Exercice 8

Problème des secrétaires

JeanDuSud, lycée Jean Durand de Castelnaudary

Résumé

Faute de temps, nous n'avons pu nous intéresser qu'à la première question de l'exercice.
L'objectif était donc de déterminer une stratégie de jeu qui maximisait la probabilité de gagner et de trouver cette probabilité maximale. Ce jeu n'autorise pas un grand éventail de stratégies, c'est pourquoi nous nous sommes rapidement aperçu que la seule formule permettant d'augmenter sciemment nos chances de gain est de refuser un nombre r de boules, puis de sélectionner la première boule qui est supérieure à toutes les précédentes. L'enjeu de cet exercice semblait donc se résumer à la détermination de r. Pour le trouver, nous avons créé une formule algébrique qui permet de calculer la probabilité de gagner en fonction du nombre total de boules (n) et du nombre de boules refusées (r). Ensuite, nous avons confronté plusieurs résultats grâce à un algorithme et constaté un concentration des valeurs maximales en un certain point du graphique présentant les résultats. Grâce à cette observation, nous avons été en mesure de proposer une stratégie grâce à laquelle on a le plus de chances de gagner et de déterminer sa probabilité de victoire.

Notations

- n est le nombre de boules dans la boîte opaque.
- r est le nombre de boules qu'Alice rejette automatiquement avant de commencer le jeu.
- A_i est l'événement « Alice sélectionne la i-ème boule ».
- B_i : « La i-ème boule porte le numéro maximum ».
- $p_n(r)$ est la probabilité qu'Alice gagne.

I – Détermination de la formule de calcul de $p_n(r)$

A) Formule générale

Il s'agit de déterminer une formule basique de calcul de $p_n(r)$.

$\forall\ n \in \mathbb{N}^*$ et $\forall\ r \in \mathbb{N}^*$ avec $r \leqslant n$.

Pour qu'Alice gagne, il faut, pour j entier positif inférieur à n :

- qu'elle sélectionne la j-ème boule (A_j)
- que la j-ème boule porte la valeur maximale (B_j)

Or, les $n - r$ dernières boules ont toutes autant de chance d'être tirées et les n boules du jeu ont toutes autant de chance de porter la valeur maximale. Ainsi, la probabilité qu'Alice tire la j-ème boule et que ce soit elle qui porte la valeur maximale se résume au produit :

$p_{B_j}(A_j) \times p(B_j)$

Il faut à présent appliquer ce raisonnement à toutes les boules qui n'ont pas été refusées : les $n - r$ dernières.

On a donc : $p_n(r) = p(A_{r+1}) \times p(B_{r+1}) + p(A_{r+1}) \times p(B_{r+1}) + \dots + p(A_n) \times p(B_n)$

Exprimé différemment, on trouve :

$$p_n(r) = \sum_{i=r+1}^{n} p(B_i) \times p_{B_i}(A_i)$$

B) Calcul des de chaque probabilité séparément

– Calcul de $p_{B_i}(A_i)$

Si la i-ème boule est sélectionnée, alors c'est la première boule portant une valeur supérieure à toutes les précédentes, parmi les boules sélectionnables. Ainsi, si la plus grande des $i - 1$ premières boules avait été sélectionnable, c'est elle qui aurait été gardée. Donc, si Alice a tiré la boule de rang i portant le numéro maximal, alors la plus grande boule parmi les précédentes fait partie des boules à refuser automatiquement, soient les r premières.

On peut donc dire que l'événement A_i sachant B_i désigne l'événement :

« parmi les $i - 1$ premières boules, celle portant le numéro maximal fait partie des r premières ».

On sait que toutes les boules ont autant de chances d'être la plus grande. Ainsi, la boule maximale de l'ensemble des $i-1$ premières boules a r chances d'être tirée parmi les l'ensemble des boules concernées.

