Atelier « Marche aléatoire vers un paradoxe du calcul des probabilités »

L'histoire du calcul des probabilités est riche de paradoxes. On peut penser que cette branche des mathématiques s'est développée notamment en surmontant ces difficultés, d'où leur intérêt.

Rappel

Si deux candidats, A et B, se présentent à une élection, autant d'électeurs étant favorables à A qu'à B, la probabilité que, lors du dépouillement d'une urne standard, l'un d'eux soit en tête pendant plus des trois quarts du temps est 0,67.

Énoncé du paradoxe

Si deux candidats, A et B, se présentent à une élection, autant d'électeurs étant favorables à A qu'à B, il n'est PAS DU TOUT PROBABLE que A et B soient aussi fréquemment en tête l'un que l'autre lors du dépouillement d'une urne.

Motivations

Le paradoxe auquel on s'intéresse ici a plusieurs atouts :

- l'énoncé est simple et évocateur, particulièrement pour les personnes qui ont déjà assisté au dépouillement d'un scrutin, notamment les scrutateurs ;
- les résultats, évidemment surprenants, sont obtenus par des méthodes élémentaires ;
- on aborde au passage les promenades aléatoires et les nombres de Catalan.

Historique

Initialement, il apparaît en 1887 un **théorème du scrutin** dans les Comptes Rendus de l'Académie des Sciences[1], sous la signature de Joseph Bertrand :

> Imaginons un scrutin entre deux candidats A et B. À la fin du scrutin, A obtient n voix, B en obtient m, avec n > m. Alors, si l'on admet que les ordres d'arrivée des voix sont équiprobables, la probabilité p_A pour que le nombre de voix pour A ait été strictement supérieur au nombre de voix pour B tout au long du scrutin vaut

1 Comptes Rendus de l'Académie des Sciences, Paris, 105 (1887) 369

$(n - m)/(n + m)$.

Une démonstration complète de ce théorème est due à Désiré André[2]. (William Feller affirmera plus tard que William Allen Whitworth, mathématicien britannique, avait déjà démontré ce résultat en 1878.)

D. Mirimanoff[3] a rendu la démonstration précédente beaucoup plus agréable en utilisant le *principe de réflexion*.

En 1957, William Feller fait paraître la deuxième édition du tome I de « An Introduction to Probability Theory and Its Applications » avec un chapitre supplémentaire « Fluctuations in Coin Tossing and Random Walks » dans lequel le théorème du scrutin est rappelé.

On y trouve cette phrase importante : « nous pouvons tirer les conclusions les plus intéressantes du théorème du scrutin si nous abandonnons la convention selon laquelle le point d'arrivée (x, y)[4] est fixé d'avance ».

Une fois cette modification instaurée, la situation est analogue, en fait, à celle d'un tirage non exhaustif et Feller peut affirmer la ressemblance entre le dépouillement d'un vote et les lancers successifs d'une pièce de monnaie. C'est dans ce cadre, (dépouillement d'une urne) qu'il est le plus intéressant de se placer pour l'étude du paradoxe[5].

Réalisation de figures

Grâce à un peu d'algorithmique et au logiciel GeoGebra, on obtient rapidement des illustrations. Ci-dessous, chaque point de l'axe des abscisses représente un instant où une enveloppe est ouverte et l'ordonnée du point correspondant de la courbe indique la différence algébrique des scores de A et B.

2 « Solution directe du problème résolu par M. Bertrand », Comptes Rendus de l'Académie des Sciences, Paris, 105 (1887) 436-437

3 « A propos de l'interprétation géométrique du problème du scrutin », L'enseignement mathématique 23 (1923) 187189

4 *x* est ici le nombre total de bulletins à dépouiller et *y* est la différence (éventuellement négative) entre le nombre de bulletins au nom de A et le nombre de bulletins au nom de B.

