

# RENDEZ-VOUS RUE DE LA PAIX

Voici une étude fascinante de Philippe Gaucher, chargé de recherche CNRS en mathématiques à l'université Paris 7, à propos du **Monopoly** :

**Quelles sont les cases les plus visitées et les moins visitées lors d'une partie ?**

<http://www.pps.jussieu.fr/~gaucher/monopoly/monopoly.html>

**Notions abordées :** simulations, programmation (implantation d'un algorithme), équiprobabilité, moyenne, écart-type, taux d'évolution, effet sur la moyenne et l'écart-type d'une transformation affine des données.

## Introduction

Le but de cet exposé est de prouver que, contrairement à une idée reçue (cf par exemple l'article de Ian Stewart dans le "Pour La Science" de Juin 1996), les différentes cases du Monopoly ne sont pas équiprobables. Nous avons fait des tests sur le Monopoly français. Nous verrons même qu'il y a des disparités entre les cases, entre les lotissements, et à l'intérieur des lotissements.

Pour prouver cela, nous avons choisi la simulation. Nous avons fait un programme en PERL, *a priori* exécutable sur n'importe quelle machine UNIX afin de simuler des suites de tours. Soyons clair, l'aspect financier n'est absolument pas pris en compte. Il n'intervient pas d'ailleurs sur la suite de déplacements que peut faire un joueur au cours d'une partie.

Nous avons par contre pris en compte tout ce qui intervenait pour les déplacements : le tirage des cartes "Chance" et "Caisse de Communauté", avec les effets sur la position, la prison au bout de trois doubles et la case "Allez en prison".

Un rappel succinct des règles est fait dans la section 2. La section 3 expose les résultats pour chaque case, et la section 4 pour chaque groupe. On conclut ensuite en section 5. Le programme proprement dit est consultable en section 6 sur le site :

<http://www.pps.jussieu.fr/~gaucher/monopoly/monopoly.html>

## Rappel des règles du jeu du Monopoly

Nous allons seulement rappeler les règles qui comptent pour les déplacements.

Par convention, on considèrera qu'il y a 41 cases : cela va de la case "Départ" (numéro 0) à la case "Rue de la Paix" (numéro 39), la case "Prison" ayant le numéro 40, la case "Simple Visite" ayant le numéro 10. Voici la liste détaillée des cases : Case Départ, Belleville, Caisse de Communauté, Lecourbe, Impôts, Gare Montparnasse, Vaugirard, Carte Chance, Courcelles, Républiques, Simple Visite, La Villette, Cie Électricité, Neuilly, Paradis, Gare de Lyon, Mozart, Caisse de Communauté, Saint-Michel, Pigalle, Parc Gratuit, Matignon, Carte Chance, Malesherbes, Henri-Martin, Gare du Nord, Saint-Honoré, La Bourse, Cie des Eaux, La Fayette, Allez en Prison, Breteuil, Foch, Caisse de Communauté, Capucines, Gare Saint-Lazare, Carte Chance, Champs-Élysées, Taxe de Luxe, La Paix, Prison.

Comme il peut y avoir des éditions différentes du Monopoly, nous allons préciser aussi les couleurs des lotissements : les bruns (Belleville, Lecourbe), les bleus clairs (Vaugirard, Courcelles,

Républiques), les violets (La Villette, Neuilly, Paradis), les oranges (Mozart, Saint-Michel, Pigalle), les rouges (Matignon, Malesherbes, Henri-Martin), les jaunes (Saint-Honoré, La Bourse, La Fayette), les verts (Breteuil, Foch, Capucines), les bleus foncés (Champs-Élysées, La Paix).