On a donc :

$$p_{B_i}(A_i) = \frac{r}{i-1}$$

– Calcul de $p(B_i)$

Parmi les n boules du jeu, toutes ont autant de chances d'être la plus grande.
On a donc :

$$p(B_i) = \frac{1}{n}$$

B) Calcul de la formule finale de calcul de $p_n(r)$

En remplaçant les données manquantes de la formule trouvée au A) par les formules de
calcul de chaque probabilité calculées au B), on aboutit à la formule suivante :

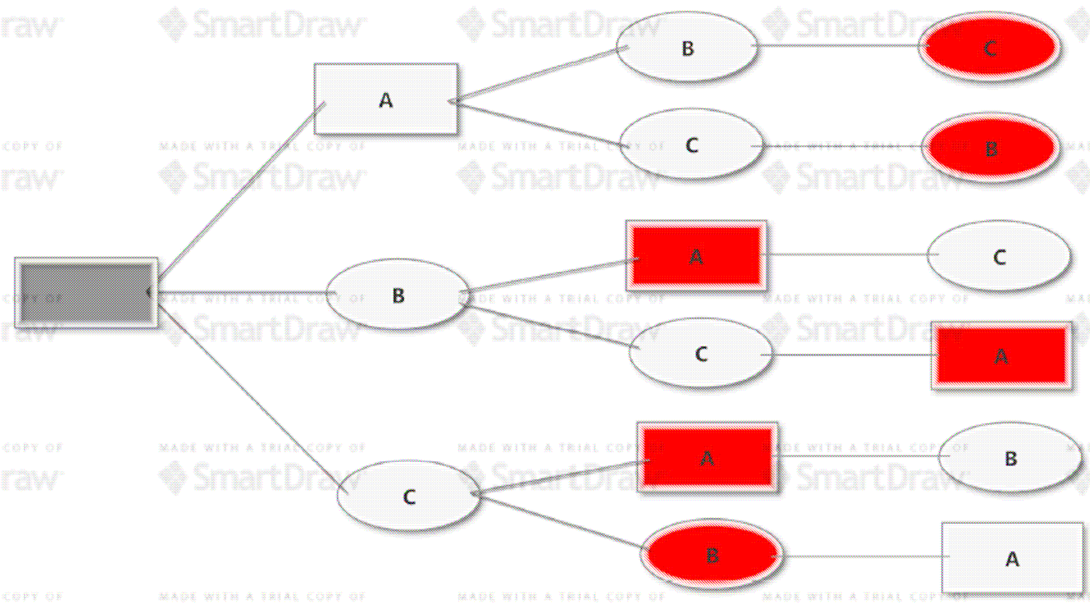
$$p_n(r) = \sum_{i=r+1}^{n} \frac{1}{n} \times \frac{r}{i-1} = \frac{r}{n} \sum_{i=r+1}^{n} \frac{1}{i-1} = \frac{r}{n} \sum_{i=r}^{n-1} \frac{1}{i}$$

