Chapitre 7

CONVERGENCE EN LOI

Les variables aléatoires X_n, X utilisées dans ce chapitre sont toutes définies sur le même espace probabilisé

7.1 CONVERGENCE EN LOI.

Souvent, en statistique, on ne connait pas les lois : on observe seulement une certaine distribution. Les théorèmes de convergence en loi permettent de justifier certaines approximations de distributions observées par des lois théoriques connues.

Définition 7.1 : On dit que la suite de variables aléatoires (X_n) **converge en loi** vers X si, pour tout x en lequel F est continue, $\lim_{n \to +\infty} F_n(x) = F(x)$. On écrit $X_n \overset{L}{\to} X$.

Théorème 8.1 : Si X_n suit la loi binomiale $\mathscr{B}(n, \frac{\lambda}{n})$ et si X suit la loi de Poisson $P(\lambda)$, alors $X_n \overset{L}{\to} X$

Loi faible des grands nombres : (admise)

Théorème 7.2 : Soit $(X_n)_n$ une suite de variables aléatoires d'ordre deux, deux à deux indépendantes, de même espérance m et de même variance $\sigma^2 > 0$. Alors, si $S_n = X_1 + \cdots + X_n$ et si $Z_n = \frac{S_n}{n}$, on a $Z_n \overset{L}{\to} m$.

8.2 CONVERGENCE VERS LA LOI NORMALE

Le théorème central limite, partie fondamentale de ce chapitre, sera admis.

Théorème 7.3 : Soit (X_n)n une suite de v.a.r. indépendantes d'ordre deux, de même loi, d'espérance m et de variance σ^2.

Alors, si $S_n = X_1 + \cdots + X_n$ et si $T_n = \frac{S_n - nm}{\sqrt{n\sigma}}$, on a $T_n \overset{L}{\to} T$ de loi normale $N(0, 1)$.

Remarque : $\mathrm{IE}(S_n) = nm$, $\mathrm{var}(S_n) = n\sigma^2$ et $T_n = \frac{S_n - \mathrm{IE}(S_n)}{\sqrt{\mathrm{var}(Sn)}}$,.

Application : Quand n est grand, c'est-à-dire en pratique, lorsque $n \geq 30$, S_n suit approximativement la loi normale N $(\mathrm{IE}(S_n), \mathrm{var}(Sn))$, c'est-à-dire la loi N $(nm, n\sigma^2)$ et $Z_n = \frac{S_n}{n}$ la loi normale N $(\mathrm{IE}(Z_n), \mathrm{var}(Z_n))$, c'est-à-dire la loi N $(m, \frac{\sigma^2}{n})$.

Correction de continuité : Lorsque l'on approxime la loi d'une v.a.r. entière X à valeurs dans IN par la loi normale N (IE(X), var(X)), de fonction de répartition F , cela revient à considérer X comme une variable aléatoire qui prend toutes les valeurs réelles, et l'intervalle]k - 0.5, k + 0.5] est l'ensemble de ces valeurs qui "s'arrondissent" à k. On remplacera donc P ([X = 0]) par F (0.5) et P ([X = k]) par F (k + 0.5) - F (k - 0.5).

(Si X est à valeur dans {0, · · · , n}, on procède de même, sauf pour P ([X = n]) que l'on remplace par 1- F (n - 0.5).)

7.3 APPROXIMATIONS

7.3.1 de la loi hypergéométrique par la loi binomiale.

Soit E un ensemble de N éléments, dont une proportion p est de type 1. On effectue dans E une série de n tirages successifs sans remise. Soit X le nombre d'éléments de type 1 obtenus : X suit la loi hypergéométrique $H(n, Np, N)$

$$\lim_{N\to+\infty} \frac{C_{Np}^{k} C_{N(1-p)}^{n-k}}{C_{Np}^{n}} = C_n^k p^k (1-p)^{n-k}$$

Application : Quand N est grand, c'est-à-dire en pratique lorsque $N \geq 10n$, on peut approximer la loi hypergéométrique $H(n, Np, N)$ par la loi binomiale $B(n, p)$

7.3.2 de la loi binomiale par la loi de Poisson.

