

Licence Creative Commons



Mis à jour le 18 septembre 2015 à 10:51

Probabilités et informatique



TABLE DES MATIÈRES

| | |
|-----------------------------------------------------|-----------|
| 1 Aléatoire ? | 5 |
| 1.1 Une machine peut-elle être aléatoire? | 6 |
| 1.2 Générateur congruentiel linéaire | 7 |
| 1.3 Tests statistiques | 8 |
| 1.3.1 Pi | 8 |
| 1.3.2 Rappel express : démonbrement for dummies | 9 |
| 1.4 RECHERCHES | 12 |
| 2 Espace probabilisé discret | 23 |
| 2.1 Probabilités? | 24 |
| 2.2 Avant la formalisation | 24 |
| 2.3 Espace probabilisable - Espace probabilisé | 25 |
| 2.4 Probabilités conditionnelles | 27 |
| 2.4.1 Un exemple pour comprendre | 27 |
| 2.4.2 Définition | 27 |
| 2.4.3 Arbre | 28 |
| 2.4.4 Formule de BAYES | 28 |
| 2.4.5 Indépendance | 29 |
| 2.5 Variables aléatoires réelles finies | 30 |
| 2.5.1 Préambule : le retour du Duc de Toscane | 30 |
| 2.5.2 Définition | 30 |
| 2.5.3 Loi de probabilité | 31 |
| 2.5.4 Variables aléatoires indépendantes | 32 |
| 2.5.5 Fonctions de répartition | 33 |
| 2.5.6 Espérance mathématique | 33 |
| 2.5.7 Variance | 34 |
| 2.5.8 Linéarité de l'espérance | 34 |
| 2.5.9 Théorème de König-Huygens | 34 |
| 2.5.10 V.a.r. centrée réduite | 35 |
| 2.5.11 Fonction indicatrice d'un évènement | 35 |
| 2.6 Quelques lois discrètes classiques | 35 |
| 2.6.1 Loi uniforme (discrète) | 35 |
| 2.6.2 Loi de Bernoulli : to be or not to be | 35 |
| 2.6.3 Loi binomiale | 36 |
| 2.6.4 Loi hypergéométrique | 36 |
| 2.6.5 Loi géométrique | 36 |
| 2.6.6 Loi de Pascal | 37 |
| 2.6.7 Loi binomiale négative | 38 |
| 2.6.8 Loi de Poisson | 38 |
| 2.7 Somme de v.a.r. indépendantes - Stabilité | 38 |
| 2.7.1 Convolution | 38 |
| 2.7.2 Somme de v.a.r. indépendantes | 39 |
| 2.7.3 Applications | 39 |
| 2.8 RECHERCHES | 40 |
| 3 Initiation aux processus aléatoires | 57 |
| 3.1 Un exemple | 58 |
| 3.2 Vision dynamique des probabilités - Automates | 58 |
| 3.2.1 Découverte | 58 |
| 3.2.2 Premier exemple - Règles de parcours | 59 |
| 3.2.3 Deuxième exemple - Valeur moyenne | 61 |
| 3.3 Comportement asymptotique des chaînes de MARKOV | 63 |
| 3.4 PageRank | 65 |
| 3.4.1 Le graphe du Web | 65 |
| 3.4.2 Les probabilités pour mieux compter | 66 |
| 3.4.3 Le modèle de téléportation de Google | 67 |
| 3.5 Exercices | 68 |

| | | |
|-------|---------------------------------|----|
| 3.5.1 | Chaînes absorbantes - automates | 68 |
| 3.5.2 | Avec le calcul matriciel | 69 |

THÈME

Aléatoire ?



1

Une machine peut-elle être aléatoire ?

L'informatique a besoin d'aléatoire dans de nombreux domaines : tri, recherche, réseau, test, infographie, jeux, cryptographie, etc.

Le problème est d'abord de savoir ce que veut dire « être aléatoire ».

Par exemple, voici une suite de 0 et de 1 :

01010101010101010101010101010101

Il y a autant de 0 que de 1. Qu'en dites-vous ?

Est-ce qu'une machine a une pièce de monnaie pour tirer à pile ou face ?

Est-ce que tirer à pile ou face est une expérience aléatoire ?

