Chapitre 20 : Convergence et approximations

1. Compléments sur les variables aléatoires quelconques

1.1 Loi d'une variable aléatoire

Rappels :
_ Soit (Ω, T) un espace probabilisable. X est une variable aléatoire réelle définie sur (Ω, T) si **X est une application de Ω dans ℝ telle que pour tout élément x de ℝ,**
$\{\omega \in \Omega \ / \ X(\omega) \leq x\} \in T$.
_ Si de plus P est une probabilité sur (Ω, T), on appelle fonction de répartition de X la fonction $\mathbf{F_X}$
$\begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto P(X \leq x) \end{cases}$.

Propriété :

Si X est une VAR et si F_X est sa fonction de répartition :
$\lim\limits_{x \to -\infty} F_X(x) = 0$ **et** $\lim\limits_{x \to +\infty} F_X(x) = 1$
F_X est croissante sur ℝ

Remarque : **la loi d'une VAR est caractérisée par sa fonction de répartition.**

Propriété :

Soit X une VAR et a un réel.
Si $\mathbf{X(\Omega) \subset [a; +\infty[}$ alors $\forall$ **$x < a$, $F_X(x) = 0$**
Si $\mathbf{X(\Omega) \subset]-\infty; b]}$ alors $\forall$ **$x > b$, $F_X(x) = 1$**

Remarque :

Donc si a et b sont deux réels et si $X(\Omega) \subset [a; b]$ alors $F_X(x) = \begin{cases} 0 \text{ si } x < a \\ ? \text{ si } a \leq x \leq b \\ 1 \text{ si } x > b \end{cases}$

Rappel :
Si X et Y sont des VAR, alors λX, X + Y, XY, max(X, Y), min(X, Y) sont des VAR.

Rappel : Soit X et Y deux VAR. On dit que X et Y sont indépendantes si pour tous intervalles I et J de ℝ, $P\big((X \in I) \cap (Y \in J)\big) = P(X \in I)P(Y \in J)$.

Variable discrète / à densité :

Soit X une VAR.
_ **S'il existe une partie I de ℕ ou ℤ telle que : $X(\Omega) = \{x_i, i \in I\}$**, on dit que X est une **variable discrète**.
_ Si **F_X est continue sur ℝ et de classe C^1 sur ℝ sauf en un nombre fini de points**, alors on dit que X est une **variable à densité**. (et dans ce cas $P(X = a) = 0 \ \forall a \in \mathbb{R}$)

Remarques :
_ Si $X(\Omega) = \{x_i, i \in I\}$, avec $I \subset \mathbb{N}$ ou ℤ, on dit que X(Ω) est dénombrable.
Les intervalles de ℝ d'intérieur non vide ne sont pas dénombrables.
_ il existe des variables aléatoires qui ne sont ni discrètes, ni à densité :

Ex : Soit X $\longrightarrow \mathcal{U}(\{0,1\})$ et Y $\longrightarrow \mathcal{U}([0;1])$ avec X et Y indépendantes. Posons Z = XY.

Alors $Z(\Omega) = [0;1]$ (non dénombrable) donc Z n'est pas une variable discrète.
et $P(Z = 0) = P(X = 0) + P((X = 1) \cap (Y = 0)) = \frac{1}{2} + P(X = 1)P(Y = 0) = \frac{1}{2} \neq 0$ donc Z n'est pas une variable à densité.

1.2 Espérance, variance

On admet qu'**on peut étendre les notions d'espérance et de variance** vues pour les variables discrètes et à densité à l'ensemble des variables aléatoires.
Les propriétés sur l'espérance et la variance peuvent alors se généraliser de la manière suivante :

Propriété :

Soit X et Y deux variables aléatoires. Si X et Y admettent une espérance, alors X + Y admet une espérance et **E(X + Y) = E(X) + E(Y)**.

Propriété : ***Croissance de l'espérance***.

Soit X et Y deux variables aléatoires. Si X et Y admettent chacune une espérance et
si $P(X \leq Y) = 1$ alors $E(X) \leq E(Y)$.

Propriété : *Espérance du produit de variables aléatoires indépendantes.*

Soit X et Y deux variables aléatoires. Si X et Y admettent une espérance et sont **indépendantes**, alors XY admet une espérance et **E(XY) = E(X)E(Y)**.

Propriété : *Variance d'une somme de variables aléatoires indépendantes.*

Soit X et Y deux variables aléatoires. Si X et Y sont **indépendantes** et admettent une variance, X + Y admet une variance et **V (X+Y) = V(X) + V(Y).**

Propriété

Soit X une variable aléatoire et a et b deux réels.
_ si X admet une espérance, alors Y = aX + b admet une espérance et
$E(Y) = \mathbf{E(aX + b) = aE(X) + b}$
_ si X admet une variance, alors Y = aX + b admet une variance et
$V(Y) = \mathbf{V(aX + b) = a^2V(X)}$
$\sigma(Y) = \boldsymbol{\sigma(aX + b) = |a|\sigma(X)}$

2. Suites de variables aléatoires discrètes

2.1 Indépendance

Définition :

Soit $n \geq 1$. Soit $(X_1, ..., X_n)$ n variables aléatoires sur un espace probabilité $(\Omega, \mathcal{T}, \mathcal{P})$.
On dit que $(X_1, ..., X_n)$ sont **indépendantes** (ou **mutuellement indépendantes**) si :
pour tous intervalles $I_1, ..., I_n$ de $\mathbb{R}$,

$$P((X_1 \in I_1) \cap ... \cap (X_n \in I_n)) = P(X_1 \in I_1) \times ... \times P(X_n \in I_n).$$

Propriété : **Lemme des coalitions**

Soit $(X_1, ..., X_n)$ n V.A.R. indépendantes. Soit $p \leq n$.
Soit $f : \mathbb{R}^p \longrightarrow \mathbb{R}$ et $g : \mathbb{R}^{n-p} \longrightarrow \mathbb{R}$.
Alors les variables aléatoires $\mathbf{Y = f(X_1, ..., X_p)}$ et $\mathbf{Z = g(X_{p+1}, ..., X_n)}$ sont **indépendantes**.

