PETITE FORMULE PROBABILISTE UNITAIRE

Cet article nous a été envoyé par Edith KOSMANEK, de l'Université de Paris I-Sorbonne, après qu'elle a lu l'article « Simulation d'un exercice de probabilités » paru dans le Petit Vert n°56 de décembre 1998.

1. Le modèle générateur

Le modèle : le tirage d'un échantillon dans une urne bicolore à N boules (N_p blanches, N_q noires), avec ou sans remise, en fixant soit le nombre n « d'essais » (la taille de l'échantillon), soit le nombre k de « succès » (tirage d'une boule blanche), génère quatre lois de probabilités dites usuelles :

- binomiale $B(n, p)$
- hypergéométrique $H(N, n, p)$
- Pascal avec remise $G(k, p)$
- Pascal sans remise $S(N, k, p)$

Les formules associées, telles qu'elles paraissent dans la littérature, peuvent sembler disparates. En fait, il s'agit de ces particuliers d'une même formule de base.

2. La formule unitaire (1)

$$P(k, n) = C_{n-\varepsilon}^{k-\varepsilon} \cdot \frac{N_p^{[k]}, N_q^{[n-k]}}{N^{[n]}}$$

Avec les notations suivantes :

$$P(k, n) = \begin{cases} P(X_n = k) \\ P(X_k = n) \end{cases}$$

où X_n est la variable aléatoire 'nombre de succès à n fixé',
et où X_k est la variable aléatoire 'nombre d'essais à k fixé' ;

$$\varepsilon = \begin{cases} 0 & \text{si } n \text{ est fixé} \\ 1 & \text{si } k \text{ est fixé} \end{cases} ;$$

$$x^{[k]} = \begin{cases} x \cdot x \cdot \dots \cdot x = x^k & \text{(puissance) si tirage avec remise} \\ x \cdot (x-1) \cdot \dots \cdot (x-k+1) = \frac{x!}{(x-k)!} & \text{(produit factoriel) si tirage sans remise} \end{cases}$$

Par ailleurs :

Le facteur $C_{n-\varepsilon}^{k-\varepsilon}$ désigne le nombre de n-suites distinctes possibles à k blanches

et (n-k) noires ; il vaut donc C_n^k pour n fixé, et $C_{n-1}^{k-1} = \frac{k}{n} \times C_n^k$ pour k fixé.

On passe donc du cas « n fixé » au cas « k fixé » en multipliant par le « taux de succès » k/n, que le tirage soit avec ou sans remise.

Le facteur $\frac{N_p^{[k]} \cdot N_q^{[n-k]}}{N^{[n]}}$ représente la probabilité associée à l'une quelconque

de ces suites. Il vaut $p^k \cdot q^{n-k}$ dans le cas avec remise (puissance), et $\frac{C_{N-p}^{Np-k}}{C_N^{Np}}$

(formule 2) dans le cas sans remise (produit factoriel). Cette dernière expression est

aussi égale à : $\frac{C_{Np}^k \cdot C_{Nq}^{n-k}}{C_n^k \cdot C_N^n}$ (formule 3)

On retrouve ainsi pour P(k,n) les quatre formules de base.

3. Tableau récapitulatif des 4 formules de base

ε	avec remise (puissance)	sans remise (produit factoriel)
0	Binomiale $C_n^k \cdot p^k \cdot q^{n-k}$	Hypergéométrique $\frac{C_n^k \cdot C_{N-n}^{Np-k}}{C_N^{Np}}$
1	Pascal avec remise $C_{n-1}^{k-1} \cdot p^k \cdot q^{n-k}$	Pascal sans remise $\frac{C_{n-1}^{k-1} \cdot C_{N-n}^{Np-k}}{C_N^{Np}}$

En utilisant la variante (3) de la formule (2), on retrouve l'autre expression de la probabilité hypergéométrique, qui manifeste la dualité entre n et N_p , et donc entre

(N - n) et N_q : $P(k, n) = \frac{C_{Np}^k \cdot C_{Nq}^{n-k}}{C_N^n}$.

Pour le cas « Pascal sans remise », on a aussi les écritures suivantes :

$$P(k,n) = \frac{k}{n} \times \frac{C_{N_p}^k \cdot C_{N_q}^{n-k}}{C_N^n} = \frac{C_{N_p}^{k-1} \cdot C_{N_q}^{n-k}}{C_N^{n-1}} \times \frac{N_p - k + 1}{N - n + 1}, \text{ la dernière expression}$$

résultant d'un raisonnement direct : (k - 1) blanches pour les (n - 1) premiers tirages, et donc (n - 1) - (k - 1) = n - k noires, suivies d'une blanche au dernier tirage.

4. Comportement asymptotique

La formule unitaire (1) a aussi l'avantage de rendre évidentes les convergences, pour $N \rightarrow \infty$, de la loi binomiale vers la loi hypergéométrique, et de la loi 'Pascal sans remise' vers la loi 'Pascal avec remise', puisque le produit factoriel $x(x-1)\dots(x-k+1)$ tend vers la puissance x^k , pour $x \rightarrow \infty$, à k fixé.

La convergence des moments se vérifie aisément sur le tableau suivant :

Moment	Loi binomiale	Loi hypergéométr	Pascal avec remise	Pascal sans remise
Espérance	np	np	$\frac{k}{p}$	$k \frac{N+1}{N_p+1}$
Variance	npq	$npq \frac{N-n}{N-1}$	$\frac{kq}{p^2}$	$kq \frac{N(N+1)(N_p-k+1)}{(N_p+1)^2(N_p+2)}$

5. Cas particuliers

- $n = 1$ dans $B(n,p)$ et $H(N,n,p) \Rightarrow$ loi de Bernoulli $B(1,p)$
- $k = 1$ dans $G(k,p) \Rightarrow$ loi géométrique $G(1,p)$
- $k = 1$ et $p = 1/N$ dans $S(N,k,p) \Rightarrow$ loi uniforme $U(n)$

On vérifie dans ce dernier cas que les formules de la dernière colonne du tableau précédent, un peu compliquées, donnent bien les expressions simples

$$\text{des moments d'une variable uniforme : } E(X) = \frac{N+1}{2} \text{ et } V(X) = \frac{N^2-1}{12}$$

6. Tableau synoptique

Voir page suivante le tableau synoptique des probabilités 'basiques' et des relations qui existent entre elles.

7. Bibliographie

[1] KOSMANEK, Edith, « Simples remarques à propos d'un tableau d'aide mémoire », bulletin APMEP n°370 de septembre 1989.

[2] KOSMANEK, Edith, « Comment mémoriser les lois de probabilités discrètes usuelles sans trop se fatiguer », revue L'OUVERT (IREM de Strasbourg), n°63, 1991.