(Ici : commentaires éclairés de votre professeur adoré)

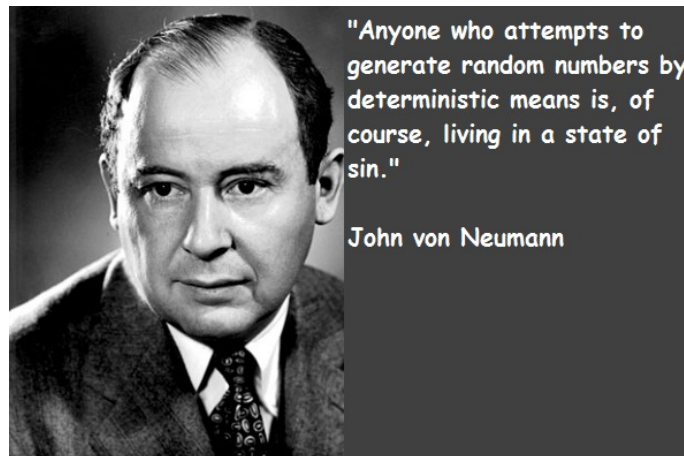
Le problème de base est d'étudier des *suites aléatoires de nombres* choisis indépendamment selon une *distribution*.

La distribution de base sur machine est la distribution *uniforme* : par exemple, vous voulez choisir au hasard un million de bits. Il se peut que les 999 999 premiers soient des 1. Cependant, la probabilité que le millionième soit encore 1 est toujours 1/2.

Depuis les années quarante, on tente de générer des suites aléatoires avec des machines.

John VON NEUMANN, qui a donné son nom à l'architecture des ordinateurs qui prévaut encore aujourd'hui, a proposé le premier un algorithme en 1946 qui consistait à prendre le carré du nombre aléatoire précédent et à en extraire le milieu.

Par exemple, si on est en train de générer une suite de nombres de 10 chiffres, que l'actuel est 5 772 156 649, son carré vaut 33 317 792 380 594 909 201 et le nouveau nombre est donc 7 923 805 949.



Pour ce module, nous allons passer du côté obscur de la force et illustrer nos propos en Python. Pour vous, je suis prêt à tous les sacrifices.

```

1 def prendMilieu(nb):
2     lnb = len(str(nb))
3     return((nb**2 // 10**(max(1,lnb // 2)) ) % 10**lnb)
4
5 def longVonNeumann(graine):
6     L = []
7     k = graine
8     while k not in L:
9         L.append(k)
10        k = prendMilieu(k)
11    return(len(L))

```

Alors :

```

1 In [45]: prendMilieu(5772156649)
2 Out[47]: 7923805949
3
4 In [48]: max([longVonNeumann(k) for k in range(900000,1000000)])
5 Out[48]: 119

```

Nous verrons en exercice quelques failles de la méthode *middle-square*.

Cependant, en cherchant bien, on a pu trouver en 1956 une suite de 750 000 nombres de 38 bits sans dégénérescence.

On peut donc avoir de bons résultats, le problème c'est qu'on ne peut le vérifier qu'a posteriori..

Jusqu'à la fin des années 70, les programmeurs ont eu tendance à éviter d'étudier ces algos et ont utilisé des générateurs sans les vérifier. Mais l'avènement du RSA et l'importance accrue de la sécurité et des algos aléatoires a changé la donne.

2 Générateur congruentiel linéaire



D. H. LEHMER
1905-1991

En 1949, D. H. LEHMER propose un générateur dépendant de quatre paramètres :

$$X_{n+1} = (aX_n + c) \pmod{m}$$

Les quatre paramètres entiers seront appelés ainsi :

- m , le module, vérifiant $0 < m$,
- a , le facteur, vérifiant $0 \leq a < m$,
- c , l'incrément, vérifiant $0 \leq c < m$,
- X_0 , la graine, vérifiant $0 \leq X_0 < m$,

Nous allons créer des générateurs congruentiels en utilisant les générateurs de Python qui sont des itérateurs particuliers car ils sont évalués paresseusement...comme en Haskell;-)

Il suffit de créer une fonction et de remplacer le `return` par `yield` : l'exécution n'est alors pas arrêtée mais seulement suspendue et prête à reprendre au même point.

```
1 def gcm(m,a,c,x):
2     while True:
3         yield x
4         x = (a*x + c) % m
```