Ex : Soient X_1, X_2, X_3, X_4 4 variables aléatoires indépendantes.
Alors $Y = X_1 + X_2$ et $Z = X_3 + X_4$ sont indépendantes.

Définition :
Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires.
Si pour tout partie finie I de $\mathbb{N}$, les V.A.R. $(X_i)_{i \in I}$ sont mutuellement indépendants, alors on dit que la suite $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires indépendantes.

Remarque :
En Python, pour simuler n variables aléatoires indépendantes qui suivent la même loi, il suffit d'ajouter le paramètre n à la fin de la liste des paramètres. On obtient ainsi un tableau de n coefficients indépendants qui suivent tous la même loi.

Ex :
Simuler 10 fois la loi binomiale B(5,1/3) :

```
import numpy.random as rd
x=rd.binomial(5,1/3,10)
```

2.2 Minimum, maximum de variables aléatoires

Propriété

Soit $(X_1, ..., X_n)$ n variables aléatoires.
_ Soit $\mathbf{Y_n = max(X_1, ..., X_n)}$. Alors :

$\forall x \in \mathbb{R}$, $\mathbf{(Y_n \leq x) = (X_1 \leq x) \cap ... \cap (X_n \leq x)}$ (idem avec <)

_ Soit $\mathbf{Z_n = min(X_1, ..., X_n)}$. Alors :

$\forall x \in \mathbb{R}$, $\mathbf{(Z_n > x) = (X_1 > x) \cap ... \cap (X_n > x)}$ (idem avec ≥)

2.3 Somme de n variables aléatoires

Remarque :
Soit n épreuves aléatoires (indépendantes ou non).

Pour tout $i \in \{1, \dots, n\}$, on définit la VAR X_i par : $X_i = \begin{cases} 1 \text{ si succès} \\ 0 \text{ si échec} \end{cases}$

Soit $X = X_1 + \dots + X_n$. Alors X représente le nombre de succès lors de ces n épreuves.

Les propriétés suivantes généralisent les propriétés vues pour n = 2 :

Propriété :

Soit $p \in]0;1[$. Soit $(X_1, \dots, X_n)$ n variables aléatoires **indépendantes** qui suivent toutes la loi de **Bernoulli** de paramètre p.
Alors $\mathbf{X = X_1 + \dots + X_n}$ **suit la loi** $\boldsymbol{\mathcal{B}(n, p)}$.

Propriété :

Soit $p \in]0;1[$ et $(n_1, \dots, n_k)$ des entiers naturels. Soit $(X_1, \dots, X_k)$ k variables aléatoires **indépendantes** telles que $\forall\, i \in \{1, \dots, k\}$, X_i suit la loi $\boldsymbol{\mathcal{B}(n_i, p)}$.
Alors $\mathbf{X = X_1 + \dots + X_k}$ **suit la loi** $\boldsymbol{\mathcal{B}(n_1 + \dots + n_k, p)}$.

Propriété :

Soit $\lambda_1, \dots, \lambda_n$ n réels strictement positifs. Soit $(X_1, \dots, X_n)$ n variables aléatoires **indépendantes** telles que $\forall\, i \in \{1, \dots, n\}$, X_i suit la loi $\boldsymbol{\mathcal{P}(\lambda_i)}$.
Alors $\mathbf{X = X_1 + \dots + X_n}$ **suit la loi** $\boldsymbol{\mathcal{P}(\lambda_1 + \dots + \lambda_n)}$.

Propriété :

Soit $(m_1, \dots, m_n) \in \mathbb{R}^n$ et $(\sigma_1, \dots, \sigma_n) \in (\mathbb{R}^+)^n$. Soit $(X_1, \dots, X_n)$ n variables aléatoires **indépendantes** telles que : $\forall\, i \in \{1, \dots, n\}$, $X_i \longrightarrow \boldsymbol{\mathcal{N}(m_i, \sigma_i^2)}$.
Alors $X = X_1 + \dots + X_n$ suit la loi $\boldsymbol{\mathcal{N}\left(m_1 + \dots + m_n, \sigma_1^2 + \dots + \sigma_n^2\right)}$

Remarque :
_ Autrement dit, la somme de lois binomiales (de Poisson, normales) indépendantes suit une loi binomiale (de Poisson, normale).
_ cette propriété est fausse pour les lois uniformes, géométriques, exponentielles, …

Rappel : Linéarité de l'espérance

Soit $(X_1, \dots, X_n)$ n variables aléatoires discrètes qui admettent chacune une espérance.
Alors $X = X_1 + \dots + X_n$ admet une espérance et

$$\mathbf{E(X) = E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n)}.$$

Propriété :

Soit $(X_1, \dots, X_n)$ n variables aléatoires **indépendantes** qui admettent chacune une variance. Alors $X = X_1 + \dots + X_n$ admet une variance et :

$$\mathbf{V(X) = V(X_1 + \dots + X_n) = \sum_{i=1}^{n} V(X_i)}$$

Exemple :
Soit $p \in]0;1[$. Soit $X_1, \dots, X_n$ n VAR indépendantes qui suivent la loi géométrique $G(p)$.
On pose $Y_n = \dfrac{X_1 + \dots + X_n}{n}$. $E(Y_n)$? $V(Y_n)$?