Une partie commence sur la case "Départ". A chaque tour, le joueur lance deux dés. Au bout de trois doubles consécutifs, le joueur va en prison. S'il tombe sur une case "Chance" ou bien "Caisse de Communauté", il tire une carte dans la pile correspondante, et ce tirage est éventuellement suivi d'un effet immédiat au niveau de la position. Quand on est en prison, on peut en sortir gratuitement en faisant un double dans les trois tours qui suivent celui de l'emprisonnement, sinon on doit payer pour sortir. On peut sortir avant la fin des trois tours en payant.

Voici la liste détaillée des cartes "Chance" : 1 envoie en prison, 1 envoie vers l'avenue Henri-Martin, 1 envoie vers boulevard de la Villette, 1 envoie vers la Rue de la Paix, 1 envoie vers la gare de Lyon, 1 envoie sur la case Départ, 1 "Reculez de trois cases". Il y a 9 autres cartes "Chance" qui n'ont aucune influence sur la position

Voici maintenant la liste détaillée des cartes "Caisse de Communauté" : 1 "Retournez à Belleville", 1 envoie en prison, 1 envoie sur la case Départ, 1 possibilité de tirer une carte "Chance" (alternative avec une amende). Il y a 12 autres cartes "Caisse de Communauté" qui n'ont aucune influence sur la position.

Nous allons considérer 4 types de stratégie pour la simulation en combinant les deux options suivantes : 1) quand on tire la carte "Caisse de Communauté" qui nous propose le choix entre le tirage d'une carte "Chance" et le paiement d'une amende, on peut choisir de tirer toujours une carte "Chance" (option C), ou bien on peut choisir de toujours payer une amende (option  $\neg C$ ) ; 2) quand on est en prison, on peut choisir de rester en prison et d'attendre de faire un double, pendant au maximum trois tours évidemment (option P), ou bien on peut choisir de sortir tout de suite (option  $\neg P$ ). Ces options combinées font bien quatre stratégies.

Quand une simulation commence, l'ordinateur ne change pas de stratégie. Et on peut objecter que tout ceci est artificiel : par exemple, en fin de partie, quand beaucoup de lotissements sont construits, on préfère rester en prison, alors qu'en début de partie, on se dépêche d'acheter, mais tout ceci n'est pas grave car nous allons voir que le choix de telle ou telle stratégie n'influe pas sur le « poids statistique » qu'a telle ou telle case.

Passons maintenant aux résultats obtenus.

### Les résultats case par case

Il y a quatre simulations correspondant aux quatre combinaisons possibles de stratégies. Les résultats concernant la simulation  $N$  ( $N$  variant de 1 à 4) sont dans les colonnes  $2N$  et  $2N + 1$ . Par convention, on a supposé qu'une partie durait 100 tours.

Fixons  $N$ . La simulation  $N$  consiste à effectuer 50 fois 1 000 parties. Pour chaque case, on obtient 50 nombres entiers  $c_1, c_2, \dots, c_{50}$  qui signifient que pendant les 1 000 premières parties, on est passé  $c_1$  fois par cette case, dans les 1 000 parties suivantes, on est passé  $c_2$  fois par cette case, etc. Soit  $C$  le nombre de tours effectivement joués dans la simulation  $N$ . Le nombre tout en bas de la colonne

$2N$  est égal à  $\frac{C}{41 \times 50}$  soit  $\frac{C}{2050}$  : c'est la valeur théorique dans le cas équiprobable des nombres  $c_1,$

$c_2, \dots, c_{50}$ .

Ces nombres ne figurent pas dans le tableau ci-dessus car ils auraient été trop parlants.

Pour chaque case, on calcule alors la moyenne arithmétique et l'écart-type de ces 50 nombres.

Rappelons que la moyenne arithmétique  $m$  est donnée par la formule : 
$$m = \frac{c_1 + c_2 + \dots + c_{50}}{50}$$

et que l'écart-type  $\sigma$  est donné par la formule : 
$$\sigma^2 = \frac{(c_1 - m)^2 + (c_2 - m)^2 + \dots + (c_{50} - m)^2}{50}$$

Pour que la formule soit plus parlante, on n'a pas inscrit directement ces deux nombres (mais on peut les retrouver avec le tableau ci-après) mais plutôt les deux nombres suivants (multipliés par 100 dans le tableau) :

$$\tilde{m} = \frac{m - \frac{C}{2050}}{\frac{C}{2050}} ; \tilde{\sigma} = \frac{\sigma}{\frac{C}{2050}}$$

Pourquoi ces deux nombres ?