II – Confrontation des résultats

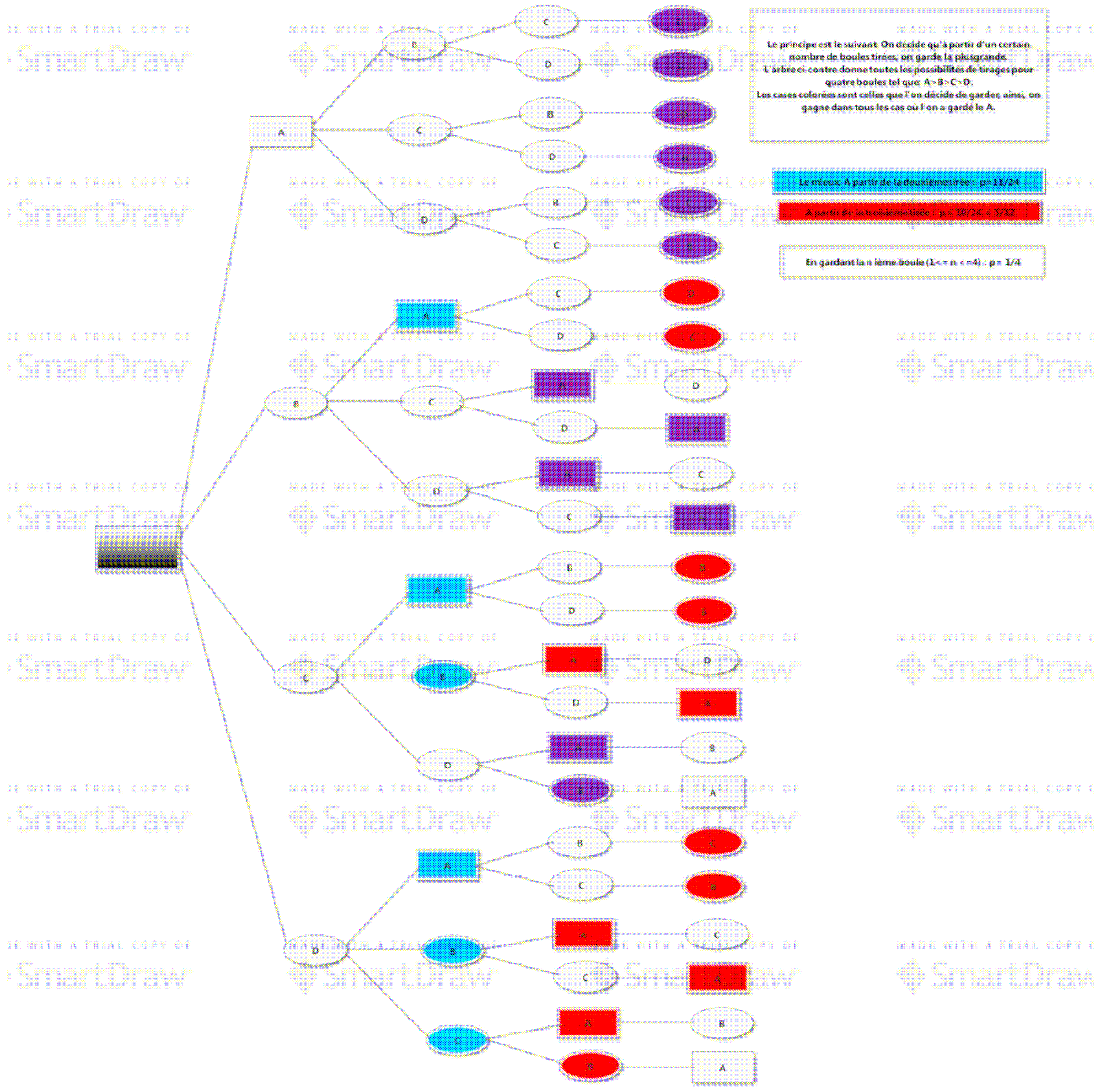
A) Confirmation de la formule

Maintenant que nous avons trouvé la formule de calcul de la probabilité de gagner, nous allons l'utiliser afin d'en tirer une logique, mais avant tout, nous allons la confirmer, en recalculant les probabilités pour $n = 3$ et $n = 4$. Nous trouvons les résultats suivants :

r n	1	2	3	4
0	1	0,5	1/3	1/4
1		0,5	1/2	11/24
2			1/3	5/12
3				1/4



Cet arbre donne toutes les possibilités de tirages de trois boules tel que A>B>C.
Si on décide de garder la n-ième boule (1<= n <=3), la probabilité de gagner est 1/3. On décide donc qu'à partir de lé deuxième boule tirée, on conservera la plus grande. Ainsi la probabilité de gagner est: p= 3/6 = 1/2



Les résultats concordent avec ceux que nous avons trouvé grâce à l'étude des arbres ci-dessus.

