Loi des grands nombres et théorème de la limite centrale

1 Loi des Grands Nombres.

La loi des grands nombres, due à Kolmogorov, est un résultat fondamental en Probabilités.

Théorème 1. Loi des grands nombres. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi. On suppose que $(X_n)_{n\geq 1}$ est intégrable et on note m son espérance. Si $S_n = X_1 + X_2 + \cdots + X_n$, alors on a

$$\frac{S_n}{n} \longrightarrow m \qquad \text{p.s.}$$

Exercice 1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi de Bernoulli $\mathcal{B}(p)$ avec $0 < p < 1$. Pour tout $n \geq 1$, on pose

$$S_n = \sum_{k=1}^{n} X_k \qquad \text{et} \qquad \overline{X}_n = \frac{S_n}{n}.$$

Montrer que $\overline{X}_n$ est un estimateur sans biais de p et $\overline{X}_n \to p$ p.s. Afin d'estimer la variance $\sigma^2 = p(1-p)$, on propose d'utiliser $U_n = \overline{X}_n(1-\overline{X}_n)$. Montrer que U_n est un estimateur biaisé de σ^2 et trouver un estimateur V_n sans biais de σ^2 tel que $V_n \to \sigma^2$ p.s. Créer un code Matlab illustrant ces deux LGN sur un n-échantillon de loi de Bernoulli $\mathcal{B}(p)$, où les paramètres n et p sont affectés par l'utilisateur.

Exercice 2. La durée de vie d'une ampoule électrique peut être modélisée par une variable aléatoire X prenant au hasard ses valeurs dans l'intervalle $[0, \theta]$ avec $\theta > 0$. Afin d'optimiser l'agenda d'un réparateur, on cherche à estimer θ à partir d'un n-échantillon $(X_1, X_2, \cdots, X_n)$ de même loi que X. On propose d'estimer θ par

$$\widehat{\theta}_n = \max_{1\leq k\leq n} X_k \qquad \text{ou} \qquad \widetilde{\theta}_n = 2\overline{X}_n.$$

Calculer la fonction de répartition puis la densité de probabilité de $\widehat{\theta}_n$ et en déduire que $\widehat{\theta}_n \to \theta$ p.s. Montrer également que $\widetilde{\theta}_n \to \theta$ p.s. Créer un code Matlab illustrant ces deux LGN sur un n-échantillon de loi uniforme $\mathcal{U}([0, \theta])$, où les paramètres n et θ sont affectés par l'utilisateur.

Exercice 3. La loi de Paréto, encore appelée loi de puissance, est utilisée pour modéliser les dépassements d'un seuil. On dit que X suit une loi de Paréto de paramètres $a, \theta > 0$ si $X = \theta \exp(Y)$ avec Y de loi exponentielle $\mathcal{E}(a)$. On propose d'estimer θ, à partir d'un n-échantillon $(X_1, X_2, \cdots, X_n)$ de même loi que X, par

$$\widehat{\theta}_n = \min_{1\leq k\leq n} X_k.$$

Calculer la fonction de répartition puis la densité de probabilité de X. En déduire la loi de probabilité associée à $\widehat{\theta}_n$ et montrer que $\widehat{\theta}_n \to \theta$ p.s. Créer un code Matlab illustrant cette LGN sur un n-échantillon de loi de Paréto $\mathcal{P}(a, \theta)$, où les paramètres n, a et θ sont affectés par l'utilisateur.

2 Théorème Limite Centrale.

Le théorème de la limite centrale est le second résultat fondamental en Probabilités. Le premier TLC, dû à De Moivre et connu sous le nom de théorème de Moivre-Laplace, concerne le cas particulier des variables aléatoires de loi de Bernoulli.

Théorème 2. Théorème limite centrale. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi. On suppose que $(X_n)_{n\geq 1}$ est de carré intégrable et on note m son espérance et $\sigma^2 > 0$ sa variance. Si $S_n = X_1 + X_2 + \cdots + X_n$, alors on a

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Remarque. En particulier, pour tout $a, b \in \mathbb{R}$ avec $a < b$, on a

$$\lim_{n\to\infty} \mathbb{P}\Big(a \leq \frac{S_n - nm}{\sigma\sqrt{n}} \leq b\Big) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp(-\frac{x^2}{2})\, dx.$$

Exercice 4. Utiliser le code Matlab suivant qui permet d'illustrer le TLC sur un n-échantillon de loi Uniforme $\mathcal{U}([0,1])$.

```
n = input('Entrez la taille de l'échantillon n (par exemple 10) : ');
N = input('Entrez le nombre de réalisations N (par exemple 2000) : ');
d = input('Préciser la discrétisation d (par exemple 100) : ');
X = rand(n, N); Z = (sum(X) - n/2)/sqrt(n/12);
dz = range(Z)/d; [Effectifs, Classes]=hist(Z, d);
Abscisses = [min(Z) - dz/2 : dz : max(Z) + dz/2];
DS =[0 Effectifs 0]/(N * dz); bar(Abscisses, DS,'b'); hold on
DN = dnorm(Abscisses); plot(Abscisses, DN,'r-');
legend('Densite simulée','Loi Normale'); hold off
```

Exercice 5. Soit $(\varepsilon_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi de Rademacher $\mathcal{R}(1/2)$. Soit $(X_n)_{n\geq 1}$ la suite définie par $X_n = n^a \varepsilon_n$ avec $a > 0$. On pose $S_n = X_1 + X_2 + \cdots + X_n$. Calculer la transformée de Laplace de S_n donnée, pour tout $t \in \mathbb{R}$, par $L_n(t) = \mathbb{E}[\exp(tS_n)]$. En déduire que

$$\frac{S_n}{n^a\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \frac{1}{2a+1}\Big).$$

Créer un code Matlab illustrant ce TLC sur un n-échantillon de loi de Rademacher $\mathcal{R}(1/2)$, où les paramètres n et a sont affectés par l'utilisateur.

Exercice 6. Pour tout $x, \lambda > 0$, on pose

$$L_n(x) = \exp(-n\lambda) \sum_{k=0}^{[nx]} \frac{(n\lambda)^k}{k!}.$$

Montrer à l'aide de la loi des grands nombres et du théorème de la limite centrale que

$$\lim_{n\to} L_n(x) = \begin{cases} 0 & \text{si } x < \lambda, \\ 1/2 & \text{si } x = \lambda, \\ 1 & \text{si } x > \lambda. \end{cases}$$

Vérifier cette convergence avec un code Matlab.