CHAPITRE 13

Convergences et approximations

13.1 Inégalité de Bienaymé-Tchebychev

Théorème 1 ***Inégalité de Bienaymé-Tchebychev***

Soit X une variable aléatoire réelle (discrète ou continue) admettant un moment d'ordre 2. Alors

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

Démonstration :

Signalons déjà que puisque X admet un moment d'ordre 2, X admet bien une espérance et une variance.

1er cas : X est une variable discrète. On pose $X(\Omega) = \{x_i \ / \ i \in I\}$ et $p_i = \mathbb{P}(X = x_i)$.
Par définition, on sait que $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$. Donc, d'après le théorème de transfert,

$$\mathbb{V}(X) = \sum_{i \in I} (x_i - \mathbb{E}(X))^2 p_i$$

Or, on a

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) = \sum_{j \in J} p_j \qquad \text{où } J = \{j \in I \ / \ |x_j - \mathbb{E}(X)| \geqslant \varepsilon\}$$

On peut donc écrire :

$$\begin{aligned}
\mathbb{V}(X) &= \sum_{j \in J} (x_j - \mathbb{E}(X))^2 p_j + \sum_{i \notin J} (x_i - \mathbb{E}(X))^2 p_i \\
&\geqslant \sum_{j \in J} (x_j - \mathbb{E}(X))^2 p_j \\
&\geqslant \varepsilon^2 \sum_{j \in J} p_j \\
&\geqslant \varepsilon^2 \mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon)
\end{aligned}$$

et donc

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}$$

2ème cas : X est une variable à densité, dont f est une densité. Alors

$$\mathbb{V}(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x)dx$$

On sépare l'intégrale en trois "bouts" à l'aide de la relation de Chasles.

$$\mathbb{V}(X) = \int_{-\infty}^{\mathbb{E}(X)-\varepsilon} (x - \mathbb{E}(X))^2 f(x)dx + \int_{\mathbb{E}(X)-\varepsilon}^{\mathbb{E}(X)+\varepsilon} (x - \mathbb{E}(X))^2 f(x)dx + \int_{\mathbb{E}(X)+\varepsilon}^{+\infty} (x - \mathbb{E}(X))^2 f(x)dx$$

De plus,

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) = \mathbb{P}(X \leqslant \mathbb{E}(X) - \varepsilon) + \mathbb{P}(X \geqslant \mathbb{E}(X) + \varepsilon) = \int_{-\infty}^{\mathbb{E}(X)-\varepsilon} f(x)dx + \int_{\mathbb{E}(X)+\varepsilon}^{+\infty} f(x)dx$$

Donc on a

$$\mathbb{V}(X) \geqslant \int_{-\infty}^{\mathbb{E}(X)-\varepsilon} (x - \mathbb{E}(X))^2 f(x)dx + \int_{\mathbb{E}(X)+\varepsilon}^{+\infty} (x - \mathbb{E}(X))^2 f(x)dx \geqslant \varepsilon^2 \left(\int_{-\infty}^{\mathbb{E}(X)-\varepsilon} f(x)dx + \int_{\mathbb{E}(X)+\varepsilon}^{+\infty} f(x)dx\right)$$

Donc

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}$$

Remarque :

Les valeurs de X sont "en principe" autour de leur moyenne qui est $\mathbb{E}(X)$. Lorsque $|X - \mathbb{E}(X)| \geqslant \varepsilon$, les valeurs de X sont éloignées de $\mathbb{E}(X)$ d'une distance au moins ε.
Ainsi, $\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon)$ mesure la probabilité que X prenne des valeurs éloignées de $\mathbb{E}(X)$. Cette probabilité est d'autant plus faible que $\mathbb{V}(X)$ est petite (la variance mesure la "dispersion" des valeurs prises par X) et que ε est grand (plus on est loin de la moyenne, moins on trouve de valeurs de X).

Exemple :

On utilise un dé cubique parfait. Combien doit-on faire de lancers pour pouvoir affirmer, avec un risque d'erreur inférieur à 5%, que la fréquence d'apparition du chiffre "1" diffère de $\frac{1}{6}$ d'au plus $\frac{1}{100}$?

Lorsqu'on a effectué n lancers, notons X_n le nombre de 1 obtenus et F_n la fréquence d'apparition du 1 au cours de ces n lancers, i.e. $F_n = \frac{X_n}{n}$.

