

Probabilités

Table des matières

1. Dénombrements	3
Arrangements, combinaisons, permutations, anagrammes	3
Appariement (devinette à un test associatif)	5
2. Nombres aléatoires	6
Généralités sur les générateurs pseudo-aléatoires	6
Générateur pseudo-aléatoire	7
Application : dessin de fractales	8
Application : Génération de clef de cryptage one time pad	9
Algorithmes probabilistes	10
3. Méthode de Las Vegas	11
Randomisation du quick-sort	11
4. Méthode d'Atlantic City	13
Test de primalité de Rabin-Miller	13
Application : cryptage RSA	16
5. Méthode de Monte Carlo	17
5.1. Les lois statistiques	17
Domaine fini : lois uniforme, triangulaire, Bernoulli, binomiale, hypergéométrique	17
Domaine infini : lois géométriques, binomiale négative, Poisson	20
Exercices TD n°0 (tirage pile ou face ; tirage ; Duc de Toscane)	22
Exercices TD n°1	23
Exercices TD n°2	25
Exercices TD n°3	28
Simuler une loi de probabilité donnée	32
5.2. Exercices d'applications	34
Des boules de toutes les couleurs	34
Somme de nombres aléatoires	35
Le paradoxe des anniversaires	36
A380	37
Chasse aux canards	38
Produits défectueux	39
Pannes de machines	40
Jeu à la foire	41
Loto	42
Poker	43
Transmission d'un signal	44
Collection de vignettes	45
Loi de Poisson	46
Sauts de puce	47
Jardinier	49
Problème de Huyghens	50
Assiettes cassées	51
Echecs successifs	52
Suite monotone	53
Un jeu curieux	54
Tournoi d'échecs	55

Rendez-vous manqué	56
Paradoxe de Parrondo	57
Faux positifs	58
Un tour de probabilité	59
Salle de devoir surveillé (probabilités récursives)	61
Statistiques	62
Sondages	64
Dénombrément à partir d'un échantillon statistique	65
5.3. Géométrie	66
Plus proche voisin	66
Triangle obtus	67
Quadrilatère dans un disque	68
L'aiguille de Buffon	69
Franc carreau	70
Diamant aztèque	71
5.4. Calcul d'intégrale	74
Calcul de π	74
Volume d'une boule de dimension n	76
Intégrale multiple	77
5.5. Marche au hasard	78
Marche au hasard	78
Calcul d'erreur : intervalle de confiance	80
Diffusion du sel (loi de Fick) et marche au hasard	84
6. Physique statistique	89
Physique statistique : fluctuations statistiques	89
Etude d'un système à deux niveaux différents (paramagnétisme)	90
Etude d'un système à deux niveaux identiques (détente Joule Gay-Lussac)	92
Modèle de Ising à 2D : transition para-ferromagnétique	94
Etude d'un système à deux niveaux différents ($N = 4$ particules se partageant une énergie E_{totale})	98
Le paradoxe du groupe de colle	100
Le paradoxe de Monty Hall	101
7. Physique quantique	102
Introduction à la mécanique quantique (interféromètre de Mach-Zehnder)	102
Heisenberg : diffraction par une fente pour 1 photon	105
Fentes d'Young à 1 photon	106
Désintégration radioactive	107

Dénombrements

ISEN-Brest. LA Croix-Rouge. F. Kany

Présentation

On donne un ensemble E de cardinal n (i.e. E possède n éléments distincts).

Comptage

1. Calculer le nombre d'arrangements de $p \leq n$ éléments de cet ensemble en énumérant tous les cas possibles. (i.e. p éléments choisis dans E , de façon ordonnée, sans répétition). Ex : tiercé dans l'ordre.
2. Calculer le nombre de combinaisons de $p \leq n$ éléments de cet ensemble en énumérant tous les cas possibles. (i.e. p éléments choisis dans E , dans un ordre quelconque, sans répétition). Ex : loto.
3. Calculer le nombre de permutations de cet ensemble en énumérant tous les cas possibles. (i.e. les n éléments de E , de façon ordonnée, sans répétition).
4. Un mot M long de n lettres est constitué de r lettres distinctes. La $j^{\text{ème}}$ lettre apparaît p_j fois dans le mot M et donc $p_1 + p_2 + \dots + p_r = n$. Calculer le nombre d'anagrammes du mot M que l'on peut constituer en énumérant tous les cas possibles (peu importe que ceux-ci ait ou non un sens).
5. Calculer le nombre de combinaisons avec répétitions de $p \neq n$ éléments de cet ensemble en énumérant tous les cas possibles. (i.e. p éléments choisis dans E , dans un ordre quelconque, sans répétition). Ex : tous les coloriages de n cartes avec une palette de p couleurs.

Calcul direct

Retrouver les résultats précédents en utilisant les formules habituelles.

Remarque

On évitera d'utiliser la fonction `math.factorial` dans les formules (qui donnent souvent des numérateurs énormes divisés par des très grands dénominateurs).

Wikipédia donne une méthode plus efficace pour calculer $\binom{n}{k}$ [à programmer en itératif et en récursif].
Source :[https://fr.wikipedia.org/wiki/Combinaison_\(math%C3%A9matiques\)](https://fr.wikipedia.org/wiki/Combinaison_(math%C3%A9matiques))

Un algorithme efficace pour calculer le nombre de combinaisons de k éléments parmi n , utilise les identités suivantes ($0 \leq k \leq n$) :

$$\binom{n}{k} = \binom{n}{n-k}, \quad \binom{n+1}{k+1} = \frac{n+1}{k+1} \binom{n}{k} \quad \text{et} \quad \binom{n}{0} = 1$$

La première permet de réduire le nombre d'opérations à effectuer en se ramenant à $k \leq n/2$. Les deux suivantes permettent de montrer que :

$$\binom{n}{k} = \frac{(n-k+1)}{1} \cdot \frac{(n-k+2)}{2} \cdot \dots \cdot \frac{n}{k}$$

À chaque étape de calcul on effectue d'abord la multiplication puis la division pour obtenir un nombre entier (c'est un coefficient binomial), c'est-à-dire que l'on peut employer la division entière. Les calculs intermédiaires restent d'un ordre de grandeur voisin du résultat final (ce ne serait pas le cas si par exemple on utilisait la première formule et la fonction factorielle).

Exemple

$$\binom{10}{7} = ? \quad 10 - 7 = 3 < 7,$$
$$\binom{7}{0} = 1 \Rightarrow \binom{8}{1} = \frac{8}{1} \times 1 = 8 \Rightarrow \binom{9}{2} = \frac{9}{2} \times 8 = 36 \Rightarrow \binom{10}{3} = \frac{10}{3} \times 36 = 120 \Rightarrow \binom{10}{7} = 120.$$

Appariements aléatoires

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soient un ensemble de M questions et un ensemble de M réponses. Chaque question correspond à une et une seule réponse. Un élève répond complètement au hasard. En moyenne, combien de questions seront associées à la bonne réponse ?

Calculer ce nombre en énumérant tous les cas possibles pour $M \in [2, 10]$.

Retrouver ce nombre en effectuant des simulations au hasard.

Les générateurs de nombres aléatoires

Il n'existe pas de méthode numérique pour tirer des nombres aléatoires selon une distribution uniforme.

On peut en revanche tirer des nombres à partir de suites pseudo-aléatoires ; ces suites doivent, en théorie, vérifier une infinité de critères : la moyenne, la variance et tous les moments de la distribution doivent également être ceux d'une distribution uniforme.

Les nombres étant représentés par un nombre fini de bits, les générateurs pseudo-aléatoires sont forcément périodiques. Il est nécessaire (mais pas suffisant) que cette période soit très grande (en tous cas très supérieure au nombre de tirages nécessaires pour effectuer une simulation de type Monte-Carlo).

Il faut également que les séquences de nombres ne soient pas corrélées entre elles.

Il existe deux types d'algorithme pour générer des suites pseudo-aléatoires : ceux basés sur la congruence linéaire et ceux basés sur le déplacement de registre.

Congruence linéaire

On initialise une suite par x_0 et on calcule $x_{n+1} = (a \cdot x_n + c) \bmod m$. On obtient des nombres entre 0 et $m - 1$ avec une période m ; il faut donc que m soit très grand.

On peut augmenter la période en mixant des suites (ex : générateur rng cmrg) : $z_n = (x_n - y_n) \bmod m_1$ avec $\begin{cases} x_n = (a_1 \cdot x_{n-1} + a_2 \cdot x_{n-2} + a_3 \cdot x_{n-3}) \bmod m_1 \\ y_n = (b_1 \cdot y_{n-1} + b_2 \cdot y_{n-2} + b_3 \cdot y_{n-3}) \bmod m_2 \end{cases}$ (P. Lecuyer).

Déplacement de registre

On initialise une suite avec 250 termes (de x_1 à x_{250}) et on calcule $x_n = x_{n-103} \oplus x_{n-250}$ où \oplus est l'opération XOR (Kirkpatrick et Stoll).

Générateur pseudo-aléatoire

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

Une méthode classique pour générer une suite aléatoire $\{r_1, r_2, \dots, r_k\}$ avec $r_i \in [0, M - 1]$ consiste à utiliser la relation de récurrence suivante : $r_i = \text{reste}\left(\frac{a \cdot r_{i-1} + c}{M}\right)$ où a et c sont des constantes. On initialise cette suite en choisissant r_1 et on obtient une séquence périodique (de $n \leq M$ termes) qui se répète à l'infini. On a donc intérêt à choisir a et M grands (par exemple $2^{31} \simeq 2.10^9$ pour une machine 32-bits) pour éviter les répétitions. (Si un programme nécessite, au total, plus de M tirages au sort, il faudra réinitialiser la suite au cours du calcul).

On peut tester le caractère aléatoire de la suite en traçant, sur un graphique à deux dimensions, les points de coordonnées $(x_i, y_i) = (r_{2.i-1}, r_{2.i})$. Le cortex étant particulièrement doué pour reconnaître les formes (à la différence de l'ordinateur), il est facile de voir si la suite est aléatoire.

1. Écrire un programme générant la suite (r_i) avec : $a = 57$, $c = 1$, $M = 256$ et $r_1 = 10$.
2. Déterminer la période de la séquence. Visualiser l'ensemble des points tels que $(x_i, y_i) = (r_{2.i-1}, r_{2.i})$ pour voir si la suite est aléatoire.
3. Reprendre les mêmes questions avec $M = 2^{48}$, $c = 11$, $a = 25\,214\,903\,917$.

Fractales

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

Dans un article célèbre¹, le mathématicien français B. Mandelbrot a montré que, dans la nature, de nombreuses structures obéissent à des règles de construction mathématiques : plantes, coquillages, polymères, colloïdes, aérosols,...

À partir du point de coordonnées $(0.5, 0.0)$, construire l'ensemble de 30 000 points $\{(x_n, y_n)\}$ tels que :

$$(x_{n+1}, y_{n+1}) = \begin{cases} (0.05.x_n, 0.6.y_n), & \text{avec une probabilité de 10\%}, \\ (0.05.x_n, -0.5.y_n + 1.0), & \text{avec une probabilité de 10\%}, \\ (0.46.x_n - 0.32.y_n, 0.39.x_n + 0.38.y_n + 0.6), & \text{avec une probabilité de 20\%}, \\ (0.47.x_n - 0.15.y_n, 0.17.x_n + 0.42.y_n + 1.1), & \text{avec une probabilité de 20\%}, \\ (0.43.x_n + 0.28.y_n, -0.25.x_n + 0.45.y_n + 1.0), & \text{avec une probabilité de 20\%}, \\ (0.42.x_n + 0.26.y_n, -0.35.x_n + 0.31.y_n + 0.7), & \text{avec une probabilité de 20\%}. \end{cases}$$

Essayer également l'algorithme de Barnsley :

$$(x_{n+1}, y_{n+1}) = \begin{cases} (0.5, 0.27.y_n), & \text{avec une proba de 2\%}, \\ (-0.139.x_n + 0.263.y_n + 0.57, 0.246.x_n + 0.224.y_n - 0.036), & \text{avec une proba de 15\%}, \\ (0.17.x_n - 0.215.y_n + 0.408, 0.222.x_n + 0.176.y_n + 0.0893), & \text{avec une proba de 13\%}, \\ (0.781.x_n + 0.034.y_n + 0.1075, -0.032.x_n + 0.739.y_n + 0.27), & \text{avec une proba de 70\%}. \end{cases}$$

Un autre objet fractal a été défini par Sierpinski de la façon suivante. On considère le triangle formé par 3 points de coordonnées : $M_1 = (a_1, b_1)$, $M_2 = (a_2, b_2)$ et $M_3 = (a_3, b_3)$. Soit un point P_0 de coordonnées arbitraires $P_0 = (x_0, y_0)$ à l'intérieur du triangle $M_1M_2M_3$. Construire 15 000 points tels que : P_{n+1} est le milieu de $[M_1P_n]$ (avec une probabilité de $1/3$), de $[M_2P_n]$ (avec une probabilité de $1/3$), de $[M_3P_n]$ (avec une probabilité de $1/3$).

1. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco (1982).

Le chiffrement par masque jetable (one-time-pad)

F. Kany. ISEN-Brest. La Croix-Rouge.

Présentation

Consulter l'article de wikipedia : https://fr.wikipedia.org/wiki/Masque_jetable.

1. Créer une fonction `generation_clef(message)` qui prend en argument un message sous forme d'une chaîne alpha-numérique et qui renvoie une chaîne de caractères aléatoires de même longueur constituée de lettres entre A et Z. Rappel sur les codes ASCII : `ord('A')`=65 et `ord('Z')`=90 ; l'opération inverse est `chr`.
2. Créer une fonction `one_time_pad(message, clef)` qui prend en argument un message sous forme d'une chaîne alpha-numérique et une clef générée par la fonction `generation_clef(message)` et qui renvoie une chaîne de caractères correspondant au message crypté. Rappel : pour effectuer l'opération \oplus (XOR), il faut utiliser l'opérateur `^` (qui n'a rien à voir avec la puissance qui s'écrit `**`), on a : $0^0=0$; $0^1=1$; $1^0=1$; $1^1=0$.
3. Vérifier qu'en appliquant la même clef et la même fonction de codage au message chiffré, on obtient bien le message déchiffré.

Algorithmes probabilistes

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Un algorithme probabiliste est un algorithme utilisant des valeurs produites par un générateur de nombres pseudo-aléatoires.

On distingue :

1. les algorithmes de type Monte Carlo :

On utilise la fonction `random` pour générer un très grand nombre N de répliques indépendantes d'un phénomène. D'après von Mises : comme les probabilités représentent la limite de la fréquence d'observation d'un événement quand N tend vers l'infini, on peut calculer une valeur approchée des probabilités avec un certain intervalle de confiance (cet intervalle diminue quand le nombre N de répliques augmente). On obtient donc des probabilités approchées avec une complexité en temps certaine.

Exemple : pour calculer le nombre π , on tire au hasard N couples (x, y) représentant les coordonnées de points dans un carré de côté 1 ; on cherche le nombre P de points qui se trouvent dans le quart de disque de rayon 1 ; on calcule $\alpha = P/N$. Si N vaut quelques millions, α tend vers la probabilité que le point se trouve dans quart de disque. On obtient une assez bonne approximation de $\pi/4$ (surface du quart de disque/surface du carré) avec un temps de calcul certain (qui est proportionnel à N).

2. les algorithmes de type Las Vegas :

On utilise la fonction `random` pour une étape interne d'un algorithme. On obtient un résultat exact (le même que l'algorithme non randomisé) mais avec une complexité qui est très probablement bien meilleure que l'algorithme non probabiliste.

Exemple : le tri rapide (quicksort) est un algorithme qui trie une liste de n nombres avec une complexité qui peut varier entre $O(n \cdot \ln(n))$ et $O(n^2)$ (dans le pire des cas, si la liste est déjà triée). En randomisant la recherche du pivot, on obtient un tri exact avec une complexité qui, de façon quasi-certaine, est $O(n \cdot \ln(n))$. (Il est en effet très improbable qu'à chaque fois, on tire au hasard un pivot qui est le plus petit ou le plus grand nombre de la sous-liste que l'on est en train de trier).

3. les algorithmes de type Atlantic City :

On utilise la fonction `random` pour effectuer un certain nombre de tests aléatoires sur un algorithme non probabiliste. Les algorithmes de type Atlantic City donnent une réponse probablement juste avec une complexité probablement efficace.

Exemple : les tests de primalité permettent de savoir rapidement si un nombre est très probablement premier. On appelle ces nombres les nombres premiers industriels ; ils sont utilisés en cryptologie (ex chiffrage RSA) où l'on a besoin de calculer rapidement des nombres premiers avec une marge de doute infime. Pour les très grands nombres, le test de Miller-Rabin est beaucoup plus rapide que le crible d'Ératosthène mais, dans des cas exceptionnels, il peut déclarer "premier" un nombre qui ne l'est pas (alors que le crible d'Ératosthène n'échoue jamais).

Randomisation du quicksort

F. Kany. ISEN-Brest. La Croix-Rouge.

Présentation

L'algorithme de tri rapide (ou quicksort) est un algorithme de tri inventé par C.A.R. Hoare en 1960. Son principe est le suivant. Dans une liste de taille n , on choisit un élément p comme pivot. On partitionne la liste en deux parties : les éléments plus petits que p et les éléments plus grands que p .

Dans le cas idéal, la partition divise la liste initiale en deux listes de taille $n/2$.

En réitérant, on obtient des sous-listes de longueurs $n/4$, puis $n/8$, puis $n/16, \dots$

Au bout de k itérations, on obtient des listes de longueurs $n/2^k$. Lorsque les listes ne contiennent plus qu'un seul élément ($\frac{n}{2^k} = 1 \Rightarrow k = \frac{\ln n}{\ln 2} = \ln_2(n)$), le tri est fini. On a donc $k = \ln_2(n)$ itérations où, à chaque fois, il faut comparer tous les éléments à leurs pivots respectifs, soit n comparaisons. Dans le meilleur des cas, il faut donc $n \cdot \ln_2(n)$ comparaisons. C'est la limite théorique minimale pour les algorithmes généraux de tri par comparaisons.

Sur le principe, le code est le suivant :

```
from sys import setrecursionlimit

setrecursionlimit(5000)

def tri(liste):
    if len(liste)>1:
        pivot = liste[0]
        petits = [v for v in liste if v<pivot]
        égaux = [v for v in liste if v==pivot]
        grands = [v for v in liste if v>pivot]
        return tri(petits)+égaux+tri(grands)
    else:
        return liste
```

Malheureusement, très souvent, on est amené à trier des listes qui sont déjà quasiment triées. La partition divise la liste initiale en deux listes : l'une de taille $n - 1$ et l'autre de taille 1. Il faut donc n itérations pour que toutes les listes soient de taille 1. À chaque fois, il faut effectuer n comparaisons. Dans le pire des cas, il faut donc n^2 comparaisons. C'est équivalent aux moins bons algorithmes : tri à bulle, tri par insertion, tri par sélection, ...

Une astuce consiste à prendre le pivot p au hasard dans la liste. Grossièrement, si on choisit p au hasard, on va avoir statiquement toutes les partitions de $(n - 1, 1)$ à $(1, n - 1)$. En moyenne, on aura des partitions de $(n/2, n/2)$.

Conclusion : même si la liste est déjà triée, la complexité reste de $n \cdot \ln_2(n)$.

Question

1. Chronométrier cet algorithme quand on l'exécute sur une liste triée.

```
def test():
    liste=list(range(1000))
    tri(liste)

from timeit import timeit
print(timeit("test()",setup="from __main__ import test",number=100))
```

2. Randomiser l'algorithme en choisissant le pivot au hasard.
3. Re-chronométrer l'algorithme dans les mêmes conditions.

Algorithme de Rabin-Miller

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soit un entier n .

- Si le test de Rabin-Miller permet d'exhiber un nombre $a \in [2, n-1]$ (appelé témoin de Miller), alors on peut en conclure avec certitude que n n'est pas un nombre premier. (Le nombre a témoigne que n est un nombre composé).
- Si le test de Rabin-Miller ne permet pas d'exhiber un témoin de Miller (après k recherches d'un nombre $a \in [2, n-1]$), alors on peut en conclure avec une très grande probabilité que n est un nombre premier. Cette probabilité peut être rendu aussi proche de 100% que l'on souhaite en augmentant le nombre de recherches k . Malheureusement, on ne peut pas exclure que le nombre n soit premier même si on n'a pas réussi à trouver de témoin de Miller.
- Si a n'est pas un témoin de Miller, alors on dit que " n est fortement probablement premier en base a ".
- Si a n'est pas un témoin de Miller, mais que n n'est pas un nombre premier, on dit que a est un "menteur fort".

Principe

Lemme¹

Dans $\mathbb{Z}/p\mathbb{Z}$, si p est premier et si $p > 2$, les seuls nombres X tels que $X^2 \equiv 1 \pmod{p}$ sont $X = +1$ et $X = -1$.

Proposition²

Soit p un nombre premier avec $p > 2$. On a $p-1$ pair et on peut toujours écrire : $p-1 = 2^s \cdot d$ où s et d sont des entiers et d impair (i.e. s est le nombre maximum de fois que l'on peut mettre 2 en facteur dans $p-1$). Pour tout $a \in (\mathbb{Z}/p\mathbb{Z})^*$, on a :

- soit $a^d \equiv 1 \pmod{p}$
- soit il existe $r \in [0, s-1]$ tel que $a^{2^r \cdot d} \equiv -1 \pmod{p}$

Par contraposée, si $a^d \not\equiv 1 \pmod{n}$ et $\forall r \in \{0, 1, \dots, s-1\} \quad a^{2^r \cdot d} \not\equiv -1 \pmod{n}$ alors n est composé et a est un témoin de Miller (que n n'est pas premier).

Exemple

Est-ce que $n = 221$ est premier ? (La réponse est non : $221 = 13 \times 17$)

On a : $n-1 = 220$ que l'on peut écrire $2^2 \times 55$ (soit $s=2$ et $d=55$).

On choisit un nombre $a \in [2, n-1]$ au hasard.

- pour $a = 174$, on calcule :
 - $(a^{2^0})^d \pmod{n} = 174^{55} \pmod{221} = 47 \neq \pm 1$
 - $(a^{2^1})^d \pmod{n} = 174^{110} \pmod{221} = 220 = n-1 = -1 \pmod{221}$.

1. Démonstration : $X^2 \equiv 1 \pmod{p} \Rightarrow (X-1)(X+1) \equiv 0 \pmod{p} \Rightarrow X = \pm 1 \pmod{p}$.

2. Démonstration : D'après le petit théorème de Fermat : $a^{p-1} = (a^d)^{2^s} \equiv 1 \pmod{p}$. En prenant de façon répétée la racine carré de a^{p-1} , on obtient :

- soit $+1 \pmod{p}$ (jusqu'à $a^d \equiv 1 \pmod{p}$)
- soit $-1 \pmod{p}$ (jusqu'à $(a^d)^2 \equiv -1 \pmod{p}$)

puisque, d'après le lemme, les seules racines possibles sont $+1$ et -1 .

Donc 174 n'est pas un témoin de Miller, 221 est fortement probablement premier en base 174. Il faudrait le confirmer en testant d'autres valeurs de a . (En fait, 221 n'est pas premier, 174 est un menteur fort).

- pour $a = 137$, on calcule :
 - $(a^{2^0})^d \bmod n = 137^{55} \bmod 221 = 188 \neq \pm 1$
 - $(a^{2^1})^d \bmod n = 137^{110} \bmod 221 = 205 \neq \pm 1$.
- Donc 137 est un témoin de Miller. 221 n'est pas premier.

Algorithme et temps d'exécution

source https://fr.wikipedia.org/wiki/Test_de_primalit%C3%A9_de_Miller-Rabin

La recherche d'un témoin de Miller peut se décrire algorithmiquement de la façon suivante, l'affectation est notée $:=$, **Témoin_de_Miller(a, n)** renvoie

- Vrai si a est un témoin de Miller que n est composé,
- Faux si n est fortement pseudo-premier en base a

```
Témoin_de_Miller(a, n):                                         entrées : n un entier impair >=3, a un entier >1
    calculer s et d tels que n - 1 = 2**s * d   avec d impair   s > 0 car n impair
    x := a*d % n      x entier reste de la division de a par n
    si x = 1 ou x = n - 1
        renvoyer(Faux)                               sortie : a n'est pas un témoin de Miller
    Tant que s > 1
        x := x**2 % n      reste de la division de x**2 par n
        si x = n - 1
            renvoyer(Faux)                           sortie : a n'est pas un témoin de Miller
            s := s - 1
        Fin de boucle tant que
        renvoyer(Vrai)                            a est un témoin de Miller, n est composé
```

Le test de Miller-Rabin peut alors être décrit comme suit, **Miller_Rabin(n, k)** renvoie

- Vrai si n est fortement pseudo-premier en base a pour k entiers a ,
- Faux s'il est composé.

```
Miller-Rabin(n,k):                                         entrées : n un entier impair >=3, k un entier >=1
    répéter k fois :
        choisir a aléatoirement dans l'intervalle [2, n-1]
        si Témoin_de_Miller(a,n)
            renvoyer(Faux)                         sortie, n est composé
        Fin de boucle répéter
        renvoyer(Vrai)                          sortie, n est probablement premier
                                                si k est suffisamment grand
```

La décomposition $n - 1 = 2^s.d$ avec d impair se calcule en $O(\log(n))$ par une boucle simple. Ce calcul pourrait être factorisé pour être effectué une seule fois dans le test de Miller-Rabin.

Le calcul du reste de la division a^d par n puis les élévations au carré successives sont des calculs d'exponentiation modulaire. Par exponentiation rapide, le calcul se fait en $O((\log d)(\log n)^2)$ pour le calcul initial, suivi de s ($\leq \log(n)$) élévations au carré en $O((\log n)^2)$. Le temps de calcul du premier algorithme, le test que a est ou non un témoin de Miller pour n est donc en $O((\log n)^3)$. Le temps de calcul du test de Miller-Rabin est donc en $O(k.(\log n)^3)$; ainsi cet algorithme est en temps polynomial et efficace.

