Projet de simulation MAP311 – X2016

X-Forum

Sujet proposé par Igor Kortchemski – igor.kortchemski@cmap.polytechnique.fr

Les courbes et graphiques demandés dans les questions de simulation doivent être rendus avec le projet, et doivent donner lieu à des commentaires, qui seront pris en compte dans la note du projet. Les paramètres utilisés doivent être précisés. Toute initiative personnelle est fortement encouragée. Les fichiers sources (.py) des programmes doivent être envoyés par email.

*Les questions **T** sont théoriques, les questions **S** relèvent de la simulation.*

À son stand de X-Forum, le CMAP distribue différents goodies : n goodies jaunes et n goodies rouges. La distribution s'arrête lorsqu'il ne reste plus que des goodies d'une seule couleur, et le CMAP souhaite distribuer le plus de goodies possibles. Le but de ce projet est d'étudier deux différentes distributions possibles :

(i) Les goodies jaunes sont dans une première boite jaune, les goodies rouges sont dans une deuxième boite rouge, et chaque visiteur du stand choisit au hasard avec probabilité $1/2$ soit la boite jaune, soit la boite rouge, et y prend un goodie.

(ii) Les goodies jaunes et les goodies rouges sont dans une même boite, et chaque visiteur du stand choisit uniformément au hasard un goodie dans la boite.

Première stratégie. On suppose ici que les n goodies jaunes sont dans une première boite jaune, les n goodies rouges sont dans une deuxième boite rouge, que chaque visiteur du stand choisit au hasard (indépendamment des autres choix) avec probabilité $1/2$ soit la boite jaune, soit la boite rouge, et y prend un goodie, et que la distribution s'arrête lorsqu'il ne reste plus que des goodies d'une seule couleur. On note G_n le nombre de goodies qui restent à la fin.

(1,**T**) Démontrer que la probabilité que les n goodies rouges disparaissent pour la première fois au bout de m tirages (avec $n \leq m \leq 2n-1$) est

$$\frac{1}{2^m}\binom{m-1}{n-1}.$$

(2,**T**) En déduire que, pour $1 \leq k \leq n$,

$$\mathbb{P}(G_n = k) = \frac{2^{k+1}}{2^{2n}}\binom{2n-k-1}{n-1}.$$

(3,**T**) Démontrer que $\frac{G_n}{\sqrt{n}}$ converge en loi lorsque $n \to \infty$ vers une variable aléatoire dont on déterminera la loi.

Indication. On pourra s'intéresser à la limite de $\mathbb{P}(G_n \leq a\sqrt{n})$ avec $a > 0$ en écrivant (avec $\lfloor x \rfloor$ désignant la partie entière de x) :

$$\begin{aligned}\mathbb{P}(G_n \leq a\sqrt{n}) = \sum_{k=1}^{\lfloor a\sqrt{n}\rfloor} \mathbb{P}(G_n = k) &= \int_1^{\lfloor a\sqrt{n}\rfloor+1} \mathbb{P}(G_n = \lfloor v \rfloor)\mathrm{d}v \\ &= \int_{1/\sqrt{n}}^{(\lfloor a\sqrt{n}\rfloor+1)/\sqrt{n}} \sqrt{n}\mathbb{P}(G_n = \lfloor u\sqrt{n}\rfloor \mathrm{d}u,\end{aligned}$$

et en combinant la formule de Stirling avec le théorème de convergence dominée.

(4,**S**) Illustrer la convergence précédente par des simulations et étudier avec l'aide de simulations la convergence de $\mathbb{E}[G_n]/\sqrt{n}$ lorsque $n \to \infty$.

Deuxième stratégie. On suppose maintenant que les n goodies jaunes et les n goodies rouges sont dans une même boite, que chaque visiteur du stand choisit uniformément au hasard un goodie dans la boite, et que la distribution s'arrête lorsqu'il ne reste plus que des goodies d'une seule couleur. On note H_n le nombre de goodies qui restent à la fin.

(5,**T**) Soit $\mathcal{E}_n$ l'ensemble des suites de $2n$ lettres contenant n fois la lettre R et n fois la lettre J. Quel est le cardinal de $\mathcal{E}_n$? En déterminer un équivalent lorsque $n \to \infty$.

Indication. On pourra utiliser la formule de Stirling.

(6,**T**) Supposons pour un instant que la distribution continue jusqu'à ce qu'il n'y ait plus de goodies à la fin. Codons le processus de distribution par une suite de $2n$ lettres $L_1, L_2, \ldots, L_{2n}$, avec $L_i = R$ si le i-ième goodie choisi est rouge et $L_i = J$ sinon.

Démontrer que $(L_1, L_2, \ldots, L_{2n})$ suit la loi uniforme sur $\mathcal{E}_n$.

(7,**T**) Démontrer que H_n converge en loi lorsque $n \to \infty$ vers une variable aléatoire qu'on déterminera.

(8,**S**) Illustrer la convergence précédente par des simulations et étudier à l'aide de simulations la convergence de $\mathbb{E}[H_n]$ lorsque $n \to \infty$.

Conclusion.

(9) Quelle stratégie vaut-il mieux utiliser ?

Question facultative :

(10,**T**,**S**) Que se passe-t-il s'il y a plus que deux couleurs ?