On cherche donc un entier n tel que $\mathbb{P}\left(\left|F_n - \frac{1}{6}\right| < \frac{1}{100}\right) \geqslant 0,95$.

On souhaite appliquer l'inégalité de Bienaymé-Tchebychev sur F_n avec $\varepsilon = \frac{1}{100}$. Il nous faut donc calculer $\mathbb{E}(F_n)$ et $\mathbb{V}(F_n)$.

On sait que $X_n \rightsquigarrow \mathcal{B}\left(n, \frac{1}{6}\right)$, donc $\mathbb{E}(X_n) = \frac{n}{6}$ et $\mathbb{V}(X_n) = \frac{5n}{36}$.
On en déduit que

$$\mathbb{E}(F_n) = \frac{1}{6}, \qquad \mathbb{V}(F_n) = \frac{5}{36n}$$

L'inégalité de Bienaymé-Tchebychev nous dit que $\mathbb{P}\left(\left|{}_n - \frac{1}{6}\right| \geqslant \frac{1}{100}\right) \leqslant \frac{5.10^4}{36n}$ autrement dit que

$$\mathbb{P}\left(\left|{}_n - \frac{1}{6}\right| < \frac{1}{100}\right) \geqslant 1 - \frac{5.10^4}{36n}$$

Il suffit donc de choisir un entier n tel que $1 - \frac{5.10^4}{36n} \geqslant 0,95$, c'est-à-dire $n \geqslant 27778$.

13.2 Convergence d'une suite de variables aléatoires

13.2.1 Loi faible des grands nombres

Théorème 2

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires qui :

- *sont indépendantes,*
- *admettent toutes une même espérance m,*
- *admettent toutes une même variance σ^2.*

Si on pose $\overline{X_n}$ la ***moyenne empirique*** *des X_i :*

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots + X_n}{n},$$

alors on a

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

On dit que $\overline{X_n}$ ***converge en probabilité*** *vers la variable constante égale à m.*

Démonstration :

(Hors programme)

On a

$$\mathbb{E}(\overline{X_n}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n} nm = m$$

et comme les variables sont indépendantes, on a

$$\mathbb{V}(\overline{X_n}) = \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{V}(X_i) = \frac{\sigma^2}{n}$$

Donc d'après l'inégalité de Bienaymé-Tchebychev,

$$0 \leqslant \mathbb{P}(|\overline{X_n} - m| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

Donc par encadrement de limites,

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

Remarque :

Un exemple pour se représenter ce théorème : on considère une suite d'épreuves indépendantes, et un événement A de probabilité p, qui peut ou non se réaliser au cours d'une des épreuves. On note X_i la variable qui vaut 1 si A s'est réalisé à la i-ième épreuve et 0 sinon. Dans ce cas, $\overline{X_n}$ représente la fréquence de réalisation de l'événement A au cours des n premières épreuves. La loi faible des grands nombre nous dit que cette fréquence "tend" vers p.

13.2.2 Convergence en loi

Définition 3
Soit (X_n) une suite de variables aléatoires réelles. On note F_{X_n} la fonction de répartition de X_n.
On dit que la suite (X_n) **converge en loi** vers la variable aléatoire réelle X, de fonction de répartition F_X, si pour tout $x \in \mathbb{R}$ dans lequel la fonction F_X est continue, soit la suite $(F_{X_n}(x))$ converge vers $F_X(x)$.

Exemple :

Pour $n \in \mathbb{N}^*$, on considère une variable aléatoire discrète X_n qui suit la loi uniforme sur l'ensemble

$$\left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\right\}$$

c'est-à-dire que l'on a

$$\forall k \in [\![0, n-1]\!],\ \mathbb{P}\left(X_n = \frac{k}{n}\right) = \frac{1}{n}$$

La suite (X_n) converge-t-elle en loi ?