Car pour une case donnée, la série

$$(\tilde{c}_1 ; \tilde{c}_2 ; \dots ; \tilde{c}_{50}) = \left( \frac{c_1 - \frac{C}{2050}}{\frac{C}{2050}} ; \frac{c_2 - \frac{C}{2050}}{\frac{C}{2050}} ; \dots ; \frac{c_{50} - \frac{C}{2050}}{\frac{C}{2050}} \right)$$

est plus parlante que la série  $(c_1 ; c_2 ; \dots ; c_{50})$  : c'est la série des taux d'évolution du nombre de passages sur cette case dans le cas équiprobable au nombre de passages réel.

Par transformation affine des données  $\left( x \mapsto \frac{x - \frac{C}{2050}}{\frac{C}{2050}} \right)$ , la moyenne  $\tilde{m}$  et l'écart-type  $\tilde{\sigma}$  de la série

$$(\tilde{c}_1 ; \tilde{c}_2 ; \dots ; \tilde{c}_{50}) \text{ vérifient donc } \tilde{m} = \frac{m - \frac{C}{2050}}{\frac{C}{2050}} ; \tilde{\sigma} = \frac{\sigma}{\frac{C}{2050}}.$$

Dans la suite de l'exposé, le couple  $(\tilde{m} ; \tilde{\sigma})$  s'appellera le *poids statistique* de la case.

Ainsi, les tableaux ci-après permettent de savoir si, par rapport au cas équiprobable, une case est plutôt plus, plutôt moins ou plutôt pareillement fréquentée. Par exemple, dans la simulation 1, on lit pour Vaugirard que son poids statistique est  $(-10,6 \% ; 1,6 \%)$ . Cela signifie que pour la stratégie CP, la case Vaugirard est 10,6 % moins fréquentée que dans le cas (imaginaire) équiprobable avec un écart-type de 1,6 %. Le poids statistique ne mesure donc pas la fréquentation statistique, mais « l'écart » par rapport à un hypothétique cas équiprobable.

Pour ce qui est du choix des valeurs 50 et 1 000 (50 fois 1 000 parties pour chaque simulation), disons que d'autres choix de valeurs ne changent pas fondamentalement les résultats des simulations sur ordinateur : la convergence est très rapide.

Voici maintenant les tableaux de résultat. Pour des questions de commodité et de clarté, on a fait un tableau par côté du plateau de jeu.

	CP		¬CP		C¬P		¬C¬P	
<b>Case Départ</b>	17.3	1.7	16	1.5	17.4	1.4	17	1.7
<b>Belleville</b>	-1.4	1.3	-1.2	1.4	-.9	1.3	-.3	1.9
<b>Caisse de Communauté</b>	-17.3	1.7	-17.1	1.7	-17.2	1.5	-16.7	1.5
<b>Lecourbe</b>	-15.9	1.9	-15.6	1.8	-16.2	1.4	-15.4	2
<b>Impôts</b>	-9.1	1.9	-9.2	1.7	-9	2.2	-9	1.8
<b>Gare Montparnasse</b>	-13.2	1.7	-13.4	1.5	-13.1	1.5	-13.5	1.7
<b>Vaugirard</b>	-10.6	1.6	-10.4	1.9	-11.2	1.5	-10.8	1.6
<b>Carte Chance</b>	-8.8	1.8	-9.1	1.8	-9.6	1.7	-9.5	1.9
<b>Courcelles</b>	-10.1	1.4	-9	1.8	-10.4	1.6	-10	1.8
<b>Républiques</b>	-11.6	1.8	-11	1.5	-11.5	1.9	-10.9	1.7
<b>Simple Visite</b>	-13.5	1.3	-12.6	1.8	-14.2	1.4	-13.5	1.7
<b>Prison</b>	61.4	2.2	59.4	2.1	60.5	2	58.7	2.3
<b>Nombre moyen par case</b>	2467.1		2465.5		2624.3		2618.9	