$Y_n = \dfrac{1}{n}(X_1 + \dots + X_n)$ donc par linéarité de l'espérance,

$$E(Y_n) = \frac{1}{n}\big(E(X_1) + \dots + E(X_n)\big) = \frac{1}{n} \times \frac{n}{p} = \frac{1}{p}$$

$$V(Y_n) = \frac{1}{n^2} V(X_1 + \dots + X_n)$$

(par indépendance) $= \dfrac{1}{n^2} \times (V(X_1) + \dots + V(X_n)) = \dfrac{1}{n^2} \times \dfrac{nq}{p^2} = \dfrac{q}{np^2}$.

Rappel :
Si (X_1, X_2) sont 2 VAR discrètes qui admettent chacune une variance et telles que $\text{cov}(X_1, X_2)$ existe, alors $X = X_1 + X_2$ admet une variance et :

$$V(X) = V(X_1 + X_2) = V(X_1) + V(X_2) + 2\text{cov}(X_1, X_2).$$

(formule qui rappelle $(a_1 + a_2)^2 = a_1^2 + a_2^2 + 2a_1a_2$)

Remarque : Soit $(a_1, \dots, a_n)$ n réels. Alors $(a_1 + \dots + a_n)^2 = \sum_{i=1}^{n} a_i^2 + \sum_{1 \le i < j \le n} 2a_ia_j$.

Propriété :

Soit $(X_1, \dots, X_n)$ n variables aléatoires **discrètes** qui admettent chacune une variance et telles que $\forall\ i, j \in \{1, \dots, n\}^2$, $\text{cov}(X_i, X_j)$ existe.
Alors $X = X_1 + \dots + X_n$ admet une variance et :

$$\mathbf{V(X) = V(X_1 + \dots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \le i < j \le n} \text{cov}(X_i, X_j)}.$$

Exemple :
$V(X_1 + X_2 + X_3) = V(X_1) + V(X_2) + V(X_3) + 2\text{cov}(X_1, X_2) + 2\text{cov}(X_1,X_3) + 2\text{cov}(X_2, X_3)$.

3. Convergences

3.1 Inégalités de Markov et de Bienaymé Tchebychev

Propriété : *Inégalité de Markov*

Soit X une variable aléatoire **à valeurs positives** et **admettant une espérance**.
Alors $\forall$ a > 0, $\mathbf{P(X \geq a) \leq \dfrac{E(X)}{a}}$.

Démonstration : (dans le cas où X est une V.A.R. discrète)
Notons $X(\Omega) = \{x_i, i \in I\}$.

Alors $E(X) = \sum_{i \in I} x_i P(X = x_i) = \sum_{i \in I / x_i < a} x_i P(X = x_i) + \sum_{i \in I / x_i \geq a} x_i P(X = x_i)$

$\sum_{i \in I / x_i < a} x_i P(X = x_i) \geq 0$ (somme de termes positifs)

$\forall i / x_i \geq a$, $x_i P(X = x_i) \geq aP(X = x_i)$ donc $\sum_{i \in I / x_i \geq a} x_i P(X = x_i) \geq \sum_{i \in I / x_i \geq a} aP(X = x_i)$

$\sum_{i \in I / x_i \geq a} x_i P(X = x_i) \geq a \sum_{i \in I / x_i \geq a} P(X = x_i)$ $\quad \sum_{i \in I / x_i \geq a} x_i P(X = x_i) \geq aP(X \geq a)$

Donc $E(X) \geq aP(X \geq a)$ $\quad$ a > 0 donc $P(X \geq a) \leq \dfrac{E(X)}{a}$.

Propriété : *Inégalité de Bienaymé-Tchebychev*

Soit X une variable aléatoire, qui admet une espérance E(X) et une variance V(X).
Alors : $\forall \varepsilon > 0$,

$$\mathbf{P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}}. \qquad \mathbf{P(|X - E(X)| \leq \varepsilon) \geq 1 - \frac{V(X)}{\varepsilon^2}}$$

Démonstration : Posons $Y = (X - E(X))^2$.
X admet une variance, donc Y admet une espérance et E(Y) = V(X)
De plus Y est à valeurs positives, donc d'après l'inégalité de Markov avec $a = \varepsilon^2$:
$P(Y \geq \varepsilon^2) \leq \dfrac{E(Y)}{\varepsilon^2}$ $\qquad P\left((X - E(X))^2 \geq \varepsilon^2\right) \leq \dfrac{V(X)}{\varepsilon^2}$ $\qquad P\left(|X - E(X)| \geq \varepsilon\right) \leq \dfrac{V(X)}{\varepsilon^2}$.
L'autre inégalité s'obtient en passant à l'événement contraire.

Remarques :
_ $|X - E(X)| \leq \varepsilon \Leftrightarrow -\varepsilon \leq X - E(X) \leq \varepsilon \Leftrightarrow E(X) - \varepsilon \leq X \leq E(X) + \varepsilon$

Donc l'inégalité peut aussi s'écrire : $P(E(X) - \varepsilon \leq X \leq E(X) + \varepsilon) \geq 1 - \dfrac{V(X)}{\varepsilon^2}$

_ l'inégalité de Bienaymé-Tchebychev donne des inégalités très larges, qui donnent des résultats pas souvent optimaux, mais qui ont l'avantage d'exister.

