X = x]\big)$ pour $x \in X(\Omega)$.

- L'idée naturelle est la suivante :
  - × on simule un grand nombre N de fois la v.a.r. X.
    (généralement, on prend $N = 10^4$ ou 10^5 ou 10^6)
  - × on représente graphiquement le diagramme en bâtons de répartition des valeurs observées.

```
N = input('Entrez le nombre d'observations souhaitées')
Obs = zeros(1, N)
for i = 1:N
    Obs(i) = simuX()
end
tab = tabul(Obs, 'i')
bar(tab(1, :), tab(2, :))
```

Rappelons que l'instruction `tabul` prend un paramètre une matrice de données (ici `Obs`) et renvoie un tableau (nommé `tab` ici) contenant deux colonnes :
- × la $1^{\text{ère}}$ colonne liste dans l'ordre croissant (grâce à l'option `'i'`) les différentes valeurs trouvées dans la matrice `Obs`.
  (on obtient un ensemble $\{x_1, \ldots, x_n\}$ de valeurs prises par X)
- × la $2^{\text{ème}}$ colonne contient le nombre d'occurences de chacune de ces valeurs x_i.

L'instruction `bar` permet alors de représenter le **diagramme des effectifs** : en abscisse les valeurs x_i et en ordonnée (hauteur du bâton d'abscisse x_i) le nombre d'occurences de x_i. Si l'on souhaite représenter le diagramme des effectifs, on remplace la ligne 7 par :

```
7   bar(tab(1, :), tab(2, :) / N)
```

- Mathématiquement parlant, pour chaque valeur x_i, cela consiste à considérer un N-échantillon $(Y_1, \dots, Y_N)$ de la v.a.r. $Y = \mathbb{1}_{[X=x_i]}$.
  Autrement dit, les v.a.r. Y_1, ..., Y_N sont :
  - × indépendantes,
  - × de même loi que $\mathbb{1}_{[X=x_i]}$.

  Comme $\mathbb{1}_{X=x_i}$ admet une variance, on se trouve dans un cadre (plus strict) permettant d'appliquer la LfGN. Simuler ce N-échantillon, c'est en obtenir un N-uplet d'observation $(y_1, \dots, y_N)$.
  La LfGN permet d'affirmer :

$$\frac{1}{N} \sum_{i=1}^{N} y_i \simeq \mathbb{E}(Y) = \mathbb{E}\big(\mathbb{1}_{[X=x_i]}\big) = \mathbb{P}\big([X = x_i]\big)$$

2) Si X est une v.a.r. à densité

Dans ce cas, la loi de X est la donnée de :

- × $X(\Omega)$,
- × la fonction de répartition $F_x : x \mapsto \mathbb{P}([X \leqslant x])$ associée à X.
  (rappel : F_X est une fonction définie sur $\mathbb{R}$)

- L'idée naturelle est la suivante :
  - × on simule un grand nombre N de fois la v.a.r. X.
    (généralement, on prend $N = 10^4$ ou 10^5 ou 10^6)
  - × on représente graphiquement l'histogramme de répartition des valeurs observées.

Il faut bien comprendre qu'on `NE` représente `PAS` la fonction de répartition de la v.a.r. X. Pour représenter un histogramme il faut définir des classes. Par exemple, si l'on s'intéresse à la répartition de valeurs prises par une v.a.r. à densité sur l'intervalle $[0, 10]$, on pourra considérer les classes :

$$c_1 = [0, 1], \ c_2 =]1, 2], \ c_3 =]2, 3], \ c_4 =]3, 4], \ \dots \ , c_{10} =]9, 10]$$

Il s'agit alors de déterminer combien d'observations on trouve dans chacune de ces classes.

```
N = input('Entrez le nombre d'observations souhaitées')
Obs = zeros(1, N)
classe = 0:10 // ou : nbClasse = 10
for i = 1:N
    Obs(i) = simuX()
end
histplot(classe, Obs) // ou : histplot(nbClasse, Obs)
```

On obtient ainsi un histogramme contenant 10 barres. Pour cet exemple, la hauteur de la $i^{\text{ème}}$ barre est une approximation de la valeur :

$$\mathbb{P}\big([X \in c_i]\big) = \mathbb{P}([i - 1 < X \leqslant i])$$

*(on ne revient pas ici sur le lien avec la LfGN - cf question **3**)*

On n'est pas forcé de décrire très précisément les classes. On peut aussi, plus simplement, précisément seulement le nombre de classes souhaitées (*cf* appels en commentaires dans le programme précédent). Dans ce cas, l'exécution de `histplot(nbClasse, Obs)` commence par la recherche des valeurs minimale $x_{\min}$ et $x_{\max}$ de la matrice `Obs`. Les `nbClasse` (10 dans l'exemple ici) sont obtenues en découpant l'intervalle $[x_{\min}, x_{\max}]$ ($[0.1, 9, 87]$ par exemple) en 10 sous-intervalles égaux. Ce qu'on pourrait obtenir par l'appel :

```
linspace(0.1, 9.87, nbClasse + 1)
```

□

Application : illustration sur un exemple classique

Rappelons tout d'abord le résultat classique suivant.

$$X \hookrightarrow \mathcal{U}([0,1[) \quad \Rightarrow \quad -\frac{1}{\lambda}\ \ln(1-X) \hookrightarrow \mathcal{E}(\lambda)$$

Illustrons toute l'étude précédente avec ce cas.