On commence tout d'abord par déterminer la fonction de répartition de X_n. On a $X_n(\Omega) \subset [0, 1[$, donc on peut déjà écrire

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 \ \text{ si } x \geqslant 1 \end{cases}$$

Soit maintenant $x \in [0, 1[$. On notera $E(x)$ la partie entière de x. On a

$$F_{X_n}(x) = \mathbb{P}(X_n \leqslant x) = \sum_{\substack{k \in [\![0, n-1]\!] \\ \frac{k}{n} \leqslant x}} \mathbb{P}\left(X_n = \frac{k}{n}\right) = \sum_{k=0}^{E(nx)} \frac{1}{n} = \frac{E(nx)+1}{n}$$

On a donc

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{E(nx)+1}{n} & \text{si } x \in [0, 1] \\ 1 \ \text{ si } x \geqslant 1 \end{cases}$$

Fixons-nous x et cherchons la limite de la suite $(F_{X_n}(x))$.
Pour $x < 0$ et $x > 1$ c'est facile.
Pour $x \in [0, 1]$, on a d'après la définition de la partie entière

$$E(nx) \leqslant nx \leqslant E(nx) + 1$$

donc

$$\frac{E(nx)}{n} \leqslant x \leqslant \frac{E(nx)+1}{n}$$

On en déduit par encadrement que

$$\frac{E(nx)+1}{n} \underset{n \to +\infty}{\longrightarrow} x$$

On a donc

$$\forall x \in \mathbb{R}, \quad \lim_{n \to +\infty} F_{X_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0, 1] \\ 1 & \text{si } x \geqslant 1 \end{cases}$$

On reconnaît ici la fonction de répartition d'une variable aléatoire suivant une loi uniforme sur $[0, 1]$. Donc (X_n) converge en loi vers une variable X qui suit la loi uniforme sur $[0, 1]$.

Proposition 4
Soit (X_n) une suite de variables aléatoire convergeant vers une variable aléatoire X. Pour tous points a et b de continuité de F_X tels que $a < b$, on a

$$\lim_{n\to+\infty} \mathbb{P}(a < X_n \leqslant b) = \mathbb{P}(a < X \leqslant b)$$

Théorème 5
Soit (X_n) une suite de variables discrètes et X une variable discrète telle que pour tout entier n, $X_n(\Omega) \subset \mathbb{N}$ et $X(\Omega) \subset \mathbb{N}$.
Alors

$$(X_n) \text{ converge en loi vers } X \iff \forall k \in \mathbb{N},\ \lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$$

Remarque :

> Attention, certaines suites de variables discrètes convergent en loi vers une variable à densité, comme dans l'exemple ci-dessus.

13.2.3 Théorème de la limite centrée

Théorème 6 — ***Théorème de la limite centrée***

Soit (X_n) une suite de variable aléatoires réelles définies sur un même espace probabilisé. On suppose que :
- *les X_n ont toutes la même loi*
- *les X_n sont indépendantes*
- *les X_n admettent une variance non nulle*

On note leur espérance commune m et leur variance commune σ^2. On note

$$\forall n \geqslant 1, \quad S_n = X_1 + X_2 + \cdots + X_n, \qquad Y_n = \frac{S_n}{n}$$

On a donc leurs variables centrées réduites associées égales :

$$S_n^* = \frac{S_n - nm}{\sigma\sqrt{n}} = \frac{\sqrt{n}(Y_n - m)}{\sigma} = Y_n^*$$

Alors, la suite (S_n^*) et la suite (Y_n^*) convergent toutes deux en loi versz une variable aléatoire de loi normale centrée réduite $\mathcal{N}(0,1)$.

Remarques :

> **R 1** – On a donc, pour $a < b$,
>
> $$\lim_{n\to+\infty} \mathbb{P}(a < S_n^* \leqslant b) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$
>
> **R 2** – Comme $S_n = \sigma\sqrt{n}S_n^* + nm$, ce théorème nous permet de dire que pour n assez grand, la variable aléatoire S_n suit approximativement la loi $\mathcal{N}(nm, n\sigma^2)$.
>
> **R 3** – Ce théorème est impressionnant car il nécessite très peu d'hypothèses sur les X_n, il montre l'importance de la loi normale en probabilités.

R 4 – En pratique, pour $n \geqslant 30$, on pourra approcher la loi de S_n^* par la loi $\mathcal{N}(0,1)$.

Exemple :

Une montre fait une erreur d'au plus une demi-minute par jour. On cherche à déterminer la probabilité que l'erreur commise au bout d'une année (non bissextile) soit inférieure ou égale à un quart d'heure.
Pour cela, on considère que l'erreur commise un jour donné, en minutes, suit une loi uniforme sur $\left[-\frac{1}{2}, \frac{1}{2}\right]$ et que les erreurs commises chaque jour sont indépendantes.

