Problème du repassage d'un point par l'origine

Nino Miyauchi-Frangville, Boris Santamaria

Collège Saint-Hubert – Watermael-Boitsfort
Enseignante en charge : Mme Swaelens

Énoncé :

Soit P un point placé sur le point d'origine d'une droite graduée. À chaque mouvement, le point P se déplace aléatoirement à gauche ou à droite, d'une unité. Combien de fois, en moyenne, le point P passe-t-il par le point d'origine après n mouvements ? Qu'en est-il si le point P se déplace aléatoirement à gauche, en haut, à droite ou en bas sur un plan cartésien ?

Résolutions :

1. Sur une droite graduée (en 1D)

Notons $P(n)$ la probabilité que le point P soit sur le point d'origine après n mouvements, et $V(n)$ la valeur attendue du nombre de fois que le point P passe par le point d'origine après n mouvements.

Pour que P soit sur le point d'origine après n mouvements, il faut qu'il ait fait $\frac{n}{2}$ mouvements à gauche et $\frac{n}{2}$ mouvements à droite. Le point P ne repasse donc par l'origine, qu'après un nombre pair de mouvements. Cela nous autorise donc à ne considérer que les n pairs.

Comme, à chaque mouvement, on a le choix entre aller à gauche et aller à droite, le nombre d'itinéraires différents sera 2^n. Sur ces 2^n itinéraires, il y en aura $C_n^{n/2}$ qui seront allés $\frac{n}{2}$ fois à gauche et $\frac{n}{2}$ fois à droite.

$P(n)$ vaut donc $\frac{C_n^{n/2}}{2^n}$.

Calculons maintenant $V(n)$.

Sur les 2^n itinéraires, il y en aura $\mathrm{P}(2) \cdot 2^n$ qui auront été sur l'origine au deuxième mouvement, $P(4) \cdot 2^n$ qui auront été sur l'origine au quatrième mouvement, etc.

Le nombre de repassages par l'origine sur tous les 2^n itinéraires vaudra donc

$$P(2) \cdot 2^n + P(4) \cdot 2^n + \cdots + P(n) \cdot 2^n.$$

$V(n)$, soit la moyenne du nombre de repassages de tous ces itinéraires vaut :

$$\frac{P(2) \cdot 2^n + P(4) \cdot 2^n + \cdots + P(n) \cdot 2^n}{2^n}$$

qui vaut, tout simplement, :

$$P(2) + P(4) + \cdots + P(n).$$

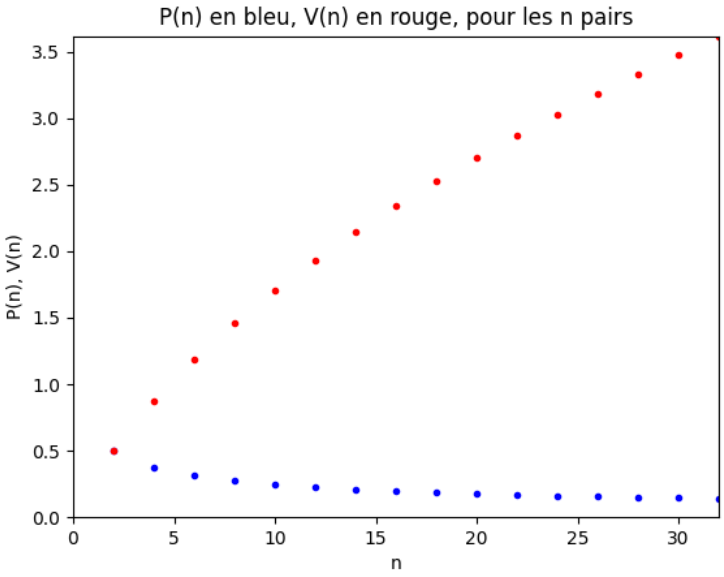
N'oublions pas que cela ne représente que $V(n)$ pour les n pairs. En généralisant cette formule, nous pouvons dire que pour tout n naturel, le nombre de fois, en moyenne, que le point P passe par le point d'origine après n mouvements vaut :

$$\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} \frac{\mathrm{C}_{2k}^{k}}{4^k} = \frac{\mathrm{C}_2^1}{2^2} + \frac{\mathrm{C}_4^2}{2^4} + \cdots + \frac{\mathrm{C}_{2\left\lfloor \frac{n}{2} \right\rfloor}^{\left\lfloor \frac{n}{2} \right\rfloor}}{2^{\left\lfloor \frac{n}{2} \right\rfloor}}.$$

Remarque : Cette formule n'est pas définie pour $n = 0, 1$. En réalité, comme P est déjà sur le point d'origine au 0ème mouvement mais que nous ne le considérons pas comme un « repassage », nous devrions ajouter le terme $\frac{\mathrm{C}_0^0}{2^0}$ et soustraire 1 mais nous préférons nier cet élément pour la beauté de la formule. (Il est de toute façon évident que l'on ne peut pas repasser par l'origine en faisant 1 mouvement, et encore moins sans en faire aucun.)