Question

1. Coder les fonctions **Témoin_de_Miller(a, n)** et **Miller_Rabin(n,k)**.
2. Tester si 4547337172376300111955330758342147474062293202868155909489 est premier
3. Tester si 4547337172376300111955330758342147474062293202868155909393 est premier

Annexe

On peut rendre le rendre le test de Miller-Rabin déterministe en testant, non pas des valeurs de a aléatoires, mais au contraire un très petit nombre de valeurs de a pré-déterminées.

En pratique :

- pour $n < 2\ 047$, il suffit de tester $a = 2$;
- si $n < 1\ 373\ 653$, il suffit de tester $a = 2$ et 3 ;
- si $n < 9\ 080\ 191$, il suffit de tester $a = 31$ et 73 ;
- si $n < 25\ 326\ 001$, il suffit de tester $a = 2, 3$ et 5 ;
- si $n < 3\ 215\ 031\ 751$, il suffit de tester $a = 2, 3, 5$ et 7 ;
- si $n < 4\ 759\ 123\ 141$, il suffit de tester $a = 2, 7$ et 61 ;
- si $n < 1\ 122\ 004\ 669\ 633$, il suffit de tester $a = 2, 13, 23$ et 1662803 ;
- si $n < 2\ 152\ 302\ 898\ 747$, il suffit de tester $a = 2, 3, 5, 7$ et 11 ;
- si $n < 3\ 474\ 749\ 660\ 383$, il suffit de tester $a = 2, 3, 5, 7, 11$ et 13 ;
- si $n < 341\ 550\ 071\ 728\ 321$, il suffit de tester $a = 2, 3, 5, 7, 11, 13$ et 17 .

Cryptage RSA

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

On propose de réaliser la méthode de cryptage due à Rivest, Shamir et Adleman (dite méthode RSA). Cet algorithme date de 1977, il repose sur le fait que, si l'on multiplie ensemble deux nombres premiers très grands (p et q), il est très difficile¹ de redécomposer le résultat obtenu ($n = p \cdot q$) en facteurs premiers (p et q).

La méthode RSA, dite à clef publique, est la suivante² :

1. à chaque lettre ℓ de l'alphabet, on affecte un nombre n_ℓ compris entre 1 et n_{max} .

(Si on ne prend que les lettres majuscules non accentuées et le caractère blanc, on a $n_{max} = 27$).

2. on transforme le texte à transmettre (formé de ℓ_{max} lettres) en un nombre : $message = \sum_{\ell=1}^{\ell_{max}} n_\ell \times n_{max}^\ell$.

3. un interlocuteur 1 choisit secrètement deux nombres premiers très grands p_1 et q_1 congrus³ à 2 modulo

3. Il publie leur produit (la clef publique) : $n_1 = p_1 \cdot q_1$. Il garde secrètement le résultat d'un autre calcul (la clef privée) : $k_1 = \frac{2 \cdot (p_1 - 1) \cdot (q_1 - 1) + 1}{3}$ (k_1 est forcément un entier).

La fonction codage de x consiste à calculer une certaine fonction $f_{n_1}(x)$;

la fonction décodage de x consiste à calculer la fonction inverse : $f_{n_1, k_1}^{-1}(x)$.

4. un interlocuteur 2 fait de même avec p_2 , q_2 , $n_2 = p_2 \cdot q_2$ et $k_2 = \frac{2 \cdot (p_2 - 1) \cdot (q_2 - 1) + 1}{3}$.

5. lorsque 1 veut transmettre le $message$ à 2, il envoie : $xxxx = f_{n_2}(f_{n_1, k_1}^{-1}(message))$.

6. lorsque 2 veut décoder $xxxx$, il calcule $f_{n_1}(f_{k_2, n_2}^{-1}(xxxx)) = message$.

7. le récepteur décompose le nombre $message$ dans la base n_{max} et transforme chaque nombre de la décomposition en lettre de l'alphabet pour retrouver le texte.

La fonction codage $f_{n_i}(yyyy)$ consiste à décomposer $yyyy$ dans la base n_i puis à éléver chaque terme de la décomposition au cube modulo n_i .

La fonction décodage $f_{n_i, k_i}^{-1}(zzzz)$ consiste à éléver $zzzz$ à la puissance $3 \cdot k_i$ modulo n_i . Cette procédure utilise le petit théorème de Fermat généralisé :

- soit m et n deux nombres premiers entre eux ($m < n$) ;

- soit $\Phi(n)$ la fonction d'Euler donnant le nombre des entiers inférieurs et premiers avec n .

Ici $\Phi(n) = (p - 1) * (q - 1)$.

On a alors : $m^{\Phi(n)} \equiv 1 [n] \Rightarrow m^{(p-1) \cdot (q-1)} \equiv 1 [n] \Rightarrow \forall k \in \mathbb{N}, m^{k \cdot (p-1) \cdot (q-1)} \equiv 1 [n]$

$\Rightarrow \forall k \in \mathbb{N}, m^{k \cdot (p-1) \cdot (q-1) + 1} \equiv m [n]$

Comme $p \equiv q \equiv 2 [3]$, on a : $2 \cdot (p - 1) \cdot (q - 1) + 1 \equiv 0 [3]$ et donc $\exists k \in \mathbb{N}$ tel que $2 \cdot (p - 1) \cdot (q - 1) + 1 = 3 \cdot k$

Code avec PYTHON

On a besoin d'un algorithme pour tester si un nombre est premier. On peut utiliser l'algorithme de Miller-Rabin ou bien la fonction `isprime` de la bibliothèque PyPRIME.

Téléchargez la bibliothèque PyPRIME à l'adresse :

<https://pypi.python.org/pypi/pyprimes/>

et installez le package `pyprimes-0.1.1a.tar.gz` à l'aide de WINPYTHON CONTROL PANEL  (Add packages suivi de Install packages).

1. Historiquement, p était un nombre de 64 chiffres et q un nombre de 65 chiffres. En 1994, le nombre $n = p \cdot q$ de 129 chiffres (RSA129) a pu être décomposé en p et q grâce à 1600 ordinateurs travaillant en parallèle pendant deux jours.

2. Les 2 premières étapes sont communes à toutes les méthodes de codage.

3. i.e. dont le reste de la division euclidienne par 3 est 2.

Lois de probabilités sur un domaine fini

F. Kany. ISEN-Brest. La Croix-Rouge.

On note $[[a, b]]$ l'intervalle d'entiers : $a, a+1, a+2, \dots, b-1, b$.

Loi uniforme

Écrire une fonction `f_uniforme(x,a,b)` qui renvoie la distribution de probabilité uniforme discrète pour x sur l'intervalle $[[a, b]]$ (avec $a < b$)

Tracer la loi uniforme `f_uniforme(x,2,5)` sur l'intervalle $[[0, 10]]$

Tracer la fonction de répartition de la loi uniforme `f_uniforme(x,2,5)` sur l'intervalle $[[0, 10]]$

Loi triangulaire

Définition (cas discret)

La loi triangulaire discrète de paramètre entier positif a est définie pour tout entier x compris entre $-a$ et a par :

$$P(x) = \frac{a+1-|x|}{(a+1)^2}$$

Écrire une fonction `f_triangle(x,a)` qui renvoie la distribution de probabilité triangulaire discrète pour x sur l'intervalle $[[-a, a]]$

Tracer la loi triangulaire `f_triangle(x,5)` sur l'intervalle $[-10, 10]$

Généralisation (cas continu)

La loi triangulaire **continue** sur le support $[a; b]$ et de mode c est définie par la densité suivante sur $[a, b]$:

$$f: x \mapsto \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{si } a \leq x < c \\ \frac{2}{b-a} & \text{si } x = c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{si } c < x \leq b \\ 0 & \text{sinon} \end{cases}$$

Écrire une fonction `f_triangle(x,a,b,c)` qui renvoie la distribution de probabilité triangulaire continue de mode c pour x sur l'intervalle $[[a, b]]$ (avec $a < b$)

Tracer la fonction `f_triangle(x,20,70,50)` sur l'intervalle $[[0, 100]]$

Lien avec la loi uniforme

Le jet d'un dé (à 6 faces) correspond à une distribution de probabilités uniforme.

Somme de deux dés

Écrire une fonction `somme_deux_des(x)` donnant la distribution de probabilités pour le jet de deux dés [on effectue la somme des dés] en effectuant une recherche de tous les cas possibles.

Tracer cette fonction sur l'intervalle $[[2,12]]$.

Refaire le même exercice en simulant (avec la fonction `random.randint`) 10 000 fois le jet de 2 dés.

Trouver les paramètres a,b,c de la distribution triangulaire qui correspond à cette distribution. Tracer la avec la même échelle que le graphe précédent.

Différence de deux dés

Ecrire une fonction `difference_deux_des(x)` donnant la distribution de probabilités pour le jet de deux dés [on effectue la valeur absolue de la différence des dés] en effectuant une recherche de tous les cas possibles.

Tracer cette fonction sur l'intervalle $[[0, 5]]$.

Refaire le même exercice en simulant (avec la fonction `random.randint`) 100 000 fois le jet de 2 dés.

Essayer de trouver les paramètres a,b,c de la distribution triangulaire qui approxime cette distribution.

Tracer la avec la même échelle que le graphe précédent.

Commenter

Loi de Bernoulli

La loi de Bernoulli décrit un tirage aléatoire à deux résultats possibles (succès et échec, numérotés 1 et 0), de probabilités respectives p et $1 - p$.

$$P(X = x) = \begin{cases} p & \text{si } x = 1, \\ 1 - p & \text{si } x = 0, \\ 0 & \text{sinon.} \end{cases}$$

À l'aide de la fonction `random.random()`, écrire une fonction `Bernoulli(p)` qui suit la loi indiquée.

Loi de Rademacher

La loi de Rademacher est une Bernoulli équiprobable ($p = 1/2$) où le succès vaut 1 et l'échec -1 .

Ecrire une fonction `Rademacher()`.

Loi binomiale

Définition 1

La loi binomiale, de paramètres n et p , est la loi de probabilité d'une variable aléatoire X égale au nombre de succès rencontrés au cours d'une répétition de n épreuves de Bernoulli, p étant la probabilité de succès dans chacune d'entre elles.

On fera des moyennes sur 10 000 essais.

Tracer sur le même graphique, pour $n = 20$, les diagrammes pour $p = 0.1$, $p = 0.5$ et $p = 0.8$ sur l'intervalle $[[0, 20]]$

Définition 2

La loi binomiale, de paramètres n et p , est la loi de probabilité discrète d'une variable aléatoire X dont la fonction de masse est donnée par :

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \text{ pour } k = 0, 1, \dots, n.$$

À l'aide de la fonction `math.factorial`, écrire une fonction `Binomiale(x,p,n)` qui suit la loi indiquée.

Tracer sur le même graphique, pour $n = 20$, les diagrammes pour $p = 0.1$, $p = 0.5$ et $p = 0.8$ sur l'intervalle $[[0, 20]]$

Loi hypergéométrique

La loi hypergéométrique décrit le résultat d'une série de tirages **Bernoulli dépendants**. Le modèle est celui d'une "urne" dont on tire des "boules" successives noires et blanches sans les remettre dans l'urne.

Définition 1

La loi hypergéométrique de paramètres associés n , p et A est décrit par le modèle suivant :

On tire simultanément n boules dans une urne contenant $p.A$ boules gagnantes et $q.A$ boules perdantes (avec $q = 1 - p$, soit un nombre total de boules valant $p.A + q.A = A$). On compte alors le nombre de boules gagnantes extraites et on appelle X la variable aléatoire donnant ce nombre.

Écrire une fonction `hypergeometrique(x,n,p,A)`.

On fera des moyennes sur 10 000 essais.

Tracer, sur le même graphique, sur l'intervalle $[0, 20]$:

Définition 2

La loi hypergéométrique, de paramètres n , p et A , est la loi de probabilité discrète d'une variable aléatoire X dont la fonction de masse est donnée par :

$$P(X = k) = \frac{\binom{pA}{k} \binom{qA}{n-k}}{\binom{A}{n}} \text{ pour } k = 0, 1 \dots, n.$$

À l'aide de la fonction `math.factorial`, écrire une fonction `hypergeometrique(x,n,p,A)` qui suit la loi indiquée.

Tracer, sur le même graphique, sur l'intervalle $[0, 20]$:

Lois de probabilité sur un domaine infini

F. Kany. ISEN-Brest. La Croix-Rouge.

Les lois étudiées ci-dessous ont pour support \mathbb{N} .

Loi de Bernoulli

La loi de Bernoulli décrit un tirage aléatoire à deux résultats possibles (succès et échec, numérotés 1 et 0), de probabilités respectives p et $1 - p$.

$$P(X = x) = \begin{cases} p & \text{si } x = 1, \\ 1 - p & \text{si } x = 0, \\ 0 & \text{sinon.} \end{cases}$$

À l'aide de la fonction `random.random()`, écrire une fonction `Bernoulli(p)` qui suit la loi indiquée.

Loi géométrique

La loi géométrique décrit le nombre d'essais nécessaires, dans une suite de tirages Bernoulli, avant d'obtenir un succès.

Définition 1

La probabilité $P(X = k)$ correspond à la probabilité d'obtenir dans une succession de k épreuves de Bernoulli, $k - 1$ échecs suivis d'un succès.

Ecrire une fonction `f_geometrique(x,p)` en simulant des tirages aléatoires.

On fera des moyennes sur 10 000 essais.

Tracer la loi géométrique `f_geometrique(x, 0.5)` sur l'intervalle $[[0, 10]]$

Définition 2

La probabilité $P(X = k)$ correspond à $P(X = k) = (1 - p)^{k-1}p$.

Redéfinir la fonction `f_geometrique(x,p)`

Tracer la loi géométrique `f_geometrique(x, 0.5)` sur l'intervalle $[[0, 10]]$

Tracer la fonction de répartition de la loi géométrique `f_geometrique(x, 0.5)` sur l'intervalle $[[0, 10]]$

Loi de binomiale négative

Définition 1

La loi **binomiale négative** est une distribution de probabilité discrète. Elle décrit la situation suivante : une expérience consiste en une série de tirages indépendants, donnant un "succès" avec probabilité p (constante durant toute l'expérience) et un "échec" avec une probabilité complémentaire. Cette expérience se poursuit jusqu'à l'obtention d'un nombre donné n de succès. La variable aléatoire représentant le nombre d'échecs (avant l'obtention du nombre donné n de succès) suit alors une loi binomiale négative. Ses paramètres sont n , le nombre de succès attendus, et p , la probabilité d'un succès.

Ecrire une fonction `f_Pascal(x,n,p)` en simulant des tirages aléatoires.

On fera des moyennes sur 10 000 essais.

Tracer la loi `f_Pascal(x, 3, .5)` sur l'intervalle $[[0, 10]]$

Définition 2

La loi binomiale négative, de paramètres n et p , est la loi de probabilité discrète d'une variable aléatoire X dont la fonction de masse est donnée par :

$$P(X = k; n, p) = \binom{k+n-1}{k} \cdot p^n \cdot q^k \text{ pour } k = 0, 1, 2, \dots$$

À l'aide de la fonction `math.factorial`, écrire une fonction `f_Pascal(x,p,n)` qui suit la loi indiquée.

Tracer la loi `f_Pascal(x,3,.5)` sur l'intervalle $[[0, 10]]$

Loi de Poisson

La loi de Poisson décrit la probabilité d'observer un certain nombre d'événements aléatoires dans un intervalle continu (durée, longueur).

Il faut que ces évènements se produisent avec une fréquence moyenne connue et indépendamment du temps écoulé depuis l'évènement précédent.

Définition 1

On rappelle que la loi binomiale, de paramètres n et p , est la loi de probabilité d'une variable aléatoire X égale au nombre de succès rencontrés au cours d'une répétition de n épreuves de Bernoulli, p étant la probabilité de succès dans chacune d'entre elles.

La loi de **Poisson** correspond à la loi binomiale lorsque $n \rightarrow \infty$ avec $n \times p = \lambda$. En pratique, il suffit que $n \geq 100$.

Ecrire une fonction `f_Poisson(x,lambda_)` en simulant des tirages aléatoires.

On fera des moyennes sur 10 000 essais.

Tracer, sur un même graphique, les diagrammes pour $p = 0.2$, $p = 0.5$ et $p = 0.8$ sur l'intervalle $[[0, 5]]$

Définition 2

Si le nombre moyen d'occurrences dans cet intervalle est λ , alors la probabilité qu'il existe exactement k occurrences (k étant un entier naturel, $k = 0, 1, 2, \dots$) est $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$

On dit alors que X suit la loi de Poisson de paramètre λ .

À l'aide des fonctions `math.exp` et `math.factorial`, écrire une fonction `f_Poisson(x,lambda_)` qui suit la loi indiquée.

Tracer, sur un même graphique, les diagrammes pour $p = 0.2$, $p = 0.5$ et $p = 0.8$ sur l'intervalle $[[0, 5]]$

TD n°0

F. Kany. ISEN-Brest & La Croix-Rouge

Exercices

Tirages pile ou face

- En utilisant la fonction `random.randint`, effectuer N tirages successifs (on considérera - par exemple - que 0 correspond à "face" et 1 correspond à "pile"). On stocke dans une liste le nombre de fois où l'on fait "pile" (exactement) 1 fois de suite, deux fois de suite, trois fois de suite, ...
- Représenter en fonction de n , la probabilité de faire "pile" n fois de suite.
- Représenter sur le même graphique le résultat théorique. Comparer.

Tirage de boules sans remise

On dispose de trois urnes, la première contenant 7 boules blanches et 4 noires, la deuxième 5 blanches et 2 noires, la troisième 6 blanches et 3 noires. On tire une boule dans chaque urne et on note le nombre de boules blanches obtenues.

- Effectuer une simulation numérique pour calculer la probabilité d'obtenir 0, 1, 2 ou 3 boules blanches.
- Comparer avec la théorie.

Tirage de boules avec remise

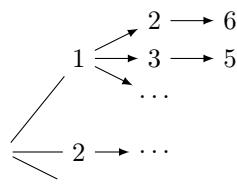
- Simuler le tirage successif de quatre boules avec remise dans une urne contenant 7 boules blanches et 3 boules noires. Compter le nombre de tirages contenant :
 - exactement deux boules blanches ;
 - au moins une boule blanche.
- Comparer avec la théorie

Problème du Duc de Toscane

Cosme II de Médicis (Florence 1590-1621), Duc de Toscane, fut le protecteur de l'illustre Galilée (né à Pise le 15 février 1564 et mort à Florence le 8 janvier 1642) son ancien précepteur. Profitant d'un moment de répit du savant entre l'écriture d'un théorème sur la chute des corps et la création de la lunette astronomique, le Grand Duc lui soumet le problème suivant : il a observé qu'en lançant trois dés cubiques et en faisant la somme des numéros des faces, on obtient plus souvent 10 que 9, alors qu'il y a autant de façons d'obtenir 9 que 10, à savoir six.

Après quelques réflexions, Galilée rédigea un petit mémoire sur les jeux de hasard en 1620 expliquant le phénomène.

- Quelles sont les six façons d'obtenir 9 avec trois dés ?
- Quelles sont les six façons d'obtenir 10 avec trois dés ?
- Simuler le lancer de 3 dés à six faces, trouver la probabilité d'avoir 9 et la probabilité d'avoir 10.
- Comparer à la théorie. On pourra s'aider du graphe suivant qui décrit les valeurs des trois dés.



TD n°1

B. Louédoc. ISEN-Brest. La Croix-Rouge.

Exercices

Exercice 1

Un joueur A lance 2 fois un dé équilibré.

On note X_1 le résultat du dé 1 et X_2 le résultat du dé 2.

Un joueur B lance également 2 fois un dé équilibré.

On note Y_1 le résultat du dé 1 et Y_2 le résultat du dé 2.

1. Simuler l'expérience aléatoire et calculer la valeur de $(X_1 + X_2) - (Y_1 + Y_2)$
2. Ré-itérer l'expérience n fois et donner une estimation numérique de $P(X_1 + X_2 = Y_1 + Y_2)$.

Exercice 2

Soit n un entier fixé supérieur ou égal à 1.

A l'instant $t_0 = 0$, on place un mobile à l'abscisse $x = 0$ de la droite réelle.

A chaque instant t_i (t_i entier), il fait au hasard un pas en avant ou un pas en arrière et ce indépendamment des pas précédents qu'il ait pu faire.

On note x_n sa position à l'instant t_n .

1. Ecrire un programme en python qui, ayant en argument l'entier n , simule l'expérience aléatoire et retourne la valeur de x_n .
2. Ecrire un programme en python qui, ayant en argument l'entier n , un entier k ($-n \leq k \leq n$) et le nombre r de fois où vous avez répété l'expérience pour votre estimation, retourne une estimation numérique de $P(x_n = k)$.

Exercice 3

On dispose d'une urne contenant initialement une boule blanche et d'une pièce de monnaie bien équilibrée.

On effectue une suite de lancers de la pièce de monnaie.

Tant que l'on obtient face lors du lancer de la pièce de monnaie, on ajoute une boule noire dans l'urne et on effectue aucun tirage.

Par contre, la première fois que l'on obtient pile, on tire alors au hasard une boule dans l'urne.

On note B : "un tirage dans l'urne a eu lieu et la boule tirée est blanche".

1. Ecrire un programme en python qui simule l'expérience aléatoire et retourne 0 si B s'est produit et 1 sinon.
2. En déduire un programme en python qui, ayant en argument le nombre r de fois où vous avez répété l'expérience pour votre estimation, renvoie une estimation numérique de $P(B)$

Exercice 4

On effectue une série de tirages dans une urne de la manière suivante :

— Au départ, l'urne contient une boule blanche et une boule noire.

— Après chaque tirage, on remet dans l'urne la boule que l'on vient de tirer ainsi qu'une autre boule de la même couleur.

On note X_n le nombre de boules blanches obtenues lors des n premiers tirages ($n \geq 1$).

1. Ecrire un programme en python qui, ayant en argument l'entier n , simule l'expérience aléatoire et retourne la valeur de X_n .
2. Ecrire un programme en python qui, ayant en argument l'entier n , un entier k ($0 \leq k \leq n$) et le nombre r de fois où vous avez répété l'expérience pour votre estimation, retourne une estimation numérique de $P(X_n = k)$. Quelle est la loi de X_n ?

Exercice 5

On considère plusieurs sacs de billes S_1, S_2, \dots, S_n tels que :

- le premier sac S_1 contient 3 billes jaunes et 2 billes vertes.
- chacun des sacs suivants S_2, \dots, S_n contient 2 billes jaunes et 2 billes vertes.

On réalise l'expérience suivante :

- on tire au hasard une bille dans S_1
- on place la bille tirée de S_1 dans S_2 , puis on tire au hasard une bille dans S_2
- on place la bille tirée de S_2 dans S_3 , puis on tire au hasard une bille dans S_3
- et ainsi de suite

Pour $n \geq 1$, on note E_n l'événement : "la bille tirée dans S_n est verte" et on note $P(E_n) = p_n$ sa probabilité.

1. Ecrire un programme en python qui, ayant en argument l'entier n , simule l'expérience aléatoire et retourne la liste des n boules tirées.
2. En déduire un programme en python qui, ayant en argument l'entier n et le nombre r de fois où vous avez répété l'expérience pour votre estimation, renvoie une estimation numérique de $P(E_n) = p_n$

Exercice 6

Soit n un entier supérieur ou égal à 3.

Une urne contient 2 boules blanches et $n - 2$ boules rouges.

On effectue des tirages sans remise dans cette urne.

On note X le rang de sortie de la première boule blanche.

1. Ecrire un programme en python qui simule l'expérience aléatoire et retourne la valeur de X .
2. Ecrire un programme en python qui, ayant en argument un entier k ($1 \leq k \leq n - 1$) et le nombre r de fois où vous avez répété l'expérience pour votre estimation, retourne une estimation numérique de $P(X = k)$.
3. Ecrire un programme en python qui, ayant en argument le nombre r de fois où vous avez répété l'expérience pour votre estimation, retourne une estimation numérique l'espérance de X .

TD n°2

B. Louédoc & F. Kany. ISEN-Brest

Exercice 1

Un objet a une probabilité $\frac{1}{2}$ de se trouver dans une commode.

Quand il s'y trouve, il a des chances égales de se trouver dans chacun des 9 tiroirs de la commode.

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne non commode ou le tiroir dans lequel se trouve l'objet.
2. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que l'objet se trouve dans le $9^{\text{ième}}$ tiroir.
3. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que l'objet ne soit pas dans les 8 premiers tiroirs.
4. Après avoir ouvert les 8 premiers tiroirs, on constate que l'objet n'y est pas.
Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que l'objet dans le $9^{\text{ième}}$ tiroir

Exercice 2

Une urne A contient quatre boules rouges et six boules noires.

Une urne B contient une boule rouge et neuf boules noires.

Un joueur dispose d'un dé à 6 faces, parfaitement équilibré, numérotés de 1 à 6.

Il le lance une fois :

- S'il obtient 1, il tire au hasard une boule dans l'urne A .
- Sinon, il tire au hasard une boule dans l'urne B .

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne l'urne dans laquelle on a tirée l'urne et la couleur de la boule tirée.
2. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que la boule tirée soit rouge.
3. Le joueur obtient une boule rouge.

Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que la boule vienne de l'urne A.

Exercice 3

Une urne contient deux boules blanches et quatre boules noires. On tire les boules une à une sans les remettre jusqu'à qu'il ne reste que des boules d'une seule couleur dans l'urne.

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne le nombre de tirages effectués.
2. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que l'on ait effectué 4 tirages.

Exercice 4

Dans mon trousseau de clés, il y a 8 clés, toutes semblables. Pour rentrer chez moi, je prends une clé au hasard et je fais ainsi des essais jusqu'à ce que je trouve la bonne, en écartant au fur et à mesure les mauvaises clés.

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne le nombre d'essais effectués pour ouvrir la porte.
2. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne la liste des estimations des probabilités d'ouvrir la porte au premier, second, ..., huitième essai.