Choisissons $m = 10$ et $X_0 = a = c = 7$:

```
1 In [105]: gen = gcm(10,7,7,7)
2
3 In [106]: gen
4 Out[106]: <generator object gcm at 0x7fc4dfd56678>
5
6 # Sur Python 2.7 next est une méthode : gen.next()
7 In [107]: next(gen)
8 Out[107]: 7
9
10 In [108]: next(gen)
11 Out[108]: 6
12
13 In [109]: next(gen)
14 Out[109]: 9
15
16 In [110]: [next(gen) for k in range(20)]
17 Out[110]: [0, 7, 6, 9, 0, 7, 6, 9, 0, 7, 6, 9, 0, 7, 6, 9, 0, 7, 6, 9]
18
19 In [111]: [x for x in itertools.islice(gen,10)]
20 Out[111]: [0, 7, 6, 9, 0, 7, 6, 9, 0, 7]
```

Danger

Tel qu'il est construit, le générateur `gen` est infini!
Il ne faudra pas demander `[x for x in gen]` !

Aparté

Les générateurs sont beaucoup plus rapides que les listes sur Python et donnent accès aux fonctions de la bibliothèque `itertools` qui sont largement inspirées de ce qui est disponible sur Haskell...;-)
Ils ont été généralisés sur Python depuis les versions 3.x

Revenons à nos moutons : cette suite n'a pas l'air très aléatoire...

Pouvez-vous faire le lien avec la [Recherche 1 - 3 page 12](#) ?

Il va falloir réfléchir pour trouver une bonne période. Malheureusement, il faudrait faire pas mal d'arithmétique.

Il existe aussi des critères d'efficacité sur machine : une division est relativement coûteuse sauf si on divise par $2^{\text{La longueur des mots du système}}$, par exemple 2^{64} sur 64 bits...

Il existe alors des théorèmes du style :

Théorème 1 - 1

Le GCM est de période m si, et seulement si :

1. c est premier avec m ;
2. $a - 1$ est un multiple de tous les diviseurs premiers de m ;
3. $a - 1$ est un multiple de 4 si m est un multiple de 4.

Mais il nous faudrait plus de temps pour pouvoir le prouver. Passons donc à la prochaine étape...

Cela nous permettrait par exemple de dire qu'avec $X_0 = 5\,772\,156\,648$, $a = 3\,141\,592\,621$, $c = 2\,718\,281\,829$ et $m = 10^{10}$ on obtient une suite de période m (ces nombres font sourire les mathématiciens...).

3 Tests statistiques

3.1 Pi

Regardez ces premières décimales de π . Y trouvez-vous une certaine symétrie ?

3,14159 26 53589 79 32 38 46 26 43 38 32 79 ...

On peut aller un peu plus loin dans l'exploration en allant voir 100 000 décimales de π et compter combien il y en a de chaque sorte. On peut utiliser la fonction `Counter` de la bibliothèque `collections`, la constante `pi` et la fonction `N` de `sympy` pour afficher des décimales de π :

```
1 In [9]: ?Counter
2 Type:          type
3 String form:   <class 'collections.Counter'>
4 File:         /usr/local/lib/python3.4/collections/__init__.py
5 Init definition: Counter(self, iterable=None, **kwds)
6 Docstring:
7 Dict subclass for counting hashable items. Sometimes called a bag
8 or multiset. Elements are stored as dictionary keys and their counts
9 are stored as dictionary values.
```

```
1 stats1Pi = Counter(str(N(pi,10**5)))
```

```
1 In [8]: stats1Pi
2 Out[8]: Counter({'1': 10137, '6': 10028, '5': 10026, '3': 10026, '7': 10025, '0': 9999,
   ... '8': 9978, '4': 9971, '2': 9908, '9': 9902, '.': 1})
```

