Chapitre11

CONVERGENCES ET APPROXIMATIONS

De nombreuses situations probabilistes sont amenées à se répéter de façons indépendantes. Par exemple la somme d'un grand nombre de dés, les choix de file d'attente à l'entrée d'un péage autoroutier pour un grand nombre de voitures ou bien les sites visités par un internaute durant une année... Dans toutes ces situations la quantité de répétitions est telle que l'on peut s'intéresser à ce que l'on va obtenir “à la limite” de ce processus. En reprenant les exemples précédents la question est de savoir (prédire ?) comment se répartissent les résultats pour un très grands nombre de dés, comment se répartissent les usagers du péage ou quels sites sont les plus visités par notre internaute ? On parle alors de “loi limite” ou plus précisément de “convergence en loi”.

11.1 Inégalité de Bienaymé-Tchebychev

LEMME 11.1 (Inégalité de Markov - Résultat non exigible)

Si X est une variable aléatoire admettant une espérance alors : $\forall a > 0, \quad P(X \geqslant a) \leqslant \dfrac{E(X)}{a}$

THÉORÈME 11.2 (Inégalité de Bienaymé-Tchebychev)

Soit X une VAR admettant $\underset{\text{(ou une variance)}}{\textbf{un moment d'ordre 2}}$. Alors :

$$\forall \varepsilon > 0, \quad P\big(|X - E(X)| \geqslant \varepsilon\big) \leqslant \frac{V(X)}{\varepsilon^2}$$

REMARQUE **11.1.**
Lorsque $|X - E(X)| \geqslant \varepsilon$, les valeurs de X sont à une distance plus grande que ε de leur moyenne $E(X)$. Donc $P\big(|X - E(X)| \geqslant \varepsilon\big)$ mesure la probabilité que X prenne des valeurs éloignées de $E(X)$. Cette probabilité est d'autant plus faible que $V(X)$ est petit (la variance mesure “l'étalement” des valeurs prises par X) et que ε est grand (plus on est loin de la moyenne moins on trouve de valeurs de X).

EXERCICE 11.1.
Soit $(X_n)_n$ une suite variables indépendantes suivant toutes la même loi de Bernoulli de paramètre $p \in]0;1[$.
On note alors :

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

On appellera cette variable, la *moyenne empirique*.

1. Justifier que $p(1-p) \leqslant \dfrac{1}{4}$.
2. Montrer que l'on peut appliquer l'inégalité de Bienaymé-Tchebychev à la variable $\overline{X_n}$.
3. En déduire que $\mathbb{P}\left(p \in \left[\overline{X_n} - \sqrt{\frac{5}{n}}, \overline{X_n} + \sqrt{\frac{5}{n}}\right]\right) \geqslant 0,95$.

   On dira que $\left[\overline{X_n} - \sqrt{\frac{5}{n}}, \overline{X_n} + \sqrt{\frac{5}{n}}\right]$ *est un intervalle de confiance de p au niveau de confiance* $0,95$.

DÉFINITION 11.3

Soit X une variable aléatoire.
• Soit $n \in \mathbb{N}^*$. On appelle **échantillon de X de taille n** toute famille $((X_1,\dots,X_n)$ de variables indépendantes et de même loi que X.
• On dit que la suite de variables $(X_n)_{n\in\mathbb{N}^*}$ est une **suite de variables indépendantes et identiquement distribuées (i.i.d.)** si pour tout $n \in \mathbb{N}^*$, les variables $(X_1,\dots,X_n)$ représentent un échantillon de taille n d'une variable X.

REMARQUE **11.2.**
Les suites de variables **i.i.d.** vont être au coeur de la suite du cours. Elles représentent en effet une succession d'expériences aléatoires identiques et indépendantes les unes des autres sur lesquelles on mesure la même grandeur (même loi). A partir d'une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables **i.i.d.** on va souvent travailler avec la variable

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

qui représente la fréquence d'apparition du phénomène étudié.

DÉFINITION 11.4

Si (X_n) est une suite de variables i.i.d., on appelle **moyenne empirique**, la variable aléatoire :

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

PROPRIÉTÉ 11.5

Soit X une variable aléatoire admettant une espérance μ et une variance σ^2.
Soit $(X_n)_n$ une suite de variables i.i.d. de même loi que X et $\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$ la moyenne empirique.

