

PROJET UE MOGPL — MU4IN200

DICE BATTLE

ARIANA CARNIELLI
IVAN KACHAIKIN

INTRODUCTION

Ce rapport présente les résultats théoriques et quelques résultats de simulations numériques pour le projet de MOGPL concernant le jeu « Dice Battle ». On suit la division en questions de l'énoncé du projet.

PROBABILITÉS

Question 1. $Q(d, k)$ est la probabilité d'obtenir k points en jetant d dés sachant qu'aucun dé a tombé sur 1. On numérote les dés de 1 à d et on note j le résultat du dernier dé, qui est donc un nombre entre 2 à 6. Alors on aura k points si et seulement si les $d - 1$ premiers dés donnent $k - j$ points. Les valeurs possibles de j étant disjointes, on peut dire que $Q(d, k)$ est la somme pour toutes les valeurs de j de $Q(d - 1, k - j)$ multiplié par la probabilité que le dernier dé donne j sachant qu'il n'est pas 1 (égale à $\frac{1}{5}$), ce qui donne donc :

$$(1) \quad Q(d, k) = \sum_{j=2}^6 \frac{Q(d-1, k-j)}{5}.$$

Plus précisément, soient X_1, \dots, X_d des variables aléatoires indépendantes et distribuées selon une loi uniforme sur $\{1, \dots, 6\}$, représentant les résultats de chaque dé. Soient $S_d = \sum_{i=1}^d X_i$ et $S_{d-1} = \sum_{i=1}^{d-1} X_i$. On peut réécrire $Q(d, k)$ comme

$$Q(d, k) = \mathbb{P}(S_d = k \mid X_1 \neq 1, \dots, X_d \neq 1).$$

Montrons d'abord que $P(d, k) = \left(\frac{5}{6}\right)^d Q(d, k)$ pour $2d \leq k \leq 6d$. On a

$$\begin{aligned} P(d, k) &= \mathbb{P}(S_d = k, X_1 \neq 1, \dots, X_d \neq 1) \\ &= \mathbb{P}(S_d = k \mid X_1 \neq 1, \dots, X_d \neq 1) \mathbb{P}(X_1 \neq 1, \dots, X_d \neq 1) \\ &= Q(d, k) \mathbb{P}(X_1 \neq 1) \cdots \mathbb{P}(X_d \neq 1) = Q(d, k) \left(\frac{5}{6}\right)^d. \end{aligned}$$

On montre maintenant la relation (1) pour $d \geq 2$ et $2d \leq k \leq 6d$. On a $S_d = S_{d-1} + X_d$. Conditionnellement à $X_1 \neq 1, \dots, X_d \neq 1$, X_d ne peut prendre que les valeurs j allant de 2 à 6, et, comme les événements $X_d = 2, \dots, X_d = 6$ forment une partition, on a

$$\begin{aligned} Q(d, k) &= \mathbb{P}(S_{d-1} + X_d = k \mid X_1 \neq 1, \dots, X_d \neq 1) \\ &= \sum_{j=2}^6 \mathbb{P}(S_{d-1} = k - j, X_d = j \mid X_1 \neq 1, \dots, X_d \neq 1) \\ &= \sum_{j=2}^6 \mathbb{P}(S_{d-1} = k - j \mid X_1 \neq 1, \dots, X_d \neq 1) \mathbb{P}(X_d = j \mid X_1 \neq 1, \dots, X_d \neq 1), \end{aligned}$$

où l'on utilise le fait que S_{d-1} et X_d sont indépendants. Comme S_{d-1} ne dépend pas de X_d , on obtient, par la définition de Q , que

$$\mathbb{P}(S_{d-1} = k - j \mid X_1 \neq 1, \dots, X_d \neq 1) = \mathbb{P}(S_{d-1} = k - j \mid X_1 \neq 1, \dots, X_{d-1} \neq 1) = Q(d-1, k-j),$$

et en plus $\mathbb{P}(X_d = j \mid X_1 \neq 1, \dots, X_d \neq 1) = \mathbb{P}(X_d = j \mid X_d \neq 1) = \frac{1}{5}$. Donc

$$Q(d, k) = \sum_{j=2}^6 \frac{Q(d-1, k-j)}{5}.$$

Les cas d'initialisation correspondent à $d = 1$ et $k \in \{2, \dots, 6\}$, auquel cas on a $Q(1, k) = \frac{1}{5}$. En plus, $Q(1, k) = 0$ pour $k > 6$.

VARIANTE SÉQUENTIELLE

Question 2. Remarquons que la formule

$$(2) \quad EP(d) = 4d \left(\frac{5}{6}\right)^d + 1 - \left(\frac{5}{6}\right)^d$$

provient du fait qu'on a une probabilité de $1 - \left(\frac{5}{6}\right)^d$ de ne marquer qu'un seul point et une probabilité de $\left(\frac{5}{6}\right)^d$ de marquer entre $2d$ et $6d$ points, et l'espérance du nombre de points marqué dans ce dernier cas est égale à $4d$.

