BRALET Antoine Groupe A

DURET Guillaume

**CPE Lyon – 3ETI**

**TP n°1 de Probabilités :**

**Jouer aux dés le jour**

**de son anniversaire**







2017 – 2018

Introduction

Ce travail a pour objectif de manipuler le dénombrement ; soit les arrangements et combinaisons ; mais aussi et surtout la notion de probabilité dans son ensemble en utilisant deux exemples concrets : la probabilité que deux personnes aient la même date de naissance dans une assemblée de personnes ainsi que le problème du chevalier de Méré consistant en une épreuve de dés qui est expliqué ci-après.

En plus de l’étude théorique de ces problèmes, une étude numérique est effectuée afin de confirmer la théorie. On peut alors observer dans tout ce travail des comparaisons entre étude empirique numérique ; soit une étude qui répète une succession d’expériences ; et une étude théorique ; soit une formule qui est applicable à chaque expérience ; afin de confirmer les intuitions et la théorie.

Une méthode numérique est à souligner en particulier au cours de cette étude, il s’agit de la méthode hist sous Matlab qui permet de répertorier et de comptabiliser les résultats identiques dans un tableau de nombres.

Joyeux anniversaire

On cherche à vérifier la validité de la phrase énoncée par un professeur lors d’une cérémonie marquant son départ à la retraite à une assemblé de 50 personnes qui est « je suis presque certain que deux d’entre vous ont le même jour d’anniversaire ».

Etude théorique du problème :

On généralise le problème à une assemblée de n personnes et on veut chercher la probabilité de l’évènement « dans un assemblée de n personnes, au moins deux personnes ont le même jour d’anniversaire ».

Pour cela on détermine l’évènement contraire « dans un assemblé de n personnes, toutes les personnes ont des jours d’anniversaires distincts ».

On obtient donc :

Pour une assemblée de 50 personnes on a : .

Etude empirique du problème :

On étudie d’abord le problème avec une assemblée de 50 personnes.

On veut simuler informatiquement l’expérience d’où le code suivant :

clear variables;

close all;

%Nombre d'expériences

N = 500;

%Probabilité que au moins deux dates soient les mêmes

Proba\_meme = [];

%Probabilité que toutes les dates soient distinctes

Proba\_theo = [];

%Variable permettant de déterminer si deux dates sont identiques

identique = 0;

%%Mise en œuvre de l'expérience

%Choix du nombre de personnes dans l'assemblée

nombre=50;

%Nombre d'expériences ou deux dates ont été identiques et distincts

Meme\_date = 0;

Distincts =0;

%Deroulement des N experiences

for k = 1:N

%Choix des dates de naissance

Dates = randi([1,365],1,nombre);

%Repertoire et denombrement des dates de naissance

h = hist(Dates,365);

%Recherche de deux dates identiques

for i = 1:365

if h(i) > 1

identique = 1;

break

end

end

%Mise à jour du nombre d'expérience ou les dates sont identiques où

%non

if identique == 1

Meme\_date = Meme\_date+1;

identique = 0;

else

Distincts = Distincts +1;

end

end

%probabilité empiriques que deux dates au moins soient identiques

%pour 50 personnes dans l’assemblé

Proba\_meme = Meme\_date/N

%probabilité théoriques que deux dates au moins soient identiques

%pour 50 personnes dans l’assemblé

Proba\_theo = 1-nchoosek(365,nombre)\*factorial(nombre)/(365^nombre)

Pour simplifier le code on modélise les différentes dates d’anniversaire à des chiffres allant de 1 à 365 et on considère que chaque personne possède une date d’anniversaire aléatoire parmi les 365 (cohérent car aucune date n’a plus de chance d’être présente dans l’assemblée plus qu’une autre).

On répète cette modélisation un très grand nombre de fois pour pouvoir évaluer le nombre de fois où deux personnes ont la même date d’anniversaire. Ainsi en comparant par rapport au nombre d’expériences réalisées on peut obtenir une approximation de la probabilité.

Pour un nombre d’expériences répétées N=500 fois on obtient :

* %une probabilité empirique que deux dates au moins soient identiques pour 50 personnes dans l’assemblé

Proba\_meme =

0.9650

* %une probabilité théorique que deux dates au moins soient identiques pour 50 personnes dans l’assemblé

Proba\_theo =

0.9704

On remarque donc que la probabilité empirique est relativement proche de la probabilité calculée théoriquement.

De plus on peut remarquer que plus on augmente le nombre de répétitions de l’expérience plus la probabilité tend vers la probabilité théorique.

En effet pour un très grand nombre d’expériences répétées N=10000000 fois on obtient :

%la probabilité empiriques que deux dates au moins soient identiques

%pour 50 personnes dans l’assemblé

Proba\_meme =

0.9704

Ce qui est parfaitement égal à la probabilité théorique, ce qui témoigne de la convergence de la probabilité empirique si on augmente le nombre de répétitions de l’expérience.