- $\overline{X_n}$ admet une espérance et $E(\overline{X_n}) = \mu$.
- $\overline{X_n}$ admet une variance et $V(\overline{X_n}) = \dfrac{\sigma^2}{n}$.

REMARQUE **11.3.**
Il faudra toujours re-démontrer cette propriété le jour J.

THÉORÈME 11.6 (Loi faible des grands nombres)

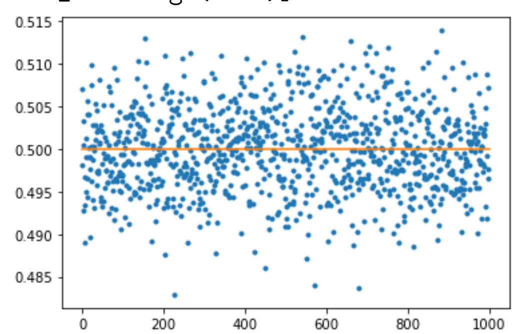
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires **i.i.d.** de même espérance μ et de même variance $\sigma^2 \neq 0$.
Alors on a :

$$\forall \varepsilon > 0, \quad \lim_{n\longrightarrow+\infty} P\left(\left|\overline{X}_n - \mu\right| \geqslant \varepsilon\right) = 0$$

Sur Python, on a tapé `plt.plot(x,'.')` après avoir entré :

```
x=[np.mean([rd.binomial(1,0.5) for _ in range(10000)]) for _ in range(1000)]
```

Ce graphique met en évidence la loi faible des grands nombres. On voit que les 1000 fréquences (moyennes empiriques) se situent toutes dans un intervalle centrée autour de la valeur théorique $p = 0.5$ d'amplitude faible($\approx 0,04$).



11.2 Convergence en loi

DÉFINITION 11.7

Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires de fonctions de répartitions respectives F_{X_n} et F_X.
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers X noté $X_n \xrightarrow{\mathscr{L}} X$ si

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x) \qquad \text{en tout point } x \text{ où } F_X \text{ est continue.}$$

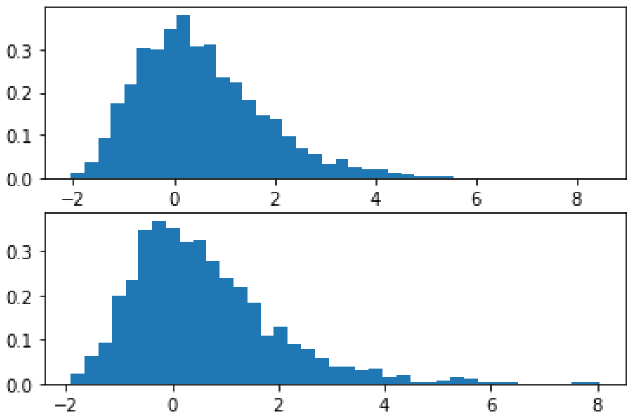
EXERCICE 11.2. Loi de Grumbel
On dit que W suit une loi de Grumbel si $W = -\ln(V)$ avec V suivant une loi exponentielle de paramètre 1.
Soit $(X_n)_n$ une suite de variables i.i.d suivant toutes la loi exponentielle de paramètre 1.
On note $M_n = \max(X_1, ..., X_n)$ et $T_n = M_n - \ln(n)$.

1. Déterminer la fonction de répartition de W.
2. En déduire que W est une variable à densité de densité $f : t \mapsto e^{-t}e^{-e^{-t}}$.
3. (a) Déterminer la fonction de répartition de M_n puis celle de T_n.
   (b) Montrer que la suite $(X_n)_n$ converge en loi vers la loi de Grumbel.

Sur Python, on tape les commandes suivantes et on obtient le graphique ci-dessous :

```
fig, axs = plt.subplots(2)
x=[-np.log(rd.exponential(1)) for _ in range(2000)]
y=[np.max([rd.exponential(1) for _ in range(1000)])-np.log(1000) for _ in range(2000)]
axs[0].hist(x,bins='auto',density='True')
axs[1].hist(y,bins='auto',density='True')
plt.show()
```

Ce graphique met en évidence la convergence en loi de la suite $(X_n)_n$ vers la loi de Grumbel. On constate en effet la similarité des histogrammes des fréquences (donc des fonctions de répartition) entre celui de la loi de Grumbel (graphique du haut) et celui de la variable X_{1000} (graphique du bas).