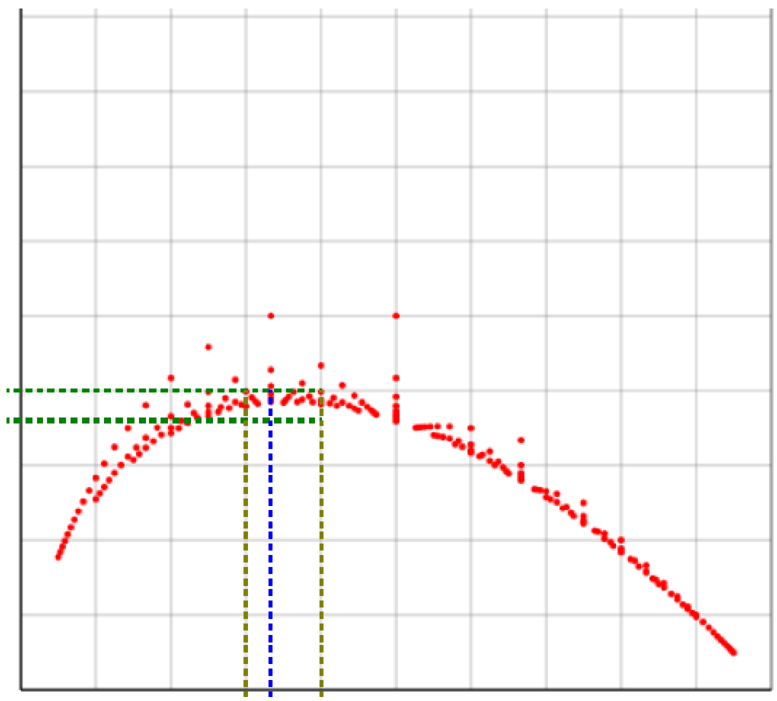
B) Elaboration d'un algorithme

A présent, nous allons confronter des résultats similaires parmi un échantillon beaucoup
plus large. En effet, nous avons créé un algorithme grâce au logiciel AlgoBox qui applique la
formule que nous avons trouvée précédemment avec toutes les valeurs de *r* possibles pour toutes les valeurs de *n* inférieures et égale à une valeur rentrée par l'utilisateur.

L'algorithme (dont vous trouverez le code intégral en annexe), a été programmé pour
afficher les résultats sous forme de texte, mais aussi sous la forme d'un graphique. Afin de pouvoir comparer toutes les probabilités sur un même document, nous avons utilisé en abscisse, pour chaque résultat, la valeur du rapport *r / n*. Ainsi, nous pouvons comparer toutes les valeurs sur une même échelle et déterminer un encadrement précis.

C) Résultats et interprétation

Ci-dessous, vous trouverez le graphique généré par l'algorithme pour une valeur entrée de 20 (les points rouges correspondent donc aux probabilités de gagner pour tous les valeurs de *r* possibles et pour tous les valeurs de *n* inférieures et égale à 20). En ordonnée est représentée $p_n(r)$, le graphique est gradué de 0 à 1 avec un intervalle de 0,1. La graduation est identique pour l'axe des abscisses où sont représentées les différentes valeurs de *r / n.*



Ainsi, le graphique nous montre que les points semblent se concentrer le long d'une courbe croissante puis décroissante. Le maximum de cette courbe paraît se situer dans l'intervalle d'abscisses [0,3 ; 0,4] (matérialisé ci-dessus par les pointillés jaunes), environs pour *r / n* = 0,33 (pointillés bleus).

On conclut donc que la probabilité maximale de gagner est atteinte lorsque le rapport entre les boules refusées et le nombre total de boules est compris entre ces valeurs.

On peut donc proposer la stratégie suivante :

« Sélectionner la première boule qui sera plus grande que toutes les précédentes et dont le rapport entre le rang et le nombre total de boules sera supérieur à un tiers »

De plus, on peut, grâce au graphique, déterminer un intervalle de la probabilité maximale de gagner en appliquant la stratégie précédente : elle semble se situer entre 0,36 et 0,4 (pointillés verts).

Conclusion

En conclusion, on peut dire que la meilleure stratégie pour gagner à ce jeu est de sélectionner la première boule supérieure à toutes les précédentes et qui n'est pas située dans le premier tiers des boules tirées. Ainsi, la Probabilité de gagner est maximale et vaut entre 0,36 et 0,4.