Généralisation avec n personnes :

On généralise le problème avec n personnes dans l’assemblée ; on obtient le code suivant :

clear variables;

close all;

%Nombre d'expériences

N = 500;

%Probabilité que au moins deux dates soient les mêmes

Proba\_meme = [];

%Probabilité que toutes les dates soient distinctes

Proba\_theo = [];

%Variable permettant de déterminer si deux dates sont identiques

identique = 0;

%%Mise en œuvre de l'expérience

%Choix du nombre de personnes dans l'assemblée

for nombre = 10:10:80

%Nombre d'expériences ou deux dates ont été identiques et distincts

Meme\_date = 0;

Distincts =0;

%Déroulement des N expériences

Code Précédent

%liste des probabilités empiriques que deux dates au moins soient identiques en

%fonction du nombre de personnes dans l'assemblée

Proba\_meme = [Proba\_meme,Meme\_date/N];

%liste des probabilités théoriques que deux dates au moins soient identiques en

%fonction du nombre de personnes dans l'assemblée

Proba\_theo= [Proba\_theo, 1-nchoosek(365,nombre)\*factorial(nombre)/(365^nombre)];

end

%Liste du nombre de personnes dans l'assemblée

X = 10:10:80;

%Tracer des probabilités empiriques et théoriques

figure(1);

plot(X,Proba\_meme,X,Proba\_theo)

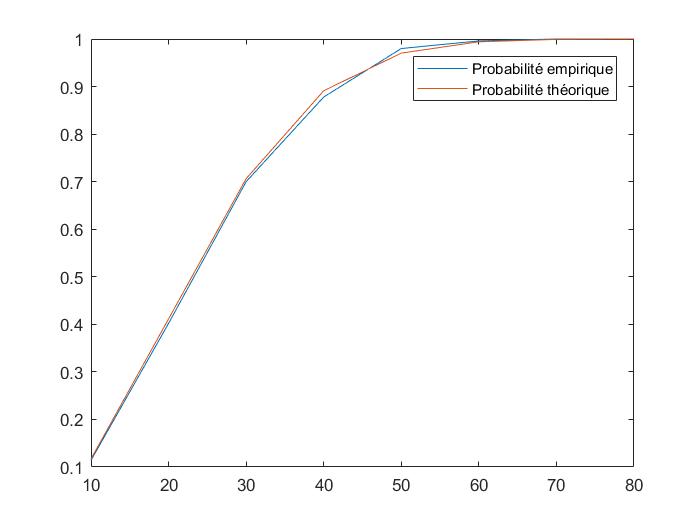
legend('Probabilité empirique', 'Probabilité théorique')

xlabel('Nombre de personnes dans l''assemblée')

ylabel('Probabilités')

On utilise le code précédent mais cette fois ci dans une boucle for qui fait varier le nombre de personne dans l’assemblé ; ainsi on peut mettre les résultats empiriques et théoriques sous forme de tableau et ainsi obtenir les courbes des probabilités empirique et théorique en fonction du nombre de personnes dans l’assemblée :

Pour un nombre d’expériences répétées N=500 fois on obtient



On remarque donc logiquement que plus le nombre de personnes augmente dans l’assemblée plus la probabilité est élevée et tend de plus en plus vers 1. De plus on peut remarquer que cette probabilité atteint exactement 1 à partir de 366.

On peut noter aussi théoriquement qu’à partir de 23 personnes dans l’assemblée la chance d’avoir 2 personnes avec 2 dates d’anniversaire identiques est supérieure à ½ ce qui peut sembler contre-intuitif à première vue devant le nombre de date d’anniversaire possible (365).

De plus à partir de 57 personnes dans l’assemblée la probabilité est supérieure à 99%.

On peut voir visuellement que la croissance de la courbe semble exponentielle.

En effet en résonant étape par étape on peut écrire :

Or en considérant () et d’après le développement limité : pour x voisin de 0.

On a donc donc qui est bien une exponentielle.

1 personne ayant un jour d’anniversaire donné :

On peut chercher la probabilité que dans une assemblée de n personnes, au moins une ait une date d’anniversaire donnée (par exemple celle du professeur).

Théoriquement on trouve en résonant par l’évènement contraire :

Ensuite tout comme précédemment on calcule numériquement la probabilité à l’aide du code ci-dessous :

clear variables;

close all;

%Nombre d'expériences

N = 5000;

%Probabilité que au moins deux dates soient les mêmes

Proba\_meme = [];

%Probabilité que toutes les dates soient distinctes

Proba\_theo = [];

%Variable permettant de déterminer si deux dates sont identiques

identique = 0;

%%Mise en œuvre de l'expérience

%Choix du nombre de personnes dans l'assemblée

for nombre = 10:140:500

%Choix d'une date choisie

Dates\_choisie = randi([1,365],1,1);

%Nombre d'expériences qui ont sont un succes ou un echec

Meme\_date = 0;

Distincts =0;

%Déroulement des N expériences

for k = 1:N

%Choix des dates de naissance

Dates = randi([1,365],1,nombre);

%Repertoire et denombrement des dates de naissance

h = hist(Dates,365);