REMARQUE 11.4.
Lorsque les variables sont discrètes et à valeurs dans $\mathbb{Z} = \{..., -2, -1, 0, 1, ...\}$ il existe un moyen plus simple de caractériser la convergence en loi :

PROPRIÉTÉ 11.8 (admis)

Si $(X_n)_{n\in\mathbb{N}}$ et X sont des variables **discrètes** à valeur dans $\mathbb{Z}$ alors,

$$X_n \xrightarrow{\mathscr{L}} X \qquad \Longleftrightarrow \qquad \forall k \in \mathbb{Z}, \quad \lim_{n\to+\infty} P(X_n = k) = P(X = k)$$

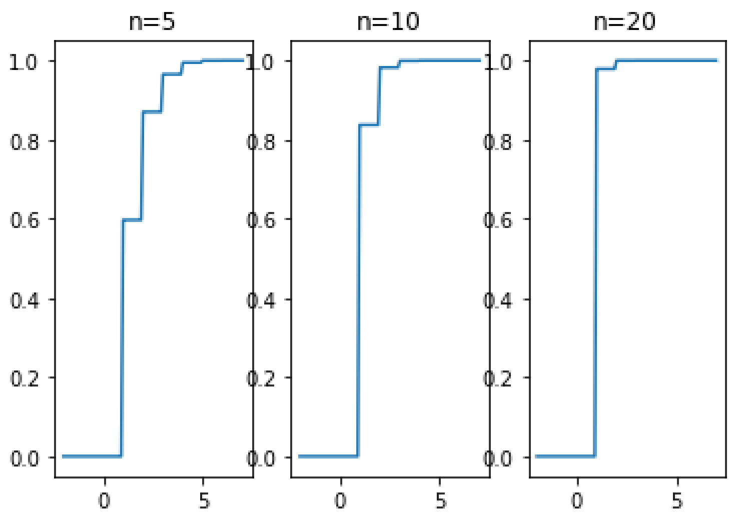
EXERCICE 11.3. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables i.i.d. suivant la loi uniforme sur $[\![1;6]\!]$.

On note $M_n = \min(X_1,...,X_n)$.
Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers loi certaine égale à 1.

Voici une fonction permettant de tracer les courbes des fonction de répartition des variables M_n pour les valeurs de $n = 5, n = 10$ et $n = 20$:

```
def frep():
    u=np.linspace(-2,7,100)
    x=[np.min([rd.randint(1,7) for _ in range(5)]) for _ in range(5000)]
    y=[np.min([rd.randint(1,7) for _ in range(10)]) for _ in range(5000)]
    z=[np.min([rd.randint(1,7) for _ in range(20)]) for _ in range(5000)]
    v=100*[0]
    w=100*[0]
    t=100*[0]
    for k in range(100):
        v[k]=np.mean(x<=u[k])
        w[k]=np.mean(y<=u[k])
        t[k]=np.mean(z<=u[k])
    plt.clf()
    fig, (ax1,ax2,ax3) = plt.subplots(1,3)
    ax1.plot(u,v)
    ax1.set_title('n=5')
    ax2.plot(u,w)
    ax2.set_title('n=10')
    ax3.plot(u,t)
    ax3.set_title('n=20')
    plt.show()
```

Ce graphique met en évidence la convergence en loi de la suite $(M_n)_n$ vers la certaine égale à 1. On constate en effet que les courbes affichées correspondent aux fonction de répartition des variables M_5, M_{10} et M_{20}. On constate une "convergence" de ces courbes vers la fonction de répartition de la loi certaine égale à 1.



Attention certaines suites de variables discrètes convergent en loi vers une variable à densité. C'est le cas de la section suivante. Dans ce cas le théorème précédent ne s'applique plus.

11.3 Théorème limite central

DÉFINITION 11.9 (et propriété)

Soit X une variable aléatoire admettant une espérance μ et une variance $\sigma^2 \neq 0$.
On dit que X est **centrée réduite** si $\mu = 0$ et $\sigma^2 = 1$.

PROPRIÉTÉ 11.10 (Centrer et réduire une variable quelconque)

Soit X une variable aléatoire admettant une espérance μ et une variance $\sigma^2 \neq 0$.
Alors $X^* = \dfrac{X-\mu}{\sigma}$ est une variable aléatoire **centrée réduite** .