Notons X_k l'erreur commise le k-ième jour de l'année. D'après les hypothèses, X_k suit la loi uniforme sur $\left[-\frac{1}{2}, \frac{1}{2}\right]$ et les X_k sont indépendantes. De plus, l'erreur commise au bout d'un an est

$$S = X_1 + X_2 + \cdots + X_{365}$$

On cherche donc à calculer $\mathbb{P}(-15 \leqslant S \leqslant 15)$.

On a $\mathbb{E}(X_k) = 0$ et $\mathbb{V}(X_k) = \frac{1}{12}$, donc $\mathbb{E}(S) = 0$ et $\mathbb{V}(S) = \dfrac{365}{12}$. Donc la variable centrée réduite associée à S est $S^* = \dfrac{S\sqrt{12}}{\sqrt{365}}$.
Comme $365 > 30$, le théorème de la limite centrée nous dit que S^* suit approximativement la loi normale centrée réduite $\mathcal{N}(0,1)$ et donc

$$\begin{aligned}
\mathbb{P}(-15 \leqslant S \leqslant 15) &= \mathbb{P}\left(-\frac{15\sqrt{12}}{\sqrt{365}} \leqslant \frac{S\sqrt{12}}{\sqrt{365}} \leqslant \frac{15\sqrt{12}}{\sqrt{365}}\right) \\
&= \mathbb{P}\left(-\frac{15\sqrt{12}}{\sqrt{365}} \leqslant S^* \leqslant \frac{15\sqrt{12}}{\sqrt{365}}\right) \\
&= \Phi\left(\frac{15\sqrt{12}}{\sqrt{365}}\right) - \Phi\left(-\frac{15\sqrt{12}}{\sqrt{365}}\right) \\
&= 2\Phi\left(\frac{15\sqrt{12}}{\sqrt{365}}\right) - 1 \\
&\simeq 2\ \Phi(2.72) - 1 \simeq 0.9934
\end{aligned}$$

13.3 Approximations

13.3.1 Approximation d'une loi binomiale par une loi de Poisson

Théorème 7

Soit λ un réel strictement positif et (X_n) une suite de variables discrètes telles que X_n suit la loi binomiale de paramètres $\left(n, \frac{\lambda}{n}\right)$. Alors, (X_n) converge en loi vers une variable X qui suit une loi de Poisson de paramètre λ.

Démonstration :

On a

$$\begin{aligned}
\mathbb{P}(X_n = k) &= \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} e^{(n-k)\ln(1-\lambda/n)}
\end{aligned}$$

Or, $n(n-1)\cdots(n-k+1) \underset{n\to+\infty}{\sim} n^k$.

De plus, comme $\dfrac{\lambda}{n} \to 0$, on a $\ln\left(1-\frac{\lambda}{n}\right) \underset{n\to+\infty}{\sim} -\frac{\lambda}{n}$, donc

$$\lim_{n\to+\infty} \mathbb{P}(X_n = k) = \frac{\lambda^k}{k!}e^{-\lambda}$$

Ainsi, (X_n) converge bien en loi vers une variable qui suit la loi de Poisson de paramètre λ.

Remarque :

En pratique, on considère que si $n \geqslant 30$ et p est petit $(p \leqslant 0.1)$, on peut approcher la loi $\mathcal{B}(n,p)$ par la loi $\mathcal{P}(np)$.

Exemple :

Soit X une variable aléatoire suivant la loi $\mathcal{B}(100, 0.05)$. On souhaite calculer $\mathbb{P}(X=2)$.

1er cas : valeur exacte par la loi binomiale.

$$\mathbb{P}(X=2) = \binom{100}{2}(0.05)^2(0.95)^{98} \simeq 0.0812$$

2ème cas : valeur approchée par la loi $\mathcal{P}(5)$

$$\mathbb{P}(X=2) \simeq \frac{5^2}{2!}e^{-5} \simeq 0.0843$$

13.3.2 Approximation d'une loi binomiale par une loi normale

Théorème 8

Soit $p \in]0,1[$ et $q = 1-p$. Soit (S_n) une suite de variables aléatoires telle que S_n suit la loi binomiale $\mathcal{B}(n,p)$. Alors, la suite de variables aléatoires $S_n^* = \dfrac{S_n - np}{\sqrt{npq}}$ converge en loi vers la loi normale centrée réduite $\mathcal{N}(0,1)$.