1. Écrire une fonction d'entête `function x = simuX(lam)` permettant de simuler la v.a.r. X.

```
function y = simuX(lam)
    x = rand()
    y = -(1 / lam) * log(1 - x)
endfunction
```

2. Écrire un programme permettant d'obtenir une valeur approchée de l'espérance de la v.a.r. X.

```
function E = approxEsp(N, lam)
    S = 0
    for i = 1:N
        S = S + simuX(lam)
    end
    E = S / N
endfunction
```

On peut aussi se servir des fonctions prédéfinies en **Scilab** en agissant comme suit.

```
function E = approxEsp(N, lam)
    Obs = zeros(1, N)
    for i = 1:N
        Obs(i) = simuX(lam)
    end
    E = mean(Obs)
endfunction
```

Considérons le cas $\lambda = \frac{1}{2}$ et testons la fonction précdente.

```
--> approxEsp(10000, 1/2)
  ans  =
     2.0144584
```

Ce qui permet de retrouver le résultat théorique :

$$\mathbb{E}(X) = \frac{1}{\lambda} = \frac{1}{\frac{1}{2}} = 2$$

3. Écrire un programme permettant d'obtenir une valeur approchée de $\mathbb{P}\big(\,[X \geqslant \ln(9)]\,\big)$.

```
function p = approxProba(N, lam)
    Obs = zeros(1, N)
    for i = 1:N
        Obs(i) = simuX(lam)
    end
    p = sum(Obs >= log(9)) / N
endfunction
```

Testons cette fonction (toujours pour $\lambda = \frac{1}{2}$) :

```
--> approxProba(10000, 1/2)
  ans  =
     0.3243
```

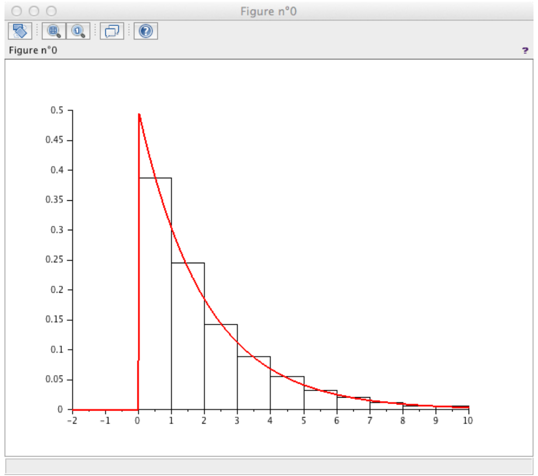
Ce qui permet de retrouver le résultat théorique :

$$\begin{aligned}
\mathbb{P}\big(\,[X \geqslant \ln(9)]\,\big) &= 1 - \mathbb{P}\big(\,[X < \ln(9)]\,\big) = \cancel{1} - \left(\cancel{1} - \exp\left(-\tfrac{1}{2}\ \ln(9)\right)\right) \\
&= \exp\left(\ln(9^{-\frac{1}{2}})\right) = 9^{-\frac{1}{2}} \\
&= \frac{1}{\sqrt{9}} = \frac{1}{3}
\end{aligned}$$

4. Écrire un programme permettant d'obtenir une représentation graphique de la loi de X.

```
N = 10 000
lam = 1/2
Obs = zeros(1, N)
classe = -2:10
for i = 1:N
    Obs(i) = simuX(lam)
end
histplot(classe, Obs)
```

On obtient la représentation graphique suivante.



Que représente exactement ce graphique ?

- La matrice `Obs` contient des valeurs (N en tout) que peut prendre la v.a.r. X. Effectuer l'histogramme consiste à visualiser graphiquement comment se répartissent les valeurs prises par X. Cependant, il faut bien comprendre que cet histogramme `NE` représente `PAS` la fonction de répartition de la v.a.r. X. En effet, $F_X(1) = \mathbb{P}\big([X \leqslant 3]\big)$ est la probabilité que X prenne une valeur plus faible que 3. On s'intéresse alors à `TOUTES` les valeurs prises par X avant x_0. Et pas seulement au nombre de valeurs prises par X dans l'intervalle $]2, 3]$ (l'une des barres du graphique est justement une approximation de $\mathbb{P}\big(\,[2 < X \leqslant 3]\,\big)$). La terminologie anglo-saxonne met en avant cette idée d'accumulation en ajoutant l'adjectif *cumulative* au terme fonction de répartion (*distribution function*).
- En réalité, cet histogramme est à rapprocher des densités de probabilité associées à X. Pour bien le comprendre, on a représenté en rouge une densité f_X de la loi $\mathcal{E}\left(\frac{1}{2}\right)$. Sur chaque classe, la hauteur de chaque barre semble être une bonne approximation de cette densité.
  Cela provient du fait que :

$$\begin{array}{ccccc} \text{aire de la } 3^{\text{ème}} \text{ barre} & \simeq & \mathbb{P}\big(\,[2 < X \leqslant 3]\,\big) & & \\ \shortparallel & & \shortparallel & & \\ \text{hauteur de la } 3^{\text{ème}} \text{ barre} & & \displaystyle\int_2^3 f_X(t)\,dt & \simeq & f_X\left(\frac{5}{2}\right) \end{array}$$

  en faisant l'approximation que f_X est constante égale à $f_X(\frac{5}{2})$ sur la classe $]2, 3]$.
- Le graphique obtenu fait penser à une représentation de la méthode des rectangles. Cela valide le fait que, pour chaque classe c_i, l'aire de chaque rectangle est une bonne approximation de l'aire sous la courbe de f_X sur l'intervalle défini par c_i.