5 On peut consulter à ce sujet le document http://jacques.faisant.pagesperso-orange.fr/JN2012/Equiprobables.pdf

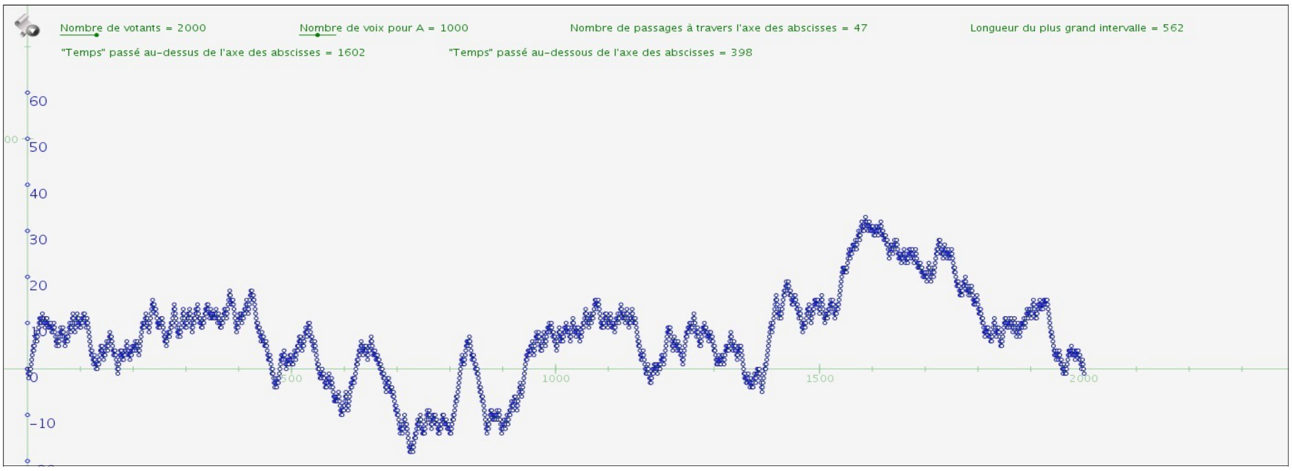

Figure 1: Chemin aléatoire construit selon un tirage exhaustif ("sans remise")

La figure 1 ci-dessus correspond au cas où le corps électoral est peu nombreux ; il n'y a qu'**une seule urne**, dans laquelle les bulletins sont tirés au sort (sans remise !), avec égalité des chances entre A et B seulement pour le premier bulletin.

Au contraire, la figure 2 qui suit décrit le dépouillement d'**une urne arbitraire**, pour laquelle il y a égalité des chances entre A et B **pour chaque bulletin** :

