DM 21 - Règle de succession de Laplace - Loi hypergéométrique

1 Règle de succession de Laplace

1. Soit $k \in \llbracket 0, N \rrbracket$. Si l'évènement $(K = k)$ est réalisé, les tirages ont lieu avec remise dans une urne composée de k boules rouges et $N - k$ boules blanches.

Ainsi, sachant $(K = k)$, S_j suit une loi binomiale de paramètres $(j, \frac{k}{N})$. On a donc :

$$\forall k \in \llbracket 0, N \rrbracket, \forall j \in \llbracket 0, n+1 \rrbracket, \forall \ell \in \llbracket 0, j \rrbracket, P_{(K=k)}(S_j = \ell) = \binom{j}{\ell} \left(\frac{k}{N}\right)^\ell \left(1 - \frac{k}{N}\right)^{j-\ell}.$$

2. Les évènements $(K = k)$, pour $k \in \llbracket 0, N \rrbracket$ forment un système complet. Par la formule des probabilités totales :

$$\forall \ell \in \llbracket 0, j \rrbracket, P(S_j = \ell) = \sum_{k=0}^N P(K = k) P_{(K=k)}(S_j = \ell).$$

Comme K suit une loi uniforme sur $\llbracket 0, N \rrbracket$:

$$\forall \ell \in \llbracket 0, j \rrbracket, P(S_j = \ell) = \frac{\binom{j}{\ell}}{N+1} \sum_{k=0}^N \left(\frac{k}{N}\right)^\ell \left(1 - \frac{k}{N}\right)^{j-\ell}.$$

3. Par définition, $p_{N,n} = P_{S_n=n}(S_{n+1} = n+1)$. Or, $(S_n = n) \cap (S_{n+1} = n+1) = (S_{n+1} = n+1)$; il faut en effet avoir tiré n boules rouges lors des n premiers tirages pour avoir une chance d'en tirer $n+1$ lors des $n+1$ premiers. Donc,

$$p_{N,n} = \frac{P(S_{n+1} = n+1)}{P(S_n = n)} = \frac{\frac{1}{N+1} \sum_{k=0}^N \left(\frac{k}{N}\right)^{n+1}}{\frac{1}{N+1} \sum_{k=0}^N \left(\frac{k}{N}\right)^n}.$$

On pourrait bien sûr simplifier les facteurs $\frac{1}{N+1}$ mais ils vous nous être utiles.

4. Pour tout entier i , on note f_i la fonction définie sur $[0, 1]$ par $f_i(x) = x^i$. On a donc

$$p_{N,n} = \frac{\frac{1}{N+1} \sum_{k=0}^N f_{n+1}\left(\frac{k}{N}\right)}{\frac{1}{N+1} \sum_{k=0}^N f_n\left(\frac{k}{N}\right)}.$$

On reconnaît deux séries de Riemann. Par opérations sur les limites :

$$\lim_{N \rightarrow +\infty} p_{N,n} = \frac{\int_0^1 x^{n+1} dx}{\int_0^1 x^n dx} = \frac{\frac{1}{n+2}}{\frac{1}{n+1}} = \frac{n+1}{n+2}.$$

Dans son *Essai philosophique sur les probabilités*, Laplace utilise ce raisonnement pour répondre à la question : *Quelle est la probabilité que le Soleil se lève demain ?*

Il part du principe que cette probabilité a été décidée une fois pour toutes (choix d'une urne de façon équiprobable), puis utilise que le Soleil s'est levé tous les jours pendant 5000 ans, soit environ 1825000 jours. L'application numérique donne une probabilité d'environ 99,99995 %.

Le lecteur pensera ce qu'il veut de ce raisonnement.

5. On doit calculer $P_{(S_n=s)}(S_{n+1} = s + 1)$. *Attention ! le calcul diffère du précédent.*

Notons A_{n+1} l'évènement : *boule rouge au n + 1-ème tirage*. Alors $(S_n = s) \cap (S_{n+1} = s + 1) = (S_n = s) \cap (A_{n+1})$. Or, conditionnellement à un évènement $(K = k)$, les deux évènements $(S_n = s)$ et (A_{n+1}) sont indépendants (*autrement dit, ces évènements sont indépendants pour la mesure de probabilité $P_{(K=k)}$*). On a donc :

$$\begin{aligned} P((S_n = s) \cap (S_{n+1} = s + 1)) &= \sum_{k=0}^N P(K = k) P_{(K=k)}((S_n = s) \cap (A_{n+1})) \\ &= \frac{1}{N+1} \sum_{k=0}^N P_{(K=k)}(S_n = s) P_{(K=k)}(A_{n+1}) \\ &= \frac{1}{N+1} \sum_{k=0}^N \binom{n}{s} \left(\frac{k}{N}\right)^s \left(1 - \frac{k}{N}\right)^{n-s} \frac{k}{N}. \end{aligned}$$