Ex : Soit $n \geq 1$. On lance n fois un dé équilibré.
Soit X le nombre de 6 obtenus, et F la fréquence d'apparition du 6.
1) Déterminer E(F) et V(F).
2) Majorer : $P\left(\left|F - \frac{1}{6}\right| \geq \frac{1}{100}\right)$

1) Il y a n lancers indépendantes, X compte le nombre de 6 obtenus et la probabilité de faire 6 est 1/6.
Donc $X \longrightarrow \mathcal{B}(n, 1/6)$. Donc $E(X) = \frac{n}{6}$ $V(X) = n \times \frac{1}{6} \times \frac{5}{6} = \frac{5n}{36}$.
Comme $F = \frac{X}{n} = \frac{1}{n} X$ on a : $E(F) = \frac{1}{n} E(X) = \frac{1}{6}$ $V(F) = \frac{1}{n^2} V(X) = \frac{5}{36n}$.
2) Inégalité de Bienaymé Tchebychev pour F : $\forall \varepsilon > 0$, $P(|F - E(F)| \geq \varepsilon) \leq \frac{V(F)}{\varepsilon^2}$

Donc $P\left(\left|F - \frac{1}{6}\right| \geq \varepsilon\right) \leq \frac{5}{36n\varepsilon^2}$

Avec $\varepsilon = \frac{1}{100}$: $P\left(\left|F - \frac{1}{6}\right| \geq \frac{1}{100}\right) \leq \frac{5}{36n} \times 100^2 = \frac{50000}{36n}$

Remarque : L'inégalité de Bienaymé-Tchebychev est intéressante d'un point de vue théorique, mais numériquement peu intéressante.

3.2 Loi faible des grands nombres

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

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables indépendantes, admettant une même espérance m, et une même variance V.

Soit $\overline{X_n} = \dfrac{X_1 + \ldots + X_n}{n}$.

Alors $\forall\, \varepsilon > 0$, $\lim\limits_{n \to +\infty} P(\left|\overline{X_n} - m\right| \geq \varepsilon) = 0$ et $\lim\limits_{n \to +\infty} P(\left|\overline{X_n} - m\right| < \varepsilon) = 1$.

Démonstration :

Par linéarité : $E(\overline{X_n}) = \dfrac{E(X_1) + \ldots + E(X_n)}{n} = \dfrac{m \times n}{n} = m$

$V(\overline{X_1}) = \dfrac{1}{n^2} V(X_1 + \ldots + X_n) = \dfrac{1}{n^2} (V(X_1) + \ldots + V(X_n))$ (indép.) $= \dfrac{1}{n^2} nV = \dfrac{V}{n}$.

D'après l'inégalité de Bienaymé-Tchebychev :

$P(\left|\overline{X_n} - E(\overline{X_n})\right| \geq \varepsilon) \leq \dfrac{V(\overline{X_n})}{\varepsilon^2}$ $\qquad 0 \leq P(\left|\overline{X_n} - m\right| \geq \varepsilon) \leq \dfrac{V}{n\varepsilon^2}$

$\lim\limits_{n \to +\infty} \dfrac{V}{n\varepsilon^2} = 0$ donc d'après le théorème des gendarmes, $\lim\limits_{n \to +\infty} P(\left|\overline{X_n} - m\right| \geq \varepsilon) = 0$.

$P(\left|\overline{X_n} - m\right| < \varepsilon) = 1 - P(\left|\overline{X_n} - m\right| \geq \varepsilon)$ donc $\lim\limits_{n \to +\infty} P(\left|\overline{X_n} - m\right| < \varepsilon) = 1$.

Remarques :

_ on dit que la suite $(\overline{X_n})_{n \in \mathbb{N}}$ converge en probabilité vers m.

_ Si on répète indéfiniment une même épreuve de manière indépendante, et si X_n est une variable aléatoire liée à cette épreuve, alors $\overline{X_n}$ est la moyenne des valeurs pour les n premières épreuves.

On a montré ici que, d'une certaine manière, **la moyenne sur un grand nombre d'épreuves "tend vers" l'espérance de X_n.**

En particulier, si X_n est une variable de Bernoulli associé à un succès de probabilité p, $\overline{X_n}$ est la fréquence d'apparition du succès, qui "tend vers" sa probabilité p.

Exemple :

En Python, on écrit le programme suivant :

```
import numpy as np
import numpy.random as rd
x=rd.geometric(1/2,1000)
y=np.mean(x)
```

A quelle valeur peut-on s'attendre environ pour y ?

x est un tableau contenant 1000 variables aléatoires indépendantes ($X_1, \ldots, X_{1000}$) qui suivent toutes la même loi G(1/2).

D'après la loi faible des grands nombres, $y = \dfrac{1}{1000}(X_1 + \ldots + X_{1000})$ est donc proche de l'espérance commune $E(X_1) = \dfrac{1}{1/2} = 2$.

3.3 Convergence en loi

Rappel : Si X suit la loi certaine de paramètre a, alors : $F_X(x) = \begin{cases} 0 \text{ si } x < a \\ 1 \text{ si } x \geq a \end{cases}$

Définition :

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire définis sur le même espace probabilisé $(\Omega, \mathcal{T}, \mathcal{P})$.
On note F_{X_n} et F_X leurs fonctions de répartition.
On dit que la suite $(X_n)_{n \in \mathbb{N}}$ **converge en loi vers X**, si en tout point x de $\mathbb{R}$ où F_X est continue, on a :
$\boldsymbol{\lim\limits_{n \to +\infty} F_{X_n}(x) = F_X(x)}$.
On note : $X_n \xrightarrow{L} X$

Remarques :
_ Attention, x est fixe, et n tend vers + ∞ !
_ Dans une inégalité entre x et n, pensez à isoler n pour trouver la limite

Exemple :
Soit $(X_n)_{n \geq 1}$ une suite de V.A.R. telle que $\forall\, n \geq 1$, $X_n \longrightarrow \mathcal{E}(n)$.
Soit X la variable certaine égale à 0. (P(X = 0) = 1).
Montrer que (X_n) converge en loi vers X.