%Recherche de 1 dates identiques à la date choisie

if h(Dates\_choisie) > 0

identique = 1;

end

%Mise à jour du nombre d'expérience ou l'experience est un succes ou

%non

if identique == 1

Meme\_date = Meme\_date+1;

identique = 0;

else

Distincts = Distincts +1;

end

end

%liste des probabilités empiriques que 1 dates au moins soient identiques à la date choisie en

%fonction du nombre de personnes dans l'assemblée

Proba\_meme = [Proba\_meme,Meme\_date/N];

%liste des probabilités théoriques que 1 dates au moins soient identiques à la date choisie en

%fonction du nombre de personnes dans l'assemblée

Proba\_theo= [Proba\_theo, 1-(364/365)^(nombre)];

end

%Liste du nombre de personnes dans l'assemblée

X = 10:140:500;

%Tracer des probabilités empiriques et théoriques

figure(1);

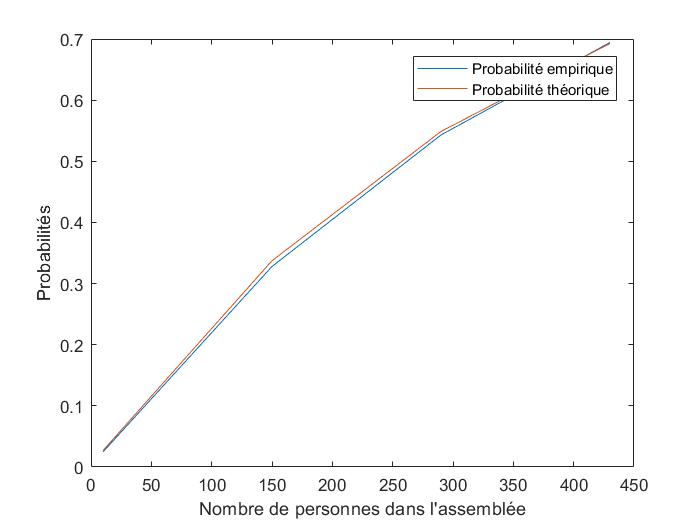
plot(X,Proba\_meme,X,Proba\_theo)

legend('Probabilité empirique', 'Probabilité théorique')

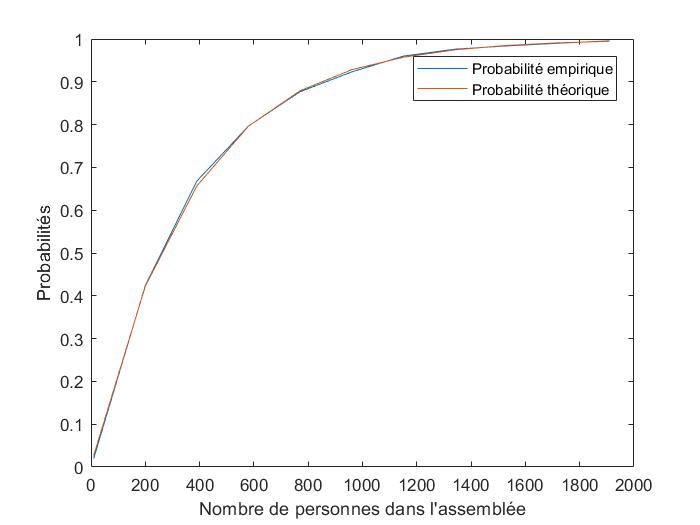
xlabel('Nombre de personnes dans l''assemblée')

ylabel('Probabilités')

On obtient pour Nmax=500 :



On voit cette fois-ci une croissance lente de la probabilité ; en effet pour avoir 1 personne ayant la date choisie avec une probabilité supérieure à ½ il faut au minimum 253 personnes dans l’assemblée.

En traçant pour Nmax=2000 on obtient :

On peut remarquer que pour avoir une probabilité supérieure à 0.9 il faut 840 personnes et pour une probabilité supérieure à 0.99 il faut 1679 personnes ce qui peut sembler contre-intuitif à première vue devant le nombre de dates d’anniversaire possibles (365).

De plus on peut faire le lien avec l’exercice précèdent ; en effet en réalité c’est la forte probabilité d’avoir 2 personnes avec une même date d’anniversaire qui réduit la probabilité d’avoir une date d’anniversaire précise. En effet on peut comprendre que la probabilité est proche de 1 quand une très grande majorité des dates d’anniversaire possibles est présente dans l’assemblée.

Dilemme du chevalier de Méré

La thématique ici est le lancer de dé. Le chevalier de Méré était convaincu que les probabilités fonctionnaient de manière proportionnelle, mais les théories probabilistes sont plus compliquées que cela. En effet, il ne suffit pas de multiplier le nombre de lancés par 6 pour avoir autant de chance de gagner au jeu (A) et au jeu (B).