D'où

$$q_{N,n,s} = \frac{P((S_n = s) \cap (S_{n+1} = s + 1))}{P(S_n = s)} = \frac{\frac{1}{N+1} \sum_{k=0}^N \left(\frac{k}{N}\right)^{s+1} \left(1 - \frac{k}{N}\right)^{n-s}}{\frac{1}{N+1} \sum_{k=0}^N \left(\frac{k}{N}\right)^s \left(1 - \frac{k}{N}\right)^{n-s}}.$$

6. a) On écrit simplement que $1 = x + (1 - x)$. Et donc, $x^a(1 - x)^b = x^{a+1}(1 - x)^b + x^a(1 - x)^{b+1}$.
Puis, on intègre sur $[0, 1]$ et on utilise la linéarité de l'intégrale.
b) Par intégration par parties :

$$\int_0^1 x^{a+1} x^b dx = \left[-x^{a+1} \frac{(1-x)^{b+1}}{b+1} \right]_0^1 - \int_0^a \frac{a+1}{b+1} x^a (-(1-x)^{b+1}) dx.$$

D'où $I(a+1, b) = \frac{a+1}{b+1} I(a, b+1)$.

c) De nouveau, on reconnaît des sommes de Riemann et

$$\lim_{N \rightarrow +\infty} q_{N,n,s} = \frac{\int_0^1 x^{s+1} (1-x)^{n-s} dx}{\int_0^1 x^s (1-x)^{n-s} dx} = \frac{I(s+1, n-s)}{I(s, n-s)}$$

Or, on sait que $I(s, n-s) = I(s+1, n-s) + I(s, n-s+1)$ et que $I(s+1, n-s) = \frac{s+1}{n-s+1} I(s, n-s+1)$. Donc, $I(s, n-s) = I(s+1, n-s) \left(1 + \frac{n-s+1}{s+1}\right) = I(s+1, n-s) \frac{n+2}{s+1}$.

Finalement, $\lim_{N \rightarrow +\infty} q_{N,n,s} = \frac{s+1}{n+2}$.

2 Sur la loi hypergéométrique

1. X a comme ensemble de valeurs $[0, 4]$. Soit $k \in [0, 4]$. Pour calculer la probabilité de $X = k$, on compte le nombre de mains comptant k As. Une telle main est obtenue en choisissant k As parmi 4, puis $5 - k$ cartes parmi les 48 restantes.

$$\text{Donc, } P(X = k) = \frac{\binom{4}{k} \binom{48}{5-k}}{\binom{52}{5}}.$$

2. Il y a dans l'urne pN boules rouges et qN boules blanches. Le nombre maximal de boules rouges tirées est donc $\min(n, pN)$ et le nombre minimal est $\max(0, n - qN)$.

Donc, l'ensemble des valeurs prises par X est $[\max(0, n - qN), \min(n, pN)]$.

3. Soit $k \in [\max(0, n - qN), \min(n, pN)]$. L'évènement $(X = k)$ correspond à choisir k boules rouges, parmi pN , et $n - k$ boules blanches, parmi qN . Et il y a au total $\binom{N}{n}$ tirages possibles.

$$\text{Donc, } P(X = k) = \frac{\binom{pN}{k} \binom{qN}{n-k}}{\binom{N}{n}}.$$

4. Si N est très grand devant n , il y a peu de chances de tirer deux fois la même boule quand on en tire n dans une urne en contenant N . Donc, l'expérience *sans remise* est bien approchable par l'expérience *avec remise*; c'est-à-dire que la loi $\mathcal{H}(n, N, p)$ est bien approchable par la loi $\mathcal{B}(n, p)$.