THÉORÈME 11.11 (Théorème limite central - admis)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires **i.i.d.** d'espérance μ et de variance $\sigma^2 \neq 0$.
Si on note $\overline{X_n} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$, alors

$$\overline{X}_n^* \quad \xrightarrow{\mathscr{L}} \quad X \hookrightarrow \mathscr{N}(0,1)$$

En particulier, si on note Φ la fonction de répartition de la loi $\mathscr{N}(0,1)$, pour $-\infty \leqslant a \leqslant b \leqslant +\infty$:

$$P(a < \overline{X}_n^* \leqslant b) \approx \Phi(b) - \Phi(a)$$

REMARQUE **11.5.**
Dans le théorème limite central on a :

$$\overline{X}_n^* = \sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right).$$

REMARQUE **11.6.**
C'est un théorème remarquable car le résultat est très fort et contient peu d'hypothèses.
De plus il montre l'importance de la loi normale en probabilités et statistiques.

11.4 Applications aux approximations de loi

COROLLAIRE 11.12 (Théorème de Moivre-Laplace)

Si $X_n \hookrightarrow \mathscr{B}(n,p)$, alors :

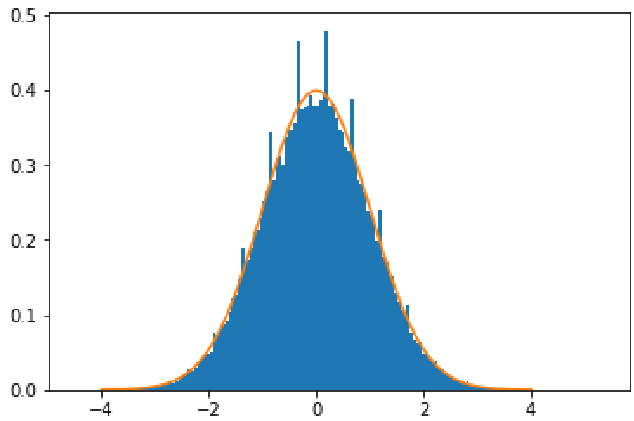
$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathscr{L}} X \hookrightarrow \mathscr{N}(0,1)$$

En pratique : on peut approcher la loi de X_n par la loi $\mathscr{N}(np, npq)$ lorsque $\begin{cases} n \geqslant 30 \\ np \geqslant 5 \\ nq \geqslant 5 \end{cases}$

Une fonction permettant d'obtenir le graphique ci-dessous avec l'instruction `convergence_binomiale(2000,0.5,100000`

```
def convergence_binomiale(n,p,N):
    x=[(rd.binomial(n,p)-n*p)/np.sqrt(n*p*(1-p)) for _ in range(N)]
    t=np.linspace(-4,4,100000)
    y=[1/np.sqrt(2*np.pi)*np.exp(-u**2/2) for u in t]
    plt.hist(x,bins='auto', density='True')
    plt.plot(t,y)
    plt.show()
```

Ce graphique met en évidence le théorème de Moivre-Laplace. On constate que l'histogramme des fréquences associés à un grand nombre de répétitions ($N = 10\,000$) de variables $\dfrac{X_n - np}{\sqrt{np(1-p)}}$ où $X_n \hookrightarrow \mathscr{B}(n,p)$ avec $n = 2000$ et $p = 0,5$ épouse bien la courbe de la d.d.p. $t \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ de la loi $\mathscr{N}(0,1)$, illustrant la convergence en loi.



COROLLAIRE 11.13 (Théorème limite central pour la loi de Poisson)

Si $X_n \hookrightarrow \mathscr{P}(n\lambda)$, alors :

$$\frac{X_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{\mathscr{L}} X \hookrightarrow \mathscr{N}(0,1)$$

En pratique : on peut approcher la loi de $X \hookrightarrow \mathscr{P}(\lambda)$ par la loi $\mathscr{N}(\lambda,\lambda)$ lorsque $\lambda \geqslant 15$

EXEMPLE.
Ici encore pour de grandes valeurs de λ, la calcul de $P(X \leqslant a)$ pourra nécessiter des calculs lourds.
Si on considère X qui suit une loi de Poisson de paramètre 64 que l'on cherche à calculer $P(X \leqslant 74)$, on a intérêt à approcher la loi de $X^* = \dfrac{X-64}{8}$ par la loi normale centrée réduite. On a donc

$$P(X \leqslant 74) = P\left(\frac{X-64}{8} \leqslant 1,25\right) \approx \Phi(1,25) \approx 0,8944$$