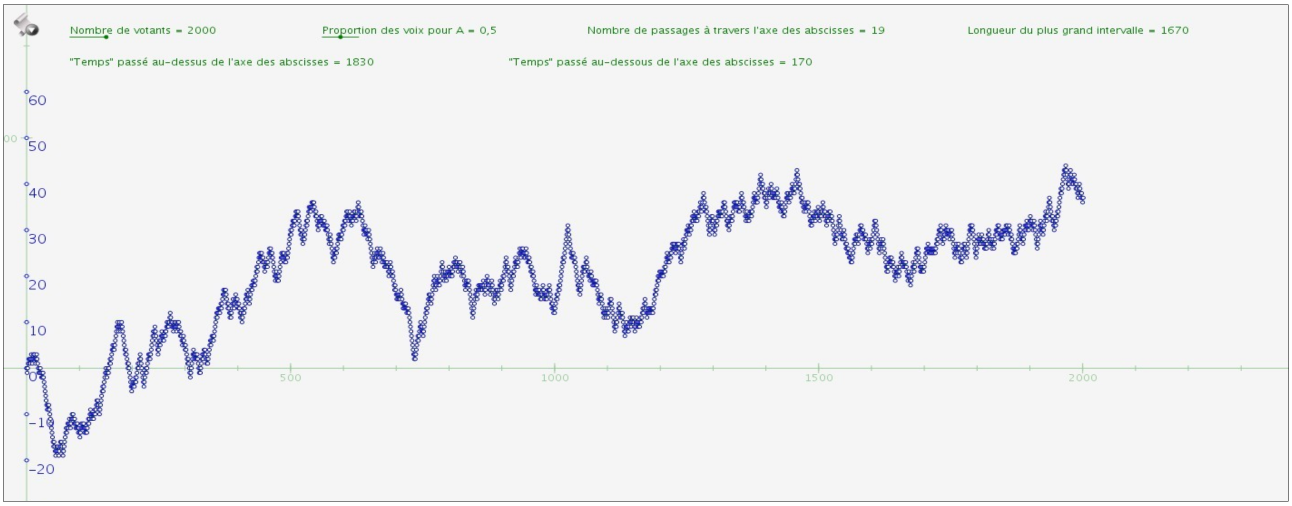

Figure 2: Chemin aléatoire construit selon un tirage non exhaustif (« avec remise »)

Simulations

Plaçons-nous en « temps discret », au rythme de l'ouverture des enveloppes. La donnée importante est le nombre total d'instants pendant lesquels le candidat A est en tête. Appelons ***dessus*** cette durée, puisqu'elle correspond aux parties des courbes précédentes qui sont *au-dessus* de l'axe des abscisses, et simulons, par exemple; 100 000 dépouillements complets de l'urne :