Exercice 5

Un joueur A et un joueur B lancent le même dé à tour de rôle.

Le joueur A commence.

La partie s'arrête quand un joueur est le premier à obtenir un six. Ce joueur est alors déclaré vainqueur de la partie.

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne le nombre de lancers effectués et le nom du vainqueur.
2. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation et un entier $n \geq 1$, et qui retourne une estimation de la probabilité que la partie s'arrête au $n^{\text{ème}}$ lancer
3. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que le joueur A gagne la partie.
4. Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que le joueur B gagne la partie.

Exercice 6

Chaque jour, un enseignant étourdi a, quand il commence sa journée avec ses notes de cours, une probabilité $\frac{1}{4}$ de les perdre au cours de la journée.

Il se rend à son lycée le lundi avec ses notes de cours.

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne le jour de la semaine où il a perdu ses notes de cours (lundi, mardi, mercredi, jeudi ou vendredi) ou non perdu s'il les a encore le vendredi soir.
2. Le vendredi, quand il rentre chez lui le soir, il ne trouve plus ses notes de cours dans son cartable.
Écrire un programme, qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité qu'il les ait perdues le lundi, le mardi, le mercredi, le jeudi, le vendredi.

Exercice 7 : gagner deux fois de suite

On considère un jeu à plusieurs manches entre trois joueurs A , B , C qui se déroulent de la manière suivante :

- Pour chaque manche, il n'y a qu'un vainqueur possible.
- Lors de chaque $n^{\text{ème}}$ manche ($n \geq 1$), quand elle a lieu, A et B ont la même probabilité $p = \frac{1}{5}$ de la remporter et C a la probabilité $\frac{3}{5}$ de la remporter .
- Le jeu s'arrête quand un des trois joueurs a remporté 2 manches consécutives et ce joueur est déclaré vainqueur du jeu.

1. Écrire un programme qui simule l'expérience aléatoire et qui retourne le nom du vainqueur.
2. Écrire un programme qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que A soit le vainqueur.

3. Écrire un programme qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que C soit le vainqueur.

TD n°3

B. Louédoc & F. Kany. ISEN-Brest

Exercice 1.

Objectif : Compréhension de la commande random()

La commande `random()` permet de tirer un réel au hasard dans l'intervalle $[0, 1]$.

Pour comprendre ce que cela signifie et l'utilisation que l'on peut en faire, écrire un programme, ayant en argument 2 réels a, b de $[0, 1]$ ($0 < a < b$) et le nombre r de fois que vous allez tirer au hasard un réel dans $[0, 1]$, à l'aide la commande `random()`, pour estimer la probabilité que le réel soit compris entre a et b . Que constatez-vous ?

Exercice 2.

Objectif : Utilisation de la commande random() pour simuler un évènement ayant une probabilité p de se produire ($0 < p < 1$).

p désigne un réel de $]0, 1[$



Soit le segment $[A, B]$ ci-contre

Un pion se déplace sur les sommets du segment $[A, B]$ selon le protocole suivant :

- Le pion est sur le sommet A au départ (instant 0)
- Lorsque le pion est à un instant donné sur le sommet A du segment, il se déplace à l'instant suivant vers le sommet B avec la probabilité $q = 1 - p$
- Lorsque le pion est à un instant donné sur le sommet B , il se déplace à l'instant suivant vers le sommet A avec la probabilité $q = 1 - p$

1. Ecrire un programme, ayant en argument le réel p et un entier n , qui simule l'expérience aléatoire et qui retourne la position du pion à l'instant n .
2. Ecrire un programme, ayant en argument le réel p , un entier n et le nombre r de fois que vous allez simuler l'expérience pour votre estimation, et qui retourne la probabilité d'être en A à l'instant n .
3. Tracer sur un même graphique des courbes représentatives de fonctions $f_n : p \rightarrow f_n(p)$ où $f_n(p)$ désigne votre estimation de la probabilité d'être en A à l'instant n .

4. Comment peut-on se servir de ce qui a été fait pour répondre au problème suivant :

Une succession d'individus A_1, A_2, \dots, A_n se transmet une information binaire du type « oui » ou « non ».

Chaque individu A_k transmet l'information qu'il a reçu avec la probabilité p à l'individu A_{k+1} ou la transforme en son inverse avec la probabilité $q = 1 - p$. ($0 < p < 1$)

Chaque individu se comporte indépendamment des autres.

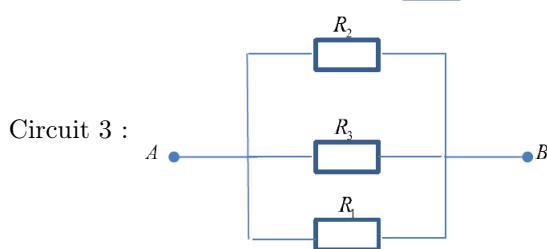
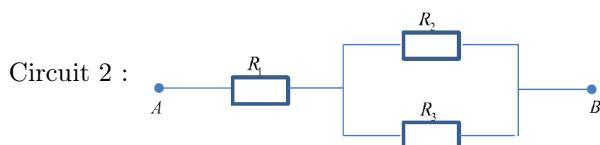
Calculer la probabilité p_n pour que l'information reçue par A_n soit identique à celle émise par A_1 .

Quelle est la limite de p_n quand $n \rightarrow \infty$?

Exercice 3.

Objectif : Comment tenir compte de l'indépendance de 2 évènements ou de la mutuelle indépendance d'une famille d'évènements dans une simulation informatique d'une expérience aléatoire ?

On dispose de 3 résistances R_1, R_2, R_3 . Chaque résistance R_i a une probabilité p_i de fonctionner. Quand une résistance ne fonctionne pas, elle fait office de coupe-circuit. Le fonctionnement ou non de l'une des résistances est indépendante du fonctionnement ou non des autres. Déterminer, en fonction de p_1, p_2, p_3 , pour chacun des 3 circuits ci-dessous, la probabilité qu'un courant puisse parcourir le dipôle AB .



1. Ecrire, pour chacun des 3 schémas, un programme, ayant en argument p_1, p_2, p_3 , et qui retourne 1 si un courant peut parcourir le dipôle AB et 0 sinon.
2. Ecrire, pour chacun des 3 schémas, un programme, ayant en argument p_1, p_2, p_3 et le nombre r de fois que vous avez simulé l'expérience, et qui retourne une estimation de la probabilité qu'un courant puisse parcourir le dipôle AB .

Exercice 4.

On considère un jeu à plusieurs manches entre trois joueurs A, B, C qui se déroulent de la manière suivante :

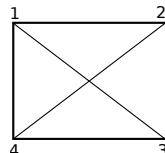
- Pour chaque manche, il n'y a qu'un vainqueur possible.
 - Lors de chaque $n^{ième}$ manche ($n \geq 1$), quand elle a lieu, A et B ont la même probabilité $p = 0.247$ de la remporter et C a la probabilité $1 - 2.p$ de la remporter.
 - Le jeu s'arrête quand un des trois joueurs a remporté 2 manches consécutives et ce joueur est déclaré vainqueur du jeu.
1. Ecrire un programme qui simule le déroulement d'une manche et qui retourne le nom du vainqueur de la manche
 2. Ecrire un programme qui simule le jeu et qui retourne le nom du vainqueur du jeu
 3. Ecrire un programme qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation, et qui retourne une estimation de la probabilité que A soit le vainqueur du jeu.

Exercice 5.

Objectif : Programmation d'une marche aléatoire sur un carré

p désigne un réel de $[0, 1[$.

Soit le carré ci-dessous où les sommets sont numérotés 1, 2, 3 et 4



Un pion se déplace sur les sommets du carré selon le protocole suivant :

- Le pion est sur le sommet 1 au départ

- Lorsque le pion est à un instant donné sur un sommet du carré, il se déplace à l'instant suivant vers un sommet voisin (relié par un côté) avec la probabilité p ou vers un sommet opposé (relié par une diagonale) avec la probabilité $1 - 2.p$.
1. Ecrire un programme, ayant en argument p et un entier n , qui simule l'expérience aléatoire et qui retourne la position du pion à l'instant n .
 2. Ecrire un programme, ayant en argument p , un entier n et le nombre de fois que vous avez simulé l'expérience pour votre estimation, et qui retourne une estimation des probabilités du pion d'être à l'instant n respectivement en 1, 2, 3, 4.

Exercice 6

Pour chaque n de \mathbb{N} la probabilité qu'une famille ait n enfants est $\frac{1}{e \cdot n!}$.

À chaque naissance, la probabilité d'avoir une fille ou un garçon est la même.

1. Ecrire un programme qui simule l'expérience aléatoire et qui retourne le nombre d'enfants de la famille.
2. Ecrire un programme qui simule l'expérience aléatoire et qui retourne la composition de la famille dans l'ordre des naissances
3. Ecrire un programme qui prend en argument le nombre r de fois où on a répété l'expérience pour notre simulation et qui retourne la probabilité que la famille ait 2 filles exactement.

Exercice 7

On considère un combat entre trois tireurs A , B et C .

Le combat se déroule en une suite de manches.

- Lors de la manche 1, les trois tireurs A , B et C sont en lice.
- Lors de chaque $n^{\text{ième}}$ manche ($n \geq 1$) :
 - le tireur A , quand il est encore en lice et qu'il a encore des adversaires, a une probabilité de $\frac{2}{3}$ de toucher sa cible et ceci indépendamment des tirs des autres tireurs.
 - le tireur B , quand il est encore en lice et qu'il a encore des adversaires, a une probabilité de $\frac{1}{2}$ de toucher sa cible et ceci indépendamment des tirs des autres tireurs.
 - le tireur C , quand il est encore en lice et qu'il a encore des adversaires, a une probabilité $\frac{1}{3}$ de toucher sa cible et ceci indépendamment des tirs des autres tireurs.
- Lors de chaque $n^{\text{ième}}$ manche ($n \geq 1$), la cible de chaque tireur est le tireur qui lui paraît le plus dangereux.
- Lorsque, lors d'une $n^{\text{ième}}$ manche, un tireur est touché, il est éliminé et ne participe plus aux manches suivantes.
- Lorsque, lors d'une $n^{\text{ième}}$ manche, un tireur se retrouve seul en lice, plus rien ne se passe lors des manches suivantes et il est déclaré vainqueur du combat.

On introduit les événements suivants :

On note ABC_n ($n \geq 1$) l'événement : « à l'issue de la $n^{\text{ième}}$ manche, les tireurs A , B et C sont encore tous en lice »

On note AB_n ($n \geq 1$) l'événement : « à l'issue de la $n^{\text{ième}}$ manche, seuls les tireurs A , B sont encore en lice »

On définit de même AC_n , BC_n .

On note A_n ($n \geq 1$) l'événement : « le jeu s'est terminé au plus tard à l'issue de la $n^{\text{ième}}$ manche et le vainqueur est le tireur A »

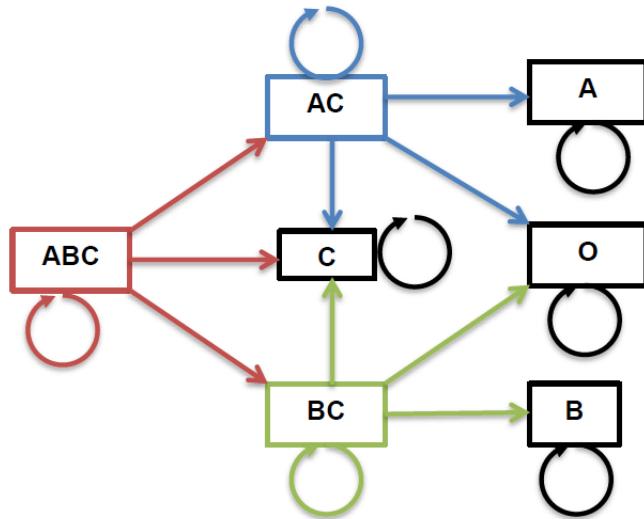
On définit de même B_n , C_n .

Enfin, on note O_n ($n \geq 1$) l'événement : « à l'issue de la $n^{\text{ième}}$ manche, les trois tireurs sont éliminés »

1. Expliquer pourquoi l'événement AB_n est l'événement impossible.

A l'issue de chaque manche, il y a donc 7 états possibles : les états ABC , AC , BC , A , B , C , O .

On peut modéliser la situation par le graphe suivant :



2. En utilisant des simulations informatiques, déterminer une estimation des probabilités qui apparaissent sur ce graphe.
3. Un pion initialement en ABC à l'instant 1 se déplace sur ce graphe avec les probabilités que vous ont donné les simulations informatiques.
Ecrire un programme, ayant un argument un entier n , et qui retourne la position du pion à l'instant n .
4. Ecrire un programme, ayant un argument un entier n et le nombre r de fois que vous avez simulé l'expérience aléatoire, et qui retourne une estimation des probabilités d'être à l'instant n respectivement en ABC, AC, BC, A, B, C, O .
5. Représenter sur un même graphe les fonctions qui à un entier $n \geq 10$ associent les probabilités d'être à l'instant n respectivement en ABC, AC, BC, A, B, C, O .
6. Quelle est une estimation de la probabilité que le vainqueur du combat soit A , soit B , soit C ?

Simuler une loi de probabilité donnée

F. Kany. ISEN-Brest. La Croix-Rouge

Simuler une loi de probabilité donnée

Comment faire pour simuler un générateur aléatoire obéissant à une certaine loi de probabilité $P(x)$?

Méthode 1 : inversion de la fonction de répartition

On se donne une loi de probabilité $P(x)$ pour $x \in [-\infty, +\infty]$.

On cherche la fonction de partition F de la loi de probabilité P définie par : $y = F(x) = \int_{-\infty}^x P(u).du$.

N.B. Si le support de la loi de probabilité est $[a, b]$, alors la fonction de partition est définie par : $y = F(x) = \int_a^x P(u).du$.

On calcule $x = F^{-1}(y)$.

Si y suit la loi de probabilité uniforme (i.e. <tt>random.random()</tt>), alors x suit la loi de probabilité $P(x)$.

Exercice 1 : loi exponentielle

Avec sympy, inverser la loi de probabilité exponentielle, à support dans \mathbb{R}^+ , de paramètre λ : $P(x) = \lambda \cdot \exp(-\lambda \cdot x)$

1. Définir la fonction $P(u) = \lambda \cdot \exp(-\lambda \cdot u)$.

2. Calculer la fonction de partition $y = F(x) = \int_0^x P(u).du$

3. Calculer $x = F^{-1}(y)$

4. Tracer l'histogramme de la loi de probabilité $F^{-1}(y)$ où y est donné par la loi uniforme sur $[0,1]$.

On fera 1000 tirages.

Comparer avec $P(x)$ pour différentes valeurs de λ : 0.1, 0.5, 0.9

Exercice 2 : loi de Laplace

Faire de même avec la loi de Laplace : $P(x) = \frac{1}{2} \cdot e^{-|x|}$.

On tracera $P(x)$ sur $[-5, 5]$

Problème

Problème : il n'est pas toujours possible d'inverser la fonction F (ex : pour la loi gaussienne : $P(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp(-x^2/2)$, on a : $F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \cdot \exp(-u^2/2).du$ qui n'est pas inversable).

Méthode 2 : algorithme de rejet

Le but de cet algorithme est de simuler un tirage suivant une loi de probabilité $P(x)$ à partir d'un générateur aléatoire suivant une loi de probabilité $g(x)$.

Exemple : à partir de $g(x)$ suivant la loi de Laplace (que l'on peut générer par la méthode d'inversion), simuler une loi de Gauss $P(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp(-x^2/2)$ (que l'on ne peut pas générer par la méthode d'inversion).

L'algorithme consiste à :

1. simuler une variable aléatoire X suivant la loi de probabilité $g(x)$
2. simuler une variable aléatoire U suivant la loi uniforme sur $[0,1]$ (i.e. `U=random.random()`)

3. — si $U \leq \frac{P(X)}{M.g(x)}$ où M est une constante, alors on accepte X comme réalisation de la variable aléatoire générée par la loi $P(X)$.
- sinon, on ré-itère depuis 1.

N.B. La constante M doit être telle que : $\forall x, P(x) \leq M.g(x)$. On a intérêt (pour minimiser les cas de rejet) à prendre M la plus petite possible.

1. A l'aide de la méthode d'inversion, créer un générateur suivant la loi de Laplace
2. Dans le cas où $P(x) = \frac{1}{\sqrt{2.\pi}}.e^{-x^2/2}$ et $g(x) = \frac{1}{2}.e^{-|x|}$, on peut montrer que la plus petite valeur de M est $\sqrt{\frac{2.e}{\pi}}$.

Dans une fonction **Gauss()** : tirer une valeur de x suivant la loi de probabilité $g(x)$, tirer une variable u suivant la loi uniforme sur $[0,1]$ et réitérer jusqu'à ce que $u \leq \frac{P(x)}{M.g(x)}$. Renvoyer alors la valeur de x .

3. Tirer au sort 10.000 valeurs en utilisant la fonction <tt>Gauss()</tt>.

Réaliser l'histogramme et comparer à $P(x)$.

Des boules de toutes les couleurs

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

On considère une urne avec 6 boules noires, 6 boules rouges, 10 boules vertes, 12 boules bleues. On tire au sort 6 boules. Quelle est la probabilité d'avoir au moins une boule de chaque couleur ?

Effectuer 100 000 simulations.

Somme de nombres aléatoires

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

On tire au hasard des nombres entre 0 et 1. En moyenne, combien de nombres faut-il tirer pour que la somme dépasse 1 ?

Le paradoxe des anniversaires

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation du problème

Le paradoxe des anniversaires est une estimation probabiliste du nombre de personnes que l'on doit réunir pour avoir une chance sur deux que deux personnes de ce groupe aient leur anniversaire le même jour.

On considérera que l'année fait 365 jours (on néglige les années bissextiles).

Faire une simulation pour trouver le nombre N de personnes qu'il faut réunir pour obtenir une coïncidence de date d'anniversaire. Ré-itérer la simulation 1000 fois. Donner la moyenne de N .

Retrouver ce résultat théoriquement.

A380

F. Kany. ISEN-Brest. Croix-Rouge.

Question

Des passagers embarquent à bord d'un Airbus A380. L'avion est complet. Idéalement, les passagers devraient embarquer un à un et s'asseoir à leur place. Cependant, le premier passager, étourdi, s'assoit à une place aléatoire. Ensuite, les passagers s'assoient à leur place si elle est libre, et à une place libre, aléatoirement, sinon. Vous embarquez le dernier. Quelle est la probabilité que votre place soit occupée ?

On prendra un A380 en configuration standard : 525 places.

Chasse aux canards

F. Kany. ISEN-Brest & La Croix-Rouge



Énoncé

Trois chasseurs A, B et C tirent simultanément sur trois canards 1, 2, 3. Chaque chasseur choisit, au hasard, indépendamment des autres, de viser un canard et ne rate jamais son coup. On note X la variable aléatoire qui correspond au nombre de canards survivants.

Faire 1000 simulations pour trouver l'espérance de X .

Retrouver ce résultat théoriquement.

Produits défectueux

F. Kany. ISEN-Brest & La Croix-Rouge

Énoncé

Une usine produit, grâce à trois machines M_1 , M_2 et M_3 , des pièces qui ont :

- pour la machine M_1 un défaut a dans 5% des cas ;
- pour la machine M_2 un défaut b dans 3% des cas ;
- pour la machine M_3 un défaut c dans 2% des cas.

Une machine M fabrique un objet en assemblant une pièce provenant de M_1 , une pièce provenant de M_2 et une pièce provenant de M_3 . Elle prend au hasard des pièces dans trois stocks comprenant un grand nombre de pièces. Les différentes pièces sont tirées au hasard et indépendamment les unes des autres.

On désigne par X la variable aléatoire qui, à chaque objet prélevé au hasard dans la production de M , associe le nombre de ses défauts. On souhaite connaître la loi de X .

Effectuer 100 000 simulations pour évaluer X .

Retrouver ce résultat théoriquement.

Pannes de machines

F. Kany. ISEN-Brest & La Croix-Rouge

Énoncé

Un atelier contient 10 machines identiques. On admet que la probabilité qu'une machine soit en panne est 0,124 et que les machines fonctionnent de manière indépendante. On appelle X la variable aléatoire qui, à chaque jour, associe le nombre de machines en panne.

Effectuer 100 000 simulations pour évaluer X .

Retrouver ce résultat théoriquement.

Jeu à la foire

F. Kany, ISEN-Brest & La Croix-Rouge



Énoncé

On propose le jeu suivant :

- on mise 1 €,
- on choisit un nombre entre 1 et 6,
- on lance 3 dés
 - si le nombre choisi sort 1 fois, on gagne la mise + 1 €
 - si le nombre choisi sort 2 fois, on gagne la mise + 2 €
 - si le nombre choisi sort 3 fois, on gagne la mise + 3 €

Simuler 10 000 parties, quelle est l'espérance de gain ?

Retrouver ce résultat théoriquement.

Loto

F. Kany. ISEN-Brest & La Croix-Rouge



Énoncé

On tire au hasard 6 boules parmi 49 (numérotées de 1 à 49).

Quelle est la probabilité d'avoir un tirage comportant au moins deux nombres consécutifs ?

(On pourra utiliser `random.shuffle` pour mélanger les boules).

Poker

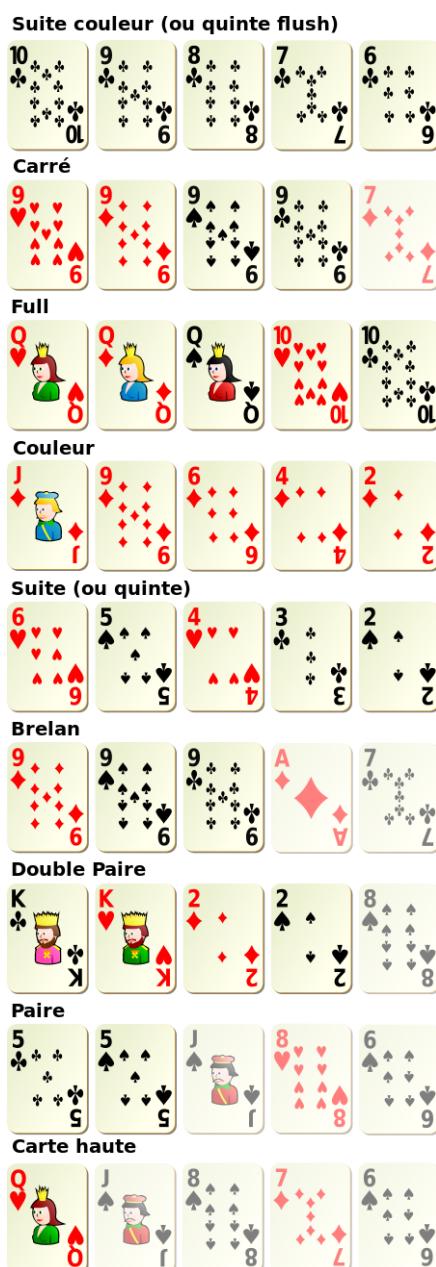
F. Kany, ISEN-Brest & La Croix-Rouge

Énoncé

On joue au poker avec un jeu de 52 cartes (4 couleurs : ♣, ♦, ♥, ♠ et 13 valeurs : As, Roi, Dame, Valet, 10, 9, 8, 7, 6, 5, 4, 3, 2).

Les différentes combinaisons sont rappelées ci-contre.

1. Créer un jeu de 52 cartes.
2. Mélanger-le (on pourra utiliser `random.shuffle`).
3. Prendre les 5 premières cartes.
4. Déterminer la plus haute combinaison que l'on peut former. (Pour cela, il sera utile de compter le nombre de cartes par couleur, par valeur et d'écrire une fonction recherchant la présence de quinte (ou *suite*)).
5. Ré-itérer l'opération 10 000 fois et établir des statistiques sur la probabilités d'avoir une certaine combinaison du premier coup.



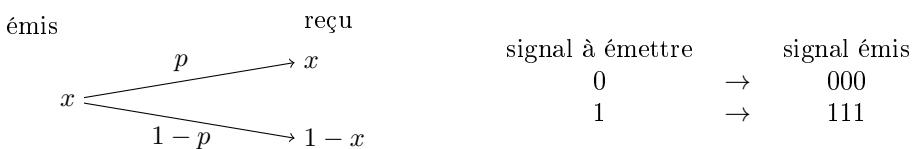
Transmission d'un signal

F. Kany, ISEN-Brest & La Croix-Rouge

Énoncé

La transmission d'un signal binaire $x \in \{0, 1\}$ se fait selon le schéma suivant :

- le signal x est transmis x avec une probabilité p
- et le signal x est transmis $1 - x$ avec une probabilité $1 - p$.



Le bruit de fond qui perturbe la transmission a des origines diverses : orages, parasites, brouillage,...

Une méthode classique pour pouvoir transmettre un message correctement malgré le bruit de fond consiste à transmettre chaque information 3 fois.

Étudier la fiabilité de cette procédure selon les valeurs de p en effectuer 10 000 simulations pour chaque valeur de $p \in [0, 1]$.

Retrouver ce résultat théoriquement.

Collection de vignettes

ISEN-Brest. Croix-Rouge. F. Kany

On souhaite collectionner des vignettes (de joueurs de foot, de héros de dessins animés,...) vendues avec des tablettes de chocolat¹. Chaque tablette contient une vignette et la collection complète comporte n vignettes.

La probabilité de tirer une vignette donnée dans une tablette est supposée uniforme² sur $[1, \dots, n]$.