Démonstration :

Toute variable binomiale de paramètre (n,p) peut être considérée comme la somme de n variables aléatoires de Bernoulli de paramètre p mutuellement indépendantes. On a donc $S_n = \sum\limits_{k=1}^{n} X_k$, où X_k est une variable de Bernoulli de paramètre p. Comme les variables de Bernoulli admettent une espérance p et une variance pq, le Théorème de la Limite Centrée s'applique bien et donne le résultat voulu.

Remarques :

- **R 1** – En pratique, on considère que si $n \geqslant 30$, $np \geqslant 5$ et $nq \geqslant 5$, on peut approcher la loi $\mathcal{B}(n,p)$ par la loi $\mathcal{N}(np, npq)$.
- **R 2** – Si S_n suit une loi $\mathcal{B}(n,p)$ avec n et p tels qu'on puisse approcher S_n par N_n qui suit la loi $\mathcal{N}(np, npq)$, on devrait écrire

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

mais comme N_n est une variable à densité, on a $\mathbb{P}(N_n = k) = 0$, donc notre approximation ci-dessus n'est pas vraiment bonne... On écrira plutôt

$$\mathbb{P}(S_n = k) \simeq \mathbb{P}(k - 0.5 < N_n < k + 0.5)$$

et on appelle cette méthode utiliser la **correction de continuité**.

Exemple :

Soit X une variable qui suit la loi $\mathcal{B}(900, 0.5)$. On cherche à calculer $\mathbb{P}(405 \leqslant X \leqslant 495)$.

Pour le calcul exact, il nous faudrait calculer des combinaisons avec de très grands nombres, ce qui nécessite un ordinateur et ne donne parfois qu'une valeur approchée.
On remarque ici que l'on est dans les conditions où l'on peut approchée notre loi binomiale par une loi normale $\mathcal{N}(450, 225)$ car $900 > 30$ et $900 \times 0.5 = 450 > 5$.
Cependant pour la loi normale $\mathcal{N}(450, 225)$, il n'est pas facile de calculer $\mathbb{P}(405 < X < 495)$. On se ramène donc à une loi normale centrée réduite.
On pose

$$X^* = \frac{X - 450}{\sqrt{225}}$$

et on a que X^* suit approximativement la loi normale centrée réduite.
De plus,

$$\mathbb{P}(405 \leqslant X \leqslant 495) = \mathbb{P}\left(-3 \leqslant \frac{X - 450}{15} \leqslant 3\right) = \mathbb{P}(-3 \leqslant X^* \leqslant 3)$$

Donc on a

$$\mathbb{P}(405 \leqslant X \leqslant 495) \simeq \Phi(3) - \Phi(-3) = 2\Phi(3) - 1 \simeq 0.9974$$

13.3.3 Approximation d'une loi de Poisson par une loi normale

Théorème 9
Soit α un réel strictement positif et soit (S_n) une suite de variables aléatoires telle que S_n suit la loi de Poisson $\mathcal{P}(n\alpha)$. Alors, la suite de variables aléatoires $S_n^* = \dfrac{S_n - n\alpha}{\sqrt{n\alpha}}$ converge en loi vers la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Démonstration :

Toute variable de Poisson de paramètre $n\alpha$ peut être considérée comme la somme de n variables aléatoires de Poisson de paramètre α mutuellement indépendantes. On a donc $S_n = \sum\limits_{k=1}^{n} X_k$, où X_k est une variable de Poisson de paramètre α. Comme les variables de Bernoulli admettent une espérance α et une variance α, le Théorème de la Limite Centrée s'applique bien et donne le résultat voulu.

Remarque :

En pratique, on considère que si $\lambda \geqslant 18$, on peut approcher la loi $\mathcal{P}(\lambda)$ par la loi $\mathcal{N}(\lambda, \lambda)$.

Exemple :

Soit X une variable aléatoire suivant une loi $\mathcal{P}(64)$ et on cherche $\mathbb{P}(X \leqslant 74)$.
Pour faire le calcul exact, il faut obligatoirement un ordinateur. On a donc intérêt à approcher la loi de X par la loi $\mathcal{N}(64, 64)$ et donc $\dfrac{X - 64}{8}$ suit la loi normale centrée réduite. On a donc

$$\mathbb{P}(X \leqslant 74) = \mathbb{P}\left(\frac{X - 64}{8} \leqslant 1.25\right) \simeq \Phi(1.25) \simeq 0.8944$$