II. Convergence en loi

II.1. Définition

Définition

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.
(pour tout $n \in \mathbb{N}^*$, on note F_{X_n} les fonctions de répartition correspondantes)
Soit X une v.a.r. définie sur ce même espace probabilisé.
(on note F_X la fonction de répartition de la v.a.r. X)

- On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers la variable X, lorsque :

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

  en tout point de continuité x de la fonction F_X.
- Lorsque c'est le cas, on note $X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X$.

Remarque

- Considérons une suite de v.a.r. $(X_n)_{n\in\mathbb{N}^*}$ telle que : $X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X$.
  On a alors, pour tout $(a,b) \in \mathbb{R}^2$ où a et b sont des points de continuité de F_X tels que $a \leqslant b$:

$$\begin{array}{ccc} \mathbb{P}\big([a < X_n \leqslant b]\big) & \underset{n\to+\infty}{\longrightarrow} & \mathbb{P}\big([a < X \leqslant b]\big) \\ \shortparallel & & \shortparallel \\ F_{X_n}(b) - F_{X_n}(a) & \underset{n\to+\infty}{\longrightarrow} & F_X(b) - F_X(a) \end{array}$$

- Rappelons par ailleurs que la fonction de répartition d'une v.a.r. X caractérise la loi de X. Autrement dit, déterminer F_X c'est déterminer la loi de X. On pourra alors retenir que la convergence en loi c'est la « convergence des fonctions de répartitions ».
- Rappelons enfin que pour toute v.a.r. X :

$$\boxed{F_X \text{ continue en } x \in \mathbb{R} \ \Leftrightarrow\ \mathbb{P}([X = x]) = 0}$$

Cas particulier des v.a.r. à densité / v.a.r. discrètes

1) Cas des v.a.r. à densité

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r.
Soit X une v.a.r. **à densité**.

Comme X est à densité, sa fonction de répartition F_X est continue sur $\mathbb{R}$. On peut alors reformuler la définition de convergence en loi :

$$\boxed{X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X \quad\Leftrightarrow\quad \forall x \in \mathbb{R},\ \lim_{n\to+\infty} F_{X_n}(x) = F_X(x)}$$

Dans ce cas (X v.a.r. **à densité**), on peut même démontrer (technique) :

$$\boxed{X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X \quad\Leftrightarrow\quad \forall (a,b) \in \mathbb{R}^2,\ \lim_{n\to+\infty} \mathbb{P}([a \leqslant X_n \leqslant b]) = \mathbb{P}([a \leqslant X \leqslant b])}$$

2) Cas des v.a.r. discrètes

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. **discrètes**.
Soit X une v.a.r. **discrète**.
On suppose de plus que les v.a.r. X_n et la v.a.r. X sont à valeurs dans le même ensemble. Plus précisément, on considère ici le cas où :

× $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. à valeurs dans $\mathbb{N}$ (resp. $\mathbb{Z}$).
× X est à valeurs dans $\mathbb{N}$ (resp. $\mathbb{Z}$).

On a alors :

$$\boxed{X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X \quad\Leftrightarrow\quad \forall k \in \mathbb{N} \text{ (resp. } \mathbb{Z}),\ \lim_{n\to+\infty} \mathbb{P}([X_n = k]) = \mathbb{P}([X = k])}$$

II.2. Illustration sur des exemples

II.2.a) Cas d'une suite de v.a.r. discrètes convergeant en loi vers une v.a.r. discrète

Théorème 5.

Soit $\lambda > 0$.

Soit $(X_n)_{n\in\mathbb{N}^*}$ *une suite de v.a.r.*

On suppose : $\forall n \in \mathbb{N}^*,\ X_n \hookrightarrow \mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$.

Alors : $X_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} X$ *où* $X \hookrightarrow \mathcal{P}(\lambda)$.

Démonstration.

- On est dans le cas où :
  - $\times$ $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. **discrètes** à valeurs dans $\mathbb{N}$.
  - $\times$ X est une v.a.r. **discrète** à valeurs dans $\mathbb{N}$.

  Il s'agit alors de démontrer :

$$\forall k \in \mathbb{N},\ \lim_{n\to+\infty} \mathbb{P}([X_n = k]) = \mathbb{P}([X = k])$$

- Soit $n \in \mathbb{N}^*$ et soit $k \in \mathbb{N}$.
  Comme on s'intéresse à un résultat lorsque $n \underset{n\to+\infty}{\longrightarrow} +\infty$, on peut considérer n aussi grand que souhaité. En particulier : $k \in [\![0, n]\!]$. On a alors :

$$\begin{aligned}
\mathbb{P}([X_n = k]) &= \binom{n}{k} \times \left(\frac{\lambda}{n}\right)^k \times \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n!}{k!(n-k)!}\ \frac{\lambda^k}{n^k}\ \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k}{k!}\ \frac{\frac{n!}{(n-k)!}}{n^k}\ \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}
\end{aligned}$$

- Intéressons-nous aux termes de ce produit.

  1) Tout d'abord :

$$\frac{\frac{n!}{(n-k)!}}{n^k} = \frac{n(n-1)\dots(n-(k-1))}{n^k} \underset{n\to+\infty}{\sim} \frac{n^k}{n^k} = 1$$

  En effet, le numérateur apparaît comme produit de k éléments qui sont tous équivalents, lorsque n tend vers $+\infty$, à n.