Questions

1. Une collection
 - Pour $n = 500$, effectuer une simulation pour déterminer le nombre $Z_{n,1}$ de tablettes à acheter pour obtenir une unique collection complète de vignettes.
 - Ré-itérer le calcul précédent une centaine de fois pour déterminer l'espérance de $Z_{n,1} : E(Z_{n,1})$
2. Deux collections
 - Pour $n = 500$, effectuer une simulation pour déterminer le nombre $Z_{n,2}$ de tablettes à acheter pour obtenir deux collections complètes de vignettes.
 - Ré-itérer le calcul précédent une centaine de fois pour déterminer l'espérance de $Z_{n,2} : E(Z_{n,2})$
3. Pour $n \in [2, 2000]$, écrire dans un fichier les valeurs n , $Z_{n,1}$ et $Z_{n,2}$ (sans refaire le calcul une centaine de fois pour chaque n).
4. À l'aide³ de `scipy.signal.savgol_filter(liste, 21, 3)`, lisser les listes de valeurs (`liste`) de $Z_{n,1}$ et $Z_{n,2}$.
5. Tracer, en bleu, sur un même graphique :

$$Z_{n,1,lissé}, \quad n \cdot \sum_{k=1}^n \frac{1}{k} \quad \text{et} \quad n \cdot \ln(n).$$

6. Tracer, en rouge, sur le même graphique :

$$Z_{n,2,lissé}, \quad n \cdot \int_0^\infty [1 - (1 - (1 + t) \cdot e^{-t})^n] \cdot dt \quad \text{et} \quad n \cdot \ln(n) + n \cdot \ln(\ln(n)).$$

1. Références : The Double Dixie Cup Problem, D.J. Newman, The American Mathematical Monthly, Vol. 67, No. 1 (Jan., 1960), pp. 58-6 <https://statistics.wharton.upenn.edu/files/?whdmaction=public:main.file&fileID=735> et G. Lavaud et A. Bégyn, Bulletin vert de l'Union des Professeurs de classes préparatoires Scientifiques. N°250 (printemps 2015), pp. 39-45.

2. Si on utilise une liste PYTHON, on pourra prendre l'intervalle $[0, \dots, n - 1]$ et utiliser `random.randint(0, n-1)`.

3. Attention : `scipy.__version__ >= 14` sinon : <http://wiki.scipy.org/Cookbook/SavitzkyGolay>.

Loi de Poisson

F. Kany, ISEN-Brest & La Croix-Rouge

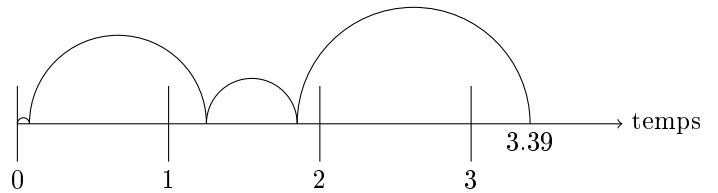
Énoncé

Une entreprise possède un parc important de machines identiques et fonctionnant de façon indépendante. Elle souhaite étudier les pannes de ces machines de façon à établir un plan de maintenance.

Simulation des temps de bon fonctionnement

On choisit une machine au hasard. On suppose que les arrivées des pannes sont indépendantes les unes des autres et que la machine est réparée instantanément. On a relevé les résultats suivants.

Temps de bon fonctionnement	Temps de bon fonctionnement cumulé
0,08	0,08
1,17	1,25
0,60	1,85
1,54	3,39



Le temps t de bon fonctionnement est donné par la loi $t = -\frac{\ln(r)}{2.5}$ où r est une variable aléatoire sur l'intervalle $[0, 1]$. Effectuer 100 000 simulations, trouver le nombre de pannes sur une durée de 2 unités. On tracera la probabilité en fonction du nombre de pannes.

Loi de Poisson

On dit que X suit une loi de Poisson de paramètre λ si la probabilité qu'il existe exactement k occurrences (k étant un entier naturel, $k = 0, 1, 2, \dots$) est donnée par : $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$. La moyenne des occurrences est alors λ .

Calculer la moyenne du nombre de pannes ; en déduire λ ; comparer les valeurs obtenues par simulation aux valeurs théoriques de la loi de Poisson.

Exploitation du modèle

En considérant que X suit la loi de Poisson précédente, déterminer :

- la probabilité qu'une machine tombe en panne au moins trois fois (i.e. 3, 4, 5 ou plus) sur une durée de 2 unités suivant sa mise en service ;
- le nombre maximum de pannes d'une machine (sur une durée de 2 unités suivant sa mise en service) avec une probabilité d'au moins 95 %.

Sauts de puce

F. Kany. ISEN-Brest & La Croix-Rouge

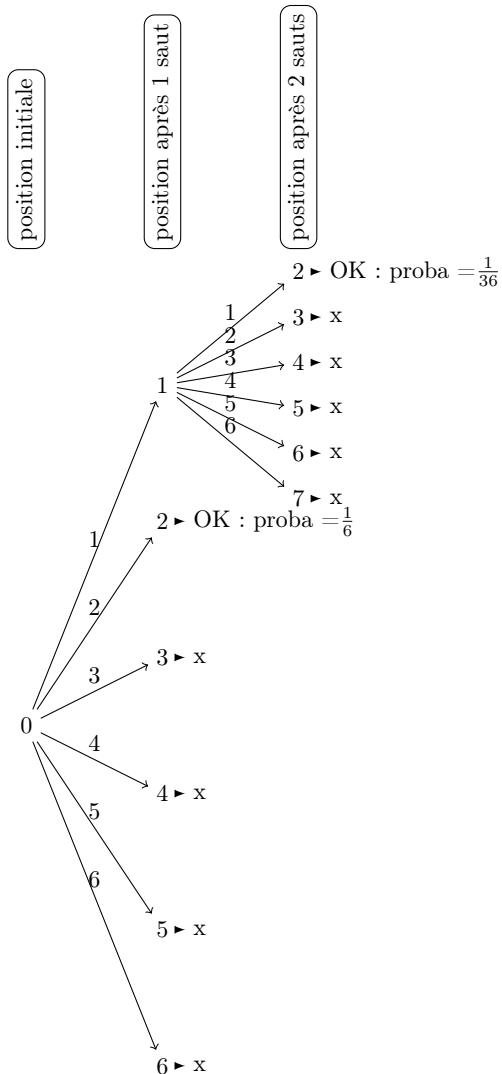
Exercice

Une puce se trouve initialement à x mètres d'un mur. Elle saute en direction de ce mur en effectuant des bonds équiprobables de 1, 2, 3, 4, 5 ou 6 mètres. Quelle est la probabilité que la puce ne s'écrase pas contre le mur, autrement dit qu'elle arrive exactement au pied de ce mur ?

Methode 1

On réalise une exploration systématique de tous les cas possibles. Lorsque - au bout de n sauts - on arrive exactement au pied du mur, on écrit la probabilité : $\frac{1}{6^n}$. On somme toutes les probabilités pour trouver la probabilité demandée.

Exemple avec $x = 2$.



Probabilité d'arriver au pied du mur : $\frac{1}{6} + \frac{1}{36} = \frac{7}{36}$.

À titre de vérification, les premières valeurs sont :

x	probabilité
1	$1/6 = 0.166666666666666666$
2	$7/36 = 0.19444444444444445$
3	$49/216 = 0.22685185185186$
4	$343/1296 = 0.2646604938271605$
5	$2401/7776 = 0.30877057613168724$
6	$16807/46656 = 0.36023233882030176$
7	$70993/279936 = 0.25360439529035206$
8	$450295/1679616 = 0.268094016727633$
9	$2825473/10077696 = 0.28036894544149776$

Méthode 2

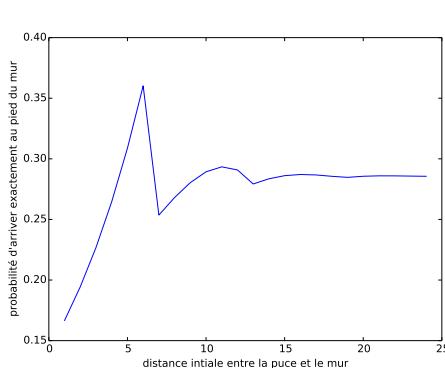
On définit le vecteur ligne $\vec{p}_n = (p_{n,0} \ p_{n,1} \ p_{n,2} \ \dots \ p_{n,x})$ comme le vecteur constitué des probabilités $p_{n,i}$ d'être à l'abscisse i au bout de n sauts.

On a la relation : $\vec{p}_{n+1} = \vec{p}_n \cdot M$ où M est la matrice de taille $(x+1) \times (x+1)$ telle que :

$$M = \frac{1}{6} \cdot \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & & & & & & & & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & & & & & & & & \\ & \ddots & & & & & & & \ddots & & & & & & & \\ & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & & & & & & \\ & & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & & & & & \\ & & & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & & & \\ & & & & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & & \\ & & & & & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & \\ & & & & & & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & \\ & & & & & & & & 0 & 1 & 1 & 1 & 1 & 1 & \\ & & & & & & & & & 0 & 1 & 1 & 1 & 1 & \\ & & & & & & & & & & 0 & 1 & 1 & 1 & \\ & & & & & & & & & & & 0 & 1 & 1 & \\ & & & & & & & & & & & & 0 & 1 & \\ & & & & & & & & & & & & & 0 & \\ 0 & & & & & & & & & & & & & & & x+1,x+1 \end{pmatrix}$$

$$\text{et } \vec{p}_0 = (1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \dots \ 0)_{x+1}$$

La probabilité recherchée est : $p = \sum_{n=x/6}^{x/1} p_{n,x}$ où le nombre n de sauts va de $x/6$ (que des sauts de 6 mètres) à $x/1$ (que des sauts de 1 mètre).



Jardinier

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

Un jardinier dispose d'un grand stock de fleurs bleues et de fleurs rouges. Les fleurs bleues sont en proportion p et les fleurs rouges en proportion $q = 1 - p$ où $p \in [0, 1]$. Il en choisit au hasard $n \times m$, avec $n, m \in \mathbb{N}^*$, et les plantes, également au hasard, dans un parterre rectangulaire avec n lignes et m colonnes. On dit qu'un ligne ou une colonne est bleue lorsqu'elle est constituée uniquement de fleurs bleues.

Quelle est la probabilité P de n'avoir ni ligne bleue, ni colonne bleue ?

Tracer P pour $p \in [0, 1]$ avec $n = m = 10$, puis 50, 100, 150, 200.

Comparer à la valeur théorique :

$$P_{n,m} = \sum_{k=0}^n (-1)^k \cdot \binom{n}{k} \cdot p^{k \cdot m} \cdot (1 - p^{n-k})^m$$

Problème de Huyghens

F. Kany. ISEN-Brest. La Croix-Rouge.

Présentation

Trois joueurs A , B et C participent au jeu suivant. Étant donné 4 jetons blancs et 8 jetons noirs, A , B et C tirent au hasard à tour de rôle un jeton (sans le remettre ensuite).

Le jeu se poursuit ainsi jusqu'à ce que le premier qui tire un jeton blanc gagne la partie.

Sachant que le joueur A commence la partie, suivi de B et C , quelles sont les chances de chacun de l'emporter ?

Simuler cette expérience 10 000 fois et calculer les probabilités demandées.

Assiettes cassées

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Cinq plongeurs lavent le même nombre d'assiettes. Si cinq assiettes sont cassées :

- quelle est la probabilité qu'un même plongeur en ait cassé quatre ?
- quelle est la probabilité qu'il ait cassé les cinq assiettes ?
- quelle est la probabilité qu'un plongeur donné ait cassé quatre assiettes ?
- quelle est la probabilité qu'un plongeur donné ait cassé cinq assiettes ?

Trouver ces probabilités de façon analytique ; puis les retrouver de façon numérique en effectuant des simulations.

Echecs successifs

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Si la probabilité qu'un tir touche la cible est $1/3$, quelle est la probabilité de toucher la cible si l'on tire 3 fois de suite ?

Généralisation : la probabilité de toucher la cible est p ; on tire jusqu'à ce qu'on touche la cible ; tracer la probabilité P de toucher la cible en fonction du nombre n de tirs.

Une autre façon de voir le problème est de chercher la probabilité Q de n échecs successifs et de tracer $P = 1 - Q$ en fonction de n .

Suite monotone

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

On tire au hasard une séquence de nombres x_i entre 0 et 1. On arrête la séquence dès que $x_{i+1} \leq x_i$.
On appelle L la longueur de la séquence (qui est donc monotone croissante stricte).

Exemple : la séquence 0.1, 0.2, 0.3, 0.4, 0.5, 0.45 a la longueur $L = 6$.

L est une valeur entière.

Tracer $p_k = P(L = k)$ en fonction de k .

Calculer $E(L)$ l'espérance de L .

Un jeu de pile ou face

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Trois joueurs A, B, C ont respectivement ℓ, m et n pièces de monnaie. Les joueurs lancent simultanément un de leurs pièces.

- Si le résultat est trois fois "pile" ou trois fois "face", alors personne ne gagne ou ne perd.
- Si le résultat est deux fois "pile" et une fois "face" (ou respectivement deux fois "face" et une fois "pile"), alors les deux joueurs qui ont fait "pile" (resp. "face") donnent leurs pièces à la troisième qui a fait "face" (resp. "pile").

Le jeu continue jusqu'à ce qu'un des joueurs n'ait plus de pièces.

Tournoi d'échecs

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soient trois joueurs d'échecs : A, B, C .

A doit rencontrer B et C dans un tournoi un peu particulier. Le tournoi se fait en trois rencontres ; A doit jouer soit contre B , puis C , puis B à nouveau soit contre C , puis B , puis C à nouveau. Le défi est le suivant : A doit remporter deux matches **successifs** ; cela signifie dans le premier exemple (B , puis C , puis B à nouveau) que A doit gagner contre B et C ou bien contre C et B (mais, dans le cas de deux victoires contre B et une défaite contre C , le défi est perdu).

A sait que le joueur C est plus fort que lui (il peut gagner contre lui avec une probabilité q) mais que le joueur B est moins fort que lui (il peut gagner contre lui avec une probabilité $p > q$).

Quel est le défi le plus simple : BCB ou bien CBC ?

Stratégies

D'un côté, le défi BCB semble plus simple puisqu'on ne rencontre le joueur C (le plus fort) qu'une seule fois.

D'un autre côté, le défi CBC permet de rencontrer le joueur C une deuxième fois (ce qui donne une deuxième occasion de le battre).

Quelle est, pour chaque stratégie, la probabilité de relever le défi ?

Rendez-vous manqué

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Deux personnes A et B se donnent rendez-vous dans un café entre 18H30 et 19H00.

Si A arrive le premier, il attend 10 minutes ; s'il ne rencontre pas B pendant cet intervalle de temps, il part.

Si B arrive le premier, il attend 5 minutes ; s'il ne rencontre pas A pendant cet intervalle de temps, il part.

Ni A , ni B n'attendent après 19H00.

Questions

1. Quelle est la probabilité que A et B se rencontrent ?
2. Si A attendait 5 minutes (au lieu de 10), quelle serait la probabilité que A et B se rencontrent ?
3. Si B attendait 10 minutes (au lieu de 5), quelle serait la probabilité que A et B se rencontrent ?

Paradoxe de Parrondo

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soit A un jeu de pile ou face avec une pièce biaisée (pile avec une probabilité de $p = \frac{1}{2} - \varepsilon$, face avec une probabilité de $\frac{1}{2} + \varepsilon$ où $\varepsilon > 0$). On lance la pièce. Si on obtient pile, on gagne un jeton (ou un euro) ; sinon on perd un jeton (ou un euro).

Soit B un jeu avec deux pièces biaisées. La pièce 1 donne pile avec une probabilité $p_1 = \frac{1}{10} - \varepsilon$ et la pièce 2 donne pile avec une probabilité $p_2 = \frac{3}{4} - \varepsilon$. Si le joueur a un capital K (de jetons ou d'euros) qui est un multiple de $m = 3$, on lance la pièce 1, sinon on lance la pièce 2. Comme dans le jeu A , si on obtient pile, on gagne un jeton (ou un euro) ; sinon on perd un jeton (ou un euro).

Soit AB un jeu avec une pièce non biaisée. On lance la pièce. Si on obtient pile, on joue à A ; sinon on joue à B .

Questions

On prend $\varepsilon = \frac{5}{1000}$.

On démarre avec un capital $K = 0$.

1. Simuler 100 000 fois, l'évolution du capital K d'un joueur qui joue k fois au jeu A avec $k \in [1, 100]$. Tracer l'évolution de la moyenne de K en fonction de k . Vérifier que A est un jeu perdant (i.e. qui provoque la ruine du joueur s'il joue suffisamment longtemps).
2. Quelle est l'espérance E_A du jeu A ? Tracer E_A en fonction de k . Vérifier que ce calcul se superpose à la simulation précédente.
3. Simuler 100 000 fois, l'évolution du capital K d'un joueur qui joue k fois au jeu B avec $k \in [1, 100]$. Tracer l'évolution de la moyenne de K en fonction de k . Vérifier que B est un jeu perdant.
4. Quelle est l'espérance E_B du jeu B ? (On calculera les probabilités par récurrence). Tracer E_B en fonction de k . Vérifier que ce calcul se superpose à la simulation précédente.
5. Simuler 100 000 fois, l'évolution du capital K d'un joueur qui joue k fois au jeu AB avec $k \in [1, 100]$. Tracer l'évolution de la moyenne de K en fonction de k .
6. Quelle est l'espérance E_{AB} de ce jeu ? (On calculera les probabilités par récurrence). Tracer E_{AB} en fonction de k . Vérifier que ce calcul se superpose à la simulation précédente.
7. A et B sont des jeux perdants. Le jeu AB est-il également perdant ?

Faux positifs

F. Kany. ISEN-Brest. La Croix-Rouge.

Présentation

Une maladie touche 0,1% de la population et l'on dispose d'un test de dépistage, fiable à 99% lorsque la personne est atteinte de la maladie et 95% quand elle ne l'est pas.

Si une personne fait le test et que ce dernier se révèle positif, quelle est la probabilité que cette personne soit réellement malade ?

Simuler cette expérience 100 000 fois et calculer le taux de faux positifs (personne saine déclarée positive).

Un tour de probabilité

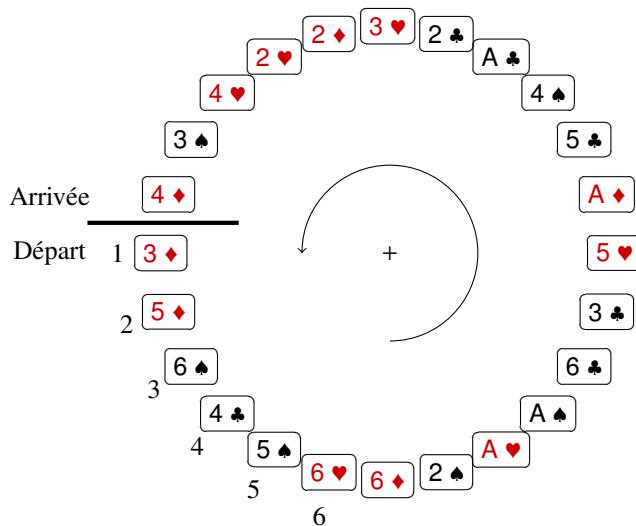
F. Kany. ISEN-Brest

Énoncé

Dans un jeu, on prend les cartes

A ♣	2 ♣	3 ♣	4 ♣	5 ♣	6 ♣
A ♦	2 ♦	3 ♦	4 ♦	5 ♦	6 ♦
A ♥	2 ♥	3 ♥	4 ♥	5 ♥	6 ♥
A ♠	2 ♠	3 ♠	4 ♠	5 ♠	6 ♠

On les mélange et on les dispose en cercle à partir d'une position de départ.



On choisit une des six premières cartes à partir de la position de départ. On avance du nombre de cartes indiquées sur la carte choisie dans le sens trigonométrique. On recommence ainsi jusqu'à franchir à nouveau la ligne de départ.

Quelle est la probabilité que, quelle que soit la carte initiale choisie, on arrive toujours sur la même carte finale.

Exemple : avec le tirage aléatoire ci-dessous, si on choisit, aléatoirement, de partir de la carte n°4. On est sur le [4♣], on avance de 4 cartes dans le sens trigonométrique, on arrive sur le [2♠], on avance de 2 cartes, on arrive sur le [A♠], on avance de 1 carte, on arrive sur le [6♠], on avance de 6 cartes, on arrive sur le [A♠], on avance de 1 carte, on arrive sur le [2♣], on avance de 2 cartes, on arrive sur le [2♦], on avance de 2 cartes, on arrive sur le [4♥], on avance de 4 cartes (après franchissement de la ligne d'arrivée) au [5♦]. Si, avec le même tirage, on était parti de la carte n°5, où serait-on arrivé ? Et en partant de la carte n°3 ?

Avec ce jeu de cartes, pour n'importe quel tirage, quelle est la probabilité que ce phénomène se reproduise ?

Dans le cas général : n séries de m cartes numérotées de 1 à m formant un cercle de $n \times m$ cartes, quelle est la probabilité que ce phénomène se reproduise ?

Bibliographie

Michaël Launay. Chaine YouTube MicMaths
<https://www.youtube.com/watch?v=1LzUcJf3E60>
Michaël Launay. Site MicMaths.
<http://micmaths.com/applis/tourproba.html>

Salle de devoir surveillé (probabilités récursives)

F. Kany. ISEN-Brest & La Croix-Rouge

Problème original

Pour un devoir surveillé, n élèves doivent s'assoir dans une salle avec n tables à leurs noms. Les élèves n'ont pas regardé le plan de la salle et s'assoient au hasard[*] dans la salle. Quelle est la probabilité qu'aucun élève ne se soit assis à la bonne place ?

[*] Au hasard signifie exactement ceci : les élèves discutent debout en étant répartis de façon quelconque dans la salle. Chaque élève est à une plus courte distance d'une des tables et, pour chaque table, cette distance la plus courte correspond à un élève distinct. À l'instant de la sonnerie, tous les élèves s'assoient, en même temps, à la table dont ils sont la plus proche.

Remarque : il peut être intéressant de résoudre le problème généralisé ci-dessous pour répondre au problème original.

Problème généralisé

On considère la généralisation suivante.

Le surveillant ne connaît pas les élèves.

k élèves, qui n'ont pas révisé, envoient un élève d'une autre classe à leur place pour faire le devoir surveillé. (On suppose que les k "remplaçants" ne sont pas des homonymes des k élèves qui se sont fait remplacés).

Il y aura toujours n élèves dans la salle mais k d'entre eux n'ont pas leur véritable nom inscrit sur le plan de la salle (et pour cause!).

On reprend le problème précédent où les n élèves s'installent au hasard sans regarder le plan de classe. Les k "remplaçants" sont forcément à une place qui n'est pas à leur véritable nom.

On note $p_{n,k}$ la probabilité qu'aucun élève ne se soit assis à la place correspondant à son véritable nom lorsqu'il y a k "remplaçants" parmi les n élèves.

Calculer $p_{n,k}$ par récurrence.

Remarque : $p_{n,0}$ correspond au problème original.

Limite

Calculer $p_{n,0}$ pour n grand et conjecturer la limite : $\lim_{n \rightarrow \infty} p_{n,0}$

Statistiques

F. Kany. ISEN-Brest & La Croix-Rouge

Exercices

1 Moyenne glissante

On donne une liste de N valeurs. On souhaite lisser la courbe en remplaçant chaque valeur par la moyenne de ses k valeurs adjacentes.

Pour k impair, on prendra la valeur centrale et les $k//2$ valeurs précédentes et suivantes.

Pour k pair, on prendra la valeur centrale; les $k//2$ valeurs précédentes et suivantes et la moitié des valeurs extrêmes.

Application : lisser la liste `liste=[5.3,5.5,4.95,5.13,5.37,4.7,5,5.37,5.3,5.2,4.8,4.62,4.5,5.05,4.45,4.85,4.4,5.25,4.9,4.26,4.2,4.14,4.55,4.05,4.12,4.3,4.05,4.06,3.75,3.8]` sur le $k = 7$ valeurs. Refaire la même opération avec $k = 8$ valeurs. Représenter la liste et les listes lissées sur le même graphique.

2 Sondages

Source : http://fr.wikiversity.org/wiki/Statistique_inférentielle/Intervalle_de_confiance_d'une_fréquence

2.1 La théorie de l'échantillonnage

En statistique, il est en général impossible d'étudier un caractère sur toute une population de taille N élevée.

La théorie de l'échantillonnage se pose la question suivante : "En supposant connus les paramètres statistiques de la population, que peut-on en déduire sur les échantillons prélevés dans la population ?"

On suppose que ces échantillons sont prélevés au hasard et que le tirage de ces échantillons est effectué avec remise.

L'ensemble de ces échantillons de taille n est appelé échantillonnage de taille n .

Étudions dans ces conditions la loi d'échantillonnage des fréquences.

2.2 Loi d'échantillonnage des fréquences

On suppose donc sur une population de taille N , un caractère de fréquence p .

Soit X la variable aléatoire valant 1 si le caractère est acquis, 0 sinon.

X suit donc une loi de Bernoulli de paramètre p , d'espérance p et de variance $p.(1-p)$.

Dans un échantillon de taille n , on répète n de ces épreuves indépendantes auxquelles correspondent n variables aléatoires : X_1, X_2, \dots, X_n de même loi que X .

La variable aléatoire représentant la moyenne de l'échantillon est : $Y_n = \frac{X_1+X_2+\dots+X_n}{n}$

Définition La loi d'échantillonnage de la fréquence est la loi de probabilité de Y_n

Elle dépend bien sûr de la taille n des échantillons.

D'après le théorème de la limite centrée¹, on déduit :

Propriété La loi d'échantillonnage de Y_n suit une loi normale d'espérance p et d'écart-type $\sqrt{\frac{p(1-p)}{n}}$.

1. voir Annexe

2.3 Intervalle de confiance de la fréquence

L'estimation ponctuelle de la fréquence dans la population à partir de celle dans l'échantillon n'indique pas le risque d'erreur.

Il s'agit de déterminer un intervalle contenant la valeur de la fréquence dans la population avec un risque d'erreur décidé à l'avance.

p et $\sqrt{\frac{p(1-p)}{n}}$ étant inconnus, on les remplace par leurs estimations ponctuelles : f et $\sqrt{\frac{f(1-f)}{n-1}}$

En posant $T_n = \frac{Y_n - f}{\sqrt{\frac{f(1-f)}{n-1}}}$, le théorème précédent implique que T_n suit une loi normale centrée réduite.

