Chapitre 6
Théorèmes de convergence

I. La convergence en loi

On a déjà rencontré une convergence en loi lors de l'approximation d'une loi binomiale par une loi de Poisson. Ce problème se place dans un cadre plus général où on souhaite remplacer la loi d'une variable aléatoire par une loi d'usage plus simple

I.1. Définition

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur un espace probabilisé (Ω, $\mathcal{P}(\Omega)$, p), F_n leurs fonctions de répartition, et X une variable aléatoire définie sur ce même espace, de fonction de répartition F .

On dit que la suite (X_n) converge en loi vers X si, et seulement si, en tout point x où F est continue, on a $\lim\limits_{n \to +\infty} F_n(x) = F(x)$

RAPPEL

Soit (X_n) une suite de variables aléatoires binomiales de paramètres n et p_n telle que $\lim\limits_{n \to +\infty} n\, p_n = \lambda$.

Alors (X_n) converge en loi vers une variable de Poison de paramètre λ.

En pratique :

Soit X une v.a. suivant la loi binomiale $\mathcal{B}(n\,;\,p)$.

Si $n \geq 30$, $p \leq 0{,}1$ et $np < 15$, alors X suit approximativement la loi de Poisson de paramètre np.

I.2. Le théorème de Moivre - Laplace

On considère une suite (X_n) de variables binomiales $\mathcal{B}(n\,;\,p)$, p étant un paramètre fixé. Alors $T_n = \dfrac{X_n - np}{\sqrt{npq}}$ converge en loi vers $\mathcal{N}(0\,;\,1)$.

En pratique, pour $n \geq 30$, $np \geq 15$ et $npq \geq 5$ (Carnec), $\mathcal{B}(n\,;\,p)$ suit approximativement $\mathcal{N}(np\,;\sqrt{npq})$.

Les conditions changent suivant les auteurs :

Saporta :	n assez grand, $np > 5$ et $nq > 5$
Grais :	$npq \geq 9$
Le programme :	$n > 30$ et $0{,}3 < p < 0{,}7$
Un document du GTD :	$n > 30$, $np > 5$ et $nq > 5$

…

Attention à la correction de continuité !

Prenons $\mathcal{B}(100\ ;\ 0{,}5)$. $npq = 25$.

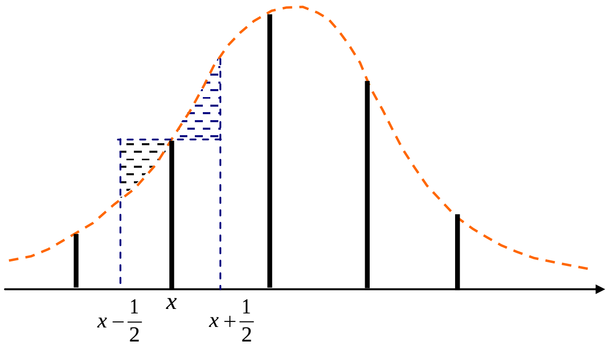
L'approximation de $T = \dfrac{X - 50}{5}$ par $\mathcal{N}(0\ ;\ 1)$ est justifiée.

On s'intéresse à $P(X \leq 52)$. La valeur exacte est 0,6914.

Par l'approximation de la loi normale on obtient :

$X \leq 52 \Leftrightarrow T \leq 0{,}4$ donc $P(X \leq 52) = \Pi(0{,}4)$

Soit $P(X \leq 52) =$ 0,6554. Erreur non négligeable !!



On corrige de la façon suivante : $P(X \leq x) \approx P\left(T \leq \dfrac{x + 0{,}5 - np}{\sqrt{npq}} \right)$

$P(X = x) \approx P\left(\dfrac{x - 0{,}5 - np}{\sqrt{npq}} \leq T \leq \dfrac{x + 0{,}5 - np}{\sqrt{npq}} \right)$

$P(y \leq X \leq x) \approx P\left(\dfrac{y - 0{,}5 - np}{\sqrt{npq}} \leq T \leq \dfrac{x + 0{,}5 - np}{\sqrt{npq}} \right)$

Pour l'exemple précédent, $P(X \leq 52) \approx \Pi(0{,}5) =$ 0,6915.