  2) Notons $u_n = \left(1 - \dfrac{\lambda}{n}\right)^n$. Alors : $u_n = \exp\left(n \ln\left(1 - \dfrac{\lambda}{n}\right)\right)$.

  Or :

$$\ln(u_n) = n \ln\left(1 - \frac{\lambda}{n}\right) \underset{n\to+\infty}{\sim} \cancel{n}\ \frac{-\lambda}{\cancel{n}} = -\lambda$$

  On en déduit : $\ln(u_n) \underset{n\to+\infty}{\longrightarrow} -\lambda$ et $u_n = \mathrm{e}^{\ln(u_n)} \underset{n\to+\infty}{\longrightarrow} \mathrm{e}^{-\lambda}$.

  3) Enfin, comme : $1 - \dfrac{\lambda}{n} \underset{n\to+\infty}{\longrightarrow} 1$, on a :

$$\left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{1}{\left(1 - \frac{\lambda}{n}\right)^k} \underset{n\to+\infty}{\longrightarrow} \frac{1}{1} = 1$$

En conclusion :

$$\boxed{\forall k \in \mathbb{N},\ \mathbb{P}([X_n = k]) \underset{n\to+\infty}{\longrightarrow} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} = \mathbb{P}([X = k])}$$

où X est une v.a.r. telle que $X \hookrightarrow \mathcal{P}(\lambda)$.

La loi de Poisson apparaît comme la limite de lois binomiales $\mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$.

Ainsi, si n grand (et donc $\frac{\lambda}{n}$ proche de 0) la loi $\mathcal{P}(\lambda)$ est une bonne approximation de la loi $\mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$. □

Remarque

- Dans la pratique, on considère qu'on peut approcher la loi $\mathcal{B}(n,p)$ par la loi $\mathcal{P}(np)$ lorsque $n \geqslant 30$ et $p \leqslant 0,1$.
- Si X une v.a.r. tel ques $X \hookrightarrow \mathcal{B}(n,p)$ et que les contraintes précédentes sur n et p sont vérifiées, on utilise alors l'approximation :

$$\forall k \in [\![0,n]\!],\ \mathbb{P}([X=k]) \ \simeq \ \mathrm{e}^{-np}\ \frac{(np)^k}{k!}$$

- La loi de Poisson est généralement utilisée comme loi de v.a.r. consistant à calculer le nombre d'événements d'un certain type se produisant sur un laps de temps donné. Cette modélisation est valide si :
  1) les événements se produisant sont indépendants,
  2) la probabilité d'apparition du phénomène dans un laps de temps donné T ne dépend que de cette durée T.

  Le résultat précédent démontre que l'on peut approcher la loi $\mathcal{P}(\lambda)$ par la loi $\mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Cette dernière loi est utilisée pour compter le nombre de succès (c'est le rôle de la v.a.r. X_n) au cours d'une expérience consistant à la succession de n épreuves de Bernoulli indépendantes et de même paramètre de succès $p = \frac{\lambda}{n}$. L'approximation est d'autant meilleure que n est grand (et donc p petit). C'est pourquoi on qualifie parfois la loi de Poisson de « loi des événements rares ».
- Illustrons enfin ce résultat sur un exemple.
  Notons X une v.a.r. telle que : $X \hookrightarrow \mathcal{B}(100, 0.05)$.
  × Par définition de la loi binomiale :

$$\mathbb{P}([X=2]) = \binom{100}{2} \left(\frac{1}{20}\right)^2 \left(\frac{19}{20}\right)^{98} \simeq 0.0812$$

  × Comme $n \geqslant 30$ et $p = 0.05 \leqslant 0.1$, on peut approcher la loi de X par la loi de Poisson $\mathcal{P}(5)$. On obtient alors :

$$\mathbb{P}([X=2]) \ \simeq \ \frac{5^2}{2!}\,\mathrm{e}^{-5} \ \simeq \ 0.0842$$

II.2.b) Cas d'une suite de v.a.r. à densité convergeant en loi vers une v.a.r. discrète

Exercice *(d'après* `ESCP 1987`*)*

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. à densité.

Pour tout $n \in \mathbb{N}^*$, la fonction de répartition de X_n est définie par :

$$F_{X_n} : x \mapsto \begin{cases} \dfrac{1}{(1-x)^n} & \text{si } x \leqslant 0 \\ 1 & \text{si } x > 0 \end{cases}$$

Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une v.a.r. X dont on précisera la loi.

Exercice

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. à densité.

Plus précisément, on considère, pour tout $n \in \mathbb{N}^*$: $X_n \hookrightarrow \mathcal{U}([-\frac{1}{n}, \frac{1}{n}])$.

Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une v.a.r. X dont on précisera la loi.

II.2.c) Cas d'une suite de v.a.r. discrètes convergeant en loi vers une v.a.r. à densité

Exercice

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r.

Plus précisément, on suppose, pour tout $n \in \mathbb{N}^*$:

$$X_n \hookrightarrow \mathcal{U}\left(\left\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}\right)$$

1. Soit $n \in \mathbb{N}^*$. Montrer que pour tout réel x :

$$F_n(x) = \begin{cases} 0 & \text{si } x < \frac{1}{n} \\ \dfrac{\lfloor nx \rfloor}{n} & \text{si } \frac{1}{n} \leqslant x < 1 \\ 1 & \text{si } x \geqslant 1 \end{cases}$$

2. Montrer l'équivalent : $\lfloor x \rfloor \underset{x\to+\infty}{\sim} x$.

3. Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une v.a.r. X telle que : $X \hookrightarrow \mathcal{U}([0,1])$.