$F_X(x) = P(X \leq x) = \begin{cases} 0 \text{ si } x < 0 \\ 1 \text{ si } x \geq 0 \end{cases}$. ($F_X$ continue sur $\mathbb{R}$ sauf en 0).
$F_{X_n}(x) = \begin{cases} 0 \text{ si } x < 0 \\ 1 - e^{-nx} \text{ si } x \geq 0 \end{cases}$
_ si x < 0 : $\lim\limits_{n \to +\infty} F_{X_n}(x) = 0 = F_X(x)$
_ si x > 0, $\lim\limits_{n \to +\infty} F_{X_n}(x) = \lim\limits_{n \to +\infty} 1 - e^{-nx} = 1 = F_X(x)$ $\left(\lim\limits_{n \to +\infty} -nx = -\infty\right)$
Donc $\forall\, x \neq 0$, $\lim\limits_{n \to +\infty} F_{X_n}(x) = F_X(x)$
Donc (X_n) converge en loi vers X.

Remarque :
Si (X_n) converge en loi vers X et si a et b sont deux points de continuité de F_X, alors
$\boldsymbol{\lim\limits_{n \to +\infty} P(a < X_n \leq b) = P(a < X \leq b)}$

Démonstration : $P(a < X_n \leq b) = F_{X_n}(b) - F_{X_n}(a)$
et $\lim\limits_{n \to +\infty} F_{X_n}(b) - F_{X_n}(a) = F_X(b) - F_X(a) = P(a < X \leq b)$

Propriété :

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire définies sur le même espace probabilisé $(\Omega, \mathcal{T}, \mathcal{P})$.
On suppose que les $(X_n)_{n \in \mathbb{N}}$ et X sont des **variables discrètes à valeurs dans $\mathbb{N}$**. Alors :
$(X_n)_{n \in \mathbb{N}}$ converge en loi vers X si et seulement si $\forall k \in \mathbb{N}$, $\displaystyle\lim_{n \to +\infty} \mathbf{P(X_n = k) = P(X = k)}$.

Remarque : Attention, **k est fixe et n tend vers + ∞** !

Ex : (Loi géométrique tronquée).
Soit $n \geq 1$. On lance n fois une pièce dont la probabilité de faire pile est $p \in]0;1[$.
Soit X_n la variable aléatoire égale au rang du premier pile s'il y en a, et à n + 1 sinon.
1) Déterminer la loi de X_n.
2) Montrer que la suite (X_n) converge en loi vers une loi usuelle dont on précisera le paramètre.

1) $X_n(\Omega) = \{1, \ldots, n + 1\}$
$\forall k \in \{1, \ldots, n\}$, $(X_n = k) = F_1 \cap \ldots \cap F_{k-1} \cap P_k$ donc par indépendance, $P(X_n = k) = q^{k-1}p$.

$P(X_n = n+1) = P(F_1 \cap \ldots \cap F_n) = q^n$. $\qquad P(X_n = k) = \begin{cases} q^{k-1}p \text{ si } k \leq n \ (n \geq k) \\ q^n \text{ si } k = n + 1 \ (n = k - 1) \\ 0 \text{ si } k > n + 1 \ (n < k - 1) \end{cases}$

2) Soit $k \in \mathbb{N}^*$. Pour tout $n \geq k$, $P(X_n = k) = q^{k-1}p$.
Donc $\displaystyle\lim_{n \to +\infty} P(X_n = k) = q^{k-1}p$.
Donc (X_n) converge en loi vers une variable aléatoire X qui suit la loi géométrique $\mathcal{G}(p)$.

4. Exemples de convergence et approximations

4.1 Loi binomiale ⟶ Loi de Poisson

Propriété

> Soit $\lambda \in [0;1]$ et $n \in \mathbb{N}^*$.
>
> Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires de loi binomiale $\mathcal{B}\left(n, \frac{\lambda}{n}\right)$.
>
> Alors $(X_n)_{n \geq 1}$ converge en loi vers une variable aléatoire X qui suit la loi de Poisson $\mathcal{P}(\lambda)$.

Démonstration :
$(X_n)_{n \geq 1}$ et X sont à valeurs dans $\mathbb{N}$. Il suffit de montrer que $\forall k \in \mathbb{N}$, $\lim\limits_{n \to +\infty} P(X_n = k) = P(X = k)$.

Soit $k \in \mathbb{N}$. Pour $n < k$, $P(X_n = k) = 0$.