Soit α la probabilité, fixée à l'avance, que T_n n'appartiennent pas à l'intervalle $[-t, t]$, alors : $P(-t \leq T_n \leq t) = 1 - \alpha$ donc $P\left(Y_n - t\sqrt{\frac{f(1-f)}{n-1}} \leq p \leq Y_n + t\sqrt{\frac{f(1-f)}{n-1}}\right) = 1 - \alpha$ on obtient donc le :

Théorème Un intervalle de confiance de la moyenne m au seuil de risque α est : $\left[f - t\sqrt{\frac{f(1-f)}{n-1}}, f + t\sqrt{\frac{f(1-f)}{n-1}}\right]$ où t est le nombre tel que $\Pi(t) = 1 - \frac{\alpha}{2}$ et se lit dans la table de la loi normale $\mathcal{N}(0; 1)$.

Définition

- α est le risque d'erreur ou seuil de risque.
- $1 - \alpha$ est le coefficient de confiance.

Application : Tirer au hasard n valeurs X dans l'intervalle $[0, 1]$ (à l'aide de `random.random()`). Si X est inférieur à une certaine probabilité p , on incrémentera un compteur S_n . On calcule ensuite la moyenne de S_n : $Y_n = \frac{S_n}{n}$. Répéter l'opération N fois et représenter graphiquement l'ensemble des points $\{Y_n\}$ et les droites horizontales d'ordonnées $p - \frac{1}{\sqrt{n}}$ et $p + \frac{1}{\sqrt{n}}$. Calculer la probabilité que Y_n soit dans l'intervalle $[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}}]$

Annexe : Théorème de la limite centrée

Source : http://fr.wikiversity.org/wiki/Théorème_de_la_limite_centrée/Enoncé_simplifié
Variables aléatoires indépendantes

Définition Deux variables aléatoires sont dites indépendantes quand le résultat de l'une n'influence pas celui de l'autre.

Énoncé simplifié

Théorème Si X_1, X_2, \dots, X_n sont des variables indépendantes de même loi de probabilité, de même espérance m et de même variance σ^2 alors lorsque n est suffisamment grand :

- La variable aléatoire : $S_n = X_1 + X_2 + \dots + X_n$ suit approximativement une loi normale d'espérance $m \times n$ et d'écart-type $\sigma\sqrt{n}$, notée : $\mathcal{N}(m.n; \sigma.\sqrt{n})$.
- La variable aléatoire : $Y_n = \frac{S_n}{n}$ converge en loi vers une variable aléatoire Y de loi normale : $\mathcal{N}(m; \frac{\sigma}{\sqrt{n}})$

Sondages

F. Kany, ISEN-Brest & La Croix-Rouge

Énoncé

Le jour d'une élection un candidat A est élu avec 51% des voix devant un candidat B (49 % des voix).

Pour pouvoir donner une estimation des résultats dès la clôture des bureaux de vote, un institut de sondage questionne N électeurs à la sortie de l'isoloir¹.

1. Effectuer une simulation de 1000 sondages sur un échantillon de $N = 1000$ électeurs. Quelle est la probabilité que l'institut de sondage se trompe (i.e. prévoit la victoire de B et non de A) ?
2. Parmi ces 1 000 simulations de sondage, quels sont les pourcentages maximum et minimum obtenus par A ?
3. En réalité, si l'élection avait eu lieu une semaine avant, le candidat A aurait eu 49 % des voix (et le candidat B 51 % des voix). Effectuer un sondage sur les électeurs de la semaine 1 ($A=49\%$) suivi d'un sondage sur les électeurs de la semaine 2 ($A=51\%$). Ré-itérer cette simulation 1000 fois. Combien de sondages prévoient correctement le changement de tendance (i.e. A perdant la semaine 1 mais gagnant la semaine 2) ?
4. Reprendre les mêmes questions avec $N = 3000$. Qu'observe-t-on ?

1. N.B. On supposera que les électeurs questionnés par l'institut de sondage répondent sincèrement (i.e. indiquent pour qui ils ont réellement voté).

Nombre de participants

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soit un événement (sportif par exemple) avec N participants : chaque participant a un numéro de carte d'invitation (ou un numéro de dossard). On prend un échantillon de $n < N$ personnes et on relève leur numéro x_i (avec $i \in [1, n]$). On cherche à estimer N à partir des valeurs des x_i .

On peut montrer que l'espérance du maximum d'un échantillon est : $E(\max x_i) = \frac{n \cdot (N+1)}{n+1}$ d'où :

$$N = \frac{n+1}{n} \cdot E(\max x_i) - 1.$$

Cette formule a le bon goût de vérifier deux propriétés essentielles :

- si on tire comme échantillon $[1, 2, 3, \dots, n]$, on obtient $N = \frac{n+1}{n} \cdot n - 1 = n$: l'estimation de N n'est pas inférieure à la plus grande valeur de l'échantillon (toutes les formules d'estimation de l'espérance ne vérifie pas cette propriété triviale !)
- si on tire comme échantillon tout l'ensemble $[1, 2, 3, \dots, N]$, on obtient $N = \frac{N+1}{N} \cdot N - 1 = N$.

Faire plusieurs simulations numériques en prélevant des échantillons de 2%, 5%, 10% et 20% de la population. Présenter les résultats sous forme d'histogrammes pour montrer la probabilité de l'écart (en %) entre la taille réelle de la population et son estimation.

Pour prélever (sans remplacement) un échantillon de taille n dans une population de taille N , on utilise la procédure suivante :

- On considère le premier élément de la population
- On le prélève avec une probabilité de $\frac{n}{N}$ pour le mettre dans l'échantillon
- Si l'élément est effectivement sélectionné, alors on diminue n de 1.
- On diminue N de 1
- On recommence avec l'élément suivant de la population jusqu'à ce que la taille de l'échantillon soit n .

Plus proche voisin

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

On place aléatoirement n points sur une droite orientée.

Ces points P_i sont classés par abscisse croissante sur la droite.

Soit P_j le plus proche voisin du point P_i (P_j peut être le point P_{i-1} ou le point P_{i+1}). Quelle est la probabilité que le plus proche voisin de P_j soit P_i ?

Même question pour des points sur un plan (on prendra la norme euclidienne).

Même question pour des points dans un espace à 3 dimensions.

Triangle obtus

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soit un rectangle de dimensions $X = 1$ et Y . On tire au hasard les coordonnées de trois points $A(x_A, y_A)$, $B(x_B, y_B)$ et $C(x_C, y_C)$ dans ce rectangle. Quelle est la probabilité que le triangle ABC soit obtus (un des angles $> 90^\circ$)?

Quadrilatère dans un disque

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Soit un disque de rayon $R = 1$. On tire au sort quatre points A, B, C et D dans le disque. Quelle est la probabilité que le quadrilatère $ABCD$ soit convexe ?

Comment tirer au sort les points ?

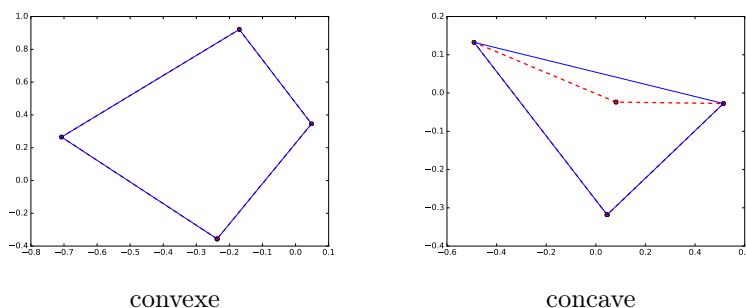
On souhaite tirer au sort les quatre points A, B, C et D en s'assurant qu'ils sont bien à l'intérieur du disque. On compare trois méthodes.

1. On tire au sort $x \in [-1, 1]$ et $y \in [-1, 1]$. On conserve le couple (x, y) si $x^2 + y^2 \leq 1$. Représenter sur un graphique 1000 points tirés par cette méthode. Quel est l'inconvénient de cette méthode ?
2. On tire au sort $r \in [0, 1]$ et $\theta \in [0, 2\pi]$. On place le point de coordonnées polaires (r, θ) . Représenter sur un graphique 1000 points tirés par cette méthode. Quel est l'inconvénient de cette méthode ?
3. On tire au sort $r \in [0, 1]$ et $\theta \in [0, 2\pi]$. On place le point de coordonnées polaires (\sqrt{r}, θ) . Représenter sur un graphique 1000 points tirés par cette méthode. Pourquoi faut-il privilégier cette méthode ?

Convexité d'un polygone

Tirer au sort quatre points.

Utiliser la fonction `scipy.spatial.ConvexHull` qui calcule les points qui forment l'enveloppe convexe correspondant à un polygone donné. (Analogie : fil le plus court délimitant l'ensemble des points du quadrilatère ; voir le tracé en ligne bleue). Si `hull = ConvexHull(points_xy)` où `points_xy` est une liste à 2 dimensions, alors : `len(hull.vertices)` donne le nombre de points formant l'enveloppe convexe. Le quadrilatère est convexe si ce nombre vaut 4 ; concave lorsqu'il vaut 3.



Quelle est la probabilité d'un quadrilatère concave ?

Aiguille de Buffon

F. Kany. ISEN-Brest. La Croix-Rouge.

Présentation

Consulter l'article de wikipedia : https://fr.wikipedia.org/wiki/Aiguille_de_Buffon.

Ecrire une fonction `Buffon(1,a)` qui place au hasard une aiguille de longueur a sur un parquet composé de lattes de largeur ℓ et qui renvoie 1 (resp. 0) si l'aiguille tombe (resp. ne tombe pas) à cheval sur au moins une rainure du parquet.

On fixe $\ell = 1$, pour $a \in [0.1, 2]$, tracer la probabilité que l'aiguille tombe à cheval sur au moins une rainure en appellant la fonction `Buffon(1,a)` 100.000 fois pour chaque valeur de a .

Sur le même graphique, représenter les probabilités théoriques :

$$\begin{cases} P = \frac{2a}{\pi\ell} & \text{si } \ell \leq a \\ P = \frac{2a}{\pi\ell} \left(1 - \sqrt{1 - \frac{\ell^2}{a^2}}\right) + \left(1 - \frac{2}{\pi} \cdot \arcsin \frac{\ell}{a}\right) & \text{si } \ell \geq a \end{cases}$$

Franc carreau

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Dans l'article Carreau de l'Encyclopédie de Diderot et d'Alembert, on peut lire :

Franc-Carreau, sorte de jeu dont M. de Buffon a donné le calcul en 1733, avant que d'être de l'Académie des Sciences. Voici l'extrait qu'on trouve de son mémoire sur ce sujet, dans le volume de l'Académie pour cette année-là.

Dans une chambre carrelée de carreaux égaux, & supposés réguliers, on jette en l'air un louis ou un écu, & on demande combien il y a à parier que la pièce ne tombera que sur un seul carreau, ou franchement.

On considère un carrelage de carré de côtés 1 et un disque de rayon $r < \frac{1}{2}$.

Calculer la probabilité que le disque ne coupe aucun bord du carrelage carré.

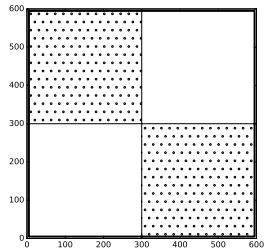
Effectuer 100 000 simulations.

Diamant aztèque

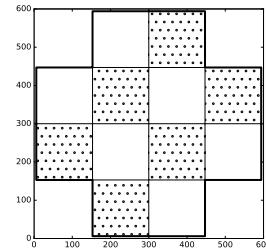
F. Kany. ISEN-Brest

Énoncé

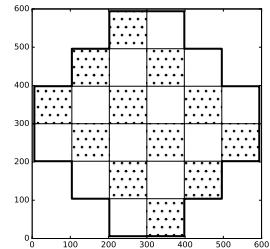
On considère un damier en forme de “diamant aztèque” de taille n (voir figure ci-dessous).



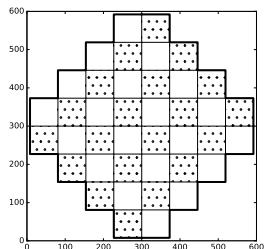
$n = 1$



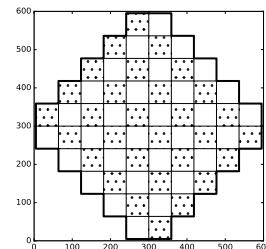
$n = 2$



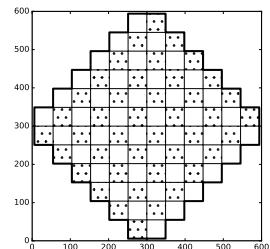
$n = 3$



$n = 4$



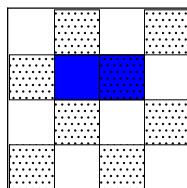
$n = 5$



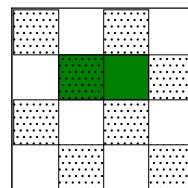
$n = 6$

On désire pavier ce damier avec des rectangles de dimensions 2×1 carreaux.

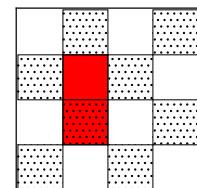
On colorie les rectangles suivant leurs positions sur le damier.



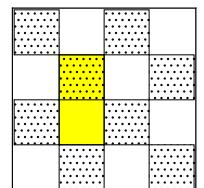
rectangle horizontal
bord droit sur case :
- sombre pour n pair
- claire pour n impair



rectangle horizontal
bord droit sur case :
- claire pour n pair
- sombre pour n impair

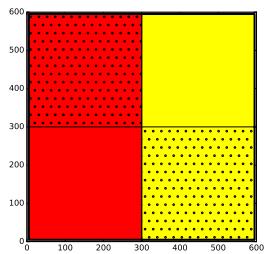


rectangle vertical
bord bas sur case :
- sombre pour n pair
- claire pour n impair

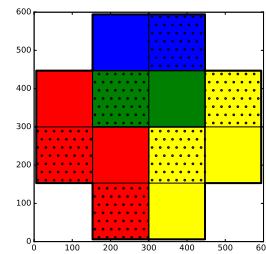


rectangle vertical
bord bas sur case :
- claire pour n pair
- sombre pour n impair

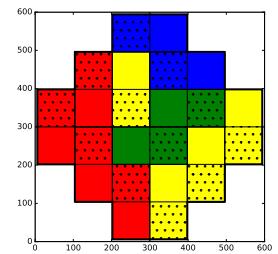
Exemples de pavage :



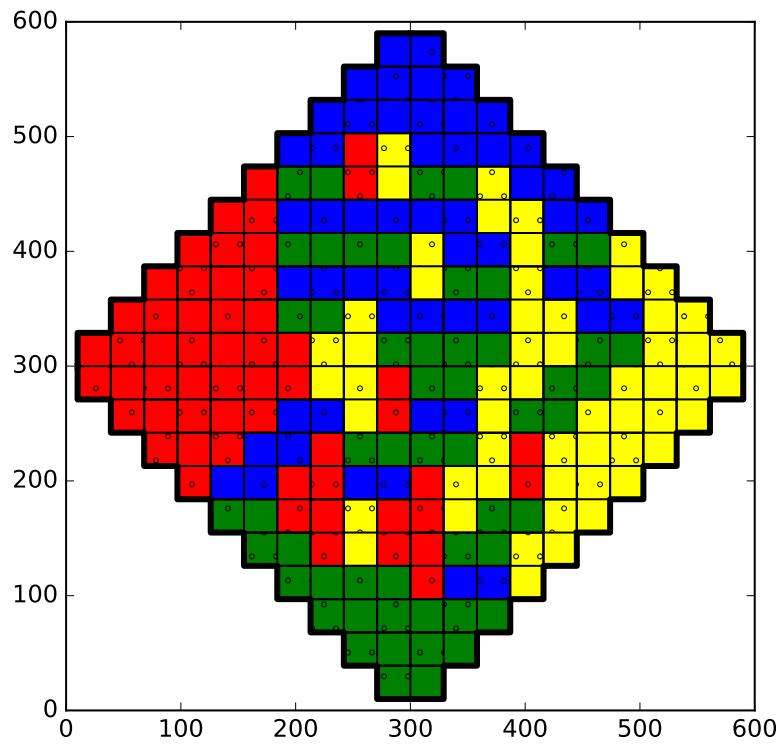
$n = 1$



$n = 2$



$n = 3$

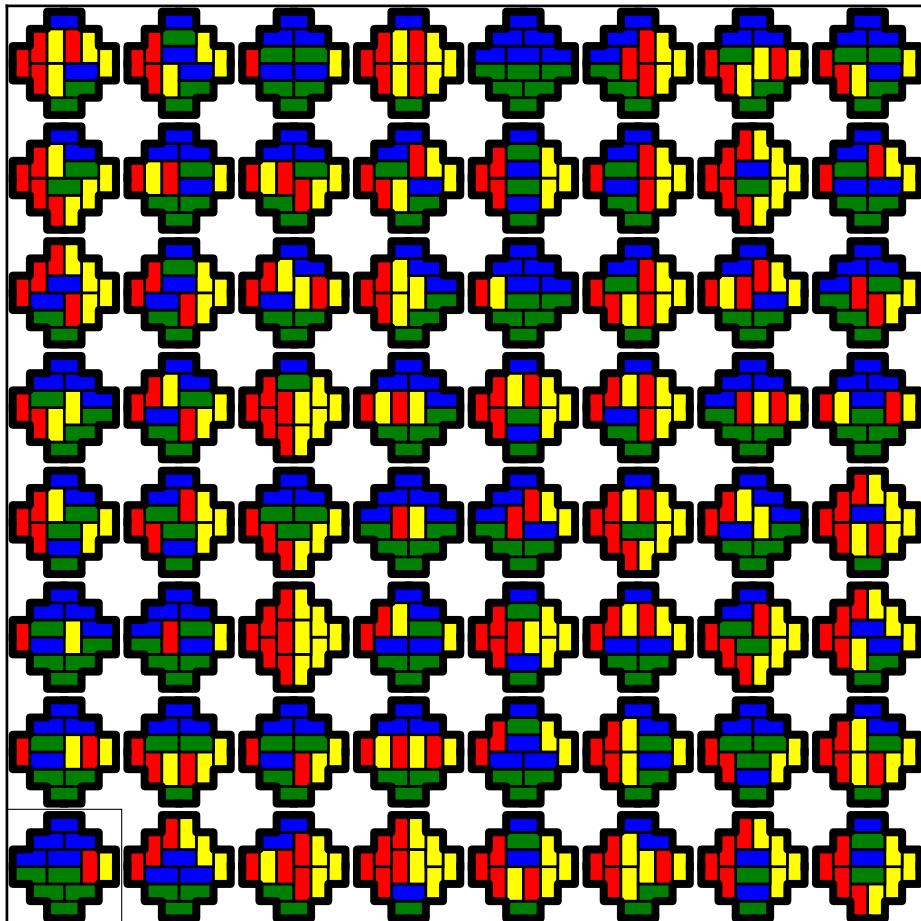


$n = 10$

Programmer l'algorithme de pavage d'un diamant aztèque (voir Bibliographie).
Que se passe-t-il lorsque $n = 1000$?

Tirer au hasard un grand nombre de pavages pour une dimension n donnée (il en existe au maximum $2^{n \cdot (n+1)/2}$) et estimer la probabilité $P(n)$ que la case “en haut à gauche” (i.e. avec la plus grande ordonnée) soit bleue.

Par exemple, pour $n = 3$, on obtient les $2^{3 \times 4/2} = 2^6 = 64$ pavages suivants. On observe que la case du haut est bleue avec une probabilité de $P(3) = \frac{56}{64} = \frac{7}{8}$



Pour $n = 4$, cette probabilité est $P(4) = \frac{960}{1024}$.
Tracer $P(n)$ pour $n \in [1, 9]$.

Bibliographie

Mickaël Launay. Chaine YouTube MicMaths
<https://www.youtube.com/watch?v=2Wq6H8GMVm0>

Mickaël Launay. Site MicMaths.
<http://micmaths.com/applications/lgn/diamantazteque.html>
 Elise Janvresse. Pavages aléatoires par touillage de dominos.
<http://images.math.cnrs.fr/pavages-aleatoires-par-touillage.html>

TD : Calcul de Π

F. Kany. ISEN-Brest & La Croix-Rouge

1 Position du problème

On se propose d'étudier différentes méthodes pour calculer les décimales du nombre π .

1.1 Convergence d'une suite

1.1.1 Méthode d'Archimète

Le premier calcul mathématique de π remonte à Archimète de Syracuse (287-212 avant J.-C.). Celui-ci reposait sur un encadrement du périmètre du cercle par ceux de polygones réguliers inscrit et circonscrits. Soit p_n le périmètre d'un polygone régulier à n côtés inscrit dans un cercle de diamètre unité et p'_n celui d'un polygone régulier à n côtés circonscrit au même cercle. En utilisant l'inégalité $p_n < \pi < p'_n$ pour $n = 6 \times 2^k$, on peut obtenir une approximation de π avec :

$$p_n = n.u_n \quad \left(\text{où } u_n = \sin \frac{\pi}{n}, \quad u_6 = \frac{1}{2} \text{ et } u_{2n} = \sqrt{\frac{1-\sqrt{1-u_n^2}}{2}} \right)$$

et $p'_n = n.u'_n \quad \left(\text{où } u'_n = \tan \frac{\pi}{n}, \quad u'_6 = \frac{1}{\sqrt{3}} \text{ et } u'_{2n} = \frac{u_n}{\sqrt{1-u_n^2}} \right)$.

Programmer le calcul de la suite u_n par une méthode itérative et par une méthode récursive. Pour évaluer la vitesse de convergence de cette suite, tracer $\log(\pi - p_n)$ en fonction de n .

1.1.2 Méthode de Monte-Carlo

Tirer au sort un couple (x_i, y_i) de nombres compris entre 0 et 1. Évaluer $r_i = \sqrt{x_i^2 + y_i^2}$. Si $r_i \leq 1$, alors incrémenter un compteur c . Réitérer le calcul pour un très grand nombre n de couples $\{(x_i, y_i)\}$. Tracer $u_n = 4 \times c/n$ en fonction de n .

1.2 Convergence d'une série

1.2.1 Série Arctan

D'après la formule de J. Gregory (1638-1675) : $\text{Arctan } x = \sum_{k=0}^{+\infty} \frac{(-1)^k \cdot x^{2k+1}}{2k+1}$. Lorsque l'on arrête la série au rang n , l'erreur commise est inférieure ou de l'ordre de $\frac{|x|^{2n+3}}{2n+3}$.

On en déduit la formule de Leibniz : $\frac{\pi}{4} = \text{Arctan } 1 = \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k+1}$.

Programmer le calcul de la série $S_n = 4 \times \sum_{k=0}^n \frac{(-1)^k}{2k+1}$.

Pour évaluer la vitesse de convergence de cette suite, tracer $\log(|\pi - S_n|)$ en fonction de n .

Cette série convergeant assez lentement (car $x = 1$), on peut - pour accélérer la convergence - utiliser des combinaisons de fonctions Arctan x avec $x \ll 1$. Programmer une série utilisant la formule de Gauss : $\pi = 48 \cdot \text{Arctan } \frac{1}{18} + 32 \cdot \text{Arctan } \frac{1}{57} - 20 \cdot \text{Arctan } \frac{1}{239}$

1.2.2 Série de Ramanujan

Au début du vingtième siècle, S. Ramanujan, mathématicien autodidacte indien, proposa la série :

$$\frac{1}{\pi} = \frac{2\sqrt{2}}{9801} \times \sum_{k=0}^{+\infty} \frac{(4k)!.(1103 + 26390.k)}{(n!)^4 \cdot 3934^n}$$

$$\text{Programmer le calcul de la série } S'_n = \frac{9801}{2\sqrt{2}} \times \sum_{k=0}^n \frac{1}{(4k)!.(1103 + 26390.k)}.$$

Cette série converge très vite : vérifier qu'au bout de 2 termes, on a déjà π à 8 décimales. Mais les termes supplémentaires ne permettent de gagner qu'une précision relativement faible. Évaluer l'évolution de la précision $P_n = \pi - S'_n$ en traçant $\log(|P_{n+1} - P_n|)$ en fonction de n .

Volume d'une boule de dimension n

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation de la méthode de Monte-Carlo

D'après Wikipédia

La méthode Monte-Carlo, désigne une famille de méthodes algorithmiques visant à calculer une valeur numérique approchée en utilisant des procédés aléatoires, c'est-à-dire des techniques probabilistes. Le nom de ces méthodes, qui fait allusion aux jeux de hasard pratiqués à Monte-Carlo, a été inventé en 1947 par Nicholas Metropolis, et publié pour la première fois en 1949 dans un article coécrit avec Stanislaw Ulam.

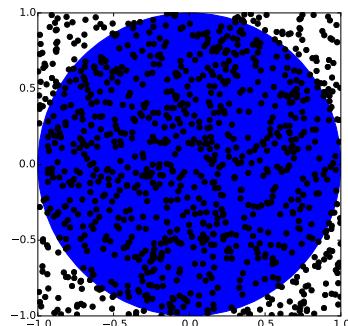
Les méthodes de Monte-Carlo sont particulièrement utilisées pour calculer des intégrales en dimensions plus grandes que 1 (en particulier, pour calculer des surfaces et des volumes).

Application

La méthode de Monte-Carlo peut être utilisée pour calculer la surface d'un disque de rayon 1.

Soit un point M de coordonnées (x, y) , où $-1 < x < 1$ et $-1 < y < 1$. On tire aléatoirement les valeurs de x et y . Le point M appartient au disque de centre $(0,0)$ de rayon 1 si et seulement si $x^2 + y^2 \leq 1$. La probabilité que le point M appartienne au disque est π .

En faisant le rapport du nombre de points dans le disque au nombre de tirages, on obtient une approximation du nombre π si le nombre de tirages est grand : $\frac{\text{Surface(disque)}}{\text{Surface(carré)}} = \frac{\text{Nombre de points dans le disque}}{\text{Nombre de points tirés au sort}}$ avec $\text{Surface(carré)} = (1 - (-1))^2 = 4$



Travail à faire

Appliquer la méthode de Monte-Carlo pour calculer le rapport du volume entre la sphère et le cube ; puis entre la sphère de dimension 4 et le cube de dimension 4 ; et ainsi de suite jusqu'à la dimension 10.