	CP		¬CP		C¬P		¬C¬P	
La Villette	3.2	1.6	2.6	1.6	2.9	1.6	2.5	1.8
Cie Électricité	-6.7	1.8	-6	2	-13.1	1.4	-12.8	1.7
Neuilly	-12.8	1.9	-12.9	1.9	-10.4	1.6	-9.8	1.8
Paradis	-.9	2.1	-1.1	1.5	-5.1	1.6	-5.5	1.5
Gare de Lyon	14	2.1	13.3	1.9	18.9	1.9	17.5	1.4
Mozart	8.2	2.1	8.1	1.6	6.4	1.9	6.5	1.8
Caisse de Communauté	5.6	2	5.4	2.2	13.2	1.9	13.3	1.9
Saint-Michel	13.9	1.7	14	2.1	11.7	1.9	11.2	2
Pigalle	12.5	1.8	12.1	1.9	16.3	1.8	16.1	1.9
Parc Gratuit	12.7	2.2	12.4	2.1	8.5	1.8	8.8	1.7
Nombre moyen par case	2467.1		2465.5		2624.3		2618.9	

	CP		¬CP		C¬P		¬C¬P	
Matignon	4.3	2.1	4.8	1.9	7.4	1.9	7.1	1.7
Carte Chance	12.4	2.2	12.6	2.1	5.7	1.7	6.1	1.8
Malesherbes	3.5	1.8	3.3	1.6	3.7	1.8	4.2	1.8
Henri-Martin	21.8	1.7	21.2	2.1	22.7	1.7	21.8	1.8
Gare du Nord	3.5	1.9	3.9	1.4	4.3	1.9	4.4	2.2
Saint-Honoré	3.8	1.9	3.3	1.9	4.2	1.9	4.1	1.5
La Bourse	2.9	1.5	2.4	1.4	2.5	1.6	3.1	1.8
Cie des Eaux	.7	1.7	.8	1.7	.9	1.6	1	1.7
La Fayette	-.8	1.8	0	2	-.5	1.6	-.1	1.7
Allez en Prison	.3	1.9	-.2	1.7	.8	1.7	.3	2
Nombre moyen par case	2467.1		2465.5		2624.3		2618.9	

	CP		¬CP		C¬P		¬C¬P	
Breteuil	1.2	1.9	.8	1.8	1.5	1.5	1.2	1.7
Foch	-1.5	1.7	-2	1.6	-1.7	2	-1.2	1.7
Caisse de Communauté	.4	2	.1	1.9	.6	1.7	.8	2
Capucines	-7.5	1.8	-7	1.6	-7.2	1.8	-7.5	1.7
Gare Saint-Lazare	-10.1	1.8	-10	1.6	-10.1	1.8	-9.9	1.8
Carte Chance	-14.5	1.8	-13.9	1.8	-13.9	1.6	-14.2	1.5
Champs-Élysées	-18.6	1.6	-18.8	1.8	-18.4	1.5	-18.4	1.8
Taxe de Luxe	-18.8	1.5	-17.8	1.6	-18.2	1.9	-17.9	1.7
La Paix	0	1.4	-.3	1.5	-.2	1.5	-.8	1.4
Nombre moyen par case	2467.1		2465.5		2624.3		2618.9	

On voit que le nombre moyen théorique de passages par une case est plus bas dans les deux premières combinaisons de stratégie que dans les deux dernières. C'est tout-à-fait normal : dans les deux premières combinaisons (CP et  $\neg$ CP), on reste en moyenne plus longtemps en prison.