On veut maximiser $EP(d)$ pour $d \in \{1, \dots, D\}$. Pour éviter de faire une recherche exhaustive, on peut faire une étude de la fonction $EP(d)$ en relaxant d'abord d à une variable réelle. Dans ce cas, on calcule

$$EP'(d) = 4 \left(\frac{5}{6}\right)^d + 4d \left(\frac{5}{6}\right)^d \ln \frac{5}{6} - \left(\frac{5}{6}\right)^d \ln \frac{5}{6} = \left(\frac{5}{6}\right)^d \left(4 + 4d \ln \frac{5}{6} - \ln \frac{5}{6}\right).$$

On cherche les valeurs d^* telles que $EP'(d^*) = 0$. Cela arrive si et seulement si

$$4 + 4d^* \ln \frac{5}{6} - \ln \frac{5}{6} = 0$$

et on calcule alors

$$d^* = \frac{\ln \frac{5}{6} - 4}{4 \ln \frac{5}{6}} = \frac{\ln \frac{6}{5} + 4}{4 \ln \frac{6}{5}} \approx 5,735.$$

On remarque aussi que $EP'(d) > 0$ pour $d < d^*$ et $EP'(d) < 0$ pour $d > d^*$, donc EP est strictement croissante sur $]-\infty, d^*[$ et strictement décroissante pour $]d^*, +\infty[$, et atteint ainsi son maximum global à $d = d^*$.

On revient maintenant à une variable discrète $d \in \{1, \dots, D\}$. Grâce à l'étude précédente, $EP(1) < EP(2) < \dots < EP(5)$ et $EP(6) > EP(7) > \dots$, ainsi les candidats à maximum global de EP dans les entiers sont 5 et 6. On calcule $EP(5) \approx 8,636$ et $EP(6) \approx 8,703$, donc le maximum global est atteint en $d = 6$. Comme $d \in \{1, \dots, D\}$, cela n'arrive que lorsque $D \geq 6$; dans le cas contraire, EP est croissante sur $\{1, \dots, D\}$ et le maximum est atteint en D . On a donc $d^*(D) = \min(D, 6)$. La méthode implémentée retourne donc cette valeur.

Question 3. On considère ici que $D = 3$ et $N = 2$, auquel cas $d^*(D) = 3$. On suppose que le joueur 1 choisit de jouer toujours 3 dés alors que le joueur 2 choisit de jouer toujours un seul dé. Soient X_i et Y_i les gains des joueurs 1 et 2, respectivement, au i -ème lancer de dés. Alors $X_i \in \{1, 6, 7, \dots, 18\}$ et $Y_i \in \{1, 2, \dots, 6\}$. En particulier, on remarque que, si $X_1 > 1$, alors le joueur 1 est sûr de gagner.

Comme $N = 2$, le jeu se finira au maximum au bout de 2 tours. Soit G la variable aléatoire donnant le gain final du joueur 1 : $G = 1$ si le joueur 1 gagne et $G = -1$ si le joueur 2 gagne (il est impossible d'avoir un match nul dans la variante séquentielle). Les événements $G = 1$ et $G = -1$ peuvent être décrits en fonction de X_i et Y_i de la façon suivante :

$$\begin{aligned} G = 1 : & \quad X_1 > 1 \cup (X_1 = 1 \cap Y_1 = 1) \\ G = -1 : & \quad X_1 = 1 \cap Y_1 > 1 \end{aligned}$$

Cela permet de calculer les probabilités

$$\mathbb{P}(G = 1) = \left(\frac{5}{6}\right)^3 + \left(1 - \left(\frac{5}{6}\right)^3\right) \frac{1}{6} \approx 0,6489$$

$$\mathbb{P}(G = -1) = \left(1 - \left(\frac{5}{6}\right)^3\right) \frac{5}{6} \approx 0,3511$$

On a donc $\mathbb{E}(G) \approx 0,2978$. Si, à la place de $d^*(D) = 3$, le joueur 1 avait choisit de jouer 1 dé comme le joueur 2, les probabilités de $G = 1$ et $G = -1$ auraient été

$$\mathbb{P}(G = 1) = \frac{5}{6} + \left(\frac{1}{6}\right)^2 \approx 0,8611$$

$$\mathbb{P}(G = -1) = \frac{1}{6} \frac{5}{6} \approx 0,1389$$

Cela donne $\mathbb{E}(G) \approx 0,7222$. Ainsi, l'espérance de gain du joueur 1 est plus grande s'il ne choisit de jouer qu'un seul dé dans ce cas.

On remarque que la situation de cet exemple n'est pas exceptionnelle : le cas $N = 2$ est équivalent, par exemple, au cas $N = 100$ lorsque les deux joueurs sont à égalité avec 98 points.

Question 4. Pour faciliter l'analyse, on représente l'état courant du jeu par un triplet (i, j, n) où i et j sont les points cumulés des joueurs 1 et 2, respectivement, et $n \in \{1, 2\}$ indique qui est le prochain joueur à jouer. On note l'espérance de gain du joueur 1 dans l'état (i, j, n) par $EG(i, j, n)$, en supposant que lui-même et son adversaire jouent toujours de façon optimale. On remarque que $EG(i, j, n)$ représente toujours l'espérance de gain du joueur 1, même lorsque $n = 2$.

Dans l'état $(i, j, 1)$, si le joueur 1 décide de jouer d dés et qu'il obtient k points, le prochain état sera $(i + k, j, 2)$. Comme il obtient k points avec probabilité $P(d, k)$, l'espérance de gain du joueur 1 lorsqu'il choisit de jouer d dés (et en supposant que lui-même et le joueur 2 jouent de façon optimale dans la suite) est

$$(3) \quad \sum_{k=1}^{6d} P(d, k) EG(i + k, j, 2).$$

Ainsi, son choix optimal est de choisir le nombre d de dés qui maximise la quantité ci-dessus, ce qui donne

$$EG(i, j, 1) = \max_{d \in \{1, \dots, D\}} \sum_{k=1}^{6d} P(d, k) EG(i + k, j, 2).$$

Comme le jeu est symétrique par rapport à la permutation des deux joueurs (en changeant le signe du gain du joueur 1), on a $EG(i + k, j, 2) = -EG(j, i + k, 1)$, et ainsi on obtient la formule

$$EG(i, j, 1) = \max_{d \in \{1, \dots, D\}} \left(- \sum_{k=1}^{6d} P(d, k) EG(j, i + k, 1) \right).$$

Comme cela ne fait intervenir que les espérances de gain lorsque c'est au joueur 1 de jouer (donc $n = 1$), on peut supprimer la troisième composante de l'état de la notation, comme à l'énoncé, pour arriver à

$$(4) \quad EG(i, j) = \max_{d \in \{1, \dots, D\}} \left(- \sum_{k=1}^{6d} P(d, k) EG(j, i + k) \right),$$

où $EG(i, j)$ doit se comprendre comme $EG(i, j, 1)$.