5. Soient $N \in \mathbb{N}$ tel que $pN \in \mathbb{N}$, soit $k \in \mathbb{N}$. On a

$$P(X_N = k) = \frac{\binom{pN}{k} \binom{qN}{n-k}}{\binom{N}{n}}.$$

Avec k fixé, on va obtenir un équivalent des trois coefficients binomiaux, quand $N \rightarrow +\infty$. On a

$$\begin{aligned} \binom{pN}{k} &= \frac{pN(pN-1)\dots(pN-k+1)}{k!} \sim \frac{p^k N^k}{k!}. \\ \binom{qN}{n-k} &= \frac{qN(qN-1)\dots(qN-n+k+1)}{(n-k)!} \sim \frac{q^{n-k} N^{n-k}}{(n-k)!}. \\ \binom{N}{n} &= \frac{N(N-1)\dots(N-n+1)}{n!} \sim \frac{N^n}{n!}. \end{aligned}$$

Par opérations élémentaires, on a donc :

$$P(X_N = k) \sim \frac{p^k N^k q^{n-k} N^{n-k} n!}{N^n k!} (n-k)! = \binom{n}{k} p^k q^{n-k}.$$

Ainsi, on a bien $\lim_{N \rightarrow +\infty} P(X_N = k) = P(Y = k)$.

6. On propose deux méthodes.

- On a $E(X) = \sum_{k=0}^n k \frac{\binom{pN}{k} \binom{qN}{n-k}}{\binom{N}{n}}$. Par la formule du chef, $k \binom{pN}{k} = pN \binom{pN-1}{k-1}$. On a donc :

$$E(X) = \frac{pN}{\binom{N}{n}} \sum_{k=0}^n \binom{pN-1}{k-1} \binom{qN}{n-k} = \frac{pN}{\binom{N}{n}} \sum_{\ell=0}^{n-1} \binom{pN-1}{\ell} \binom{qN}{(n-1)-\ell}.$$

Par la formule de Vandermonde, la somme vaut $\binom{pN-1+qN}{n-1} = \binom{N-1}{n-1}$. D'où

$$E(X) = \frac{pN \binom{N-1}{n-1}}{\binom{N}{n}} = pN \frac{n}{N} = pn.$$

- (Après avoir vu le cours sur l'espérance.) On numérote les boules rouges de 1 à pN . Pour chaque $k \in \llbracket 1, pN \rrbracket$, on note A_k l'évènement : la boule rouge numéro k est tirée lors d'un des n tirages.

Alors, $X = \sum_{k=1}^{pN} \mathbb{1}_{A_k}$. Donc, par linéarité de l'espérance, $E(X) = \sum_{k=1}^{pN} E(\mathbb{1}_{A_k}) = \sum_{k=1}^{pN} P(A_k)$.

Il s'agit donc de calculer $P(A_k)$. C'est la probabilité qu'une boule quelconque se trouve dans une partie de n éléments. Or, il y a $\binom{N-1}{n-1}$ parties à n éléments contenant une

boule fixée (on ajoute les autres éléments) et un total de $\binom{N}{n}$ parties au total. Ainsi,

$$\text{pour tout } k, P(A_k) = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}.$$

On retrouve bien $E(X) = pn$.

7. Soit $k \in \llbracket 0, n_1 \rrbracket$.

$$\begin{aligned} P_{X_1+X_2=n}(X_1 = k) &= \frac{P(X_1 = k) \cap (X_2 = n - k)}{P(X_1 + X_2 = n)} \\ &= \frac{P(X_1 = k)P(X_2 = n - k)}{P(X_1 + X_2 = n)} \text{ par indépendance} \\ &= \frac{\binom{n_1}{k} p^k (1-p)^{n_1-k} \binom{n_2}{n-k} p^{n-k} (1-p)^{n_2-n+k}}{\binom{n_1+n_2}{n} p^n (1-p)^{n_1+n_2-n}} \\ P_{X_1+X_2=n}(X_1 = k) &= \frac{\binom{n_1}{k} \binom{n_2}{n-k}}{\binom{n_1+n_2}{n}}. \end{aligned}$$

Pour le passage de la deuxième à la troisième ligne, on a utilisé que $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$, car X_1 et X_2 sont indépendantes.

Ainsi, la loi de X_1 , conditionnellement à l'évènement $(X_1 + X_2 = n)$ est la loi $\mathcal{H}(n, n_1 + n_2, \frac{n_1}{n_1 + n_2})$.

Interprétation : On peut se représenter ainsi le résultat. On dispose de deux urnes : une urne avec n_1 boules rouges et une urne avec n_2 boules blanches. On tire $n_1 + n_2$ fois une pièce qui

a probabilité p de tomber sur Pile ; pour les n_1 premiers lancers, si la pièce tombe sur Pile, on prend une boule rouge de la première urne ; pour les n_2 suivants, on prend une boule blanche de la deuxième urne.

On collecte toutes les boules et on constate qu'il y en a n ; on cherche à déterminer la loi du nombre de boules rouges. Il est à *peu près clair* que tout se passe comme si on avait dès le départ réuni le contenu des deux urnes et pioché N boules (sans remise) ; d'où le résultat.