Théorème 7.5 : Si $(p_n)_n$ est une suite de réels de [0, 1] telle que $\lim_{n\to+\infty} np_n = \lambda$, alors, pour tout k fixé,

$$\lim_{n\to+\infty} C_n^k p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

Application : En pratique, lorsque $p \leq 0.1$, $n \geq 30$ et $np < 15$, on peut approximer la loi binomiale $B(n, p)$ par la loi de Poisson $P(np)$

Exemple : Si X suit la loi binomiale $B(40, 0.03)$, $P([X = 2])= C_{40}^2 (0{,}03)^2 (0{,}97)^{38} \approx 0{,}2206$

Or $40 \times 0.03 = 1.2 < 15$, donc $B(40, 0.03)$ peut être approximée par $P(1.2)$ et

$$P([X = 2]) \simeq e^{-1,2} \frac{(1{,}2)^2}{2!} \approx 0{,}2169$$

7.3.3 de la loi de Poisson par la loi normale

Théorème 7.6 : Si $(X_n)_n$ est une suite de variables aléatoires indépendantes, de même loi de Poisson P(λ) et si $S_n = X_1 + \cdots + X_n$, alors S_n suit la loi de Poisson P(nλ).

Preuve : La démonstration se fait par récurrence sur n.

• Pour n = 1, X_1 suit bien la loi de Poisson P(λ).

• Si $S_n = X_1 + \cdots + X_n$ suit la loi de Poisson P(nλ), $S_{n+1} = S_n + X_{n+1}$ et, vu que S_n et X_{n+1} sont indépendantes, que S_n suit la loi de Poisson P(nλ) et X_{n+1} la loi de Poisson P(λ), d'après la propriété 6.2.2), $S_n + X_{n+1} = S_{n+1}$ suit la loi de Poisson de paramètre nλ + λ = (n + 1)λ.

Application : Sn suit la loi de Poisson P(nλ) donc $IE(S_n) = n\lambda$ et $var(S_n) = n\lambda$.

D'après le théorème central limite, la loi de Sn peut être approximée par la loi normale $N(IE(S_n), var(S_n))$, c'est-à-dire N (nλ, nλ).

En pratique, lorsque $\lambda \geq 15$, on peut approximer la loi de Poisson P(λ) par la loi normale N (λ,λ).

Exemple : Si X suit la loi de Poisson P(16), $P([X = 16]) \simeq e^{-16}\frac{16^2}{16!} \approx 0{,}0992$

En approximant la loi P(16) par la loi N (16, 16) de fonction de répartition F, on obtient : $P([X = 16]) \approx F(16.5) - F(15.5) + \Phi\left(\frac{0.5}{4}\right) - \Phi\left(-\frac{0.5}{4}\right) = 2\Phi(0.125) - 1 \approx 0.0995$

Le gain de temps est surtout sensible pour le calcul des valeurs de la fonction de répartition :

$$P([X \leq 20]) \approx F(20.5) = \Phi\left(\frac{20.5 - 16}{4}\right) = \Phi(1.125) \approx 0.8697$$

7.3.4 de la loi binomiale par la loi normale

Théorème 7.7 : Si $(X_n)_n$ est une suite de variables aléatoires indépendantes de même loi de Bernoulli $B(p)$ et si $S_n = X_1 + \cdots + X_n$, alors S_n suit la loi binomiale $B(n, p)$.

Application : S_n suit la loi binomiale $B(n, p)$ donc $IE(S_n) = np$ et $var(S_n) = np(1 - p)$.

D'après le théorème central limite, la loi de S_n peut être approximée par la loi normale $N(IE(S_n), var(S_n))$, c'est-à-dire la loi $N(np, np(1 - p))$.