Il faut rappeler que :

* Jeu (A) : Parier sur l’apparition d’au moins un 6 après 4 lancers d’un dé
* Jeu (B) : Parier sur l’apparition d’au moins un double 6  après 24 lancer de 2 dés

Etude théorique des jeux :

Considérons les événements :

* G : “Obtenir au moins un 6 après 4 lancers d’un dé”
* G2 : “Obtenir au moins un double 6 après 24 lancers de deux dés”

Ainsi donc pour calculer théoriquement les probabilités de ces évènements il faut passer par les évènements contraires :

* N : “N’obtenir aucun 6 lors du lancer d’un dé”
* N2 : “Ne pas obtenir un double 6 lors du lancer de deux dés”

On obtient donc aisément :

Soit alors :

A la fin de cette étude théorique on peut alors bien remarquer que le jeu (B) est donc moins avantageux que le jeu (A) ce qui va à l’encontre de l’intuition du chevalier de Méré.

Mais tout ceci est essentiellement théorique, qu’en est-il de l’étude dite empirique qui consiste à répéter l’expérience de nombreuses fois ?

Etude empirique des jeux :

Pour mettre en œuvre cette étude, il est plus simple de passer par l’outil numérique, c’est pourquoi dans la suite de cette étude on observe essentiellement des codes pour mener à bien cette étude. Ici il s’agira de l’outil Matlab, mais tout autre outil peut être mis en place, il suffit d’adapter les notations.

Afin de simplifier le code en question, il est préférable de passer par une étude mettant en œuvre des fonctions et en particulier une fonction permettant de simuler n lancers d’un dé. Celle-ci peut prendre la forme suivante :

function y = LancerDeSixFaces(n)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Fonction permettant de simuler n lancers d'un de%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Initialisation de la liste des resultats obtenus

y = [];

%Repetition de l'experience du lancer de de

for k = 1:n

%Simulation du lancer du de par l'outil rand

y = [y,ceil(6\*rand())];

end

end

On peut aisément vérifier la véracité de cette fonction simplement à l’aide d’un lancer de 10000 dé et de vérifier que l’ensemble des 6 résultats soit bien équiprobable. Soit mettre en œuvre le programme suivant :

function main()

%%%%%%%%%%%%%%%%%%%%%

%Fonction principale%

%%%%%%%%%%%%%%%%%%%%%

%Lancer des des

y = LancerDeSixFaces(10000);

%Liste des resutats possibles

X = 1:6;

%Repertoire et comptabilisation des resultats obtenus et

%probabilisation

Y = hist(y,1:6) /10000;

%Trace des probabilités de chaque resultat

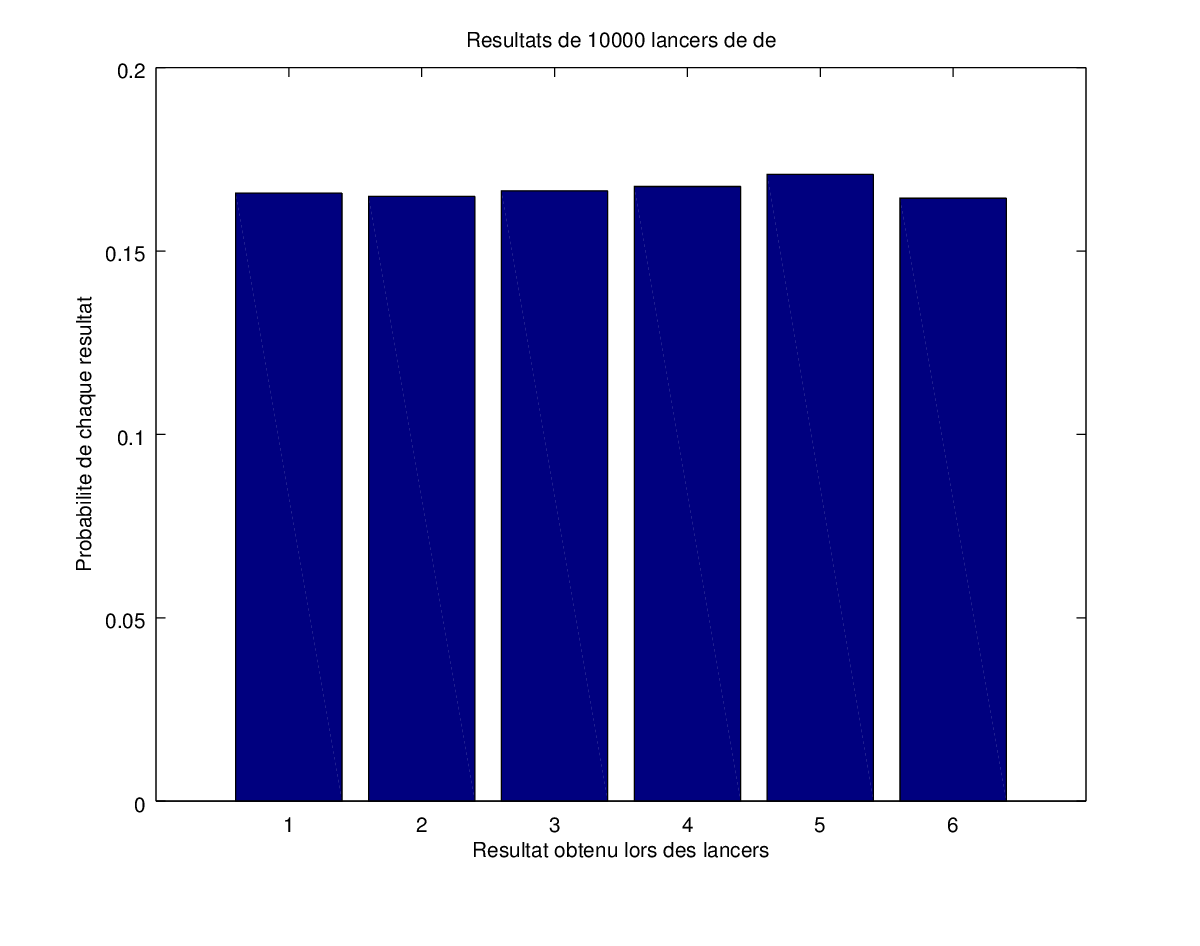
bar(X,Y);

