DM 21 - Règle de succession de Laplace – Loi hypergéométrique – Théorème de Weierstrass

1 Règle de succession de Laplace

Soient N et n deux entiers strictement positifs. On considère $N+1$ urnes, notées $U_0,\dots,U_N$. Pour chaque $k\in[\![0,N]\!]$, l'urne U_k est composée de k boules rouges et $N-k$ boules blanches. On tire au hasard uniformément un entier $K\in[\![0,N]\!]$, puis on tire uniformément et avec remise $n+1$ boules dans l'urne U_K. Pour tout $j\in[\![0,n+1]\!]$, on note S_j le nombre de boules rouges tirées à l'issue du j-ème tirage.

1. Pour tout $j\in[\![0,n+1]\!]$, déterminer la loi conditionnelle de S_j sachant $(K=k)$.
2. En déduire la loi de S_j.
3. On suppose qu'on a tiré n boules rouges lors des n premiers tirages. Quelle est la probabilité $p_{N,n}$ de tirer encore une boule rouge au $(n+1)$-ème tirage ?
4. Déterminer la limite $\lim\limits_{N\to+\infty} p_{N,n}$.
5. On suppose maintenant qu'on a tiré s boules rouges lors des n premiers tirages, où $s\in[\![0,n]\!]$. Quelle est la probabilité $q_{N,n,s}$ de tirer une boule rouge au $(n+1)$-ème tirage ?
6. Pour tous entiers naturels a,b, on note $I(a,b)=\displaystyle\int_0^1 x^a(1-x)^b\,dx$.
   a) Déterminer une relation entre $I(a,b+1)$ et $I(a+1,b)$ pour tous $a,b\in\mathbb{N}$.
   En déduire la valeur de $I(a,b)$.
   b) En déduire la limite $\lim\limits_{N\to+\infty} q_{N,n,s}$.

2 Sur la loi hypergéométrique

1. On pioche une main de 5 cartes dans un jeu de 52 cartes et on note X le nombre d'As obtenus. Déterminer la loi de X.

On généralise en considérant une urne qui contient N boules, rouges ou blanches. On note p la proportion de boules rouges et $q=1-p$ la proportion de boules blanches. On tire aléatoirement $n\leq N$ boules dans l'urne, *sans remise*, et on note X le nombre de boules rouges tirées.

2. Quel est l'ensemble des valeurs prises par X ?
   On distinguera selon la position relative de n, pN et qN.
3. Déterminer la loi de X.

On dit que X suit la *loi hypergéométrique* de paramètres n,N,p et on note $X\sim\mathscr{H}(n,N,p)$.
On souhaite montrer que si $n\ll N$, la loi hypergéométrique $\mathscr{H}(n,N,p)$ est bien approchée par la loi binomiale $\mathscr{B}(n,p)$.

4. Justifier heuristiquement ce résultat.

On fixe $n \in \mathbb{N}^*$ et $p \in [0,1] \cap \mathbb{Q}$. Pour tout $N \geq n$ tel que pN est entier, on note X_N une variable aléatoire suivant une loi $\mathcal{H}(n, N, p)$ et Y une variable aléatoire suivant une loi $\mathcal{B}(n,p)$.

5. Montrer le résultat de *convergence en loi*[1] suivante : $\forall k \in \mathbb{N}, \lim\limits_{\substack{N \to +\infty \\ pN \in \mathbb{N}}} P(X_N = k) = P(Y = k)$.

6. Calculer l'espérance de $X \sim \mathcal{H}(N, n, p)$.

   On pourra utiliser la formule de Vandermonde.

7. Soient X_1 et X_2 deux variables aléatoires indépendantes, de lois $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$, où $n_1, n_2 \in \mathbb{N}^*$ et $p \in]0,1[$. Soit $n \in [\![0, n_1 + n_2]\!]$. Montrer que la loi de X_1 conditionnellement à l'évènement $(X_1 + X_2 = n)$ est une loi hypergéométrique dont on précisera les paramètres.

3 Théorème de Weierstrass par les polynômes de Bernstein

On cherche à montrer le théorème de Weierstrass suivant : si $f \in \mathcal{C}\big([a,b], \mathbb{R}\big)$, il existe une suite $(P_n)_{n\in\mathbb{N}}$ de fonctions polynomiales convergeant uniformément vers f sur $[a,b]$.

Soit f une fonction continue sur $[0,1]$. Pour tout $x \in [0,1]$ et tout $n \in \mathbb{N}^*$, on note $X_{n,x}$ une variable aléatoire de loi $\mathcal{B}(n,x)$.

1. Montrer que l'application $x \mapsto E\left(f\left(\dfrac{X_{n,x}}{n}\right)\right)$ est polynomiale de degré au plus n.

On note $B_n(f) \in \mathbb{R}_n[X]$ le polynôme ainsi défini et on fixe $\varepsilon > 0$.

2. Justifier l'existence de $\delta > 0$ tel que : $\forall y, z \in [0,1], |y - z| \leq \delta \implies |f(y) - f(z)| \leq \dfrac{\varepsilon}{2}$.

Pour tout $x \in [0,1]$ et tout $n \in \mathbb{N}^*$, on note $A_{n,x}$ l'évènement $\left(\left|\dfrac{X_{n,x}}{n} - x\right| \leq \delta\right)$.

3. Justifier les égalités et l'inégalité suivantes :

$$\begin{aligned}|B_n(f)(x) - f(x)| = \left|E\left(f\left(\frac{X_{n,x}}{n}\right) - f(x)\right)\right| \leq E\left(\left|f\left(\frac{X_{n,x}}{n}\right) - f(x)\right|\right) &= E\left(\left|f\left(\frac{X_{n,x}}{n}\right) - f(x)\right| \mathbb{1}_{A_{n,x}}\right) \\ &+ E\left(\left|f\left(\frac{X_{n,x}}{n}\right) - f(x)\right| \mathbb{1}_{\overline{A_{n,x}}}\right)\end{aligned}$$

4. Montrer que $E\left(\left|f\left(\dfrac{X_{n,x}}{n}\right) - f(x)\right| \mathbb{1}_{A_{n,x}}\right) \leq \dfrac{\varepsilon}{2}$.

5. Montrer que $E\left(\left|f\left(\dfrac{X_{n,x}}{n}\right) - f(x)\right| \mathbb{1}_{\overline{A_{n,x}}}\right) \leq 2\|f\|_\infty P(\overline{A_{n,x}})$.

6. En déduire que pour tout $x \in [0,1]$ et tout $n \in \mathbb{N}^*$, $|B_{n,f}(x) - f(x)| \leq \dfrac{\varepsilon}{2} + \dfrac{\|f\|_\infty}{2n\delta^2}$.

7. Conclure.

[1] Remarquons que l'énoncé a un sens même si les variables aléatoires X_N et Y sont définies sur des univers différents.