On initialise la récurrence en remarquant que $EG(i, j) = 1$ si $i \geq N$ et $j < N$ et que $EG(i, j) = -1$ si $i < N$ et $j \geq N$. Comme le jeu est séquentiel, il n'est pas nécessaire d'initialiser $EG(i, j)$ pour $i \geq N$ et $j \geq N$: il est impossible que les deux joueurs aient une quantité supérieure ou égale à N points car le jeu s'arrête dès que le premier joueur atteint N points ou plus.

Question 5. Notons $OPT(i, j)$ la stratégie optimale du joueur 1 dans l'état $(i, j, 1)$, c'est-à-dire le nombre de dés qu'il doit lancer pour maximiser l'espérance de son gain. Cela revient à maximiser (3) par rapport à d , ce qui donne

$$OPT(i, j) = \operatorname{argmax}_{d \in \{1, \dots, D\}} \left(- \sum_{k=1}^{6d} P(d, k) EG(j, i + k) \right).$$

Cela peut être calculé en même temps que le calcul récursif de EG par la formule (4).

Question 6. Dans ce cas, la somme dans (4) doit commencer à $k = 0$ à la place de $k = 1$. Le calcul de $EG(i, j)$ par (4) utilise en particulier, dans le second membre, la valeur de $EG(j, i)$. Or, le calcul de $EG(j, i)$ par (4) utilise dans son second membre la valeur de $EG(i, j)$. Ainsi, (4) ne permet pas de calculer $EG(i, j)$ de façon explicite, mais donne uniquement une relation implicite où $EG(i, j)$ dépend de lui-même. Il faudrait alors implémenter une méthode pour être capable de calculer $EG(i, j)$ à partir de cette relation implicite.

Mise en œuvre : questions 7, 8 et 9. On a implémenté une stratégie aveugle qui retourne, comme nombre de dés à joueur, toujours le minimum entre D et 6. La stratégie optimale implémentée calcule au début les tableaux contenant les espérances de gain $EG(i, j)$ et les choix optimaux $OPT(i, j)$ et ainsi, à chaque tour de jeu, elle ne fait que consulter le tableau OPT dans la case correspondante aux nombres des points des deux joueurs et retourner le nombre de dés correspondant. Outre ces deux stratégies, on a implémenté une stratégie aléatoire qui choisit à chaque tour un nombre de dés selon une loi uniforme dans $\{1, \dots, D\}$ et une stratégie qui permet à l'utilisateur de jouer en lui demandant à chaque tour le nombre de dés souhaité.

		Joueur 2		
		Aveugle	Optimale	Aléatoire
Joueur 1	Aveugle	0,0672	-0,0316	0,3199
	Optimale	0,1684	0,0688	0,3988
	Aléatoire	-0,2038	-0,2866	0,0558

TABLE 1. Gain moyen du joueur 1 pour différentes stratégies dans le jeu séquentiel.

La Table 1 présente une estimée de l'espérance de gain du joueur 1 lorsque les joueurs 1 et 2 utilisent les trois différentes stratégies implémentées, aveugle, aléatoire et optimale. Chaque espérance de gain a été estimée comme la moyenne empirique du gain de 10^6 jeux avec $D = 10$ et $N = 100$. On y remarque que l'espérance de gain est très petite quand une stratégie joue contre elle-même : en effet, on s'attend intuitivement à ce que deux stratégies identiques s'équilibrent, mais le jeu séquentiel donne un léger avantage au premier joueur. Parmi les trois stratégies, c'est la stratégie optimale qui a la plus grande espérance de gain lorsqu'elle joue contre elle-même, même si cette espérance est du même ordre que dans le cas de la stratégie aveugle.

Lorsqu'on fait jouer deux stratégies différentes, la stratégie optimale montre plus clairement sa supériorité, avec une espérance de gain de presque 0,40 contre la stratégie aléatoire et 0,17 contre la stratégie aveugle lorsqu'elle est le joueur 1. Lorsqu'elle est le joueur 2, la stratégie optimale arrive aussi à gagner en moyenne des stratégies aveugle et aléatoire : son avantage est très léger contre l'aveugle (espérance de gain de $-0,03$ pour l'aveugle), mais beaucoup plus important contre l'aléatoire (espérance de gain de $-0,29$ de l'aléatoire). Entre l'aveugle et l'aléatoire, l'aveugle a un avantage important.