title("Resultats de 10000 lancers de de")

xlabel("Resultat obtenu lors des lancers")

ylabel("Probabilité de chaque résultat")

end



On observe alors bien ici qu’il y a équiprobabilité des résultats ce que l’on souhaitait bien obtenir, par conséquent, cet exemple confirme bien la véracité et le bon fonctionnement de la fonction mise en œuvre plus haut.

Maintenant que la fonction principale est posée, les deux jeux peuvent êtres tester et ce à répétition. Pour ce faire il faut alors faire marcher le programme suivant :

function main()

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Fonction permettant de mettre en oeuvre les deux jeux%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all ; close all ; clc;

%Nombres de fois ou l'experience est realisee

N = 10000;

%Nombre de fois ou on a gagne le jeu A

gagne6 = 0;

%Nombre de fois ou on a gagne le jeu B

gagne2x6 = 0;

%Liste des probabilites de gagner au jeu A et B en fonction du %nombre de simulations

GagneA = [];

GagneB = [];

%Repetition de l'experience

for simulation = 1:N

%%%JEUX A

%Lance du de 4 fois

A = LancerDeSixFaces(4);

%Repertoire et denombrement des resultats

h = hist(A,[1:6]);

%Recherche si le jeu a ete gagne ou non

if h(6) > 0

gagne6 = gagne6 +1;

end

%%%JEUX B

%Lance des des 24 fois

B1 = LancerDeSixFaces(24);

B2 = LancerDeSixFaces(24);

%Somme des resultats

B = B1 + B2;

%Repertoire et denombrement des resultats obtenus

hB = hist(B,[1:12]);

%Recherche si on a gagne le jeu (donc 2 x 6 donc 12)

if hB(12) >0

gagne2x6 = gagne2x6 +1;

end

%Repertoire des probabilites de gagner aux deux jeux

GagneA = [GagneA,gagne6 / simulation];

GagneB = [GagneB,gagne2x6 / simulation];

end

%Trace de l'etude empirique

figure(1)

plot(1:N,GagneA,1:N,GagneB)

legend('Probabilite du jeu A', 'Probabilite du jeu B')

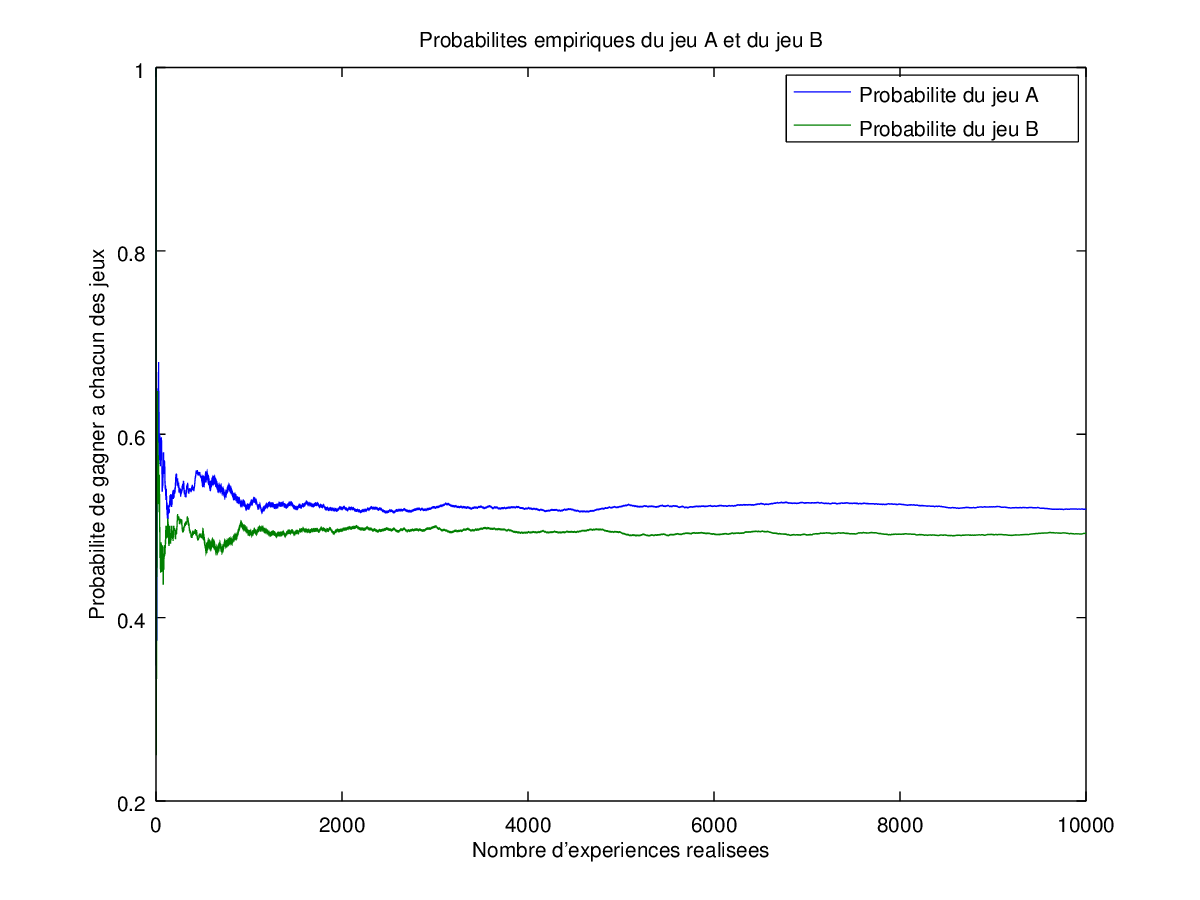
title("Probabilites empiriques du jeu A et du jeu B")