$\forall\ n \geq k$, $P(X_n = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n!}{k!(n-k)!} \times \frac{\lambda^k}{n^k} \times \exp\left((n-k)\ln\left(1 - \frac{\lambda}{n}\right)\right)$

$\frac{n!}{k!(n-k)!} = \frac{n(n-1)\ldots(n-k+1)}{k!}$ $\quad n - 1 \sim_{+\infty} n \quad n - 2 \sim_{+\infty} n \quad \ldots\ n - k + 1 \sim_{+\infty} n$

$\frac{n(n-1)(n-k+1)}{k!} \sim \frac{n^k}{k!}$ $\quad \frac{n!}{k!(n-k)!} \times \frac{\lambda^k}{n^k} \sim_{+\infty} \frac{\lambda^k}{k!}$

$\lim\limits_{n \to +\infty} \frac{\lambda}{n} = 0$ donc $\ln\left(1 - \frac{\lambda}{n}\right) \sim_{+\infty} -\frac{\lambda}{n}$ $\quad n - k \sim_{+\infty} n$ donc $(n-k)\ln\left(1 - \frac{\lambda}{n}\right) \sim_{+\infty} -\lambda$

Donc $\lim\limits_{n \to +\infty} \exp\left((n-k)\ln\left(1 - \frac{\lambda}{n}\right)\right) = e^{-\lambda}$ $\quad \lim\limits_{n \to +\infty} P(X_n = k) = \frac{\lambda^k e^{-\lambda}}{k!} = P(X = k)$.

Remarque :

Pour n assez grand, on peut donc approcher loi $\mathcal{B}(n, \lambda/n)$ par $\mathcal{P}(\lambda)$. De plus $p = \frac{\lambda}{n} \Leftrightarrow \lambda = np$

Conclusion :

> Pour n assez grand et p assez petit, on peut approcher la loi $\boldsymbol{\mathcal{B}(n,p)}$ par la loi $\boldsymbol{\mathcal{P}(\lambda)}$ avec $\boldsymbol{\lambda = np}$

Remarques :

_ On considère en général qu'on peut faire cette approximation lorsque : $\begin{cases} p \leq 0{,}1 \\ n \geq 30 \\ np < 15 \end{cases}$

_ si $X \longrightarrow \mathcal{B}(n,p)$ alors $E(X) = np$ $\quad$ si $X \longrightarrow \mathcal{P}(np)$ alors $E(X) = np$. **On conserve l'espérance.**

_ La loi de Poisson est parfois appelée "loi des événements rares" : on considère une population de grande taille (n grand) et on étudie un événement rare (p petit). Dans ce cas, on peut approcher $\mathcal{B}(n,p)$ par $\mathcal{P}(np)$.

Ex : Soit $X \longrightarrow \mathcal{B}(100, 0{,}05)$. $P(X = 3)$?

$\begin{cases} p \leq 0{,}1 \\ n \geq 30 \\ np = 5 < 15 \end{cases}$ On peut approcher X par $Y \longrightarrow \mathcal{P}(\lambda)$ avec $\lambda = np = 5$.

$P(X = 3) \approx P(Y = 3) = \frac{5^3 e^{-5}}{3!} = 0{,}140373\ldots$

Erreur commise : $P(X = 3) = \binom{100}{3} 0{,}05^3 0{,}95^{97} = 0{,}139575...$ $\quad$ Erreur : 0,57 %

4.2 Théorème de la limite centrée

Théorème de la limite centrée :

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires mutuellement indépendantes, de même loi, et admettant une espérance m et une variance non nulle σ^2.

On note $\mathbf{\overline{X_n} = \dfrac{X_1 + \ldots + X_n}{n}}$ et $\mathbf{\overline{X_n}^* = \dfrac{\overline{X_n} - E(\overline{X_n})}{\sigma(\overline{X_n})}}$.

Alors $\mathbf{\overline{X}_n^*}$ **converge en loi** vers une variable aléatoire Y de **loi normale $\mathcal{N}$(0,1)**.

De même, si on pose $\mathbf{S_n = X_1 + \ldots + X_n}$ et $S_n^* = \dfrac{S_n - E(S_n)}{\sigma(S_n)}$

Alors S_n^* converge en loi vers une variable aléatoire Y de **loi normale $\mathcal{N}$(0,1)**.

Remarques :

_ Posons $S_n = \sum_{k=1}^{n} X_k$. Alors $E(S_n) = \sum_{k=1}^{n} E(X_k) = mn$

$V(S_n) = \sum_{k=1}^{n} V(X_k)$ (car $X_1, \ldots, X_n$ indépendants) $= n\sigma^2$

$\overline{X_n} = \dfrac{S_n}{n}$ donc $E(\overline{X_n}) = \dfrac{E(S_n)}{n} = m$ et $V(\overline{X_n}) = \dfrac{1}{n^2} V(S_n) = \dfrac{\sigma^2}{n}$. $\sigma(\overline{X_n}) = \sqrt{\dfrac{\sigma^2}{n}} = \dfrac{\sigma}{\sqrt{n}}$.

Donc $\mathbf{\overline{X_n}^* = \sqrt{n}\dfrac{\overline{X_n} - m}{\sigma}}$.

_ On a $S_n = n\overline{X_n}$ donc $S_n^* = \dfrac{S_n - E(S_n)}{\sigma(S_n)} = \dfrac{n\overline{X_n} - nE(\overline{X_n})}{n\sigma(\overline{X_n})} = \dfrac{\overline{X_n} - E(\overline{X_n})}{\sigma(\overline{X_n})} = \overline{X_n}^*$.

_ Φ est continue sur $\mathbb{R}$,

on a donc $\forall$ a < b, $\mathbf{\lim\limits_{n \to +\infty} P(a < \overline{X_n}^* \leq b) = P(a \leq Y \leq b) = \int_a^b \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \Phi(b) - \Phi(a)}$

$\lim\limits_{n \to +\infty} (a < \sqrt{n}\dfrac{\overline{X_n} - m}{\sigma} \leq b) = \Phi(b) - \Phi(a)$

$\lim\limits_{n \to +\infty} \left(m + \dfrac{a\sigma}{\sqrt{n}} < \overline{X}_n \leq m + \dfrac{b\sigma}{\sqrt{n}} \right) = \Phi(b) - \Phi(a)$

$\lim\limits_{n \to +\infty} \left(mn + a\sigma\sqrt{n} < S_n \leq mn + b\sigma\sqrt{n} \right) = \Phi(b) - \Phi(a)$.