Pour voir comme les gains du joueur 1 dépendent des paramètres D et N , on a fixé un des cas de la Table 1, celui où le joueur 1 joue avec la stratégie optimale et le joueur 2 avec la stratégie aveugle, et fait varier D à $N = 100$ fixé et ensuite fait varier N à $D = 10$ fixé. Les résultats sont donnés dans la Figure 1, dans laquelle chaque point a été obtenu avec une moyenne de 10^5 simulations. On y remarque que, à N fixé, l'augmentation du nombre de dés a tendance à faire légèrement augmenter le gain, ce que l'on peut interpréter en remarquant qu'avoir un plus grand D implique avoir plus de choix par rapport au choix des 6 dés la stratégie aveugle. À D fixé, l'augmentation de N fait diminuer le gain et, à $N = 2000$, le gain n'est que de l'ordre de 0,06, soit le même ordre que la diagonale de la Table 1, et qui peut donc être imputé à l'avantage que le jeu séquentiel

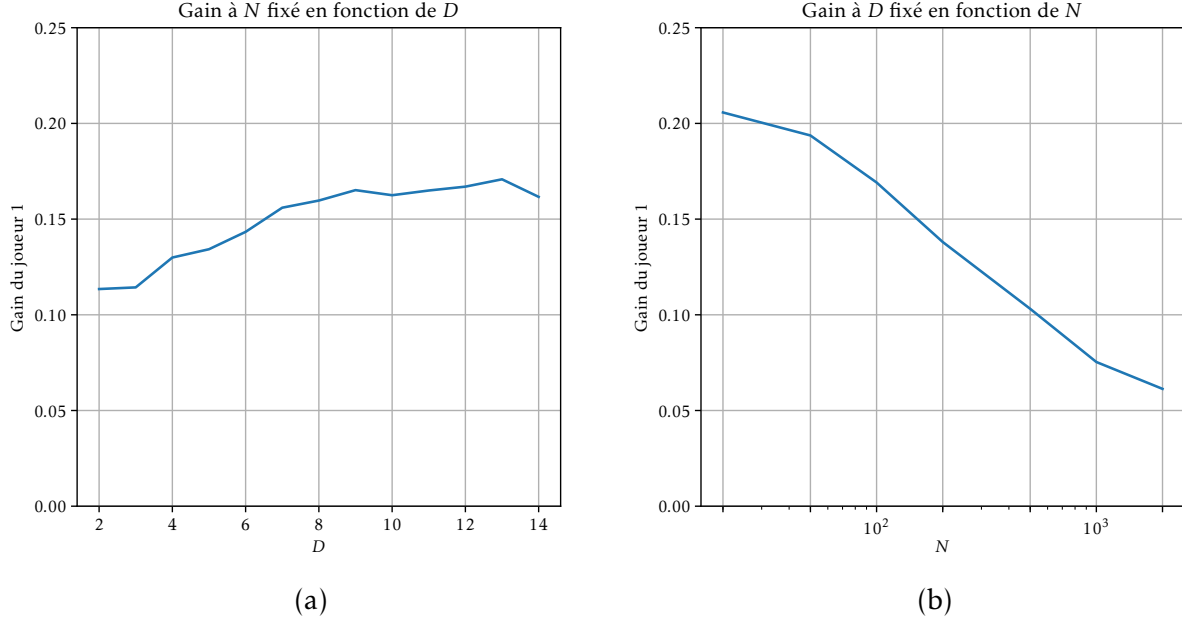


FIGURE 1. Gain moyen de la stratégie optimale (joueur 1) contre la stratégie aveugle (joueur 2) dans le jeu séquentiel (a) en fonction de D à $N = 100$ fixé et (b) en fonction de N à $D = 10$ fixé. Moyennes sur 10^5 simulations.

donne au joueur 1. L'augmentation de N accentue ainsi les effets aléatoires du jeu et rend les choix stratégiques moins efficaces.

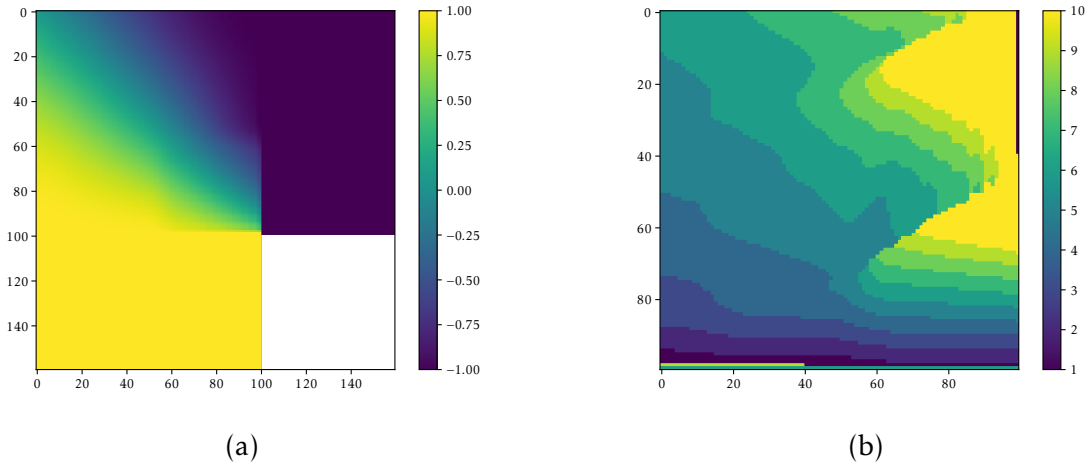


FIGURE 2. (a) Espérance de gain et (b) nombre de dés optimal à joueur en fonction du nombre de points du joueur 1 (ordonnée) et du joueur 2 (abscisse) lorsque c'est au joueur 1 de jouer.