Exemple 1

On lance une pièce de monnaie « honnête » 1 000 fois. Quelle est la probabilité d'obtenir au moins 548 piles ?

Solution

X désigne le nombre de piles obtenus. On cherche $P(X \geq 548)$.

$P(X \geq 548) = 1 - P(X \leq 547)$.

On peut approcher par la loi normale. $n = 1\,000$; $np = 500$; $npq = 250$

Soit $\sqrt{npq} = 5\sqrt{10} \approx 15{,}8$.

$T = \dfrac{X - 500}{5\sqrt{10}}$ suit approximativement $\mathcal{N}(0\,;1)$

$P(X \leq 547) \approx P\left(T \leq \dfrac{547{,}5 - 500}{5\sqrt{10}} \right)$

$P(T \leq 3) \approx 0{,}998\,65$ $\boxed{\text{D'où } P(X \geq 548) \approx 0{,}001\,35 \text{ (une chance sur 1 000)}}$

Remarque

1. La valeur exacte de $P(X \geq 548$ est $1 - 0{,}998\,92 = 0{,}001\,18$
2. $P(X \in [\ m - 3\,\sigma\ ;\ m + 3\,\sigma\ ]) \approx 0{,}997$

Exemple 2

On organise un QCM de 100 questions. Pour chaque question il y a trois réponses possibles dont une et une seule est exacte.

Trouver un entier k tel que si un candidat a au moins k réponses justes, il y a moins de 5% de chances que toutes les réponses soient dues au hasard.

Solution

On note X la variable aléatoire donnant le nombre de réponses exactes par un candidat répondant à toutes les questions au hasard. Il s'agit de trouver un entier k tel que : $P(X \geq k) < 0{,}05$ ou $P(X \leq k-1) \geq 0{,}95$.

X suit $\mathcal{B}(100\ ;\ 1/3)$. Les conditions d'approximation par la loi normale sont vérifiées et $np = \frac{100}{3}$, $\sqrt{npq} = \frac{10}{3}\sqrt{2}$.

$T = \dfrac{X - \frac{100}{3}}{\frac{10}{3}\sqrt{2}}$ suit approximativement $\mathcal{N}(0\ ;\ 1)$ et

$P(X \leq k-1) \approx P\left(T \leq \dfrac{k - 1 + 0{,}5 - np}{\sqrt{npq}}\right) \geq 0{,}95 = \Pi(1{,}645)$

D'où $\dfrac{k - 1 + 0{,}5 - np}{\sqrt{npq}} \geq 1{,}645$ soit $k \geq 41{,}59$ donc $k \geq 42$.

Les « fourchettes » d'une fréquence binomiale

Soit X_n une variable aléatoire qui suit $\mathcal{B}(n\ ;\ p)$. On suppose les conditions d'approximation par la loi normale vérifiées ;

On pose $F_n = \frac{X}{n}$ la fréquence associée. Alors : $P\left(|F_n - p| \leq \frac{1}{\sqrt{n}}\right) > 0{,}95$

Fourchette au niveau 0,95

Démonstration

$T_n = \dfrac{X_n - np}{\sqrt{npq}}$ converge en loi vers $\mathcal{N}(0\ ;\ 1)$. Donc $P(|T_n| \leq 1{,}96) \approx 0{,}95000435$

$|T_n| \leq 1{,}96 \Leftrightarrow \left|\dfrac{X_n - np}{\sqrt{npq}}\right| \leq 1{,}96 \Leftrightarrow |F_n - p| \leq \dfrac{1{,}96}{\sqrt{n}}\sqrt{pq}$

Or $pq = p(1-p) \leq 0{,}25$ donc $|F_n - p| \leq \frac{1}{\sqrt{n}}$. Ainsi $|T_n| \leq 1{,}96 \Rightarrow |F_n - p| \leq \frac{1}{\sqrt{n}}$ et

$P\left(|F_n - p| \leq \frac{1}{\sqrt{n}}\right) \geq P(|T_n| \leq 1{,}96) \approx 0{,}95000435 > 0{,}95$ cqfd.