II.2.d) Cas d'une suite de v.a.r. à densité convergeant en loi vers une v.a.r. à densité

Exercice

Soit X une v.a.r. à densité.

Pour tout $n \in \mathbb{N}^*$, on pose $X_n = \mathrm{e}^{\frac{1}{n}} X$.

Montrer que (X_n) converge en loi vers X.

III. Approximation de v.a.r. par le TCL

III.1. Théorème central limite

Théorème 6 (Théorème central limite)**.**
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. :
$\times$ *indépendantes,*
$\times$ *de même loi,*
$\times$ *de même espérance m,*
$\times$ *et de même variance σ^2* **non nulle***.*
Pour tout $n \in \mathbb{N}^*$, on note :

$$S_n = X_1 + \cdots + X_n \quad \text{et} \quad \overline{X}_n = \frac{X_1 + \cdots + X_n}{n} = \frac{S_n}{n}$$

$$S_n^* = \frac{S_n - nm}{\sigma\sqrt{n}} \quad \text{et} \quad \overline{X}_n^* = \frac{\overline{X}_n - m}{\frac{\sigma}{\sqrt{n}}}$$

Alors : $\boxed{\overline{X_n}^* \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} Z \quad \text{où} \quad Z \hookrightarrow \mathcal{N}(0,1)}$

Autrement dit, la suite $(\overline{X}_n^*)_{n\in\mathbb{N}^*}$ converge en loi vers une v.a.r. de loi normale centrée réduite $\mathcal{N}(0,1)$. En particulier, pour tout $(a,b) \in \mathbb{R}^2$ avec $a \leqslant b$:

$$\boxed{\lim_{n\to+\infty} \mathbb{P}\left(\left[a \leqslant \overline{X}_n^* \leqslant b\right]\right) = \mathbb{P}([a \leqslant Z \leqslant b])}$$

Remarque

- Insistons sur le fait que, pour tout $n \in \mathbb{N}^*$: $\overline{X}_n^* = S_n^*$.
  Le TCL peut donc s'exprimer avec l'une ou l'autre de ces variables. Le choix dépendra du contexte. Dans un exercice d'estimation, on préférera introduire la v.a.r. $\overline{X}_n$ (moyenne empirique) ; dans un contexte d'approximation de lois (à suivre), on préférera travailler sur la v.a.r. S_n.
- En pratique, on considère que l'on peut approcher la loi de $\overline{X}_n^*$ par la loi normale centrée réduite dès que : $n \geqslant 30$.

III.2. Idée générale de l'approximation de lois par TCL

- On se place dans le cadre d'utilisation du théorème précédent. On a alors :

$$\overline{X_n}^* \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} Z \quad \text{où} \quad Z \hookrightarrow \mathcal{N}(0,1)$$

  On se pose alors la question de savoir ce que résultat permet de conclure quant à la loi de S_n. Pour ce faire, déterminons F_{S_n}.
- Soit $x \in \mathbb{R}$.

$$\begin{array}{rcll}
F_{S_n}(x) & = & \mathbb{P}\big([S_n \leqslant x]\big) & \\
& = & \mathbb{P}\big([S_n - nm \leqslant x - nm]\big) & \\
& = & \mathbb{P}\Big(\Big[\dfrac{S_n - nm}{\sigma\sqrt{n}} \leqslant \dfrac{x-nm}{\sigma\sqrt{n}}\Big]\Big) & \\
& = & \mathbb{P}\Big(\Big[S_n^* \leqslant \dfrac{x-nm}{\sigma\sqrt{n}}\Big]\Big) & \\
& \simeq & \mathbb{P}\Big(\Big[Z \leqslant \dfrac{x-nm}{\sigma\sqrt{n}}\Big]\Big) & \textit{(pour une valeur de n suffisamment grande)} \\
& = & \mathbb{P}\Big([\sigma\sqrt{n}\, Z + nm \leqslant x]\Big) \ = \ F_{\sigma\sqrt{n}\, Z+nm}(x) &
\end{array}$$

  On peut donc considérer que, pour n suffisamment grand, les v.a.r. S_n et $Z_n = \sigma\sqrt{n}\, Z + nm$ ont même loi.
- Mais quelle est la loi de $Z_n = \sigma\sqrt{n}\, Z + nm$?
  Comme Z_n est une transformée affine de la v.a.r. Z qui suit une loi normale, Z_n suit une loi normale. On détermine les caractéristiques de cette loi normale en déterminant espérance et variance de Z_n :

$$\begin{array}{rcccccl}
\mathbb{E}(Z_n) & = & \mathbb{E}(\sigma\sqrt{n}\, Z + nm) & = & \sigma\sqrt{n}\, \cancel{\mathbb{E}(Z)} + nm & = & nm \\
\mathbb{V}(Z_n) & = & \mathbb{V}(\sigma\sqrt{n}\, Z + nm) & = & (\sigma\sqrt{n})^2\, \mathbb{V}(Z) & = & n\sigma^2
\end{array}$$

$$\boxed{\text{Ainsi, pour } n \text{ suffisamment grand, la v.a.r. } S_n \text{ suit approximativement la loi } \mathcal{N}\left(nm, n\sigma^2\right).}$$

III.3. Illustration sur des exemples

III.3.a) Approximation d'une loi binomiale par une loi normale

Théorème 7.

Soit $p \in \,]0,1[$.