La Figure 2 représente les matrices EG et OPT utilisées par la stratégie optimale dans le cas $D = 10$ et $N = 100$. La matrice EG contient les cas d'initialisation et la partie en blanc au coin inférieur gauche correspond aux cases non-initialisées. On y remarque une structure presque symétrique, même si, en regardant en détail, on observe qu'elle n'est pas symétrique : sur la diagonale, par exemple, l'espérance de gain est toujours strictement positive et augmente lorsque le nombre de points déjà obtenus augmente, à cause de l'avantage que le jeu donne au joueur 1. La visualisation de la matrice OPT montre la complexité des frontières de décision entre les différents choix possibles de nombre de dés. Elle démontre que, lorsque le joueur 1 a un avantage important, il est

plus intéressant de jouer moins de dés, et, lorsque le joueur 1 perd d'une différence importante, il est plus intéressant de jouer plus de dés.

On remarque aussi quelques lignes horizontales et verticales de couleur assez différente, comme la ligne bleue foncée verticale à la dernière colonne. L'explication que l'on donne à cette ligne est que, lorsque le joueur 1 a moins de 40 points et le joueur 2 en a 99, même si c'est au tour du joueur 1, il n'a aucune chance de gagner : peu importe le nombre de dés qu'il choisit de jouer, sa probabilité de gagner sera 0. Le choix du nombre de dés ne lui a aucune importance et la méthode que nous avons implémentée prend le premier nombre de dés correspondant au maximum de la probabilité, qui est donc 1 dé.

VARIANTE SIMULTANÉE

Question 10. Soit G la variable aléatoire donnant le gain du joueur 1 lorsqu'il a jeté d_1 dés et le joueur 2 a jeté d_2 dés. Alors, par la définition de l'espérance,

$$EG_1(d_1, d_2) = 1 \cdot \mathbb{P}(G = 1 \mid d_1, d_2) + 0 \cdot \mathbb{P}(G = 0 \mid d_1, d_2) + (-1) \cdot \mathbb{P}(G = -1 \mid d_1, d_2).$$

Soient K_1 et K_2 les variables aléatoires représentant le nombre de points obtenus par les joueurs 1 et 2, respectivement. Les événements $G = 1$, $G = 0$ et $G = -1$ peuvent s'écrire en termes de K_1 et K_2 comme $K_1 > K_2$, $K_1 = K_2$ et $K_1 < K_2$, respectivement. On a

$$\{K_1 > K_2 \mid d_1, d_2\} = \bigcup_{j=1}^{6d_2} \bigcup_{i=j+1}^{6d_1} \{K_1 = i, K_2 = j \mid d_1, d_2\}.$$

On remarque que certains des ensembles du membre de droite peuvent être vides, par exemple pour $1 < i < 2d_1$ ou $1 < j < 2d_2$. En plus, l'union sur i peut être vide, par exemple dans le cas $j \geq 6d_1$, auquel cas il n'y a aucun i possible entre $j+1$ et $6d_1$. On a donc des ensembles vides qui ne changent pas l'union finale. De la même manière, on peut facilement obtenir que

$$\{K_1 < K_2 \mid d_1, d_2\} = \bigcup_{i=1}^{6d_1} \bigcup_{j=i+1}^{6d_2} \{K_1 = i, K_2 = j \mid d_1, d_2\}.$$

Avant d'obtenir la formule de $EG_1(d_1, d_2)$, on remarque que tous les événements dans les unions ci-dessus sont deux à deux disjoints, et donc la probabilité de leur union est la somme de leurs probabilités. En plus, comme K_1 et K_2 sont indépendantes et K_i ne dépend que de d_i , on a $\mathbb{P}(K_1 = i, K_2 = j \mid d_1, d_2) = P(d_1, i)P(d_2, j)$. On peut finalement obtenir une formule pour $EG_1(d_1, d_2)$:

$$\begin{aligned} EG_1(d_1, d_2) &= \mathbb{P}(G = 1 \mid d_1, d_2) - \mathbb{P}(G = -1 \mid d_1, d_2) \\ &= \sum_{j=1}^{6d_2} \sum_{i=j+1}^{6d_1} \mathbb{P}(K_1 = i, K_2 = j \mid d_1, d_2) - \sum_{i=1}^{6d_1} \sum_{j=i+1}^{6d_2} \mathbb{P}(K_1 = i, K_2 = j \mid d_1, d_2) \\ &= \sum_{j=1}^{6d_2} \sum_{i=j+1}^{6d_1} P(d_1, i)P(d_2, j) - \sum_{i=1}^{6d_1} \sum_{j=i+1}^{6d_2} P(d_1, i)P(d_2, j) \\ &= \sum_{j=1}^{6d_2} P(d_2, j) \sum_{i=j+1}^{6d_1} P(d_1, i) - \sum_{i=1}^{6d_1} P(d_1, i) \sum_{j=i+1}^{6d_2} P(d_2, j) \end{aligned}$$

En utilisant cette formule, on calcule, pour $D = 3$, la matrice de gains

$d_1 \backslash d_2$	1	2	3
1	0	-0,375	-0,227
2	0,375	0	-0,199
3	0,227	0,199	0

Comme le jeu est à somme nulle, la matrice est anti-symétrique.