Indiquer pour quelle dimension le rapport des volumes est maximum ; donner une estimation de ce rapport.

On pourra utiliser `random.uniform(-1,1)` pour tirer au sort un nombre entre -1 et 1 .

On utilisera au moins 100.000 points pour déterminer le rapport.

Intégrale multiple

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

Si l'on veut calculer certaines propriétés de l'atome de magnésium (12 électrons), on est amené à intégrer des fonctions par rapport aux 3 coordonnées de chaque électron. On doit donc réaliser des intégrales à $3 \times 12 = 36$ dimensions. Si l'on utilise 64 points pour calculer numériquement chaque intégrale, il faudra réaliser $64^{36} \simeq 10^{65}$ évaluations de la fonction à intégrer. Même avec un ordinateur rapide (10^6 opérations.s $^{-1}$), il faudrait 10^{59} s pour faire cette intégrale (c'est-à-dire beaucoup plus que l'âge de l'Univers $\simeq 10^{17}$ s).

Une méthode plus rapide (et plus précise !) consiste à tirer au sort N valeurs de la fonction f à intégrer pour calculer sa valeur moyenne $\langle f \rangle$ et à calculer : $I = \int_{x_1_{min}}^{x_1_{max}} \int_{x_2_{min}}^{x_2_{max}} \dots f(x_1, x_2, \dots) \cdot x_1 \cdot x_2 \dots$ sous la forme approchée : $I_{approx} = (x_{1_{max}} - x_{1_{min}}) \cdot (x_{2_{max}} - x_{2_{min}}) \dots \langle f \rangle$.

Appliquer cette méthode pour calculer l'intégrale à 10 dimensions suivante :

$$I = \int_{x_1=0}^1 \int_{x_2=0}^1 \dots \int_{x_{10}=0}^1 (x_1 + x_2 + \dots + x_{10})^2 \cdot x_1 \cdot x_2 \dots \cdot x_{10}$$

Pour cela :

1. tirer au sort les valeurs de x_1 à x_{10} (dans l'intervalle d'intégration),
2. calculer $f(x_1, x_2, \dots, x_{10})$ et en déduire I_{approx} ,
3. réitérer N fois ($N = 2, 4, 8, \dots, 8192$) les étapes 1 et 2
calculer $\langle I_{approx_N} \rangle$: la moyenne des N évaluations de I_{approx} .
4. tracer $\langle I_{approx_N} \rangle = f(N)$ et montrer que la précision du calcul est proportionnelle à $\frac{1}{\sqrt{N}}$ en traçant $|I - \langle I_{approx_N} \rangle| = f(\frac{1}{\sqrt{N}})$.

Marche au hasard

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

On se propose de simuler la marche au hasard d'une particule dans un espace à une ou plusieurs dimensions. Ce phénomène correspond en physique au mouvement Brownien ; il s'applique au mouvement d'un électron dans un fil métallique, à l'adsorption d'une particule sur une surface, à la diffusion d'un parfum dans un volume,... On suppose que la particule M part de l'origine O et, qu'à chaque pas, elle se déplace d'une petite quantité par rapport à sa position précédente. On veut trouver, au bout de N pas, la position finale $\overrightarrow{OM_{(N)}}$ de la particule et sa distance $d_{(N)}$ à l'origine. On refait le calcul K fois. On souhaite vérifier que, si toutes les directions sont équiprobables, la moyenne des $d_{(N)}$ est nulle et que l'écart-type dépend de \sqrt{N} .

Déplacement sur une droite

On suppose que la particule se déplace, à chaque itération, de Δx sur un axe (Ox).

1. À chaque pas, la particule effectue un déplacement entier $\Delta x = \pm 1$. (On a donc : $\langle (\Delta x)^2 \rangle = 1$).
 - (a) Se fixer une valeur de N (par exemple $N = 100$).
 - (b) Calculer $OM_{(N)}$ au bout de N pas ($OM_{(N)} \in \mathbb{Z}$) ; incrémenter un compteur $c_{(N)}[OM_{(N)}]$ pour comptabiliser le nombre de fois où la position $OM_{(N)}$ a été atteinte.
 - (c) Réitérer K fois l'opération b. (K de l'ordre de \sqrt{N}).
 - (d) Tracer un graphique indiquant, pour chaque position $x \in \mathbb{Z}$, le nombre de fois $c_{(N)}(x)$ où cette position a été atteinte lors des K essais.
 - (e) Calculer la moyenne et l'écart-type de $c_{(N)}(x)$.
 - (f) Reprendre tout le calcul pour différentes valeurs de N et tracer l'écart-type de $c_{(N)}(x)$ en fonction de \sqrt{N} .
2. À chaque pas, la particule effectue un déplacement dans \mathbb{R} tel que $\Delta x \in [-\sqrt{3}, +\sqrt{3}]$. (On a donc : $\langle (\Delta x)^2 \rangle = 1$ car $\langle (\Delta x)^2 \rangle = \int_{-x_{max}}^{+x_{max}} x^2 \cdot p$ avec $p = \frac{x}{2 \cdot x_{max}}$).
 - (a) Se fixer une valeur de N .
 - (b) Calculer $OM_{(N)}$ au bout de N pas ($OM_{(N)} \in \mathbb{R}$).
 - (c) Réitérer K fois le calcul et calculer σ^2 : la moyenne de $(OM_{(N)})^2$ (i.e. la variance de la distance à l'origine).
 - (d) Reprendre tout le calcul pour différentes valeurs de N et tracer σ (i.e. l'écart-type de la distance à l'origine) en fonction de \sqrt{N} .

Déplacement dans un plan

On suppose que la particule se déplace, à chaque itération, de $(\Delta x, \Delta y)$ dans un plan (Oxy).

En théorie, on a : $d_{(N)}^2 = (\Delta x_1 + \Delta x_2 + \dots + \Delta x_N)^2 + (\Delta y_1 + \Delta y_2 + \dots + \Delta y_N)^2$.

D'où : $d_{(N)}^2 = \Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_N^2 + 2 \cdot \Delta x_1 \cdot \Delta x_2 + 2 \cdot \Delta x_1 \cdot \Delta x_3 + \dots + 2 \cdot \Delta x_2 \cdot \Delta x_3 + \dots + (x \rightarrow y)$.

Pour un déplacement équiprobable dans toute les directions, les termes croisés s'annulent en moyenne ; il reste : $d_{(N)}^2 \simeq \Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_N^2 + \Delta y_1^2 + \Delta y_2^2 + \dots + \Delta y_N^2 = N \cdot (\langle \Delta x^2 \rangle + \langle \Delta y^2 \rangle) = N \cdot \langle d^2 \rangle$.

Finalement : $d_{(N)} \simeq \sqrt{N} \cdot d_{rms}$.

1. À chaque pas, la particule effectue un déplacement entier vers le haut ($\Delta y = +1$), le bas ($\Delta y = -1$), la droite ($\Delta x = +1$) ou la gauche ($\Delta x = -1$). Cela revient à choisir une direction parmi 4 avec la probabilité $1/4$ et à avoir : $\langle (\Delta x)^2 \rangle = \langle (\Delta y)^2 \rangle = 1/2$.
 - (a) Se fixer une valeur de N .
 - (b) Tracer l'évolution de $\overrightarrow{OM_{(N)}}$ en fonction de N .
 - (c) Réitérer K fois le calcul et calculer σ^2 : la moyenne de $(OM_{(N)})^2$.
 - (d) Reprendre tout le calcul pour différentes valeurs de N et tracer σ en fonction de \sqrt{N} .
2. À chaque pas, la particule effectue un déplacement $\Delta x \in [-\sqrt{3/2}, +\sqrt{3/2}]$ et $\Delta y \in [-\sqrt{3/2}, +\sqrt{3/2}]$. (Ainsi : $\langle (\Delta x)^2 \rangle = \langle (\Delta y)^2 \rangle = 1/2$). Reprendre les mêmes questions.

Calcul d'erreur

F. Kany. ISEN-Brest & La Croix-Rouge

1 Problème de la marche au hasard. Loi de Gauss

1.1 Position du problème

Exemple :

On mesure 100 m avec 1 mètre que l'on reporte 100 fois avec, à chaque fois, une incertitude de 1 cm.
Quelle est l'erreur totale commise ? $100 \times 1 \text{ cm} = 1 \text{ m}$?

FAUX. Il est peu probable qu'à chaque fois, on fasse l'erreur de 1 cm dans le même sens.

Probabilité $\frac{1}{2}$ de déplacer le mètre de 1 cm vers la droite.

Probabilité $\frac{1}{2}$ de déplacer le mètre de 1 cm vers la gauche.

Probabilité de déplacer 100 fois de suite le mètre dans le même sens : $(\frac{1}{2})^{100} = 7,9 \cdot 10^{-31}$.

La probabilité d'avoir une incertitude de 1 m est ridicule. Il serait exagéré d'écrire : il est probable que $X \in [99 \text{ m} ; 101 \text{ m}]$.

1.2 Marche au hasard

Généralisation du problème : la marche au hasard.

Soit p la probabilité de se déplacer vers la droite.

Soit $q = 1 - p$ la probabilité de se déplacer vers la gauche.

Au bout de N itérations, quelle est la probabilité P de faire au total : n_1 pas à droite et n_2 pas à gauche ?

Probabilité de faire n_1 pas à droite : p^{n_1} ; probabilité de faire n_2 pas à gauche : q^{n_2} ; nombre de façons de réaliser ce déplacement : $C_N^{n_1} = C_N^{n_2} = \frac{N!}{n_1! \cdot n_2!}$

D'où : $P = p^{n_1} \cdot q^{n_2} \cdot \frac{N!}{n_1! \cdot n_2!}$

Soit m : le déplacement total vers la droite : $m = n_1 - n_2$.

On a : $N = n_1 + n_2$. D'où : $n_1 = \frac{N+m}{2}$ et $n_2 = \frac{N-m}{2}$

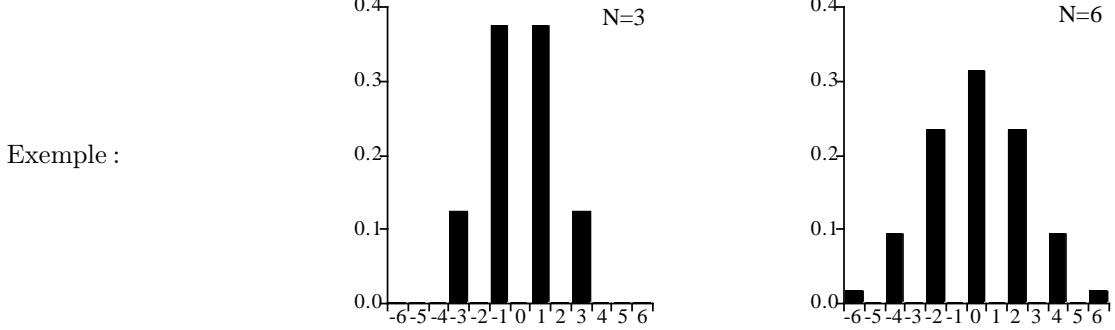
$$P_N(m) = p^{\frac{N+m}{2}} \cdot q^{\frac{N-m}{2}} \cdot \frac{N!}{(\frac{N+m}{2})! \cdot (\frac{N-m}{2})!}$$

C'est la loi binomiale (cf. développement de $(p+q)^N = \sum_m P_N(m)$).

Calcul des C_j^i .

$$\begin{array}{ccccccccc} 1 & & 1 \\ 1 & 2 & & 1 \\ 1 & 3 & 3 & & 1 \\ 1 & 4 & 6 & 4 & & 1 \\ 1 & 5 & 10 & 10 & 5 & & 1 \\ 1 & 6 & 15 & 20 & 15 & 6 & & 1 \end{array}$$

Cas particulier $p = q = 1/2$. $P_N(m) = (\frac{1}{2})^N \cdot \frac{N!}{(\frac{N+m}{2})! \cdot (\frac{N-m}{2})!}$



Exemple :

1.3 Cas où $N \rightarrow +\infty$

Lorsque N est très grand, on montre¹ que la probabilité de se trouver à la position m est : $P_N(m) \approx \sqrt{\frac{2}{\pi \cdot N}} \cdot e^{-\frac{m^2}{2 \cdot N}}$

- Probabilité de se trouver entre m et $m + \Delta m$: $\Delta P = \frac{1}{2}P_N(m) \cdot \Delta m$
(car $P_N(m)$ est à peu près constante sur $[m, m + \Delta m]$; le facteur $\frac{1}{2}$ vient du fait que $P_N(m)$ n'est défini que pour une valeur de m sur deux : les m pairs (resp. impairs) si N est pair (resp. impair) ; il faut donc éliminer une valeur sur deux par rapport à une fonction $P_N(m)$ qui serait définie pour tout m .)
- Nombre d'événements ΔN tel que l'on se trouve entre m et $m + \Delta m$: $\Delta N = N \cdot P_N(m)$
- Écart-type : $\sigma^2 = \frac{1}{N} \cdot \sum ((m - \bar{m})^2 \cdot \Delta N) = \sum m^2 \cdot \Delta P \approx \sqrt{\frac{2}{\pi \cdot N}} \cdot \sum m^2 \cdot e^{-\frac{m^2}{2 \cdot N}} \cdot \frac{\Delta m}{2}$

Si le pas Δm est petit, alors, par passage à la limite, on a : $\sigma^2 = \sqrt{\frac{1}{2 \cdot \pi \cdot N}} \cdot \int_{-\infty}^{+\infty} x^2 \cdot e^{-\frac{x^2}{2 \cdot N}} \cdot dx \cdot \frac{2 \cdot N}{2 \cdot N} \cdot \frac{\sqrt{2 \cdot N}}{\sqrt{2 \cdot N}}$

$$\sigma^2 = \frac{2 \cdot N}{\sqrt{\pi}} \cdot \int_{-\infty}^{+\infty} u^2 \cdot e^{-u^2} \cdot du \text{ avec } u = \frac{x}{\sqrt{2 \cdot N}}$$

Intégration par parties : $f = u/2$ et $g' = 2 \cdot u \cdot e^{-u^2}$

$$\Rightarrow f' = 1/2 \text{ et } g = -e^{-u^2}$$

$$\int_{-\infty}^{+\infty} u^2 \cdot e^{-u^2} \cdot du = \left[\frac{u}{2} \cdot e^{-u^2} \right]_{-\infty}^{+\infty} + \frac{1}{2} \cdot \int e^{-u^2} \cdot du$$

$$= \frac{1}{2} \cdot \int e^{-u^2} \cdot du$$

- Calcul de $I = \int_{-\infty}^{+\infty} e^{-x^2} \cdot dx$.

Astuce : poser $J = \int_{-\infty}^{+\infty} e^{-y^2} \cdot dy$

et calculer $I \cdot J = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} \cdot dx \cdot dy$.

On passe en polaire : $I \cdot J = \int_{r=0}^{r=+\infty} \int_{\theta=0}^{\theta=2\pi} e^{-r^2} \cdot r \cdot dr \cdot d\theta$.

On pose : $\alpha = r^2 \Rightarrow d\alpha = 2 \cdot r \cdot dr$

$$I \cdot J = 2 \cdot \pi \cdot \frac{1}{2} \cdot \int_{\alpha=0}^{\alpha=+\infty} e^{-\alpha} \cdot d\alpha = \pi \cdot [-e^{-\alpha}]_0^{+\infty} = \pi \Rightarrow I = \sqrt{\pi}$$

$$\text{D'où : } \sigma^2 = \frac{2 \cdot N}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = N$$

$$\boxed{\sigma = \sqrt{N}}$$

1.4 Autre démonstration

σ_N^2 représente la moyenne du carré de la distance (m^2) parcourue au bout de N itérations.
 σ_N est donc une mesure de la distance parcourue au bout de N itérations.

Si $m = 1$, alors il est évident que $\sigma_1 = 1$.

Si, au bout de N itérations, la distance parcourue est σ_N ; alors à la $(N+1)^{\text{ème}}$ itération, on a :

$$\sigma_{N+1} = \sigma_N + 1 \quad \text{ou} \quad \sigma_{N+1} = \sigma_N - 1$$

$$\text{D'où : } \begin{cases} \sigma_{N+1}^2 = \sigma_N^2 + 2 \cdot \sigma_N + 1 \\ \text{ou} \\ \sigma_{N+1}^2 = \sigma_N^2 - 2 \cdot \sigma_N + 1 \end{cases}$$

1. Ce résultat se démontre en utilisant la formule de Stirling $\ln(n!) \simeq \frac{1}{2} \ln(2\pi) + (n + \frac{1}{2}) \cdot \ln n - n$ et en faisant : $\frac{m}{N} \ll 1$

et donc en moyenne : $\sigma_{N+1}^2 = \sigma_N^2 + 1$.

Comme $\sigma_1 = 1$, il est évident que $\sigma_{N+1}^2 = N + 1$;

$$\text{d'où : } \boxed{\sigma_N = \sqrt{N}}$$

1.5 Retour sur l'exemple initial

Application :

1 mètre que l'on reporte 100 fois $\Rightarrow N = 100 \Rightarrow \sigma = 10$.

L'incertitude est donc de $10 \times 1 \text{ cm} = 10 \text{ cm}$ (et non 100 cm !)

Donc, il est probable que $\ell \in [99,9 \text{ m} ; 100,1 \text{ m}]$

Nouvelle question :

Puisque l'on a considérablement réduit l'intervalle, quelle est la probabilité que ℓ soit réellement dans cet intervalle ?

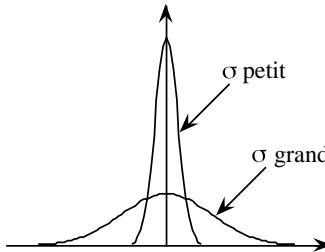
1.6 Loi de Gauss

— Passage en variable continue :

$$\Delta P = \frac{1}{2} P_N(m) \cdot \Delta m = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} \cdot \Delta m$$

devient : $dP = p(x) \cdot dx$

$$\text{avec : } \boxed{p(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}}}$$



On a : $p(0) = \frac{1}{\sqrt{2\pi}\sigma}$ et $p(\sigma) = p(0) \cdot e^{-1/2} \approx 0,6 \cdot p(0)$

— **Probabilité que** $x \in [-\sigma, +\sigma]$

$$P = \int_{-\sigma}^{+\sigma} p(x) \cdot dx = \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_{-\sigma}^{+\sigma} e^{-\frac{x^2}{2\sigma^2}} \cdot dx$$

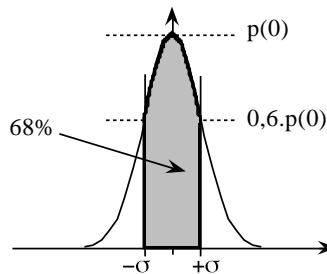
On pose : $t = \frac{x}{\sqrt{2\sigma}}$

$$\Rightarrow P = \frac{1}{\sqrt{\pi}} \cdot \int_{-1/\sqrt{2}}^{+1/\sqrt{2}} e^{-t^2} \cdot dt = \frac{2}{\sqrt{\pi}} \cdot \int_0^{+1/\sqrt{2}} e^{-t^2} \cdot dt$$

On appelle $\theta(u) = \frac{2}{\sqrt{\pi}} \cdot \int_0^u e^{-t^2} \cdot dt$ la fonction d'erreur (fonction tabulée dans la plupart des ouvrages et qui existe sur certaines calculatrices : UTPN sur "Hewlett-Packard").

$$\text{D'où : } P = \theta(u = \frac{\sqrt{2}}{2}) = \boxed{68\%}$$

1.7 Conclusion



On adopte σ comme valeur de l'incertitude absolue :

il y a $\boxed{68\%}$ de chances que $X \in [\bar{x} - \sigma; \bar{x} + \sigma]$

il y a $\boxed{95.4\%}$ de chances que $X \in [\bar{x} - 2\sigma; \bar{x} + 2\sigma]$

et $\boxed{99.7\%}$ de chances que $X \in [\bar{x} - 3\sigma; \bar{x} + 3\sigma]$
Remarque : si $\bar{x} \neq 0$, alors : $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$

2 Question

1. Simuler 100 000 marches au hasard de $N = 2500$ pas de ± 1 .
2. Calculer la proportion de mesures dans l'intervalle $[\bar{x} - \sigma; \bar{x} + \sigma]$
3. Calculer la proportion de mesures dans l'intervalle $[\bar{x} - 2\sigma; \bar{x} + 2\sigma]$
4. Calculer la proportion de mesures dans l'intervalle $[\bar{x} - 3\sigma; \bar{x} + 3\sigma]$

Loi de Fick

F. Kany. ISEN-Brest & La Croix-Rouge

1 Énoncé

En 1855, Adolph Fick réalise une expérience pour vérifier la loi de diffusion qui porte son nom. Fick place une couche de chlorure de sodium (NaCl) dans le fond d'un récipient d'eau. Au sommet, ce récipient est surmonté par un réservoir d'eau non salée (voir fig 1).

Le sel diffuse du fond vers le sommet du récipient et Fick mesure la concentration C en sel en fonction de l'altitude z (par densitométrie). Suivant la géométrie du récipient, la section $S(z)$ change et les profils de concentrations varient.

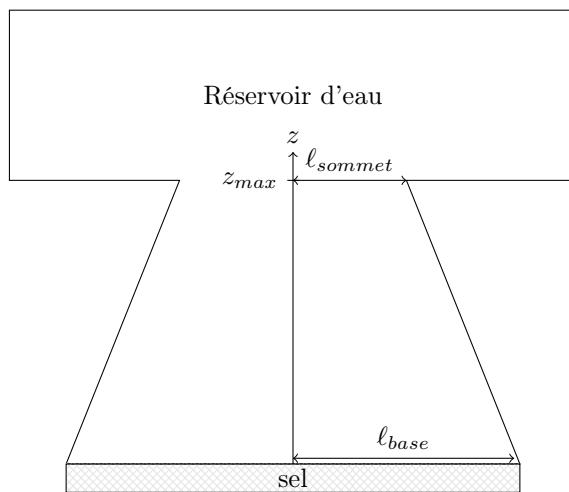


Figure 1

1.1 Calcul théorique

En notant D le coefficient de diffusion du sel dans l'eau, l'équation de diffusion (seconde loi de Fick) s'écrit :

$$D \cdot \left[\frac{\partial^2 C}{\partial z^2} + \frac{\partial C}{\partial z} \cdot \frac{1}{S(z)} \cdot \frac{dS}{dz} \right] = \frac{\partial C}{\partial t}$$

En régime permanent ($\frac{\partial}{\partial t} \equiv 0$), l'équation devient :

$$\left[\frac{\partial^2 C}{\partial z^2} + \frac{\partial C}{\partial z} \cdot \frac{1}{S(z)} \cdot \frac{dS}{dz} \right] = 0$$

En posant $C' = \frac{\partial C}{\partial z}$, on obtient :

$$\left[\frac{dC'}{dz} + C' \cdot \frac{1}{S(z)} \cdot \frac{dS}{dz} \right] = 0$$

On note ℓ_{base} la demi-largeur de la base du récipient (en $z = 0$) et ℓ_{sommet} la demi-largeur du sommet (en $z = z_{max}$). Soit $m = \frac{\ell_{sommet} - \ell_{base}}{z_{max}}$ la pente de la paroi du récipient.

On étudie deux géométries :

- à deux dimensions, le problème est un prisme d'épaisseur h
- à trois dimensions, le problème est invariant par rotation autour de l'axe (Oz).

À deux dimensions, on a : $S(z) = \ell(z) \times h$; on obtient : $\frac{dC'}{C'} = \frac{-m}{\ell_{base} + m.z} dz$
ce qui donne : $C(z) = \frac{\alpha_2}{m} \cdot \log(\ell_{base} + m.z) + \gamma_2$ si $m \neq 0$

À trois dimensions, on a : $S(z) = \pi \times (\ell(z))^2$; on obtient : $\frac{dC'}{C'} = \frac{-2.m}{\ell_{base} + m.z} dz$
ce qui donne : $C(z) = -\frac{\alpha_3}{m} \cdot (\ell_{base} + m.z)^{-1} + \gamma_3$ si $m \neq 0$

Les conditions limites sont $C(z = z_{max}) = 0\%$ et $C(z = 0) = 100\%$ ce qui permet de calculer les constantes d'intégration α_i et γ_i .

On pose $\beta = \frac{\ell_{sommet}}{\ell_{base}}$ et $\mathcal{Z} = \frac{z}{z_{max}}$.

À deux dimensions, la solution est :

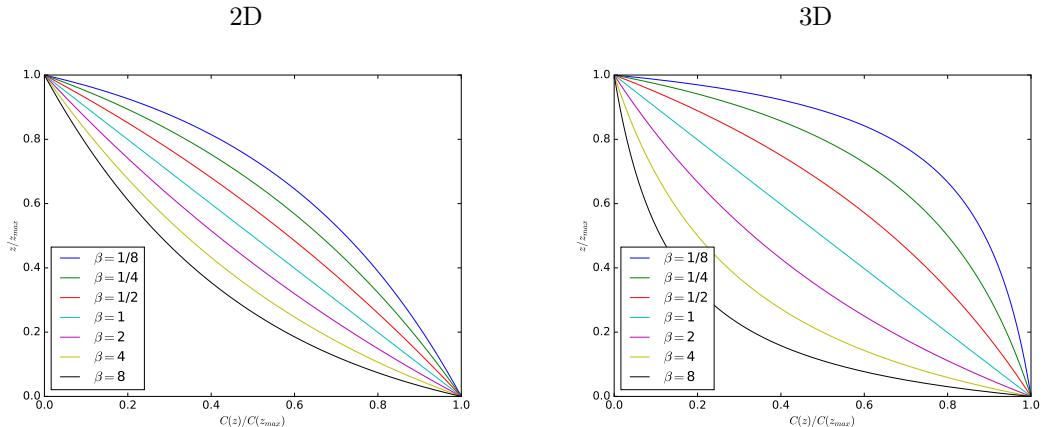
$$\frac{C(z) - C(0)}{C(z_{max}) - C(0)} = \frac{1}{\log(\beta)} \cdot \log(1 + (\beta - 1) \cdot \mathcal{Z}) \quad \text{si } \beta \neq 1$$

$$\frac{C(z) - C(0)}{C(z_{max}) - C(0)} = \mathcal{Z} \quad \text{si } \beta = 1$$

À trois dimensions, la solution est :

$$\frac{C(z) - C(0)}{C(z_{max}) - C(0)} = \frac{\beta \cdot \mathcal{Z}}{1 + (\beta - 1) \cdot \mathcal{Z}}$$

Graphiquement, on obtient :



$\beta = 1$ correspond à $\ell_{sommet} = \ell_{base}$; $\beta < 1$ correspond à $\ell_{sommet} < \ell_{base}$ (comme sur la figure 1).