Soit $(X_n)_{n \in \mathbb{N}^*}$ *une suite de v.a.r. indépendantes.*

On suppose : $\forall n \in \mathbb{N}^*, \ X_n \hookrightarrow \mathcal{B}(p)$.

Pour tout $n \in \mathbb{N}^*$, *on note* $S_n = \sum\limits_{k=1}^{n} X_k$.

1) *Alors pour tout* $n \in \mathbb{N}^*$ *:* $S_n \hookrightarrow \mathcal{B}(n,p)$.

2) $\boxed{S_n^* = \dfrac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow[n \to +\infty]{\mathscr{L}} Z, \quad \text{où } Z \hookrightarrow \mathcal{N}(0,1)}$

3) *Pour* n *suffisamment grand, la v.a.r.* S_n *suit approximativement la loi* $\mathcal{N}(np, np(1-p))$.

Démonstration.

1) Déjà vu lors du chapitre sur les couples de v.a.r.

2) C'est une application directe du TCL.

3) Soit $x \in \mathbb{R}$.

$$\begin{aligned}
F_{S_n}(x) &= \mathbb{P}\big([S_n \leqslant x]\big) \\
&= \mathbb{P}\left(\left[\frac{S_n - np}{\sqrt{np(1-p)}} \leqslant \frac{x - np}{\sqrt{np(1-p)}}\right]\right) \\
&= \mathbb{P}\left(\left[S_n^* \leqslant \frac{x - np}{\sqrt{np(1-p)}}\right]\right) \\
&\simeq \mathbb{P}\left(\left[Z \leqslant \frac{x - np}{\sqrt{np(1-p)}}\right]\right) \qquad \textit{(pour une valeur de } n \textit{ suffisamment grande)} \\
&= \mathbb{P}\left(\left[\sqrt{np(1-p)}\, Z + np \leqslant x\right]\right) = F_{\sqrt{np(1-p)}\, Z + np}(x)
\end{aligned}$$

Ainsi, les v.a.r. S_n et $Z_n = \sqrt{np(1-p)}\, Z + np$ suivent approximativement la même loi pour n suffisamment grand.
Comme Z_n est une transformée affine de la v.a.r. Z qui suit une loi normale, Z_n suit une loi normale. On détermine les caractéristiques de cette loi normale en déterminant espérance et variance de Z_n :

$$\begin{aligned}
\mathbb{E}(Z_n) &= \mathbb{E}(\sqrt{np(1-p)}\, Z + np) = \sqrt{np(1-p)}\, \cancel{\mathbb{E}(Z)} + np = np \\
\mathbb{V}(Z_n) &= \mathbb{V}(\sqrt{np(1-p)}\, Z + np) = (\sqrt{np(1-p)})^2\, \mathbb{V}(Z) = np(1-p)
\end{aligned}$$

Ainsi, pour n suffisamment grand, la v.a.r. S_n suit approximativement la loi $\mathcal{N}(np, np(1-p))$. $\square$

Remarque

Considérons n et p définis de telle sorte qu'on peut approcher la v.a.r. S_n (où $S_n \hookrightarrow \mathcal{B}(n,p)$) par la v.a.r. Z_n (avec $Z_n \hookrightarrow \mathcal{N}(np, np(1-p))$).
On devrait alors écrire :

$$\forall k \in [\![0,n]\!], \ \mathbb{P}([S_n = k]) \ \simeq \ \mathbb{P}([Z_n = k])$$

Or, comme Z_n est une v.a.r. à densité, on a : $\mathbb{P}([Z_n = k]) = 0$.
Ainsi, l'approximation ci-dessus n'est pas bonne. On écrira plutôt :

$$\mathbb{P}([S_n = k]) \ \simeq \ \mathbb{P}([k - 0,5 < Z_n < k + 0,5])$$

On appelle cela utiliser la **correction de continuité**.

Exercice

Soit X une v.a.r. qui suit la loi $\mathcal{B}(900, 0.5)$.
Donner une valeur approchée de $\mathbb{P}([405 \leqslant X \leqslant 495])$.

III.3.b) Approximation d'une loi de Poisson par une loi normale

Théorème 8.

Soit $\alpha > 0$.

Soit $(X_n)_{n\in\mathbb{N}^*}$ *une suite de v.a.r. indépendantes.*

On suppose : $\forall n \in \mathbb{N}^*,\ X_n \hookrightarrow \mathcal{P}(\alpha)$.

Pour tout $n \in \mathbb{N}^*$, *on note* $S_n = \sum\limits_{k=1}^{n} X_k$.

1) *Alors pour tout* $n \in \mathbb{N}^*$: $S_n \hookrightarrow \mathcal{P}(n\,\alpha)$.

2) $\boxed{S_n^* = \dfrac{S_n - n\alpha}{\sqrt{n\alpha}} \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} Z, \quad \text{où } Z \hookrightarrow \mathcal{N}(0,1)}$.

3) *Pour n suffisamment grand, la v.a.r.* S_n *suit approximativement la loi* $\mathcal{N}(n\alpha, n\alpha)$.

Démonstration.

1) Déjà vu lors du chapitre sur les couples de v.a.r.

2) C'est une application directe du TCL.