Pour quelques valeurs de n, nous obtenons le tableau suivant :*

n	1	2	3	4	6	8	10	14	24	38	56
$P(n)$	0	0.5	0	0.375	0.3125	0.273 …	0.246 …	0.209 …	0.161 …	0.129 …	0.106 …
$V(n)$	0	0.5	0.5	0.875	1.1875	1.461 …	1.707 …	2.142 …	3.03 …	4.015 …	5.05 …



**

Pour $V(n)$, les valeurs pour les n impairs sont évidemment identiques que pour les n pairs qui les précèdent, étant donné que P ne repasse jamais par l'origine à un mouvement impair. Nous remarquons aussi que $P(n)$ est décroissant et s'approche de 0, ce qui fait que $V(n)$, qui est l'accumulation de $P(n)$, croît avec une dérivée décroissante. En effet, $V(n)$ dépasse rapidement 1 avec $n = 6$ mais il faudra attendre $n = 14$ pour $V(n) > 2$, puis $n = 24$ pour $V(n) > 3$, $n = 38$ pour $V(n) > 4$, etc. Ce ne sera que pour $n = 190$ que $V(n)$ dépassera 10.

2. Sur un plan cartésien

Le point P peut maintenant se déplacer dans 4 directions : à droite, à gauche, en haut et en bas. Appelons le cas sur la droite graduée d_1 et celui sur le plan cartésien d_2, respectivement pour 1 dimension et 2 dimensions. Nous noterons plus tard, d_t pour t dimensions.

Intuitivement, il semble naturel d'appliquer le même raisonnement en d_2 qu'en d_1.

Notons d'abord la probabilité que P soit sur l'origine après n mouvements en d_2, $P_2(n)$, et la valeur attendue du nombre de fois que P passe par l'origine après n mouvements en d_2, $V_2(n)$.

Nous remarquons d'abord que comme en d_1, P ne peut passer par l'origine qu'après un nombre pair de mouvements, et que le nombre d'itinéraires pour n mouvements est 4^n, puisqu'à chaque mouvement, on a le choix entre 4 directions. A nouveau, nous ne considérons plus que les n pairs.

Calculons maintenant $P_2(n)$.

Pour que P soit sur l'origine, il faut qu'il ait fait le même nombre de mouvements à droite qu'à gauche, et le même nombre de mouvements en haut qu'en bas. Pour n mouvements, il doit donc avoir fait j mouvements en haut, j mouvements en bas, $\frac{n}{2} - j$ mouvements à droite et $\frac{n}{2} - j$ mouvements à gauche, avec $0 \leq j \leq \frac{n}{2}$. (En faisant la somme de ces éléments, nous obtenons bien n mouvements au total.)

Le nombre de façons de faire $\frac{n}{2} - j$ mouvements à droite et à gauche vaut $\mathrm{C}_{n-2j}^{n/2-j}$, et le nombre de façons de faire j mouvements en haut et en bas vaut C_{2j}^{j}.

Nous devons aussi prendre en compte le nombre de façons de faire $2j$ mouvements sur la droite verticale (en haut et en bas) et $n - 2j$ mouvements sur la droite horizontale (à droite et à gauche), qui vaut C_{n}^{2j}.

$P_2(n)$ vaudra donc la somme du nombre de manières de faire j mouvements en haut, j mouvements en bas, $\frac{n}{2} - j$ mouvements à droite et $\frac{n}{2} - j$ mouvements à gauche, pour tous les j pairs compris entre 0 et $\frac{n}{2}$, sur les 4^n itinéraires totaux.

Nous obtenons donc :

$$P_2(n) = \frac{\sum_{j=0}^{n/2} \mathrm{C}_{n}^{2j} \cdot \mathrm{C}_{2j}^{j} \cdot \mathrm{C}_{n-2j}^{n/2-j}}{4^n}.$$

Comme, pour rappel, C_{a}^{b} vaut $\frac{a!}{b!(a-b)!}$, nous remarquons que l'expression peut se simplifier en :

$$\frac{\sum_{j=0}^{n/2} \frac{n!}{\cancel{(2j)!}\,\cancel{(n-2j)!}} \cdot \frac{\cancel{(2j)!}}{j!\,j!} \cdot \frac{\cancel{(n-2j)!}}{(n/2-j)!\,(n/2-j)!}}{4^n} = \sum_{j=0}^{n/2} \frac{n!}{j!^2\,(n/2-j)!^2\,4^n}.$$

.