On peut observer des disparités d'une case à une autre, y compris à l'intérieur d'un même lotissement. Les tableaux ci-avant parlent d'eux-mêmes mais disons que parmi les cases appartenant à des lotissements, la plus intéressante est l'avenue Henri-Martin qui est 20 % plus fréquentée qu'une case équiprobable, et la moins intéressante est l'avenue des Champs-Élysées qui est presque 20 % moins fréquentée qu'une case équiprobable. Selon que l'on ait choisi l'option P ou l'option  $\neg$ P, la probabilité de sortir de prison avec un double diffère. Les doubles possibles sont 2, 4, 6, 8, 10 et 12. En partant de la prison, on tombe sur les cases Cie d'Électricité, Paradis, Mozart, Saint-Michel, Parc Gratuit et la case "Chance" juste après Matignon : toutes ces cases ont un poids statistique égal ou plus élevé dans la stratégie P.

Toutes les cases des lotissements orange et rouge sont intéressantes, et on peut d'ores et déjà dire que ce sont les lotissements les plus intéressants. Cela sera confirmé ci-après quand on présentera les résultats par groupe.

### Les résultats par groupe

Puisqu'il s'agit d'un article mathématique, nous allons quand même caser un théorème. Pour cela, nous allons commencer par introduire la notion de poids statistique d'un groupe de cases.

Soient  $N$  cases fixées ( $N$  étant compris entre 1 et 41) :  $N = 3$  ou  $N = 2$  par exemple pour un lotissement. Imaginons que nous fassions une simulation sur ordinateur (donc 50 fois 1 000 parties de 100 tours). Supposons qu'au cours de cette simulation, la case numéro  $i$  du groupe ( $i$  variant de 1 à  $N$ ) soit fréquentée  $c_1^i$  fois, puis  $c_2^i$  fois, ... puis  $c_{50}^i$  fois (les exposants ne sont pas des puissances ici mais des indices). Notre groupe de  $N$  cases a été fréquenté  $\sum_{i=1}^N c_1^i$  fois, puis  $\sum_{i=1}^N c_2^i$  fois, ... puis  $\sum_{i=1}^N c_{50}^i$  fois. Supposons que le nombre total de tours joués pendant cette simulation soit de  $C$  : dans le cas équiprobable, chacun des nombres précédents devrait alors être égal à  $\frac{NC}{2050}$  fois. Notons  $m$  la moyenne arithmétique de cette série de 50 nombres et  $\sigma$  son écart-type. Posons alors :

$$\tilde{m} = \frac{m - \frac{NC}{2050}}{\frac{NC}{2050}} ; \quad \tilde{\sigma} = \frac{\sigma}{\frac{NC}{2050}}$$

Le couple  $(\tilde{m} ; \tilde{\sigma})$  sera appelé le *poids statistique* du groupe. Il a la même interprétation que celle donnée ci-avant dans le cas d'une case et c'est une grandeur qui a le mérite d'être parlante pour le joueur (et le lecteur). Par exemple, le poids statistique du lotissement rouge est dans la simulation 1 (9,9 % ; 1,1 %) (cf. ci-après). Cela signifie que le lotissement rouge est 9,9 % de fois plus fréquenté qu'un lotissement de 3 cases dans le cas équiprobable.

Voici maintenant l'unique théorème de ce papier :

**Théorème** : soit un groupe de  $N$  cases. Soit  $(\tilde{m}_i ; \tilde{\sigma}_i)$  le poids statistique de la  $i$ -ème case du groupe ( $i$  variant de 1 à  $N$ ) et soit  $(\tilde{m} ; \tilde{\sigma})$  le poids statistique du groupe. Alors  $m$  est la moyenne arithmétique des  $\tilde{m}_i$ .