xlabel("Nombre d'experiences realisees")

ylabel("Probabilite de gagner a chacun des jeux")

hold on;

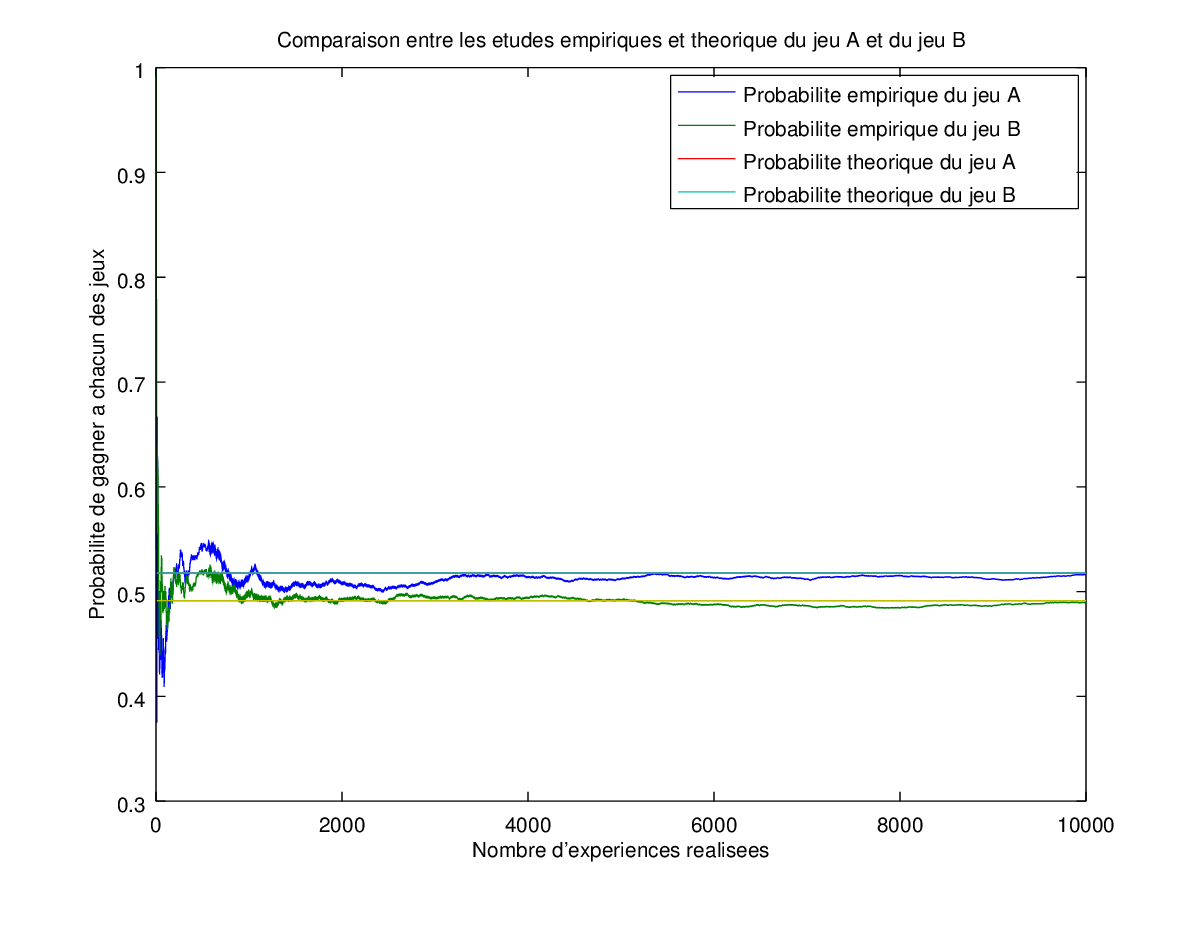
end



On peut alors observer ici qu’après un moment de flottement lors des 1000 premières expériences, on atteint une stabilisation de la probabilité de gain après 2000 expériences, probabilité qui semble suffisamment significative pour définir le jeu (A) comme étant plus bénéfique que le jeu (B). Ce qui confirmerait donc les attentes issues de l’étude théorique précédente.