Exemple :

Soit $(X_n)_{n \geq 1}$ une suite de VAR indépendantes qui suivent toutes la loi uniforme sur [0;1].

On pose $Y_n = \dfrac{X_1 + \ldots + X_n}{n}$.

1) Déterminer $E(Y_n)$, $\sigma(Y_n)$.

2) Déterminer $\lim\limits_{n \to +\infty} P\left(\dfrac{1}{2} - \dfrac{1}{\sqrt{12n}} \leq Y_n \leq \dfrac{1}{2} + \dfrac{1}{\sqrt{12n}}\right)$.

1) $E(X_1) = \dfrac{0+1}{2} = \dfrac{1}{2}$ $\quad V(X_1) = \dfrac{(1-0)^2}{12} = \dfrac{1}{12}$.

$E(Y_n) = \dfrac{1}{n} E(X_1 + \ldots + X_n) = \dfrac{1}{n} \times n \times \dfrac{1}{2} = \dfrac{1}{2}$. (par linéarité)

$V(Y_n) = \dfrac{1}{n^2}\big(V(X_1) + \ldots + V(X_n)\big)$ (par indépendance) $= \dfrac{1}{n^2}\left(n \times \dfrac{1}{12}\right) = \dfrac{1}{12n}$.

2) (X_n) est une suite de VAR indépendantes, de même loi, et qui admettent une espérance et une variance.

D'après le théorème de la limite centrée : $Y_n^* = \dfrac{Y_n - E(Y_n)}{\sigma(Y_n)} = \sqrt{12n}\left(Y_n - \dfrac{1}{2}\right)$ converge en loi vers $X \longrightarrow \mathcal{N}(0,1)$.

$P\left(\dfrac{1}{2} - \dfrac{1}{\sqrt{12n}} \leq Y_n \leq \dfrac{1}{2} + \dfrac{1}{\sqrt{12n}}\right) = P\left(- \dfrac{1}{\sqrt{12n}} \leq Y_n - \dfrac{1}{2} \leq \dfrac{1}{\sqrt{12n}}\right) = P(-1 \leq Y_n^* \leq 1)$

$\lim\limits_{n \to +\infty} P(-1 \leq Y_n^* \leq 1) = P(-1 \leq X \leq 1) = 2\Phi(1) - 1 \approx 2\times0{,}8413 - 1 = 0{,}6826$

Approximation :

_ Pour $n \geq 30$, on considère qu'on peut donc approcher la loi de $\overline{X_n}^* = \dfrac{\overline{X_n} - E(\overline{X_n})}{\sigma(\overline{X_n})}$ par la loi $\mathcal{N}(0,1)$.

_ Rappel : $X \longrightarrow \mathcal{N}(m, \sigma^2) \Leftrightarrow X^* = \dfrac{X - m}{\sigma} \longrightarrow \mathcal{N}(0,1)$.

Donc on peut approcher $\overline{X_n}$ par la loi $\mathcal{N}(E(\overline{X_n}), V(\overline{X_n}))$.

S_n par la loi $\mathcal{N}(E(S_n), V(S_n))$.

Ex :

Soit $(X_n)_{n \geq 1}$ une suite de V.A. indépendantes de même loi exponentielle de paramètre $\lambda = 1$.

Soit $S_n = \sum_{k=1}^{n} X_k$. $P(105 < S_{100} \leq 110)$?

$E(X_n) = 1$ $V(X_n) = 1$ donc $E(S_n) = \sum_{k=1}^{n} E(X_k) = \sum_{k=1}^{n} 1 = n$.

Par indépendance, $V(S_n) = \sum_{k=1}^{n} V(X_k) = \sum_{k=1}^{n} 1 = n$ donc $\sigma(S_n) = \sqrt{n}$

Alors $S_n^* = \dfrac{S_n - E(S_n)}{\sigma(S_n)} = \dfrac{S_n - n}{\sqrt{n}}$ converge en loi vers une V.A. X de loi normale $\mathcal{N}(0,1)$.

$100 \geq 30$ donc on peut approcher $S_{100}^* = \dfrac{S_{100} - 100}{10}$ par X.

$P(105 < S_{100} \leq 110) = P(5 < S_{100} - 100 \leq 10) = P(0,5 < S_{100}^* \leq 1)$
$\approx P(0,5 < X \leq 1) = \Phi(1) - \Phi(0,5) = 0,8413 - 0,6915 = 0,1498$

4.3 Loi binomiale ⟶ Loi normale

Principe :
Soit $p \in]0;1[$. Soit $(X_n)_{n \geq 1}$ une suite de V.A.R. indépendantes de même loi de Bernoulli $\mathcal{B}(p)$. Leur espérance est $m = p$, leur écart-type est $\sigma = \sqrt{p(1-p)}$.

Pour $n \geq 1$, soit $S_n = \sum_{k=1}^{n} X_k$.

On sait que S_n suit la loi binomiale de paramètres n et p. Donc $E(S_n) = np$ $V(S_n) = np(1-p)$

D'après le théorème de la limite centrée, $S_n^* = \dfrac{S_n - E(S_n)}{\sigma(S_n)} = \dfrac{S_n - np}{\sqrt{p(1-p)}\sqrt{n}}$ converge en loi vers une V.A.R. X de loi normale $\mathcal{N}(0,1)$.

On peut approcher S_n^* par une V.A.R. de loi normale $\mathcal{N}(0,1)$.