*démonstration* :

Avec les notations ci-avant, on note  $m_i$  (resp.  $\sigma_i$ ) la moyenne arithmétique (resp. l'écart-type) de la série  $(c_1^i, c_2^i, \dots, c_{50}^i)$ .

On a donc pour chaque égalité  $\tilde{m}_i = \frac{m_i - \frac{NC}{2050}}{\frac{NC}{2050}}$ .

$$\text{Comme } \sum_{i=1}^N 50m_i = 50m, \text{ alors } \tilde{m} = \frac{m - \frac{NC}{2050}}{\frac{NC}{2050}} = \frac{\sum_{i=1}^N 50m_i - \frac{50NC}{2050}}{\frac{50NC}{2050}} = \frac{\sum_{i=1}^N \tilde{m}_i}{N}$$

Il n'y a malheureusement pas de relations simples entre  $\tilde{\sigma}$  et les  $\tilde{\sigma}_i$  : en général, la variance (le carré de l'écart-type) d'une somme de variables aléatoires indépendantes est la somme des variances de chaque variable aléatoire : mais ici, chaque variable aléatoire correspondrait au nombre de passages par une case donnée au bout de 100 000 tours et ces variables aléatoires ne sont clairement pas indépendantes.

Pour calculer le poids statistique d'un groupe, donc pas seulement les moyennes arithmétiques mais aussi les écarts-type, il faut donc reprendre les résultats bruts et faire un calcul direct. Le programme donné en annexe (section 6) se débrouille très bien pour faire tout ça. Voici finalement le tableau des résultats par groupe (voir la section 2 pour la signification des couleurs, qui peuvent varier d'une édition du Monopoly à une autre) :

	CP		∩CP		C∩P		∩C∩P	
<b>Les gares</b>	-1.4	1	-1.5	.8	0	.8	-.3	.9
<b>Les Cies</b>	-3	1.4	-2.6	1.3	-6.1	1.2	-5.8	1.2
<b>Brun</b>	-8.7	1.1	-8.4	1.1	-8.6	.9	-7.9	1.2
<b>Bleu clair</b>	-10.8	.9	-10.1	1	-11	.8	-10.6	.9
<b>Violet</b>	-3.4	.9	-3.8	.9	-4.2	.7	-4.2	.9
<b>Orange</b>	11.6	.9	11.4	1	11.5	1	11.3	1
<b>Rouge</b>	9.9	1.1	9.8	.8	11.3	.9	11	.9
<b>Jaune</b>	2	.9	1.9	.8	2.1	.8	2.3	.8
<b>Vert</b>	-2.6	.8	-2.7	.8	-2.4	.9	-2.4	.8
<b>Bleu foncé</b>	-9.3	1	-9.6	1	-9.3	.8	-9.6	1
<b>Nombre moyen par case</b>	<b>2467.1</b>		<b>2465.5</b>		<b>2624.3</b>		<b>2618.9</b>	

Faisons quelques commentaires maintenant. Le groupe des quatre gares est très légèrement moins fréquenté qu'un groupe idéal qui serait composé de quatre cases équiprobables. Les lotissements les plus intéressants (parmi les lotissements de trois cases) sont clairement les lotissements rouge et orange : tous les joueurs de Monopoly en ont, je pense, l'intuition. Parmi les groupes de deux cases, les bruns et les bleus foncés sont au même niveau, et le groupe formé des deux compagnies (eau et électricité) est légèrement meilleur.

## Conclusion

Contrairement à ce que déclare Ian Stewart dans le numéro de Juin 1996 de "Pour La Science", les cases du Monopoly ne sont donc pas équiprobables. En fait, dans son article, Ian Stewart ne tient compte que des jets de dés. Il ne tient pas compte des deux piles de cartes ("Chance" et "Caisse de Communauté") et des déplacements qu'elles engendrent, ni de la règle des trois doubles successifs. Il y a donc de légères fluctuations d'une case à une autre et cela fait finalement tout le pigment du jeu (avec l'aspect financier et l'aspect jeu de société).