Comparaison des deux études :

Afin de voir si les deux études précédentes se complètent il suffit alors de tracer les deux droites théoriques sur le graphique précédent comme suit :



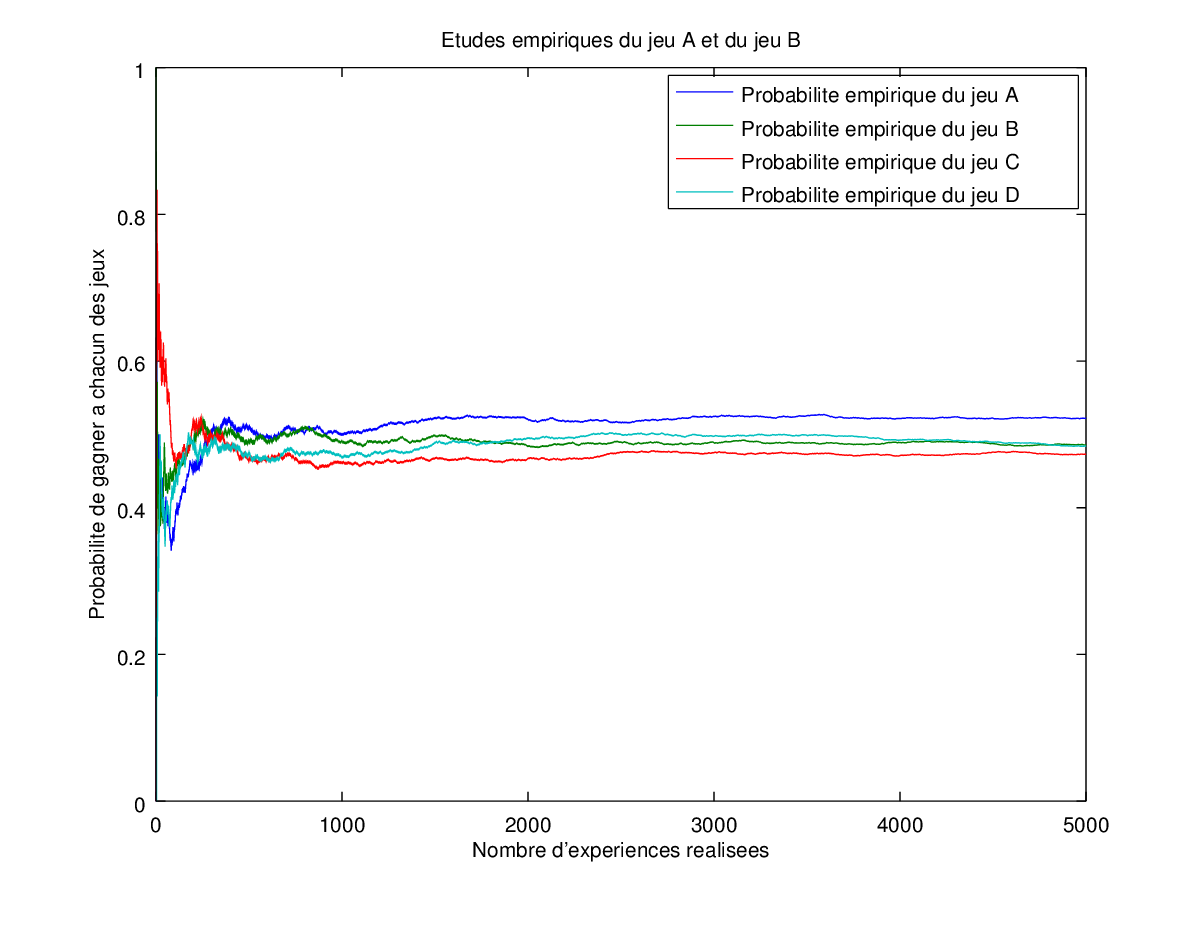
Cette comparaison reflète bien le fait que l’étude théorique et l’étude empirique se rejoignent bien et donc ceci confirme tous nos dires : Le chevalier de Méré a donc bel et bien tort et l’on ne peut pas considérer de proportionnalité entre ces deux épreuves.

Et avec plus de dés ?