Calculons maintenant $V_2(n)$.

Avec le même raisonnement qu'en d_1, pour n pair, $V_2(n)$ vaut $P_2(2) + P_2(4) + \cdots + P_2(n)$, soit :

$$\sum_{j=0}^{2/2} \frac{2!}{j!^2\,(2/2-j)!^2\,4^2} + \sum_{j=0}^{4/2} \frac{4!}{j!^2\,(4/2-j)!^2\,4^4} + \ldots + \sum_{j=0}^{n/2} \frac{n!}{j!^2\,(n/2-j)!^2\,4^n}.$$

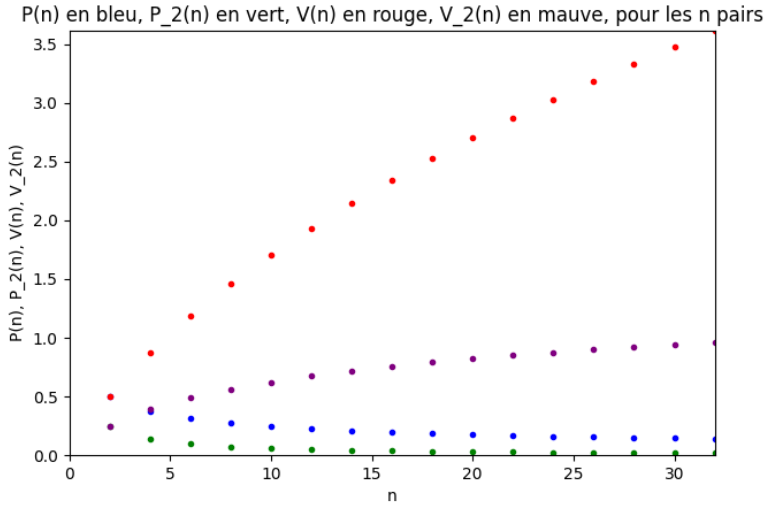
Pour tout n naturel supérieur ou égal à 2, on obtient donc :

$$V_2(n) = \sum_{k=1}^{\lfloor n/2 \rfloor} \sum_{j=0}^{\mathrm{k}} \frac{(2k)!}{j!^2\,(k-j)!^2\,16^k}.$$

.

A nouveau, nous pouvons établir le tableau suivant, en le comparant avec celui de la d_1.

n	0	2	4	6	8	10	14	38	870
$P(n)$	0	0.5	0.375	0.3125	0.273 …	0.246 …	0.209	0.129 …	0.0270…
$P_2(n)$	0	0.25	0.141 …	0.098 …	0.075 …	0.061 …	0.044 …	0.017 …	0.0007…
$V(n)$	0	0.5	0.875	1.1875	1.461 …	1.707 …	2.142…	4.015 …	22.555…
$V_2(n)$	0	0.25	0.391 …	0.488 …	0.563 …	0.624 …	0.718	1.016 …	2.0007…



Nous remarquons que $P_2(n)$ décroît beaucoup plus vite que $P(n)$, ce qui semble logique, étant donné que P peut aller dans plus de directions, ce qui ajoute des conditions pour qu'il revienne au point d'origine. Par conséquent, $V_2(n)$ croît bien plus lentement que $V(n)$.

3. Généralisation à t dimensions

Le point P peut maintenant se déplacer dans $2t$ directions : 2 directions par dimension.

Nous notons $P_t(n)$ la probabilité que P soit sur l'origine après n mouvements et $V_t(n)$ la valeur attendue du nombre de fois que P passe par l'origine après n mouvements, sur t dimensions.

A chaque mouvement, le point P a le choix entre $2t$ directions. Le nombre d'itinéraires total pour n mouvements vaut donc $(2t)^n$.

Pour que P soit sur l'origine après n mouvements, il doit avoir fait le même nombre de mouvements dans les 2 directions de chaque dimension. Nous considérons donc les n pairs seulement.

Concrètement, pour revenir au point d'origine, P doit avoir fait a_1 mouvements dans chaque direction de la première dimension, a_2 mouvements dans chaque direction de la deuxième dimension,…, a_t mouvements dans chaque direction de la t-ième direction, tel que $a_1 + a_2 + \cdots + a_t = \frac{n}{2}$.