On a, de même, les autres fourchettes :

$P\left(|F_n - p| \leq \frac{0{,}825}{\sqrt{n}}\right) > 0{,}90$; $P\left(|F_n - p| \leq \frac{1{,}5}{\sqrt{n}}\right) > 0{,}99$

Pour $n = 1000$ on obtient : $P(|F_{1000} - p| \leq 2{,}61\%) > 90\%$

$P(|F_{1000} - p| \leq 3{,}2\%) > 95\%$

$P(|F_{1000} - p| \leq 4{,}75\%) > 99\%$

1.3. Convergence de la loi de Poisson vers la loi normale

Soit une suite (X_n) de variables de Poisson $\mathcal{P}(n\lambda)$ où λ est un réel positif fixé.

Alors $T_n = \dfrac{X_n - n\lambda}{\sqrt{n\lambda}}$ converge en loi vers $\mathcal{N}(0\ ;\ 1)$ lorsque n tend vers $+\infty$.

L'approximation est satisfaisante dès que $n\lambda > 15$.

1.4. Le théorème central-limite

Soit X_1, X_2, ... , X_n des variables aléatoires mutuellement indépendantes ayant la même loi de moyenne m et d'écart type σ.

On pose $\overline{X}_n = \dfrac{X_1 + X_2 + ... + X_n}{n}$

Alors $(\overline{X}_n)$ converge en loi vers $\mathcal{N}(m\ ;\ \dfrac{\sigma}{\sqrt{n}})$

Pour $\boldsymbol{n \geq 30}$, $\overline{X}$ suit approximativement $\mathcal{N}(m\ ;\ \dfrac{\sigma}{\sqrt{n}})$.

Les deux théorèmes de convergence précédents sont des cas particuliers de ce théorème plus général.

1.5. Résumé

Approximation d'une loi par une autre

Épreuve : n tirages sans remise

Loi hypergéométrique

$\mathcal{H}\left(N\,;n\,;\frac{a}{N}\right)$ où $a+b=N$

$$P(X=k)=\frac{C_a^k\,C_b^{n-k}}{C_{a+b}^n}$$

Si $a+b=N$ et si $N \geq 10\,n$

approximation par la **loi binomiale**

$\mathcal{B}(n;p)$ où $p=\frac{a}{N}$

Épreuve : n tirages avec remise

Loi binomiale $\mathcal{B}(n\,;p)$

$P(X=k)=C_n^k\,p^k q^{n-k}$ avec $q=1-p$

Si $n \geq 30$, $p \leq 0{,}1$ et $np < 15$

approximation par la

loi de Poisson $\mathcal{P}(np)$

Si $n \geq 30$, $np \geq 15$ et npq > 5

approximation par la

loi normale $\mathcal{N}(np\,;\sqrt{npq}\,)$

Loi de Poisson $\mathcal{P}(\lambda)$ où $\lambda > 0$

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$$

S $\lambda > 15$, approximation par la

loi normale $\mathcal{N}(\lambda\,;\sqrt{\lambda}\,)$

Loi normale $\mathcal{N}(m\,;\sigma)$ de densité de probabilité f avec :

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right)$$

La variable $T=\frac{X-m}{\sigma}$ suit la **loi normale centrée réduite** $\mathcal{N}(0\,;1)$ de densité de probabilité f avec :

$$f(t)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right)$$

2. La convergence en probabilité

2.1. Définition

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur un espace probabilisé (Ω, $\mathcal{P}(\Omega)$, p) et X une variable aléatoire définie sur ce même espace.

On dit que la suite (X_n) converge en probabilité vers X si, et seulement si, quel que soit ε strictement positif : $\lim_{n\to+\infty} P(|X_n - X| < \varepsilon) = 1$

2.2. L'inégalité de Bienaymé-Tchébychev

Soit X une variable aléatoire de moyenne m et d'écart type σ.