Question 11. Si on suppose que le joueur 2 connaît la probabilité $p_1(i)$, $i = 1, 2, \dots, D$, alors il peut calculer le gain espéré du joueur 1. En fait, quand le joueur 2 joue j dés ($j = 1, 2, \dots, D$), le joueur 1 gagne en espérance $EG_1(1, j)$ avec probabilité $p_1(1)$, $EG_1(2, j)$ avec probabilité $p_1(2)$, ..., $EG_1(D, j)$ avec probabilité $p_1(D)$. L'espérance du gain du joueur 1 est donc donnée par la formule suivante :

$$\sum_{i=1}^D p_1(i)EG_1(i, j),$$

quand le joueur 2 joue j dés. Par conséquent, il ne reste au joueur 2 qu'à minimiser cette espérance de gain. On suppose qu'il choisit une stratégie mixte $p_2(1), p_2(2), \dots, p_2(D)$ (on remarque que les stratégies mixtes contiennent les pures). Le joueur 2 perdra ainsi $\sum_{i=1}^D p_1(i)EG_1(i, 1)$ avec probabilité $p_2(1)$, $\sum_{i=1}^D p_1(i)EG_1(i, 2)$ avec probabilité $p_2(2)$, ..., $\sum_{i=1}^D p_1(i)EG_1(i, D)$ avec probabilité $p_2(D)$. Ainsi, l'espérance de sa perte est

$$\sum_{j=1}^D p_2(j) \sum_{i=1}^D p_1(i)EG_1(i, j) = \sum_{i=1}^D \sum_{j=1}^D p_1(i)EG_1(i, j)p_2(j).$$

Notons $p_1 = (p_1(1), p_1(2), \dots, p_1(D))$, $p_2 = (p_2(1), p_2(2), \dots, p_2(D))$, et EG_1 la matrice dont l'élément à la ligne i et colonne j est $EG_1(i, j)$, $i, j = 1, 2, \dots, D$. La perte moyenne du joueur 2 s'écrit donc :

$$p_1^\top EG_1 p_2,$$

où on a supposé que p_1 et p_2 sont des vecteurs colonne.

Pour le joueur 2, il faut donc minimiser son espérance de perte. C'est pourquoi il prendra une stratégie mixte $p_2^* = (p_2^*(1), p_2^*(2), \dots, p_2^*(D))$ telle que

$$(5) \quad p_1^\top EG_1 p_2^* = \min_{p_2} p_1^\top EG_1 p_2.$$

Question 12. Le joueur 1 veut maximiser de son côté son espérance de gain. Supposant que le joueur 2 joue optimalement, le joueur 1 peut déduire que, pour toute stratégie p_1 qu'il choisit, son adversaire jouera selon (5). Sachant ce fait, il doit donc maximiser (5) de son côté, c'est-à-dire, il doit choisir une stratégie mixte $p_1^* = (p_1^*(1), p_1^*(2), \dots, p_1^*(D))$ telle que

$$(6) \quad \min_{p_2} p_1^{*\top} EG_1 p_2 = \max_{p_1} \min_{p_2} p_1^\top EG_1 p_2.$$

Le joueur 1 doit donc résoudre le problème de maximisation à droite dans (6) pour déterminer une stratégie p_1^* qui maximisera son espérance de gain sachant que son adversaire joue optimalement.

Pour trouver une solution du problème (6), on remarque que, comme vu en cours (Lemme 1 du Cours 4),

$$\max_{p_1} \min_{p_2} p_1^\top EG_1 p_2 = \max_{p_1} \min_j p_1^\top EG_1^j,$$

où EG_1^j est la colonne d'indice j de la matrice EG_1 ($j = 1, 2, \dots, D$). En plus, d'après les propriétés vues en cours de ce type de problème linéaire, le problème $\max_{p_1} \min_j p_1^\top EG_1^j$ peut être résolu en résolvant

$$(7) \quad \begin{aligned} & \max \quad z' - z'' \\ & -p_1^\top EG_1^j + z' - z'' \leq 0 \quad \forall j \in \{1, \dots, D\}, \\ & \sum_{i=1}^D p_1(i) \leq 1 \\ & -\sum_{i=1}^D p_1(i) \leq -1 \\ & z' \geq 0, z'' \geq 0, p_1(i) \geq 0 \quad \forall i \in \{1, \dots, D\} \end{aligned}$$

Donc, en résolvant le problème (7) on peut obtenir un vecteur p_1^* qui sera aussi une solution du problème (6). Finalement, le joueur 1 prend donc p_1^* comme sa stratégie optimale.

On remarque que, comme ce jeu est à somme nulle, la stratégie p_1^* est aussi la stratégie optimale du joueur 2.

Mise en œuvre : questions 13 et 14. Pour l’implémentation du jeu simultané, on considère qu’une stratégie retourne toujours la quantité de dés à jouer tiré de façon aléatoire selon le vecteur $p = (p(1), \dots, p(D))$ des probabilités de choisir chaque quantité de dés possible. La stratégie optimale calcule au début la résolution du programme linéaire de la question précédente, en utilisant pour cela la fonction `linprog` de la bibliothèque d’optimisation `optimize` de *SciPy*, et ne fait que utiliser cette même solution à chaque appel.

D	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$
2	0.000	1.000				
3	0.000	0.000	1.000			
4	0.000	0.000	0.000	1.000		
5	0.000	0.176	0.053	0.000	0.771	
6	0.000	0.176	0.053	0.000	0.771	0.000

TABLE 2. Stratégie mixte optimale pour le jeu simultané à un tour en fonction de la valeur de D .

La stratégie optimale retournée par la fonction pour les valeurs de D allant de 2 à 6 est donnée dans la Table 2. On y remarque que, pour D entre 2 et 4, la stratégie optimale est pure et consiste à jouer le maximum de dés possible. Pour D à partir de 5, on a une stratégie mixte avec une grande probabilité de jouer 5 dés mais aussi une probabilité de jouer 2 ou 3 dés. On a pu constater que, pour les valeurs de $D \geq 5$ que nous avons testé, cette stratégie reste toujours la stratégie optimale. Augmenter la quantité maximale de dés possibles au-delà de 5 n’a donc pas d’effet sur la stratégie optimale, ce qui rappelle le résultat correspondant pour la stratégie aveugle montré à la Question 2 (mais avec $D = 6$).