En pratique, lorsque $n \geq 30$, $np \geq 15$ et $np(1 - p) > 5$, la loi binomiale $B(n, p)$ peut être approximée par la loi normale $N(np, np(1 - p))$

Si X suit la loi $B(n, p)$, et si F est la fonction de répartition de la loi $N(np, np(1 - p))$, on remplacera donc $P([X = 0])$ par $F(0.5)$; $P([X = k])$ par $F(k + 0.5) - F(k - 0.5)$ pour $1 \leq k \leq n - 1$ et $P([X = n])$ par $1 - F(n - 0.5)$

Exemple **:** Combien de lancers d'un dé faut-il effectuer pour pouvoir affirmer avec un risque d'erreur inférieur à 5% que la fréquence d'apparition de l'as au cours de ces lancers diffère de $\frac{1}{6}$ d'au plus $\frac{1}{100}$?

Solution :

Soit S_n le nombre d'as obtenus après n lancers : S_n suit la loi binomiale B (n, $\frac{1}{6}$) et la fréquence d'apparition de l'as au cours de ces lancers est $\frac{S_n}{n}$

.On cherche n tel que $P(\left|\frac{S_n}{n}-\frac{1}{6}\right| \geq \frac{1}{100}) \leq \frac{5}{100}$.

Lorsque $n \geq 30$, $\frac{n}{6} \geq 15$ et $\frac{5n}{36} > 5$ on peut approximer la loi de $\frac{S_n}{n}$ par la loi normale B $N\left(\frac{1}{6}, \frac{5n}{36}\right)$ et la loi $\frac{S_n}{n}-\frac{1}{6}$ par la loi normale $N\left(0, \frac{5}{36n}\right)$.

$$P\left(\left|\frac{S_n}{n}-\frac{1}{6}\right| \geq \frac{1}{100}\right) = P\left(\left|\frac{6\sqrt{n}}{\sqrt{5}}\left(\frac{S_n}{n}-\frac{1}{6}\right)\right| \geq \frac{6\sqrt{n}}{\sqrt{5}}\frac{1}{100}\right) \approx 2\left(1-\Phi\left(\frac{6\sqrt{n}}{\sqrt{5}}\frac{1}{100}\right)\right) \approx 0{,}05$$

D'après le théorème 3.4. On a donc $\Phi\left(\frac{6\sqrt{n}}{\sqrt{5}}\frac{1}{100}\right) \approx 0{,}975 \approx \Phi(1{,}96)$ et

$$n = 1{,}96^2 \times \frac{5000}{36} \approx 5536$$

On a bien $5536 \geq 90,$ ce qui légitime l'approximation

Exercice : Sur une autoroute, la proportion des camions par rapport à l'ensemble des véhicules est 0.07.

1) Soit X le nombre de camions parmi 100 véhicules choisis au hasard. Calculer $P([X \geq 5])$.

2) Soit Y le nombre de camions parmi 1000 véhicules choisis au hasard. Calculer $P([65 \leq Y \leq 75])$.

Solution :

1) Soit X une v.a.r. de loi binomiale B(100, 0.07).

$100 \geq 30$, $100 \times 0.07 = 7 < 15$, $0.07 \leq 0.1$ donc l'approximation à utiliser est celle par la loi de Poisson $P(16)$,

$$P([X \geq 5]) \simeq 1 - e^{-7}\sum_{k=0}^{4}\frac{7^k}{k!} \approx 0{,}827$$

2) Y suit la loi binomiale B(1000, 0.07).

$1000 \geq 30$, $1000\times0.07 = 70 \geq 15$, $70\times0.93 = 65.1 > 5$ donc l'approximation à utiliser est celle par la loi normale N (70, 65.1) et si F désigne la fonction de répartition de la loi N (70, 65.1)

$$P([65 \leq Y \leq 75]) \approx F(75{,}5) - F(64{,}5) = \Phi\left(\frac{5{,}5}{\sqrt{65{,}1}}\right) - \Phi\left(-\frac{5{,}5}{\sqrt{65{,}1}}\right)$$

$$= 2\Phi\left(\frac{5,5}{\sqrt{65,1}}\right) - 2 \approx \Phi(0,68) - 1 \approx 0,5$$