1.2 Résolution par marche au hasard

On raisonne avec une unité de longueur u arbitraire.

On souhaite retrouver les résultats du cas à deux dimensions (avec une épaisseur $h = 1$ u) en effectuant une marche au hasard.

1.2.1 Méthode 1

On fixe la taille des pas à $\lambda = 0.1$ u (correspondant à l'écart-type du processus aléatoire gaussien). On prend $z_{max} = 10$ u. On impose un nombre de pas $n = 100\,000$. On a donc $n.\lambda \gg z_{max}$: les particules peuvent faire plusieurs allers-retours entre la base et le sommet du récipient.

Méthode :

1. On fixe une géométrie en choisissant β .

Soient $\ell_{max} = 1$ u et $\ell_{min} = \min(\ell_{base}, \ell_{sommet}) = \frac{1}{8}$ u.

On considère un rectangle de hauteur $z_{max} = 10$ u et largeur totale $2.\ell_{max} = 2$ u (soit une surface de 20 u²).

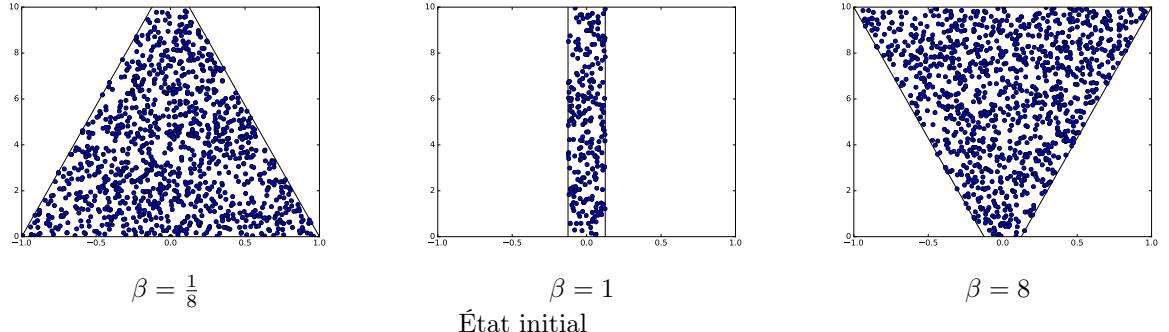
2. On tire au sort les positions (x_i, z_i) de 2000 molécules d'eau dans le rectangle (ce qui correspond à une densité moyenne de 100 molécules.u⁻³ pour $h = 1$ u)

- si la position de la molécule est à l'intérieur du récipient, on garde la molécule ;

- sinon, on l'élimine.

Cette méthode permet de garantir que, quelle que soit la valeur de β , la concentration en molécules sera toujours d'environ 100 molécules.u⁻³.

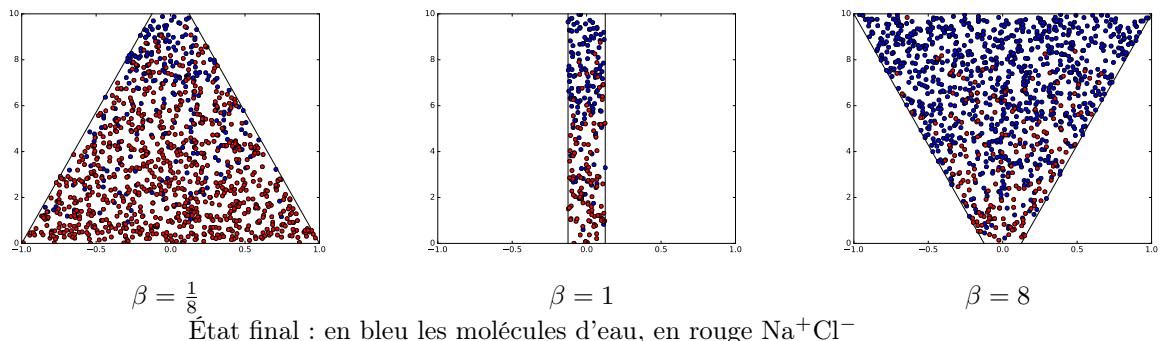
Exemple : dans le cas où $\ell_{base} = \ell_{sommel} = \ell_{min} = \frac{1}{8}$ u, on a un rectangle de surface 2,5 u² avec environ 250 molécules.



3. On ne prend pas en compte les interactions entre les molécules.

Pour chaque molécule, on réitère n fois l'opération suivante :

- on tire au sort un angle θ dans $[0, 2\pi]$
- on déplace la molécule de $\Delta x = \lambda \cdot \cos \theta$ et $\Delta z = \lambda \cdot \sin \theta$
- si la molécule sort des limites latérales du récipient, on la fait rebondir sur la paroi
- si la molécule atteint $z > z_{max}$, on la transforme en molécule d'eau (et on la fait rebondir vers le bas)
- si la molécule atteint $z < 0$, on la transforme en sel (et on la fait rebondir vers le haut)



4. On découpe l'intervalle $[0, z_{max}]$ en 50 intervalles égaux $Z_{i \in [1, 50]}$.

On classe chaque molécule dans l'intervalle Z_i qui correspond à son altitude z finale en notant son état (salé ou non).

On obtient un histogramme¹ des concentrations C_i allant de 100% (pour l'intervalle $i = 1$ débutant à $z = 0$) à 0% (pour l'intervalle $i = 50$ finissant à $z = z_{max}$).

5. On réitère 100 fois les opérations de 2 à 4 et on effectue la moyenne $\overline{C_i}$ des histogrammes C_i .

On trace $Z_i = f(\overline{C_i})$

1.2.2 Méthode 2

On fixe la taille des pas à $\lambda = 0.1$ u (correspondant à l'écart-type du processus aléatoire gaussien). On prend $z_{max} = 10$ u.

1. Attention : dans le cas particulier où $\ell_{base} = \ell_{sommel} = \ell_{min} = \frac{1}{8}$ u, il n'y a au total que 250 molécules ; ce qui fait, en moyenne, 5 molécules (salée ou non) par intervalle. Mais, en raison des fluctuations statistiques, certains intervalles Z_i peuvent ne contenir aucune molécule. Dans ce cas, on reprendra la concentration C_{i-1} de l'intervalle précédent.

On suppose que le système est ergodique². On considère l'évolution d'une seule particule qui fait $n = 100\ 000\ 000$ pas.

Méthode :

1. On fixe une géométrie en choisissant β .

Soient $\ell_{max} = 1$ u et $\ell_{min} = \min(\ell_{base}, \ell_{sommet}) = \frac{1}{8}$ u.

On considère un rectangle de hauteur $z_{max} = 10$ u et largeur totale $2.\ell_{max} = 2$ u (soit une surface de 20 u²).

2. On découpe l'intervalle $[0, z_{max}]$ en 50 intervalles égaux $Z_{i \in [1, 50]}$.
 3. On tire au sort la position (x, z) d'une seule particule (dans les limites du récipient).
 4. Pour cette molécule, on réitère n fois l'opération suivante :
 - on tire au sort un angle θ dans $[0, 2\pi]$
 - on déplace la molécule de $\Delta x = \lambda \cdot \cos \theta$ et $\Delta z = \lambda \cdot \sin \theta$
 - si la molécule sort des limites latérales du récipient, on la fait rebondir sur la paroi
 - si la molécule atteint $z > z_{max}$, on la transforme en molécule d'eau (et on la fait rebondir vers le bas)
 - si la molécule atteint $z < 0$, on la transforme en sel (et on la fait rebondir vers le haut)
 - Après chaque déplacement élémentaire, on classe la molécule dans l'intervalle Z_i qui correspond à son altitude z finale en notant son état (salé ou non).
- On obtient un histogramme des concentrations C_i allant de 100% (pour l'intervalle $i = 1$ débutant à $z = 0$) à 0% (pour l'intervalle $i = 50$ finissant à $z = z_{max}$).

1.3 Aide pour le rebond sur les parois latérales

1.3.1 Méthode rigoureuse

On appelle A le point de la paroi de coordonnées $(\pm \ell_{base}, 0)$ et B le point de la paroi de coordonnées $(\pm \ell_{sommet}, z_{max})$.

Si un point $M(x, z)$ dépasse la droite (AB) , il faut prendre son symétrique $M'(x', z')$ correspondant à une réflexion sur la paroi (AB) .

Soit H la projection de M sur \overrightarrow{AB} .

$$H \text{ est défini par : } \overrightarrow{AH} = \frac{(\overrightarrow{AB} \cdot \overrightarrow{AM})}{\|\overrightarrow{AB}\|} \cdot \vec{u}_{AB} = (\overrightarrow{AB} \cdot \overrightarrow{AM}) \cdot \frac{\overrightarrow{AB}}{\|\overrightarrow{AB}\|^2}$$

$$\text{On a : } \overrightarrow{HM} = \overrightarrow{AM} - \overrightarrow{AH}.$$

$$\text{D'où : } \overrightarrow{HM'} = -\overrightarrow{HM} = \overrightarrow{AH} - \overrightarrow{AM}.$$

$$\text{On calcule : } \overrightarrow{OM'} = \overrightarrow{OA} + \overrightarrow{AH} + \overrightarrow{HM'} = \overrightarrow{OA} + \overrightarrow{AH} + (\overrightarrow{AH} - \overrightarrow{AM}) = \overrightarrow{OA} + 2\overrightarrow{AH} - \overrightarrow{AM}$$

$$\overrightarrow{OM'} = \overrightarrow{OA} + \overrightarrow{MA} + 2(\overrightarrow{AB} \cdot \overrightarrow{AM}) \cdot \frac{\overrightarrow{AB}}{\|\overrightarrow{AB}\|^2} = \overrightarrow{OA} + (\overrightarrow{MO} + \overrightarrow{OA}) + 2(\overrightarrow{AB} \cdot \overrightarrow{AM}) \cdot \frac{\overrightarrow{AB}}{\|\overrightarrow{AB}\|^2}$$

$$\overrightarrow{OM'} = 2\overrightarrow{OA} - \overrightarrow{OM} + 2(\overrightarrow{AB} \cdot \overrightarrow{AM}) \cdot \frac{\overrightarrow{AB}}{\|\overrightarrow{AB}\|^2}$$

On pose : $u = x_B - x_A$; $v = z_B - z_A$

$$x' = 2x_A - x + 2((x_B - x_A)(x - x_A) + (z_B - z_A)(z - z_A)) \cdot \frac{(x_B - x_A)}{(x_B - x_A)^2 + (z_B - z_A)^2}$$

$$x' = 2x_A - x + 2(u(x - x_A) + v(z - z_A)) \cdot \frac{u}{u^2 + v^2}$$

$$x' = \frac{(2x_A - x)(u^2 + v^2) + 2(u^2 \cdot (x - x_A) + u \cdot v \cdot (z - z_A))}{u^2 + v^2}$$

$$x' = \frac{x(u^2 - v^2) + 2x_A \cdot v^2 + 2u \cdot v \cdot (z - z_A)}{u^2 + v^2}$$

De même pour z' :

$$z' = 2z_A - z + 2((x_B - x_A)(x - x_A) + (z_B - z_A)(z - z_A)) \cdot \frac{(z_B - z_A)}{(x_B - x_A)^2 + (z_B - z_A)^2}$$

$$z' = 2z_A - z + 2(u(x - x_A) + v(z - z_A)) \cdot \frac{v}{u^2 + v^2}$$

$$z' = \frac{(2z_A - z)(u^2 + v^2) + 2(u \cdot v \cdot (x - x_A) + v^2 \cdot (z - z_A))}{u^2 + v^2}$$

$$z' = \frac{z(v^2 - u^2) + 2z_A \cdot u^2 + 2u \cdot v \cdot (x - x_A)}{u^2 + v^2}$$

2. hypothèse ergodique : à l'équilibre, la valeur moyenne d'une grandeur calculée de manière statistique est égale à la moyenne d'un très grand nombre de mesures prises dans le temps.

1.3.2 Méthode simpliste

Si l'on voit que le déplacement $(\Delta x, \Delta y)$ fait sortir la molécule des limites latérales, on annule simplement le déplacement.

2 Bibliographie

- A. Fick, Poggendorff's Annalen, 94, 3, (1855) 59-86.
- Simulation of Fick's Verification of the 2nd Law. Richard DiDomizio, Afina Lupulescu, Martin E. Glicksman. Diffusion Fundamentals 4 (2006) 2.1 - 2.14

Physique statistique : fluctuations statistiques

F. Kany. ISEN-Brest. La Croix-Rouge.

Télécharger, à l'adresse ci-dessous, le programme qui simule l'évolution d'un ensemble de particules subissant des collisions élastiques.

<https://jakevdp.github.io/blog/2012/08/18/matplotlib-animation-tutorial/>

1. Modifier le code pour :
 - enlever la gravité
 - augmenter le nombre de particules ($50 \rightarrow 200$)
 - augmenter la durée de la simulation ($6 \text{ s} \rightarrow 30 \text{ s}$)
2. Modifier la fonction `animate(i)` pour qu'à chaque itération elle mesure le nombre de particules dans le carré $ABCD$: $A(-0.5, -0.5)$ $B(-0.5, +0.5)$ $C(+0.5, +0.5)$ $D(+0.5, -0.5)$. Stocker ce nombre dans une liste `sous_ensemble`.

Tracer `sous_ensemble` et observer les fluctuations statistiques.

Entropie statistique : système à 2 états de niveaux d'énergie différents

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

On considère un système Σ , isolé, constitué de n_{tot} particules pouvant se répartir entre deux états 1 et 2 de d'énergies respectives $-\varepsilon$ et $+\varepsilon$.

Exemple : n_{tot} spins sur un réseau (i.e. ayant une position fixe) pouvant avoir deux états (état 1 noté \uparrow et état 2 noté \downarrow), sans interaction entre eux¹, avec champ magnétique extérieur (dans le sens \uparrow).
On appelle n_1 le nombre de spins dans l'état \uparrow et n_2 le nombre de particules dans l'état \downarrow .
On a : $n_1 + n_2 = n_{tot}$.

L'énergie e_{tot} d'une distribution (n_1, n_2) est : $e_{tot} = n_1 \times (-\varepsilon) + n_2 \times (+\varepsilon) = (n_2 - n_1).\varepsilon$.

On suppose que les particules sont discernables (c'est le cas dans l'exemple du réseau de spins car les positions des spins sont fixes).

Questions

- Combien y a-t'il de micro-états particuliers Ω_{n_1} avec n_1 particules dans l'état 1 (et n_2 particules dans l'état 2) ?
- Avec Python, tracer $\ln(\Omega_{n_1})/n_{tot}$ en fonction de e_{tot}/n_{tot} pour $n_{tot} \in [10, 200]$.
(Pour chaque n_{tot} : on fait varier n_1 de 0 à n_{tot} ; on calcule e_{tot} et Ω_{n_1}).
Ce tracé correspond à l'entropie du système (voir Annexe).
Expliquer pourquoi la fonction s'annule en $e_{tot}/n_{tot} = \pm\varepsilon$.
- On définit la dérivée numérique d'une liste de points par :

```
liste_derivee[i] = (liste_y[i+1]-liste_y[i]) / (liste_x[i+1]-liste_x[i])
```

Pour $n_{tot} = 1000$, calculer $\frac{d \ln(\Omega_{n_1})}{d e_{tot}}$ et tracer cette quantité en fonction de e_{tot}/n_{tot} .

- La quantité précédente représente l'inverse de la température du système (voir Annexe).
Tracer $T = f(e_{tot}/n_{tot})$. Commenter.

Annexe

Ensemble micro-canonique

En physique statistique, on définit l'ensemble micro-canonique comme l'ensemble des répliques fictives d'un système réel dont l'énergie (e_{tot}), le volume (V) et le nombre de particules (n_{tot}) sont fixés.

L'**hypothèse micro-canonique** consiste à supposer que, quand un système est **isolé** et en équilibre, celui-ci se trouve avec probabilités égales dans chacun de ses micro-états accessibles.

1. Si le système est isolé, les spins peuvent avoir une interaction entre eux ; si un spin se retourne sous l'action d'une telle interaction, un autre spin doit se retourner en sens inverse pour assurer la conservation de l'énergie.

Lien avec l'entropie

On appelle entropie statistique, dans un état macroscopique donné, la quantité :

$$S = -k_B \cdot \sum_{\ell} P_{\ell} \cdot \ln(P_{\ell})$$

D'après l'hypothèse micro-canonical $P_{\ell} = \frac{1}{\Omega}$ où Ω est le nombre de micro-états accessibles.
Donc : $S = -k_B \cdot \sum_{\ell} P_{\ell} \cdot \ln(P_{\ell}) = -k_B \cdot \sum_{\ell=1}^{\Omega} \frac{1}{\Omega} \cdot \ln\left(\frac{1}{\Omega}\right) = k_B \cdot \Omega \cdot \frac{1}{\Omega} \cdot \ln(\Omega) = k_B \cdot \ln(\Omega)$

Définition de la température

La température est définie comme $\frac{1}{T} = \frac{\partial S}{\partial e_{tot}}$.

Entropie statistique : système à 2 états de même énergie

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

On considère un système Σ , isolé, constitué de n_{tot} particules pouvant se répartir entre deux états 1 et 2 de même énergie ε .

Exemple 1 : n_{tot} molécules de gaz parfait dans une enceinte de volume V que l'on sépare en deux compartiments de volume $V/2$; on appelle n_1 le nombre de particules dans le compartiment de gauche (état 1) et n_2 le nombre de particules dans le compartiment de droite (état 2). On a : $n_1 + n_2 = n_{tot}$.

Exemple 2 : n_{tot} spins sur un réseau (i.e. ayant une position fixe) pouvant avoir deux états (état 1 noté \uparrow et état 2 noté \downarrow), sans interaction entre eux, sans champ magnétique extérieur ; on appelle n_1 le nombre de spins dans l'état \uparrow et n_2 le nombre de particules dans l'état \downarrow . On a : $n_1 + n_2 = n_{tot}$.

On suppose que les particules sont discernables (c'est le cas dans l'exemple 2 [où les positions des spins sont fixes] ; c'est le cas dans l'exemple 1, si on peut voir - pour une particule donnée à l'avance - si elle se trouve dans le compartiment de gauche ou de droite).

Questions

1. Au total, combien y a-t'il de micro-états possibles ? On notera Ω_{tot} ce nombre.

2. Tous ces micro-états sont, a priori, équiprobables.

Combien y a-t'il de micro-états particuliers Ω_{n_1} avec n_1 particules dans l'état 1 (et n_2 particules dans l'état 2) ?

3. On note $x = \frac{n_1}{n_{tot}}$.

Avec Python, tracer $\ln(\Omega_{n_1})$ en fonction de x pour $n_{tot} \in [10, 200]$.

(Pour chaque n_{tot} : on fait varier n_1 de 0 à n_{tot} ; on calcule x et Ω_{n_1}).

Ce tracé correspond à l'entropie du système (voir Annexe).

Pour quelle valeur $x = x_{max}$, Ω_{n_1} est-il maximum ?

Cette valeur correspond à l'état d'équilibre du système (i.e. celui qui maximise l'entropie dans l'ensemble micro-canonical).

4. On note $P_{n_1} = \frac{\Omega_{n_1}}{\Omega_{tot}}$ la probabilité que Σ se trouve dans un des états Ω_{n_1} .

— Avec Python, tracer P_{n_1} en fonction de x pour $n_{tot} \in [10, 200]$.

— Normalement $\sum_{n_1} P_{n_1} = 1$ (ou $\int_0^1 P(x).dx = 1$ par passage à la limite), pourtant la surface sous la courbe pour $n_{tot} = 10$ semble plus grande que la surface sous la courbe pour $n_{tot} = 200$. Expliquer.

— Avec Python, tracer P_{n_1} en fonction de x pour $n_{tot} \in [1000, 2000]$.

— Lorsque n augmente, que peut-on dire de P_{max} le maximum de P_{n_1} ?

— Tracer $\log(P_{max})$ en fonction de $\log(n_{tot})$ pour $\log(n_{tot}) \in [1, 5]$; faire une régression linéaire et en déduire un équivalent de P_{max} et de Ω_{max} .

Annexe

Ensemble micro-canonical

En physique statistique, on définit l'ensemble micro-canonical comme l'ensemble des répliques fictives d'un système réel dont l'énergie (e_{tot}), le volume (V) et le nombre de particules (n_{tot}) sont fixés.

L'**hypothèse micro-canonique** consiste à supposer que, quand un système est **isolé** et en équilibre, celui-ci se trouve avec probabilités égales dans chacun de ses micro-états accessibles.

Lien avec l'entropie

On appelle entropie statistique, dans un état macroscopique donné, la quantité :

$$S = -k_B \cdot \sum_{\ell} P_{\ell} \cdot \ln(P_{\ell})$$

D'après l'hypothèse micro-canonique $P_{\ell} = \frac{1}{\Omega}$ où Ω est la nombre de micro-états accessibles.

Donc : $S = -k_B \cdot \sum_{\ell} P_{\ell} \cdot \ln(P_{\ell}) = -k_B \cdot \sum_{\ell=1}^{\Omega} \frac{1}{\Omega} \cdot \ln\left(\frac{1}{\Omega}\right) = k_B \cdot \Omega \cdot \frac{1}{\Omega} \cdot \ln(\Omega) = k_B \cdot \ln(\Omega)$

La transition de phase ferromagnétique-paramagnétique.

F. Kany. ISEN-Brest & La Croix-Rouge

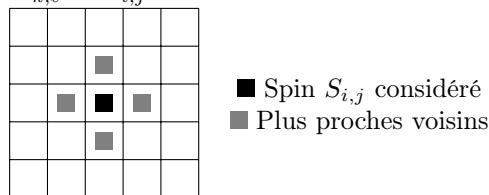
Présentation

Il existe une température critique appelée température de Curie (ex : $T_{Curie}(\text{Fer}) = 770 \text{ }^{\circ}\text{C}$) au-dessus de laquelle l'aimantation spontanée d'un matériau ferromagnétique disparaît.

- Le modèle le plus simple à deux dimensions est le modèle d'Ising. On considère un réseau carré $N \times N$ composé de N^2 spins $S_{i,j}$ pouvant prendre les valeurs ± 1 . En l'absence de champ magnétique extérieur, l'énergie du système est :

$$E = -J \cdot \sum_{i,j} S_{i,j} \cdot \left(\sum_{k,\ell} S_{k,\ell} \right)$$

(J est la constante de couplage, $J > 0$ pour une interaction ferromagnétique). La somme s'effectue sur les plus proches voisins $S_{k,\ell}$ de $S_{i,j}$.



- Nous allons utiliser un algorithme très général : l'algorithme de Monte-Carlo. Il peut s'appliquer à tout modèle pour lequel il est possible :

- de faire des modifications aléatoires du modèle (en général à partir d'un nombre fini de variables continues ou discrètes qui sont changées de manière aléatoire)
- d'associer une variation d'énergie à chacune de ces modifications. L'algorithme procède alors de la manière suivante :
Soit une modification aléatoire envisagée qui transforme le modèle A_t en modèle A_{t+1} , associé à une variation d'énergie ΔE .
 - Si $\Delta E < 0$, la modification est acceptée.
 - Si $\Delta E > 0$, la modification est acceptée avec la probabilité $P = e^{-\frac{\Delta E}{k_B T}}$ où T est la température du système et k_B la constante de Boltzmann.

Il est facile de voir qu'une fois un état stationnaire atteint, la distribution des états générés $\{A_i\}$ correspond à une distribution de Boltzmann. N.B. : Un système est dans un état stationnaire si les variables le décrivant n'évoluent plus au cours du temps, ce n'est pas forcément un état d'équilibre.

Un réseau carré de dimension $N \times N$ spins (moments magnétiques) est traité comme un automate cellulaire, ainsi le spin $S_{i,j}$ est repéré par les coordonnées i,j . De plus, pour éviter les effets de bords dus à la taille finie du réseau, le système obéit à des conditions aux limites périodiques (toriques) de telle façon que :

$$\forall (k, \ell) \in \mathbb{Z}^2, \quad S_{k,\ell} = S_{k \% N, \ell \% N}$$

1 Etat initial du système

Pour simplifier on choisira $k_B = 1$ (constante de Boltzmann) et $J = 1$ (interaction ferromagnétique).

On prendra un réseau carré 300×300 et $T = 5$ K. La configuration de départ est fixée en choisissant tous les spins égaux à +1.

1. Quelle est l'énergie initiale E_0 du système ?
2. Quelle est l'aimantation moyenne par spin initiale m_0 avec l'aimantation moyenne $m = \frac{1}{N^2} \cdot \sum_{i,j} S_{i,j}$?
3. Coder l'obtention de la matrice de spin de l'état initial.