		Joueur 2		
		Aveugle	Optimale	Aléatoire
Joueur 1	Aveugle	$5,0 \cdot 10^{-5}$		
	Optimale	0,0152	$-2,7 \cdot 10^{-4}$	
	Aléatoire	0,0138	-0.0626	$-4,7 \cdot 10^{-4}$

TABLE 3. Espérance de gain du joueur 1 pour différentes stratégies dans le jeu simultané à un tour.

La Table 3 présente une estimée de l’espérance de gain du joueur 1 lorsque les joueurs 1 et 2 utilisent les trois différentes stratégies implémentées. Chaque espérance de gain a été estimée comme la moyenne empirique du gain de 10^6 jeux avec $D = 10$ (et $N = 1$ pour ne faire qu’un seul tour). Comme attendu pour un jeu à somme nulle, la diagonale est essentiellement nulle. On remarque que la stratégie optimale a un avantage sur les deux autres, même si celui-ci est assez petit. La stratégie aléatoire semble aussi avoir un avantage sur la stratégie aveugle.

Pour voir comme les gains dépendent de D , on a fixé le cas d’une stratégie optimale (joueur 1) contre une stratégie aveugle (joueur 2) et fait varier D . Les résultats sont donnés dans la Figure 3, dans laquelle chaque point a été obtenu avec une moyenne de 10^6 simulations. On voit que la stratégie optimale gagne plus souvent que la stratégie aveugle dès que $D \geq 6$, mais son espérance de gain est essentiellement nulle pour $D \leq 5$ car, comme vu à la Table 2, ces deux stratégies sont identiques pour $D \leq 4$ et très proches pour $D = 5$. La probabilité d’avoir un match nul augmente avec D et dépasse les probabilités des victoires des deux joueurs lorsque $D = 5$. Toutes les quantités se stabilisent pour $D \geq 6$ car les stratégies ne changent plus. On peut en déduire que la stratégie optimale est plus efficace qu’une stratégie aveugle si on joue un seul tour et $D \geq 6$.

Par contre, il semble intéressant de tester et comparer ces stratégies si on joue plus qu’un seul tour. En prenant $N = 100$, $D = 10$ et en faisant une moyenne sur 10^5 parties, on obtient 0,4% de matchs nuls, 42,9% de victoires de la stratégie optimale et 56,7% de victoires de la stratégie aveugle. La stratégie “optimale” est donc moins efficace ici qu’une stratégie aveugle. Pour l’améliorer, il faut donc considérer aussi dans l’optimisation combien de points les joueurs ont déjà accumulé.

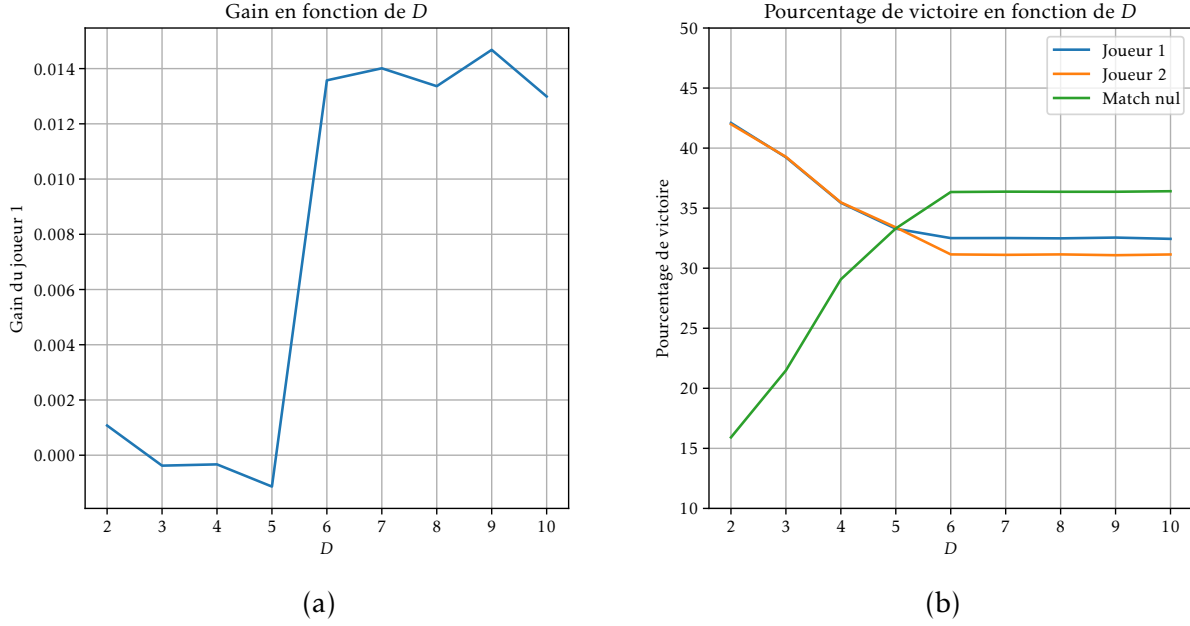


FIGURE 3. (a) Gain moyen et (b) pourcentages de victoire de la stratégie optimale (joueur 1) contre la stratégie aveugle (joueur 2) dans le jeu simultané à un tour en fonction de D . Moyennes sur 10^6 simulations.