3) Soit $x \in \mathbb{R}$.

$$\begin{aligned}
F_{S_n}(x) &= \mathbb{P}\big([S_n \leqslant x]\big) \\
&= \mathbb{P}\Big(\Big[\frac{S_n - n\alpha}{\sqrt{n\alpha}} \leqslant \frac{x - n\alpha}{\sqrt{n\alpha}}\Big]\Big) \\
&= \mathbb{P}\Big(\Big[S_n^* \leqslant \frac{x - n\alpha}{\sqrt{n\alpha}}\Big]\Big) \\
&\simeq \mathbb{P}\Big(\Big[Z \leqslant \frac{x - n\alpha}{\sqrt{n\alpha}}\Big]\Big) \qquad \textit{(pour une valeur de n suffisamment grande)} \\
&= \mathbb{P}\Big([\sqrt{n\alpha}\, Z + n\alpha \leqslant x]\Big) = F_{\sqrt{n\alpha}\, Z + n\alpha}(x)
\end{aligned}$$

Ainsi, les v.a.r. S_n et $Z_n = \sqrt{n\alpha}\ Z + n\alpha$ suivent approximativement la même loi pour n suffisamment grand.
Comme Z_n est une transformée affine de la v.a.r. Z qui suit une loi normale, Z_n suit une loi normale. On détermine les caractéristiques de cette loi normale en déterminant espérance et variance de Z_n :

$$\begin{aligned}
\mathbb{E}(Z_n) &= \mathbb{E}(\sqrt{n\alpha}\ Z + n\alpha) = \sqrt{n\alpha}\ \cancel{\mathbb{E}(Z)} + n\alpha = n\alpha \\
\mathbb{V}(Z_n) &= \mathbb{V}(\sqrt{n\alpha}\ Z + n\alpha) = (\sqrt{n\alpha})^2\ \mathbb{V}(Z) = n\alpha
\end{aligned}$$

Ainsi, pour n suffisamment grand, la v.a.r. S_n suit approximativement la loi $\mathcal{N}(n\alpha, n\alpha)$. □

Exercice

Soit X une v.a.r. qui suit la loi $\mathcal{P}(64)$.

Donner une valeur approchée de $\mathbb{P}([X \leqslant 74])$.

Exercice (d'après `ESSEC 2007 - Maths II`)

Soit X une v.a.r. telle que $X \hookrightarrow \mathcal{P}(5k)$, avec $k \in \mathbb{N}^*$.
On note α l'unique réel vérifiant $\Phi(\alpha) = 0,99$.

Justifier que $0,01$ est une valeur approchée de $\mathbb{P}\left(\left[X - 5k > \alpha\sqrt{5k}\right]\right)$, lorsque k est grand.

Exercice

On considère un dé équilibré.

L'expérience consiste à effectuer une succession de $n = 3\,600$ lancers.

Pour tout $i \in [\![1, n]\!]$, on note X_i la v.a.r. :

× égale à 1 si on obtient 1 lors du $i^{\text{ème}}$ lancer,

× égale à 0 sinon.

On note enfin $S_n = \sum\limits_{i=1}^{n} X_i$ la v.a.r. qui compte le nombre de 1 obtenu au cours de l'expérience.

1. ***a)*** Soit $i \in [\![1, n]\!]$. Déterminer la loi de X_i.

b) Déterminer la loi de S_n. Rappeler les valeurs de $\mathbb{E}(S_n)$ et $\mathbb{V}(S_n)$.

2. **À l'aide de l'inégalité de Bienaymé-Tchebychev**

a) Démontrer que, pour tout $\varepsilon > 0$:

$$\mathbb{P}([|S_n - 600| > \varepsilon]) \leqslant \frac{500}{\varepsilon^2}$$

b) En déduire que, pour tout $\varepsilon > 0$:

$$\mathbb{P}([600 - \varepsilon \leqslant S_n \leqslant 600 + \varepsilon]) \geqslant 1 - \frac{500}{\varepsilon^2}$$

c) Démontrer enfin : $\mathbb{P}([540 \leqslant S_n \leqslant 660]) \geqslant \dfrac{31}{36}$.

3. **À l'aide de l'inégalité du TCL**

a) Soit $(a, b) \in \mathbb{R}^2$. Expliquer pourquoi on peut considérer :

$$\mathbb{P}\left(\left[nm + a\sigma\sqrt{n} \leqslant S_n \leqslant nm + b\sigma\sqrt{n}\right]\right) \simeq \Phi(b) - \Phi(a)$$

b) En déduire une approximation de $\mathbb{P}([540 \leqslant S_n \leqslant 660])$.
On donne $\dfrac{6}{\sqrt{5}} \simeq 2,68$.

4. Comparer les résultats obtenus par les deux méthodes.

Démonstration.

1. ***a)*** $X_i \hookrightarrow \mathcal{B}\left(\dfrac{1}{6}\right)$.

b)
- L'expérience consiste en la succession de n épreuves de Bernoulli (1 lancer du dé) indépendantes, de même paramètre de succès $\dfrac{1}{6}$ (correspondant à la probabilité d'obtenir 1).
- S_n est la v.a.r. qui compte le nombre de 1 obtenu au cours de l'expérience.