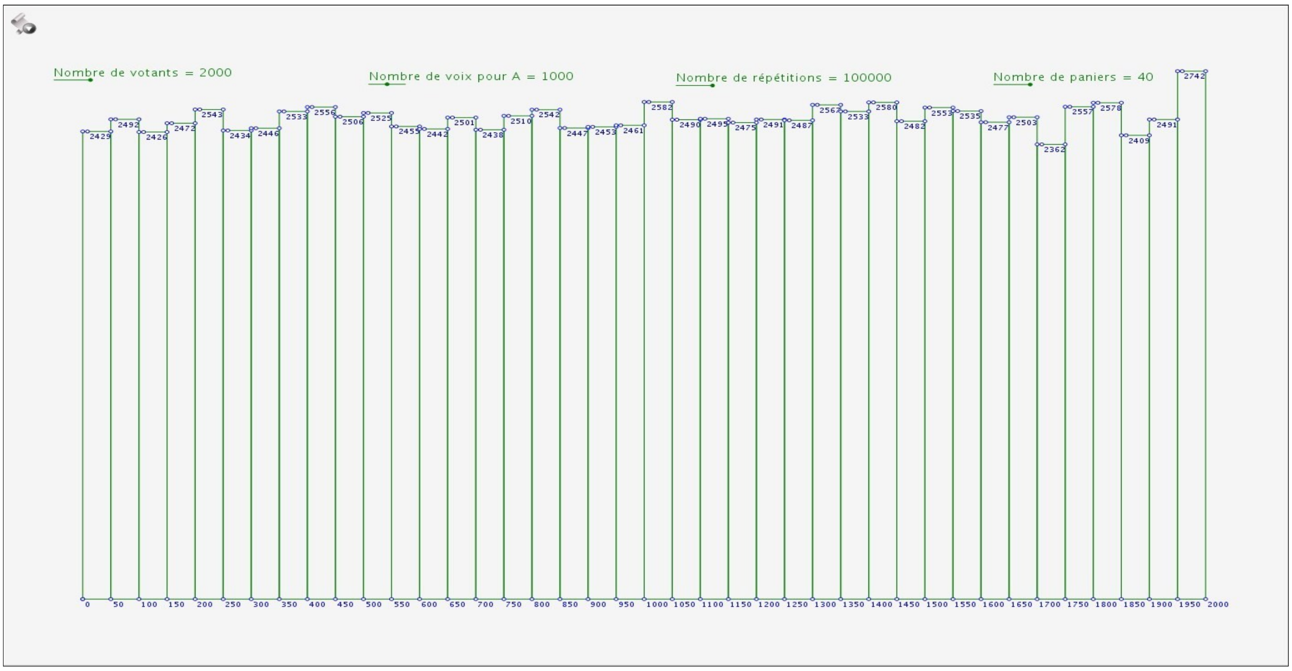

Figure 3: Loi empirique pour **dessus***, cas exhaustif*

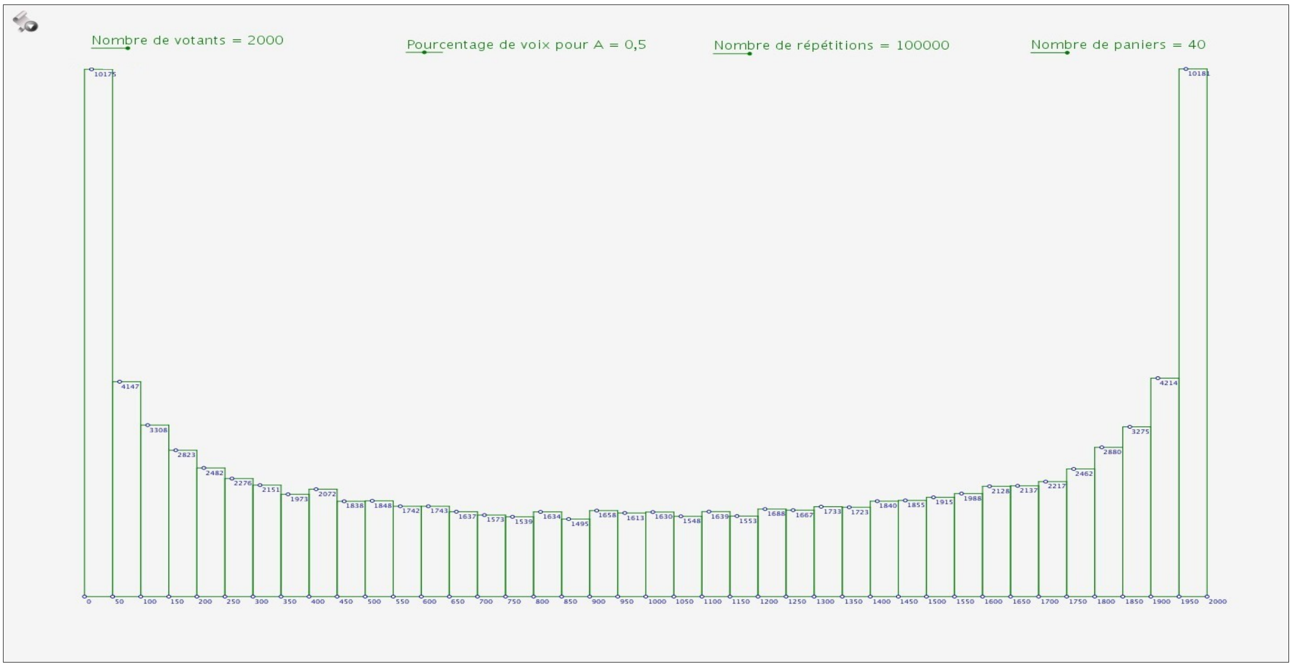

Figure 4: Loi empirique pour **dessus***, cas non exhaustif*

La différence d'aspect entre les figures 3 et 4 est vraiment stupéfiante … La figure 3 est déjà fort surprenante, et la figure 4 correspond bien à l'énoncé initial tout en le précisant :

> il n'est PAS DU TOUT PROBABLE que A et B soient aussi fréquemment en tête l'un que l'autre lors du dépouillement d'**une** urne.

Conclusion

Dans le cadre de l'atelier,

- nous étudierons les démonstrations les plus importantes
- et nous construirons et mettrons en œuvre avec GeoGebra les algorithmes qui permettent de construire les illustrations et les simulations.

Adresse de téléchargement du logiciel GeoGebra

http://www.geogebra.org/cms/fr/installers

Site WEB de l'atelier

http://jacques.faisant.pagesperso-orange.fr/JN2012/