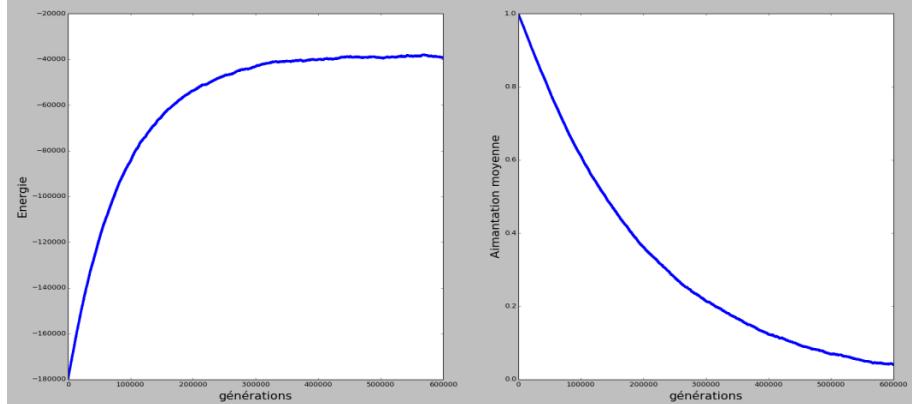
2 Evolution du système

Durant la dynamique, il faut modifier la configuration du système de façon modérée afin que la nouvelle configuration puisse être acceptée avec une probabilité raisonnable pour ne pas la piéger dans un puits de potentiel non optimal.

Pour réaliser cette condition, un seul spin, choisi de manière aléatoire, est possiblement modifié à chaque itération selon le principe de Monte-Carlo.

Soit E_n l'énergie du système à la $n^{ième}$ génération.

4. En supposant que le spin $S_{i,j}$ est retourné (passage de +1 à -1 ou de -1 à +1), Calculer la variation d'énergie ΔE telle que $E_{n+1} = E_n + \Delta E$. Calculer également Δm , telle que $m_{n+1} = m_n + \Delta m$.
5. A l'aide de la fonction `random()` de la bibliothèque `random` et de l'algorithme de Monte-Carlo, trouver une inégalité faisant intervenir ΔE qui, si elle est satisfaite, aboutit à un retournement du spin $S_{i,j}$ désigné aléatoirement et à une absence de modification du spin $S_{i,j}$ sinon.
6. Ecrire un script en important les modules nécessaires et en effectuant toutes les initialisations nécessaires permettant d'obtenir pour 600 000 générations (itérations) la liste des énergies E_n et la liste des aimantations m_n . *Attention à bien prendre garde aux conditions périodiques aux limites.*
7. Coder le tracé de l'énergie d'une part et de l'aimantation d'autre part en fonction des générations.
8. On obtient les évolutions ci-dessous. Commenter ces courbes. Pourquoi un temps de calcul si important ?



9. On fait évoluer la température de 5 K à 0,1 K par pas de 0,05 K. Le nombre de générations pour chaque température est de 10^6 . Pour éviter de remplir trop l'espace mémoire, les matrices concernant les valeurs de spins sont enregistrées (via la bibliothèque `pickle`) pour être rappelées ensuite et représentées graphiquement.

Pourquoi choisir 10^6 comme nombre de générations ?

10. On obtient les trois figures ci-dessous. Commenter précisément mais concisément chacune d'elles.

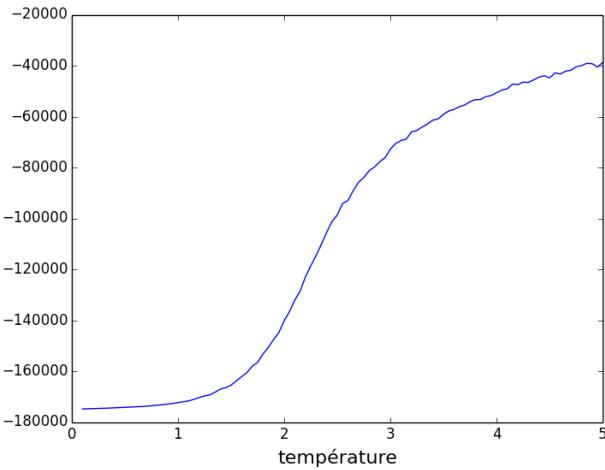


Figure A : Evolution de l'énergie en fonction de la température d'un réseau carré de 9.10^4 spins obtenu après 10^6 générations suivant l'algorithme de Monte-Carlo en partant initialement de spins tous égaux à +1.

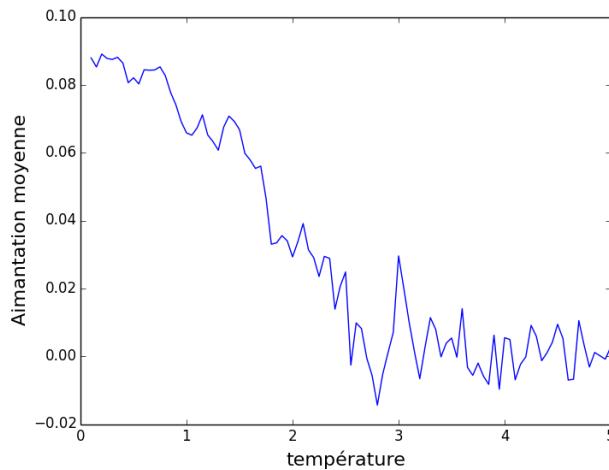


Figure B : Evolution de l'énergie en fonction de la température d'un réseau carré de 9.10^4 spins obtenu après 10^6 générations suivant l'algorithme de Monte-Carlo en partant initialement de spins tous égaux à +1.

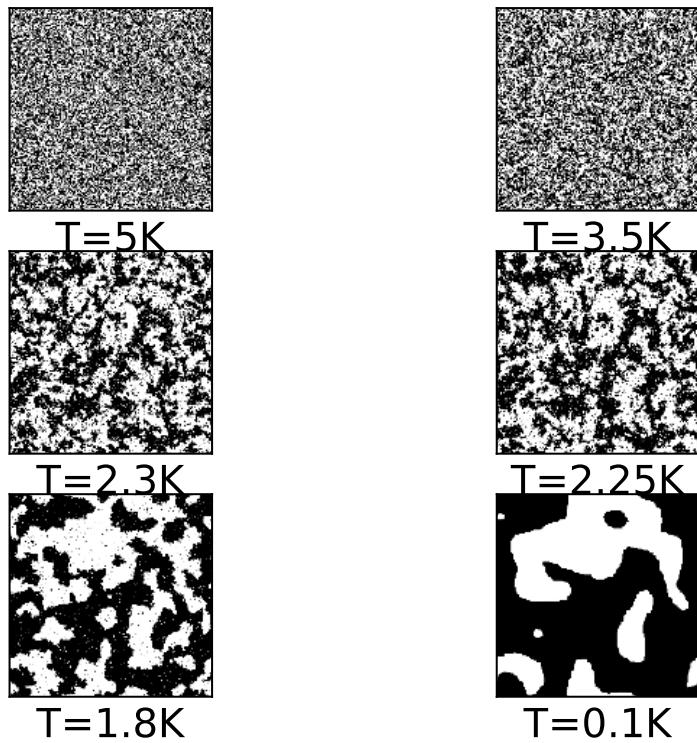


Figure C : Evolution en fonction de la température d'un réseau carré de 9.10^4 spins obtenu après 10^6 générations suivant l'algorithme de Monte-Carlo en partant initialement de spins tous égaux à +1.

Les zones noires correspondent aux spins haut de valeur +1 et les blanches aux spins bas de valeur -1.

11. La résolution analytique du problème de la transition ferromagnétique-paramagnétique fait appel à des outils développés par la mécanique statistique. En 1944, Onsager a résolu ce problème à deux dimensions. Pour un réseau carré, il a montré qu'il existe une température de transition de phase et que sa valeur est :

$$T_c = J \cdot \frac{2}{\ln(1 + \sqrt{2})} \simeq 2,27J$$

Cette valeur est-elle en accord avec notre résolution numérique ?

Entropie statistique : système à N niveaux d'énergie

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Principe

On cherche tous les N -uplets $\{n_0, n_1, \dots, n_{N-1}\}$ tel que $n_i \in \mathbb{N}$ vérifiant

$$\begin{cases} n_0 + n_1 + \dots + n_{N-1} = n_{tot} \\ n_0 \cdot e_0 + n_1 \cdot e_1 + \dots + n_{N-1} \cdot e_{N-1} = e_{tot} \end{cases}$$

avec n_{tot} , e_i et e_{tot} des constantes données.

Contexte

Soit un système Σ , isolé, d'énergie totale e_{tot} , constitué de n_{tot} particules pouvant occuper N niveaux d'énergie e_i (notés e_0, e_1, \dots, e_{N-1}).

Le niveau d'énergie e_i est occupé par n_i particules.

La conservation de la matière impose : $n_0 + n_1 + \dots + n_{N-1} = n_{tot}$.

La conservation de l'énergie impose : $n_0 \cdot e_0 + n_1 \cdot e_1 + \dots + n_{N-1} \cdot e_{N-1} = e_{tot}$.

Exemple : $N = 4$ avec $e_0 = 0, e_1 = \varepsilon, e_2 = 2\varepsilon, e_3 = 3\varepsilon, n_{tot} = 3$ particules notées A, B, C et $e_{tot} = 3\varepsilon$.

On a :

Niveaux d'énergie	Configurations											
		A	B	C	A	A	B	B	C	C		
3. ε					A	A	B	B	C	C		
2. ε					B	C	A	C	A	B		
ε	ABC	BC	AC	AB	C	B	C	A	B	A		
0												

Si les particules A, B et C sont **discernables** (i.e. on peut différencier les particules), on a $\Omega = 10$ N -uplets possibles : $\{0,1,0,0\}, \{2,0,0,1\}$ trois fois et $\{1,1,1,0\}$ six fois.

Si les particules A, B et C sont **indiscernables** (i.e. rien ne peut différencier les particules), on a $\Omega = 3$ N -uplets possibles : $\{0,1,0,0\}, \{2,0,0,1\}$ et $\{1,1,1,0\}$.

Les Ω N -uplets représentent les états microscopiques compatibles avec les conditions macroscopiques que l'on a imposées (ici : $n_{tot} = 3$ et $e_{tot} = 3\varepsilon$).

Ensemble micro-canonical

En physique statistique, on définit l'ensemble micro-canonical comme l'ensemble des répliques fictives d'un système réel dont l'énergie (e_{tot}), le volume (V) et le nombre de particules (n_{tot}) sont fixés.

L'**hypothèse micro-canonique** consiste à supposer que, quand un système est **isolé** et en équilibre, celui-ci se trouve avec probabilités égales dans chacun de ses micro-états accessibles.

Dans l'exemple ci-dessus, pour des particules discernables, $\Omega = 10$ et la probabilité de chaque micro-état est $P_\ell = \frac{1}{10}$.

Lien avec l'entropie

On appelle entropie statistique, dans un état macroscopique donné, la quantité :

$$S = -k_B \cdot \sum_{\ell} P_{\ell} \cdot \ln(P_{\ell})$$

D'après l'hypothèse micro-canonique $P_{\ell} = \frac{1}{\Omega}$ où Ω est la nombre de micro-états accessibles.

Donc : $S = -k_B \cdot \sum_{\ell} P_{\ell} \cdot \ln(P_{\ell}) = -k_B \cdot \sum_{\ell=1}^{\Omega} \frac{1}{\Omega} \cdot \ln\left(\frac{1}{\Omega}\right) = k_B \cdot \Omega \cdot \frac{1}{\Omega} \cdot \ln(\Omega) = k_B \cdot \ln(\Omega)$

Dans l'exemple ci-dessus : $S = k_B \cdot \ln(10)$ pour des particules discernables.

Algorithme

Ecrire une fonction `microetats(niveaux_energie,nbre_particules,energie_totale,indiscernable)` qui prend en arguments : `niveaux_energie` : la liste des niveaux d'énergie, `nbre_particules` : l'entier n_{tot} , `energie_totale` l'entier e_{tot} et `indiscernable` : un booléen indiquant le type de particules.

La fonction renvoie la liste des ΩN -uplets (sous la forme de listes) triée à l'aide de la fonction `sort`.

Exemple :

entrée

`[0,1,2,3],3,3,True`

sortie

`[[0, 3, 0, 0], [1, 1, 1, 0], [2, 0, 0, 1]]`

entrée

`[0,1,2,3],3,3,False`

sortie

`[[0, 3, 0, 0], [1, 1, 1, 0], [1, 1, 1, 0], [1, 1, 1, 0], [1, 1, 1, 0], [1, 1, 1, 0], [1, 1, 1, 0], [2, 0, 0, 1], [2, 0, 0, 1], [2, 0, 0, 1]]`

Paradoxe du groupe de colle

Il ne faut pas tricher !

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Un groupe de colle est constitué de trois élèves A , B et C .

Le collieur pose vingt questions binaires (la réponse est oui ou non) ; les élèves répondent de façon indépendante ; l'interrogateur note de la façon suivante : si, à une question, deux ou trois élèves donnent la bonne réponse, il accorde un point au groupe ; si, à une question, deux ou trois élèves donnent la mauvaise réponse, il enlève un point au groupe. La note est collective : tous les élèves ont la même note.

Le niveau des trois élèves n'est pas le même : A et B répondent correctement dans 80% des cas ; C répond correctement dans 60% des cas.

Questions

1. Quelle est l'espérance de la note de colle de ce groupe ? (Faire le calcul à la main et vérifier avec une simulation numérique).
2. Clairement, la note du groupe est plombée par l'élève C (ce qui énerve les élèves A et B). L'élève C décide de tricher : il donne systématiquement la même réponse que A . Son taux de réponse correcte passe ainsi de 60% à 80%. Quelle est la nouvelle espérance de la note de colle ? La note de ce groupe augmente-t-elle ?
3. Pour comprendre ce paradoxe, reprendre les mêmes questions avec un élève C qui répond correctement dans 50% des cas (c'est-à-dire comme un singe qui répond au hasard!), puis qui répond comme l'élève A .
4. Reprendre les mêmes questions avec un élève C qui répond correctement dans 40% des cas (c'est-à-dire moins bien qu'un singe!), puis qui répond comme l'élève A .
5. Expliquer le "paradoxe". Interpréter le résultat en terme d'entropie.

Paradoxe de Monty Hall

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation

Consulter l'article de wikipedia : https://fr.wikipedia.org/wiki/Probl%C3%A8me_de_Monty_Hall.
Ecrire une fonction `MontyHall(strategie)` qui simule l'expérience décrite dans l'article et qui renvoie 1 (resp. 0) si le candidat gagne (resp. perd) suivant la valeur de la variable `strategie`. (On prendra par exemple `strategie=0` si le candidat garde son choix initial et `strategie=1` si le candidat change son choix initial).

Exécuter la fonction `MontyHall(strategie)` 100 000 fois et donner les probabilités de gain des deux stratégies.

Expliquer le "paradoxe". Interpréter le résultat en terme d'entropie.

Initiation à la mécanique quantique

Interféromètre de Mach-Zehnder

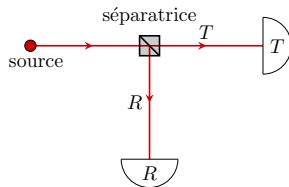
F. Kany. ISEN-Brest & La Croix-Rouge

1 Présentation

- La lumière est constituée de particules élémentaires indivisibles : les photons.
- On sait réaliser des miroirs semi-transparents (ou des lames semi-réfléchissantes) qui laissent passer la moitié de la lumière et qui réfléchissent la moitié de la lumière. (C'est assez facile à faire en jouant sur l'épaisseur du dépôt métallique). Très exactement, cela veut dire ceci : si on éclaire ce miroir semi-transparent avec un laser de 100 mW : 50 mW est réfléchi et 50 mW est transmis.

1.1 Dispositif à une séparatrice

Dans le cas où un photon unique arrive sur un miroir semi-transparent, on ne peut pas prédire si celui-ci sera transmis ou réfléchi. Le phénomène est aléatoire (complètement imprévisible).



Simuler l'arrivée d'un photon sur un miroir semi-transparent

- en utilisant des probabilités classiques
- en utilisant le formalisme quantique.

Ce formalisme repose sur trois principes.

- Premier principe : la probabilité P pour qu'une particule, émise initialement en I , soit détectée finalement en F , est le module carré d'un nombre complexe qui représente l'amplitude de probabilité pour que la particule aille de I à F . On écrit : $P = |\langle F|I \rangle|^2$.

Avec cette notation : le symbole $\langle \dots \rangle$ signifie “amplitude de probabilité pour que ...” et le contenu des crochets signifie “condition finale | condition initiale”.

Pour une particule libre (i.e. soumise à aucune force), d'impulsion \vec{p} , l'amplitude de probabilité d'aller de I à F est (à un facteur numérique près) : $\langle F|I \rangle = \frac{1}{IF} e^{i \cdot \vec{p} \cdot IF / \hbar}$
avec $p^2 \cdot c^2 = E^2 - (m_0 \cdot c^2)^2$ pour une particule relativiste
ou $p^2 = 2 \cdot m \cdot E_c$ (i.e. $\vec{p} = m \cdot \vec{v}$) pour une particule non-relativiste.

- Deuxième principe : lorsqu'une particule a à sa disposition deux chemins indiscernables ① ou ② pour aller de I à F , on **somme** les amplitudes de probabilité : $\langle F|I \rangle = \langle F|I \rangle_1 + \langle F|I \rangle_2$.

On a alors : $P = |\langle F|I \rangle_1 + \langle F|I \rangle_2|^2$.

En revanche, si les deux chemins sont discernables, on somme les probabilités (et non les amplitudes) : $P = |\langle F|I \rangle_1|^2 + |\langle F|I \rangle_2|^2 = P_1 + P_2$.

- Troisième principe : lorsqu'un chemin particulier se décompose en deux étapes, par exemple s'il faut que la particule aille de I à E et de E à F , on **multiplie** les amplitudes de probabilité : $\langle F|I \rangle = \langle F|E \rangle \cdot \langle E|I \rangle$.

Ici, la source émet une particule dans l'état $|S\rangle$.

On montre, en électromagnétisme, que la réflexion sur un miroir introduit un déphasage qui se traduit ici par une multiplication par $e^{i\cdot\pi/2} = i$.

L'état initial $|S\rangle$ doit s'écrire : $|S\rangle = \frac{1}{\sqrt{2}}(|T\rangle + i|R\rangle)$; calculer $\langle T|S\rangle$ et $\langle R|S\rangle$; en déduire les probabilités d'atteindre les détecteurs R et T .

On pourra utiliser la bibliothèque `sympy.physics.quantum` et créer une classe de Bra et Ket orthogonaux :

```
from sympy.physics.quantum import HilbertSpace, Ket, Bra

class OrthogonalKet(Ket):

    def __new__(cls, n):
        return Ket.__new__(cls, n)

    @property
    def n(self):
        return self.label[0]

    @classmethod
    def dual_class(self):
        return OrthogonalBra

    @classmethod
    def _eval_hilbert_space(cls, label):
        return HilbertSpace()

    def _eval_innerproduct_OrthogonalBra(self, bra, **hints):
        if self.n == bra.n:
            return 1
        else:
            return 0

class OrthogonalBra(Bra):

    def __new__(cls, n):
        return Bra.__new__(cls, n)

    @property
    def n(self):
        return self.label[0]

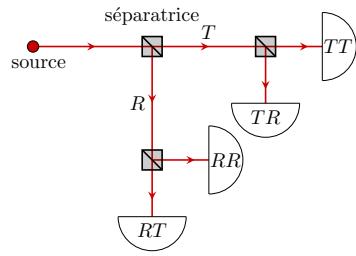
    @classmethod
    def dual_class(self):
        return OrthogonalKet
```

1.2 Deux séparatrices en cascade

Simuler le comportement d'un photon sur deux miroirs semi-transparents en cascade (voir figure).

- en utilisant des probabilités classiques
- en utilisant le formalisme quantique.

Calculer $\langle TT|S\rangle$, $\langle TR|S\rangle$, $\langle RR|S\rangle$ et $\langle RT|S\rangle$; en déduire les probabilités d'atteindre les différents détecteurs.



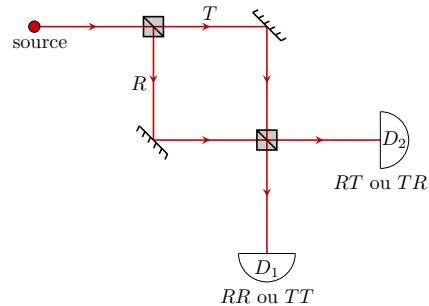
1.3 Interféromètre de Mach Zehnder

Simuler le comportement d'un photon dans un interféromètre de Mach Zehnder (voir figure).

- en utilisant des probabilités classiques
- en utilisant le formalisme quantique.

Les miroirs parfaits transforment $|T\rangle$ en $|T'\rangle = i|T\rangle$ et $|R\rangle$ en $|R'\rangle = i|R\rangle$. Calculer $\langle D_1|S\rangle$ et $\langle D_2|S\rangle$; en déduire les probabilités d'atteindre les détecteurs D_1 et D_2 .

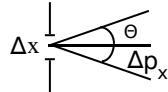
Comparer les deux méthodes et conclure.



Mécanique quantique : principe d'indétermination d'Heisenberg

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation



On considère une source de lumière émettant des photons de longueur d'onde λ et d'impulsion $\vec{p} = \hbar \cdot \vec{k}$. Ces photons arrivent sur une plaque opaque possédant une fente de largeur $a = \Delta x$.

Un calcul purement quantique permet de montrer¹ que la probabilité qu'un photon soit diffracté dans la direction θ est donnée par :

$$P(\theta) = \frac{a}{2\pi} \cdot \text{sinc}^2(\alpha)$$

avec $\alpha = p.a. \sin(\theta)/(2\hbar)$ et $\text{sinc}(x) = \frac{\sin(x)}{x}$.

Questions

1. Représenter la fonction $P(\theta)$ pour $\theta \in [-\pi/2, \pi/2]$.

Le fait d'avoir $P \neq 0$ pour $\theta \neq 0$ s'interprète par les relations d'indétermination d'Heisenberg $\Delta x \cdot \Delta p_x \gtrsim \hbar$: le fait d'imposer à un photon d'avoir une position x , à Δx près, entraîne une indétermination sur la projection de l'impulsion, dans la direction x , de Δp_x .

Rien ne permet de prédire quel sera l'angle θ du photon après le passage par la fente (on peut seulement estimer la probabilité de cet angle).

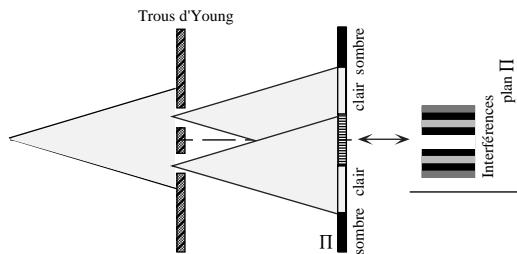
2. Simuler le passage de 50 000 photons à travers les deux fentes d'Young. On prendra $a = 4\lambda$.

1. <https://arxiv.org/ftp/quant-ph/papers/0703/0703126.pdf>

Mécanique quantique : fentes d'Young

F. Kany. ISEN-Brest & La Croix-Rouge

Présentation



On considère une source de lumière émettant des photons de longueur d'onde λ et d'impulsion $\vec{p} = \hbar \cdot \vec{k}$.

Ces photons arrivent sur une plaque opaque possédant deux fentes de largeur a , parallèles, distantes de d .

Un calcul purement quantique permet de montrer¹ que la probabilité qu'un photon soit diffracté dans la direction θ est donnée par :

$$P(\theta) = \frac{2.a}{\pi} \cdot [\cos^2(\phi/2) \cdot \text{sinc}^2(\alpha)]$$

avec $\phi = p.d \cdot \sin(\theta)/\hbar$, $\alpha = p.a \cdot \sin(\theta)/(2.\hbar)$ et $\text{sinc}(x) = \frac{\sin(x)}{x}$.

Questions

1. Représenter la fonction $P(\theta)$ pour $\theta \in [-\pi/2, \pi/2]$.
2. Simuler le passage de 50 000 photons à travers les deux fentes d'Young. On prendra $a = \lambda$ et $d = 4.\lambda$.

1. <https://arxiv.org/ftp/quant-ph/papers/0703/0703126.pdf>

Désintégration radioactive

F. Kany. ISEN-Brest & La Croix-Rouge

Position du problème

Soit $N(t)$ le nombre de particules radioactives à l'instant t . Pendant un intervalle de temps Δt , $-\Delta N$ particules se désintègrent radioactivement. On appelle p la probabilité, pour chaque particule radioactive, de se désintégrer pendant un intervalle de temps unité. On a : $p = -\frac{\Delta N(t)}{N(t)} = \frac{\Delta t}{\tau}$ où τ est une constante. D'où : $\frac{\Delta N}{\Delta t} = -\frac{N(t)}{\tau}$

À la limite, lorsque N tend vers l'infini et Δt vers zéro, on a : $\frac{dN(t)}{dt} = -\frac{N(t)}{\tau}$ et donc : $N(t) = N(0).e^{-t/\tau}$. τ est donc une constante de temps qui représente une sorte de "durée de vie" des particules radioactives¹.

On veut vérifier que la description probabiliste du phénomène de désintégration radioactive correspond bien : à une loi exponentielle pour N grand ; à un phénomène chaotique pour N petit.

On prendra $N(0) = 10^5$ particules, $\Delta t = 1$ s et $\lambda = \frac{1}{\tau} = 0,3$ s⁻¹. Tracer $N(t_k)$ avec $t_k = k.\Delta t$.

1. Le processus de désintégration radioactif suit une loi exponentielle. Mais cela ne veut pas dire que les atomes "vieillissent". Un atome de ¹⁴C vieux de 3000 ans a rigoureusement la même probabilité de se désintégrer qu'un atome de ¹⁴C apparu il y a 5 minutes. Il ne faut donc pas en déduire que le temps de demi-vie désigne un "âge" des atomes radio-actifs. Pour l'homme, au contraire, le taux de mortalité augmente avec l'âge : plus on est vieux, plus on a de chance de mourir. Si nous suivions une loi exponentielle (i.e. un taux de mortalité constant comme les atomes radio-actifs), une demi-vie de 75 ans correspondrait à : 25% d'une classe d'âge atteignant l'âge de 150 ans et 0,1% d'une classe d'âge atteignant 750 ans !