Question 15. Si $i \geq N$ ou $j \geq N$, alors le jeu est déjà fini, et on aura

$$EG_1(i, j) = \begin{cases} 1 & \text{si } i \geq N \text{ et } i > j, \\ 0 & \text{si } i = j \geq N, \\ -1 & \text{si } j \geq N \text{ et } j > i. \end{cases}$$

Question 16. À partir de l'état (i, j) avec $i < N$ et $j < N$, en jetant d_1 dés, le joueur 1 a une probabilité $P(d_1, k)$ de se retrouver avec $i + k$ points au prochain état et, en jetant d_2 dés, le joueur 2 a une probabilité $P(d_2, \ell)$ de se retrouver avec $j + \ell$ points au prochain état. Comme les deux lancers sont indépendants, la probabilité que le prochain état soit $(i + k, j + \ell)$ est donc $P(d_1, k)P(d_2, \ell)$ et, comme le joueur 1 a une espérance de gain de $EG_1(i + k, j + \ell)$ dans ce prochain état, son espérance de gain à l'état (i, j) est donc

$$(8) \quad E_1^{d_1, d_2}(i, j) = \sum_{k=1}^{6d_1} \sum_{\ell=1}^{6d_2} EG_1(i + k, j + \ell) P(d_1, k) P(d_2, \ell).$$

Question 17. Cette méthode a été implémentée à l'intérieur d'une classe dédiée à la stratégie optimale pour le jeu séquentiel et remplit le tableau $E_1(i, j)$ à l'aide de (8) en utilisant, pour chaque valeur de d_1 et d_2 , des produits matriciels pour faire le calcul du membre de droite de (8).

Question 18. Soient $p_1^{i,j}$ et $p_2^{i,j}$ les stratégies mixtes des joueurs 1 et 2 dans l'état (i, j) , c'est-à-dire, $p_1^{i,j}(d)$ est la probabilité avec laquelle le joueur 1 choisit de lancer $d \in \{1, \dots, D\}$ dés et de même pour le joueur 2. On simplifie leurs notations à p_1 et p_2 dans cette question car on raisonne à état (i, j) fixé. En procédant comme à la Question 11, on en déduit que le gain du joueur 1 dans ce cas sera, sous écriture matricielle,

$$p_1^\top E_1(i, j) p_2.$$

Toujours comme à la Question 11, on considère que le joueur 2 choisit sa stratégie mixte p_2 de façon à minimiser la valeur ci-dessus, c'est-à-dire, il va résoudre $\min_{p_2} p_1^\top E_1(i, j) p_2$. Le joueur 1, anticipant ce comportement de son adversaire, joue selon le problème de maximisation

$$\max_{p_1} \min_{p_2} p_1^\top E_1(i, j) p_2.$$

Si les joueurs jouent de cette façon, on a donc

$$EG_1(i, j) = \max_{p_1} \min_{p_2} p_1^\top E_1(i, j) p_2.$$

Par conséquent, pour calculer $EG_1(i, j)$ si $i < N$ et $j < N$ il faut:

1. Calculer la matrice $E_1(i, j)$ à partir des valeurs de $EG_1(k, \ell)$ ($k > i, \ell > j$) et de $P(d, k)$ par (8).
2. Résoudre le programme linéaire correspondant, identique à (7) mais avec la matrice EG_1 remplacée par $E_1(i, j)$.

Question 19. Dans l'état (i, j) on prend comme stratégie optimale pour le joueur 1 la valeur

$$p_1^* = \operatorname{argmax}_{p_1} \min_{p_2} p_1^\top E_1(i, j) p_2,$$

qui est une solution du problème linéaire de la Question 18. La stratégie optimale peut donc être calculée en même temps que $EG_1(i, j)$ et stockée dans un tableau.

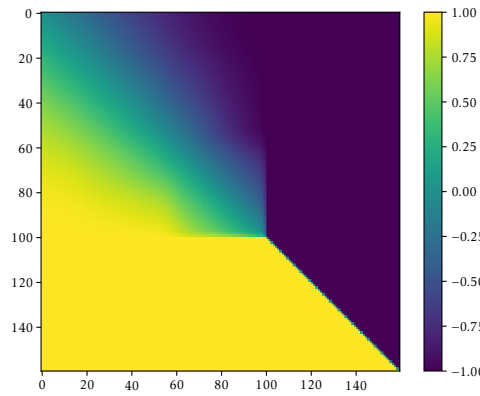


FIGURE 4. Espérance de gain du joueur 1 en fonction du nombre de points du joueur 1 (ordonnée) et du nombre de points du joueur 2 (abscisse).

On a implémenté la méthode calculant le tableau EG_1 et les stratégies optimales dans la classe dédiée à la stratégie optimale pour le jeu simultané. La matrice EG_1 pour $D = 10$ et $N = 100$ est représentée sur la Figure 4, où l'on représente la matrice jusqu'à la ligne et la colonne $N + 6D - 1$. On y remarque que, même si elle ressemble à la Figure 2(a), il y a maintenant une différence importante : elle est symétrique, ce que l'on peut vérifier numériquement en faisant la somme $EG_1 + EG_1^\top$ et constatant que le plus grand élément en valeur absolue de cette somme est de l'ordre de 10^{-16} .

QUESTION 20

La Table 4 présente le gain moyen obtenu par la stratégie optimale simultanée contre quatre stratégies. On remarque qu'elle a un avantage, même si petit, contre les autres stratégies, et fait en moyenne match nul contre elle-même, comme attendu.

Aveugle	Aléatoire	Optimale séquentielle	Optimale simultanée
0,0142	0,0606	0,0144	0,000882

TABLE 4. Gain moyen du joueur 1 pour la stratégie optimale simultanée (joueur 1) jouant contre différentes stratégies (joueur 2).