DM 21 - Règle de succession de Laplace – Loi hypergéométrique

Quelques notions.

Probabilité conditionnelle. Soit A un évènement dans un espace probabilisé fini (Ω, P) tel que $P(A) > 0$. Si B est un évènement quelconque, on note $P_A(B)$ ou $P(B \mid A)$ la probabilité de B sachant A, définie comme $P_A(B) = \dfrac{P(A \cap B)}{P(A)}$. Intuitivement, c'est la probabilité que l'évènement B ait lieu *sachant que* l'évènement A s'est réalisé.

Loi conditionnelle. Soit A un évènement dans (Ω, P) tel que $P(A) > 0$. Si X est une variable aléatoire sur Ω, la loi conditionnelle de X sachant A est la donnée, pour tout $x \in X(\Omega)$, de la probabilité conditionnelle $P_A(X = x)$.

Formule des probabilités totales. Soit $(A_1, \dots, A_n)$ un système complet d'évènements dans (Ω, P). Si B est un évènèment quelconque, on a $P(B) = \sum_{k=1}^{n} P(A_k) P_{A_k}(B)$, où $P(A_k) P_{A_k}(B)$ vaut par définition 0 si $P(A_k) = 0$.

1 Règle de succession de Laplace

Soient N et n deux entiers naturels non nuls. 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 ?
   Indication. En notant A_{n+1} l'évènement *Une boule rouge est tirée au* $n+1$*-ème tirage*, on admettra[1] que $\forall k \in [\![0, N]\!], P_{(K=k)}\big((S_n = s) \cap A_{n+1}\big) = P_{(K=k)}(S_n = s) P_{(K=k)}(A_{n+1})$.
6. Pour tous entiers naturels a, b, on note $I(a, b) = \displaystyle\int_0^1 x^a (1-x)^b \, dx$.

[1] On dira que les deux évènements $(S_n = s)$ et A_{n+1} sont indépendants conditionnellement à l'évènement $(K = k)$.

a) Montrer que $I(a,b) = I(a+1,b) + I(a,b+1)$.

b) Montrer que $I(a+1,b) = \dfrac{a+1}{b+1} I(a,b+1)$.

c) 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 cette situation en considérant une urne qui contient N boules, rouges ou blanches. On note $p \in [0,1]$ 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 approchable par la loi binomiale $\mathscr{B}(n,p)$.

4. Justifier heuristiquement ce résultat.

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

5. Montrer le résultat de *convergence en loi*[2] suivante : $\forall k \in \mathbb{N}, \lim\limits_{\substack{N\to+\infty \\ pN\in\mathbb{N}}} P(X_N = k) = P(Y = k)$.

L'espérance $E(X)$ d'une variable aléatoire X à valeurs dans $[\![0,n]\!]$ est $E(X) = \sum\limits_{k=0}^{n} kP(X=k)$.

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

   On pourra utiliser la formule de Vandermonde.

7. Soient X_1 et X_2 deux variables aléatoires indépendantes[3], de lois $\mathscr{B}(n_1,p)$ et $\mathscr{B}(n_2,p)$, 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.

[2]Remarquons que l'énoncé a un sens même si les variables aléatoires X_N et Y sont définies sur des univers différents.
[3]On utilisera que $X_1 + X_2 \sim \mathscr{B}(n_1+n_2,p)$ et que $\forall k,\ell \in \mathbb{N}, P\big((X_1=k)\cap(X_2=\ell)\big) = P(X_1=k)P(X_2=\ell)$.