Pour chaque $\{a_1, a_2, \dots, a_t\}$, nous calculons le nombre de façons d'arranger ces mouvements : cela nous donne le nombre d'itinéraires pour lequel P revient sur l'origine après n mouvements. Pour calculer $P_t(n)$, nous divisons ensuite le tout par le nombre total d'itinéraires, c'est-à-dire $(2t)^n$. Ainsi, nous obtenons la formule suivante :

$$P_t(n) = \sum_{a_1=0}^{n/2} \sum_{a_2=0}^{n/2-a_1} \cdots \sum_{a_{t-1}=0}^{n/2-a_1-a_2-\cdots-a_{t-2}} \frac{n!}{{a_1!}^2\, {a_2!}^2 \dots {a_{t-1}!}^2\, (n/2 - a_1 - a_2 - \cdots - a_{t-1})!^2\, (2t)^n}$$

Par le même raisonnement que précédemment, $V_t(n) = P_t(2) + P_t(4) + P_t(6) + \cdots + P_t(n)$.

Ainsi, pour tout n naturel supérieur à 1, nous obtenons la merveilleuse formule

$$V_t(n) = \sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} \sum_{a_1=0}^{k} \sum_{a_2=0}^{k-a_1} \cdots \sum_{a_{t-1}=0}^{k-a_1-a_2-\cdots-a_{t-2}} \frac{(2k)!}{a_1!^2\, a_2!^2 \ldots a_{t-1}!^2\, (k-a_1-a_2-\cdots-a_{t-1})!^2\, (2t)^{2k}}.$$

Conclusions :

Ainsi, nous pouvons conclure de ce travail de recherche qu'à chaque fois que l'on rajoute une dimension, il devient plus rare que le point P passe par l'origine, et que les calculs deviennent monstrueusement plus grands. P a également moins de chances de se trouver sur le point d'origine au fur et à mesure qu'il fait des mouvements.

À l'avenir, il serait intéressant de rechercher sur ces différents points :

- Est-il possible de simplifier ces t-uples sommes ? Peut-être sous la forme d'une unique somme ou même d'une unique opération ? Ou peut-être, pourrait-on trouver des approximations de $P_t(n)$ et $V_t(n)$ pour des n assez grands ?

- Qu'en est-il si l'on ne se concentre pas sur l'origine mais sur un certain point X ? (Nombre de fois moyen que P repasse par X, probabilité qu'il repasse par X, etc.)

- Si le nombre de dimension t tend vers l'infini, combien de mouvements le point P doit-il faire pour que la probabilité qu'il passe par l'origine dépasse une certaine valeur ?

*Calculs réalisés à l'aide de SymPy en Python.

Exemple pour calculer le plus petit n pour lequel $V_2(n)$ dépasse 1 :

```
import sympy
from sympy import symbols,summation,binomial,N
k,jv=symbols('k jv')       #jv au lieu de j pour éviter la confusion avec les nombres complexes.
prob=summation(binomial(2*k,2*jv)*binomial(2*jv,jv)*binomial(2*k-2*jv,k-jv),(jv,0,k))/4**(2*k)
n=0
val=0
for nino in range(1,200) :
 n+=1
 val+=N(prob.subs(k,n))
 if val>1 :
  print(2*n)
  break
```

**Graphiques réalisé à l'aide de Matplotlib en Python.

Exemple pour les graphiques de $P(n)$, $P_2(n)$, $V(n)$ et $V_2(n)$:

```
import matplotlib.pyplot as plt
from matplotlib.pyplot import plot,show,xlabel,ylabel,xlim,ylim,title
def factorial(number):
   result=1
   for t in range(1,number+1):
      result*=t
   return result
limit=16
x=[0]*limit
y1=[0]*limit
y2=[0]*limit
y3=[0]*limit
y4=[0]*limit
for i in range(1,limit+1):
   x[i-1]=2*i
   y1[i-1]=factorial(2*i)/factorial(i)**2/4**i
   for k in range(0,i+1):
      y3[i-1]+=factorial(2*i)/factorial(k)**2/factorial(i-k)**2/16**i
for i in range(1,limit+1):
   for k in range(1,i+1):
      y2[i-1]+=y1[k-1]
      y4[i-1]+=y3[k-1]
print(y3,y4)
xlim(0,2*limit)
ylim(0,y2[limit-1])
xlabel("n")
ylabel("P(n), P_2(n), V(n), V_2(n)")
title("P(n) en bleu, P_2(n) en vert, V(n) en rouge, V_2(n) en mauve, pour les n pairs")
plot(x,y1, linestyle="None", marker="o", markersize=3, c="blue", label="P(n)")
plot(x,y2, linestyle="None", marker="o", markersize=3, c="red", label="V(n)")
plot(x,y3, linestyle="None", marker="o", markersize=3, c="green", label="P_2(n)")
plot(x,y4, linestyle="None", marker="o", markersize=3, c="purple", label="V_2(n)")
show()
```