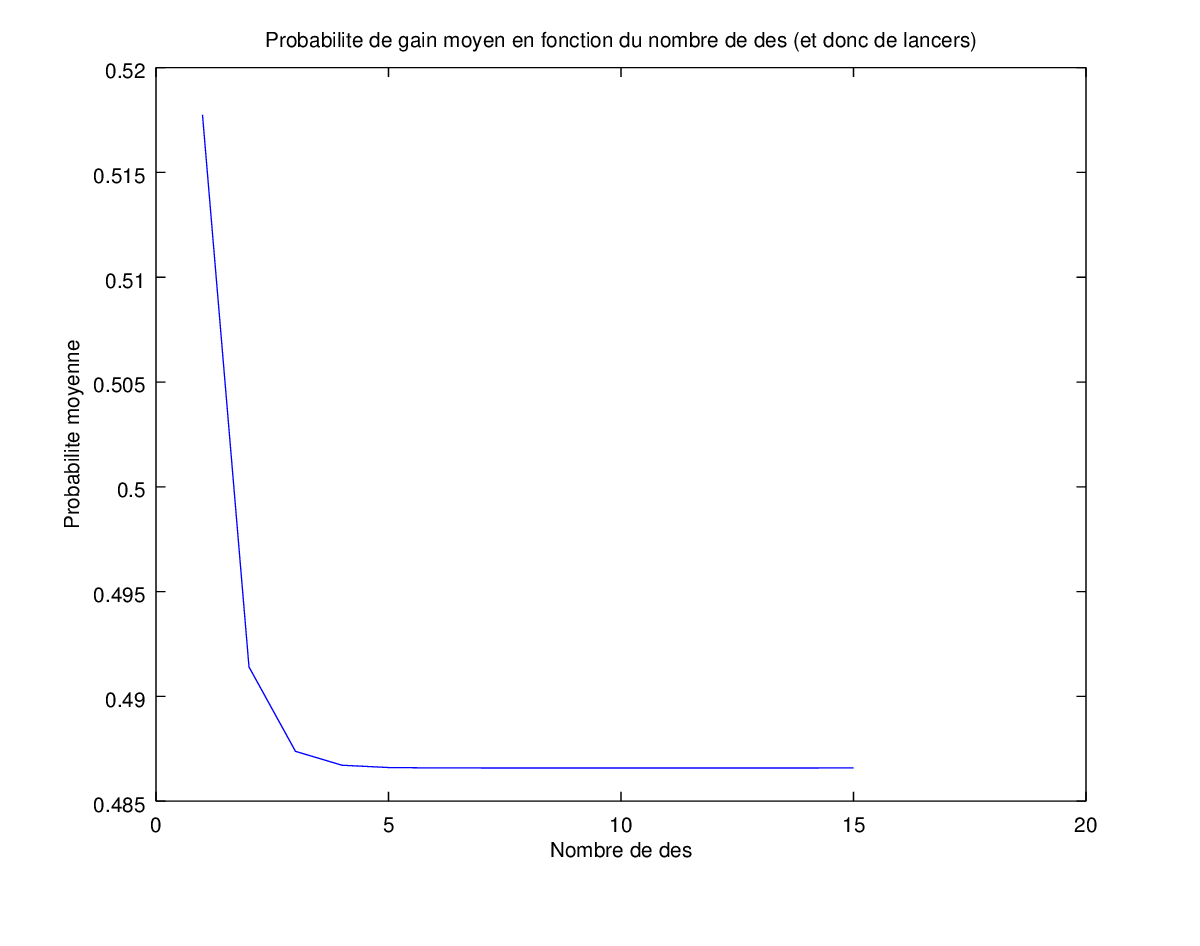
On peut également observer l’évolution des probabilités en fonction du nombre de dés, pour cela il suffit de considérer la même expérience avec 3,4 et 5 dés et considérer les événements suivants :

* G3 : “Obtenir au moins un triple 6 au cours de 144 lancers de trois dés”
* G4 : “Obtenir au moins un quadruple 6 au cours de  864 lancers de quatre dés”

On obtient alors les simulations suivantes :



On peut alors avec cette étude constater que les lancers semblent converger vers une valeur intermédiaire, pour confirmer ou infirmer cette affirmation qui paraît intuitive au regard de ce graphique, passons par une étude théorique du phénomène. Il suffit alors de rassembler dans un même graphique les probabilités de gain en fonction du nombre de dés que l’on lance.



On peut alors remarquer que l’on a bien une stabilisation de la probabilité de gagner si l’on augmente le nombre de dés et donc le nombre de lancers car rappelons que le nombre de lancers (L) est déterminé par le nombre de dés (D) par la formule qui suit :

La stabilisation évoquée ci-dessus peut alors être assimilée à la valeur :

P(G) = 0.48659

Conclusion

Au fil de ces études il a alors pu être constaté deux phénomènes bien distincts :

Dans un premier temps, on a pu voir que la probabilité d’avoir 2 personnes ayant deux dates d’anniversaire identique est élevé même pur un nombre des personnes relativement faible.

Et dans un second temps que si l’on lance peu de dé, la probabilité de gagner en ayant une combinaison de 6 est d’autant plus élevée qu’il y a peu de dé mais lorsque l’on lance un grand nombre de dé un grand nombre de fois comme précisé ci avant, alors la probabilité de gagner reste constante et ce peu importe le nombre de dés.

On donc pu remarquer que l’étude numérique des probabilité est d’autant plus proche de la valeur théorique que le nombre d’itérations est élevé, de plus dans ces 2 phénomènes on prend conscience que les probabilité peuvent parfois être contre intuitif et que seule le vrai calcul de la probabilité peut attester son fonctionnement.