Pour tout réel t strictement positif, $P(|X - m| \leq t) > 1 - \frac{\sigma^2}{t^2}$.

2.3. Le théorème de Bernoulli

Soit X une variable aléatoire qui suit $\mathcal{B}(n\ ;\ p)$.

On pose $F_n = \frac{X}{n}$ la fréquence associée.

Alors pour tout $\varepsilon > 0$, $P(|F_n - p| \leq \varepsilon) > 1 - \frac{pq}{n\varepsilon^2}$

C'est l'inégalité précédente avec $t = \varepsilon$ et $\sigma^2 = \frac{pq}{n}$.

Corollaire :

Soit X une variable aléatoire qui suit $\mathcal{B}(n\ ;\ p)$.

On pose $F_n = \frac{X}{n}$ la fréquence associée.

Alors, pour tout $\varepsilon > 0$, $\lim_{n\to+\infty} P\left(|F_n - p| \leq \varepsilon\right) = 1$

Autrement dit (F_n) converge en probabilité vers la variable aléatoire certaine p.

> Ce corollaire (version « allégée » de la loi faible des grands nombres) justifie le point de vue des « fréquentistes » qui attribuent comme probabilité d'un événement une valeur autour de laquelle la fréquence d'apparition de cet événement se stabilise lorsque le nombre d'expériences indépendantes devient très grand.

2.4. Un exemple d'application

Trouver un entier n_0 à partir duquel P($|F_n - 1/6| \leq 0{,}01$) > 0,95 :

– avec Bienaymé Tchebychev.

– avec la « fourchette » .

– avec la loi normale.

On lance un dé parfait n fois. On note F_n la fréquence de sortie de l'as.

Solution

1. $P\left(\left|F_n - \frac{1}{6}\right| \leq 10^{-2}\right) > 1 - \dfrac{\frac{1}{6}\times\frac{5}{6}}{n \times 10^{-4}}$

d'où $1 - \dfrac{5\times 10^{-4}}{36n} \geq 0{,}95$ ou $\dfrac{5\times 10^{-4}}{36n} \leq 5\times 10^{-2}$ soit $n \geq \dfrac{10^6}{36}$

$n_0 = 27\,778$

2. Fourchette (conditions d'approximation supposées satisfaites)

$P\left(\left|F_n - \frac{1}{6}\right| \leq \frac{1}{\sqrt{n}}\right) > 0{,}95.$

Il suffit de prendre $\dfrac{1}{\sqrt{n}} \leq 10^{-2}$ soit $n_0 = 10^4$

3. Approximation par la loi normale

On suppose $n > 30$. Ainsi les conditions sont vérifiées.

nF_n suit $\mathcal{B}\left(n\,;\frac{1}{6}\right)$ et $T_n = \dfrac{nF_n - \frac{n}{6}}{\sqrt{\frac{5n}{36}}}$ suit approximativement $\mathcal{N}(0\,;1)$.

$$T_n = \frac{F_n - \frac{1}{6}}{\frac{1}{6}\sqrt{\frac{5}{n}}}$$

$$P\left(\left|F_n - \frac{1}{6}\right| \leq 10^{-2}\right) \approx P\left(|T_n| \leq \frac{6\times 10^{-2}}{\sqrt{\frac{5}{n}}}\right)$$

$P\left(|T_n| \leq \dfrac{6\times 10^{-2}}{\sqrt{\frac{5}{n}}}\right) > 0{,}95 = P(|T_n| \leq 1{,}96)$ d'où $\dfrac{6\times 10^{-2}}{\sqrt{\frac{5}{n}}} > 1{,}96$

soit $\dfrac{\sqrt{n}}{\sqrt{5}} > \dfrac{196}{6}$ donc $n > \dfrac{5\times 196^2}{36} \approx 5\,325{,}55$

$n_0 = 5\,336$

Comparer les trois méthodes ! !