On peut donc approcher S_n par une V.A.R. de loi normale $\mathcal{N}(np, np(1-p))$.

A retenir :

On peut approcher **$\mathcal{B}(n,p)$ par $\mathcal{N}(np, np(1-p))$** (condition : si $\begin{cases} n \geq 30 \\ np \geq 15 \\ np(1-p) > 5 \end{cases}$)

Remarques :

_ si $X \longrightarrow \mathcal{B}(n,p)$ $E(X) = np$ $V(X) = np(1-p)$
si $X \longrightarrow \mathcal{N}(np, np(1-p))$, $E(X) = np$ $V(X) = np(1-p)$.

On conserve l'espérance et la variance.

_ Correction de continuité :

Si $X \longrightarrow \mathcal{B}(n,p)$, X prend des valeurs entières entre 0 et n.
Pour approcher X, on remplace :
$P(X = 0)$ par $P(X \leq 0{,}5)$
$P(X = k)$ (pour $k \in \{1, \dots, n-1\}$ par $P(k - 0{,}5 < X \leq k + 0{,}5)$
$P(X = n)$ par $P(X > n - 0{,}5)$.

Ex :
On lance 100 fois une pièce équilibrée. Soit X le nombre de piles obtenus. $P(X = 53)$?

$X \longrightarrow \mathcal{B}(100, 1/2)$. $n = 100 \geq 30$;$np = 50 \geq 15$ $np(1-p) = 50 \times \dfrac{1}{2} = 25 > 5$.

$E(X) = 50$. $V(X) = 25$. Donc on peut approcher X par Y qui suit la loi normale $N(50, 25)$.
$\sigma = 5$
$P(X = 53) = P(52{,}5 < X \leq 53{,}5)$ est approché par $P(52{,}5 < Y \leq 53{,}5)$.

$Y^* = \dfrac{Y - 50}{5} \longrightarrow N(0,1)$

$P(52{,}5 < Y \leq 53{,}5) = P(2{,}5 < Y - 50 \leq 3{,}5) = P(0{,}5 < Y^* \leq 0{,}7) = \Phi(0{,}7) - \Phi(0{,}5)$
$= 0{,}7580 - 0{,}6915 = 0{,}0665$ $P(X = 23) \approx 0{,}0665$

(remarque $P(X = 53) = \dbinom{100}{53} \left(\dfrac{1}{2}\right)^{53} \left(\dfrac{1}{2}\right)^{47} \approx 0{,}0665905$)

4.4 Loi de Poisson -> Loi normale

Principe :
Soit $\lambda > 0$. Soit $(X_n)_{n \geq 1}$ une suite de V.A.R. indépendantes de même loi de Poisson $\mathcal{P}(\lambda_0)$. Leur espérance est $m = \lambda_0$, leur écart-type est $\sigma = \sqrt{\lambda_0}$.

Pour $n \geq 1$, soit $S_n = \sum_{k=1}^{n} X_k$.

On sait que S_n suit la loi de Poisson de paramètre $n\lambda_0$. $E(S_n) = n\lambda_0$ $V(S_n) = n\lambda_0$

D'après le théorème de la limite centrée, $S_n^* = \dfrac{S_n - E(S_n)}{\sigma(S_n)} = \dfrac{S_n - n\lambda}{\sqrt{\lambda}\sqrt{n}}$ converge en loi vers une V.A.R. X de loi normale $\mathcal{N}(0,1)$.

On peut approcher S_n^* par une V.A.R. de loi normale $\mathcal{N}(0,1)$.
Par transformation affine, on peut approcher S_n par une V.A.R. de loi normale $\mathcal{N}(n\lambda_0, n\lambda_0)$.
En posant $\lambda = n\lambda_0$, on obtient :

A retenir :

Pour λ suffisamment grand, on peut approcher **$\mathcal{P}(\lambda)$ par $\mathcal{N}(\lambda, \lambda)$** (Exemple de condition : $\lambda \geq 18$)

Remarques :
_ si $X \longrightarrow \mathcal{P}(\lambda)$ $E(X) = \lambda$ $V(X) = \lambda$ si $X \longrightarrow \mathcal{N}(\lambda, \lambda)$, $E(X) = \lambda$ $V(X) = \lambda$.
On conserve l'espérance et la variance.
_ comme pour la loi binomiale, on peut utiliser la correction de continuité, en remplaçant
$P(X = 0)$ par $P(X \leq 0,5)$ et $\forall\, k \geq 1$, $P(X = k)$ par $P\left(k - \dfrac{1}{2} \leq X \leq k + \dfrac{1}{2}\right)$

Exemple :
Soit $X \longrightarrow \mathcal{P}(100)$. $P(X \leq 90)$?

$\lambda \geq 18$, donc on peut approcher X par $Y \longrightarrow \mathcal{N}(100,100)$ $Y^* = \dfrac{Y - 100}{10} \longrightarrow \mathcal{N}(0,1)$

$P(X \leq 90) \approx P(Y \leq 90) = P(Y - 100 \leq -10) = P(Y^* \leq -1) = \Phi(-1) = 1 - \Phi(1)$
$= 1 - 0,8413 = 0,1587$ (0,1713 avec Poisson)

Conclusion :

Loi :	Approchée par :	Conditions (selon énoncé !)
$\mathcal{B}(n,p)$	$\mathcal{P}(np)$	n grand, p petit
$\mathcal{B}(n,p)$	$\mathcal{N}(np, np(1-p))$	n grand
$\mathcal{P}(\lambda)$	$\mathcal{N}(\lambda, \lambda)$	λ grand