Ainsi : $S_n \hookrightarrow \mathcal{B}\left(n, \dfrac{1}{6}\right)$. On a alors :

$$\mathbb{E}(S_n) = 3\,600 \times \frac{1}{6} = 600 \quad \text{et} \quad \mathbb{V}(S_n) = 3\,600 \times \frac{1}{6} \times \frac{5}{6} = 500$$

2. ***a)*** Soit $\varepsilon > 0$. D'après l'inégalité de Bienaymé-Tchebychev :

$$\begin{array}{ccc} \mathbb{P}([|S_n - \mathbb{E}(S_n)| > \varepsilon]) & \leqslant & \dfrac{\mathbb{V}(S_n)}{\varepsilon^2} \\ \| & & \| \\ \mathbb{P}([|S_n - 600| > \varepsilon]) & & \dfrac{500}{\varepsilon^2} \end{array}$$

b) Remarquons :

$$\begin{array}{rcl} \mathbb{P}([|S_n - 600| > \varepsilon]) & = & 1 - \mathbb{P}([|S_n - 600| \leqslant \varepsilon]) \\ & = & 1 - \mathbb{P}([-\varepsilon \leqslant S_n - 600 \leqslant \varepsilon]) \\ & = & 1 - \mathbb{P}([600 - \varepsilon \leqslant S_n \leqslant 600 + \varepsilon]) \end{array}$$

D'où : $1 - \mathbb{P}([600 - \varepsilon \leqslant S_n \leqslant 600 + \varepsilon]) \leqslant \dfrac{500}{\varepsilon^2}$.

Ainsi : $\mathbb{P}([600 - \varepsilon \leqslant S_n \leqslant 600 + \varepsilon]) \geqslant 1 - \dfrac{500}{\varepsilon^2}$.

c) On choisit alors $\varepsilon = 60$ et on obtient :

$$\mathbb{P}([540 \leqslant S_n \leqslant 660]) \geqslant 1 - \frac{500}{(60)^2} = 1 - \frac{500}{3\,600} = 1 - \frac{5}{36} = \frac{31}{36}$$

Donc la probabilité que le nombre de 1 soit compris entre 540 et 660 est supérieure à $\frac{31}{36}$.

3. a) La suite $(X_i)_{i\in\mathbb{N}^*}$ est constituée de v.a.r. :

× indépendantes (le résultat d'un lancer ne dépend pas des précédents),

× de même espérance $p = \frac{1}{6}$,

× de même variance $\sigma^2 = p\,(1-p) = \frac{5}{36} \neq 0$.

On a de plus :

$$S_n^* = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{E}(S_n)}} = \frac{S_n - n\,m}{\sigma\,\sqrt{n}}$$

Soit $(a,b) \in \mathbb{R}^2$ avec $a < b$. D'après le théorème central limite :

$$\lim_{n\to+\infty} \mathbb{P}([a \leqslant S_n^* \leqslant b]) = \Phi(b) - \Phi(a)$$

On a alors :

$$\begin{array}{cll} & \mathbb{P}\left([nm + a\sigma\sqrt{n} \leqslant S_n \leqslant nm + b\sigma\sqrt{n}]\right) & \\ = & \mathbb{P}\left([a\sigma\sqrt{n} \leqslant S_n - n\,m \leqslant b\sigma\sqrt{n}]\right) & \\ = & \mathbb{P}\left(\left[a \leqslant \dfrac{S_n - n\,m}{\sigma\,\sqrt{n}} \leqslant b\right]\right) & \textit{(car } \sigma\sqrt{n} > 0\textit{)} \\ = & \mathbb{P}\left([a \leqslant S_n^* \leqslant b]\right) & \\ \simeq & \Phi(b) - \Phi(a) & \textit{(pour } n \textit{ suffisamment grand)} \end{array}$$

(on peut aussi remarquer : $S_n = \sigma\,\sqrt{n}\;S_n^* + nm$, remplacer S_n par cette écriture en première ligne puis isoler S_n^*)

b) Ici : $n = 3\,600$, $m = \frac{1}{6}$ et $\sigma = \sqrt{\frac{5}{36}} = \frac{\sqrt{5}}{6}$.
On obtient alors :

$$\Phi(b) - \Phi(a) \simeq \mathbb{P}\left(\left[600 + 10\sqrt{5}a \leqslant S_n \leqslant nm + 10\sqrt{5}b\right]\right)$$

On souhaite connaître une valeur approchée de : $\mathbb{P}\left([540 \leqslant S_n \leqslant 660]\right)$.
On choisit donc a tel que $10\sqrt{5}a = -60$ et $10\sqrt{5}b = 60$.
On obtient $a = -\frac{6}{\sqrt{5}}$ et $b = \frac{6}{\sqrt{5}}$. Finalement, on a donc :

$$\begin{array}{rcl} \mathbb{P}([540 \leqslant S_n \leqslant 660]) & \simeq & \Phi\left(\dfrac{6}{\sqrt{5}}\right) - \Phi\left(-\dfrac{6}{\sqrt{5}}\right) \\ & = & \Phi\left(\dfrac{6}{\sqrt{5}}\right) - \left(1 - \Phi\left(\dfrac{6}{\sqrt{5}}\right)\right) = 2\,\Phi\left(\dfrac{6}{\sqrt{5}}\right) - 1 \end{array}$$

Or, par lecture de la table de la loi $\mathcal{N}(0,1)$:

$$\Phi\left(\frac{6}{\sqrt{5}}\right) \simeq \Phi(2,68) \simeq 0,9963$$

D'où : $\mathbb{P}([540 \leqslant S_n \leqslant 660]) \simeq 0,9926$.

4. En utilisant l'inégalité de Bienaymé-Tchebychev, on a obtenu :

$$\mathbb{P}([540 \leqslant S_n \leqslant 660]) \geqslant \frac{31}{36} \simeq 0,8611$$

On constate que :

× le théorème central limite donne une estimation plus précise de la valeur de $\mathbb{P}([540 \leqslant S_n \leqslant 660])$.

× l'inégalité de Bienaymé-Tchebychev fournit cependant une minoration valable pour toute réalisation de S_n. □