On peut aussi chercher par paquets de 2. Pour cela, voici une pure pythonerie qui permet de grouper efficacement les éléments d'une liste par groupe de longueur n :

```
1 def parPaquets(iterable, n):
2     """
3     iter transforme un itérable en itérateur : on utilise son côté pile
4     * n répète n fois la liste
5     * iters découpe la liste iters en une suite d'arguments
6     zip crée une liste de n-uplets à partir de n listes
7     """
8     iters = [iter(iterable)] * n
9     return zip_longest(* iters)
10
11 stats2Pi = Counter( parPaquets( str( N(pi, 10**5) ), 2 ) ).values()
```



```

1 In [59]: stats2Pi
2 Out[59]: dict_values([506, 440, 483, 499, 523, 469, 522, 517, 556, 491, 469, 511, 491,
→ 538, 498, 489, 510, 488, 503, 509, 558, 536, 509, 511, 494, 523, 505, 499,
→ 504, 498, 474, 459, 518, 492, 485, 494, 462, 507, 487, 491, 536, 558, 529,
→ 499, 493, 537, 549, 465, 522, 487, 472, 498, 517, 521, 482, 483, 482, 494,
→ 503, 498, 494, 452, 501, 490, 491, 512, 514, 470, 456, 484, 515, 527, 479,
→ 512, 502, 490, 473, 468, 530, 509, 484, 1, 488, 513, 520, 556, 445, 491, 1,
→ 505, 475, 494, 463, 463, 468, 507, 501, 510, 551, 512, 534, 507])

```

C'est assez homogène.

Explorer les décimales de π est assez rigolo : on y trouve tout ce qu'on veut.

Voici un code naïf :

```

1 def chercheMotif(motif,chaîne):
2     en = enumerate(chaîne)
3     deb = motif[0]
4     n = len(motif)
5     for i,k in en:
6         if k == deb and chaîne[i:i+n] == motif:
7             return i
8     return 'Pas trouvé'

```

qui permet de trouver James *Bond* :

```

1 In [10]: chercheMotif('007',str(N(pi,10**7)))
2 Out[10]: 2807

```

Votre code secret :

```

1 In [12]: chercheMotif('1234567',str(N(pi,10**7)))
2 Out[12]: 9470345

```

Mais tout ceci est un peu du bricolage au petit bonheur : nous avons besoin d'effectuer des tests avec un peu plus de méthode.

N'oubliez pas que passer des tests signifie qu'on n'a pas encore trouvé la faille.

Nous nous contenterons d'étudier des tests *empiriques* qui implique la manipulation de groupes de nombres par la machine. Les tests théoriques sont un peu plus compliqués à étudier : vous les verrez quand vous serez plus grands...

Aparté

Quand on peut reconnaître des motifs dans une série de nombre, cela nous permet de compresser l'information. La notion moderne permettant de mesurer le degré de hasard d'une suite de nombre fait référence à cette notion informatique de compression : une suite est aléatoire si on ne peut pas la compresser plus mais c'est assez compliqué à mettre en œuvre. Nous allons explorer des voies plus classiques.

3 2 Rappel express : démonbrement for dummies

3 2 1 Nombre de permutations

Definition 1 - 1 (Permutation) Soit E un ensemble. Une permutation de E est une application bijective de E sur E . On note $\mathfrak{S}(E)$ l'ensemble des permutations de E .

Généralement, on note les permutations sous forme d'une matrice où la première ligne correspond aux éléments de E et où la deuxième ligne correspond aux images des éléments de E .

Par exemple, les permutations de $\{a, b, c\}$ sont :

$$\begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix} \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix} \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix} \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix} \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix} \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

On peut même se contenter d'écrire (a, b, c) , (a, c, b) , (b, a, c) , (b, c, a) , (c, a, b) , (c, b, a) si l'on est sûr que la première ligne est toujours (a, b, c) .

Théorème 1 - 2

Nombre de permutations

Il y a $n!$ permutations d'un ensemble de n éléments.

La démonstration s'effectue par récurrence. Il y a bien sûr $1!$ façon de permuter 1 élément.

On peut donc supposer qu'il existe au moins un entier k tel qu'il y ait $k!$ permutations d'un ensemble à k éléments.

Considérons un ensemble à $k + 1$ éléments. On en choisit 1 : il y a $k + 1$ choix possibles. Ensuite, il s'agit d'ordonner les k éléments restant ce qui laisse $k!$ possibilités.

Il y a donc finalement $(k + 1) \times k! = (k + 1)!$ permutations au total.

On retrouve dans l'exemple précédent qu'il y a $3! = 6$ permutations de $\{a, b, c\}$.

On conviendra que $0! = 1$.

On a déjà rencontré les permutations lors de notre découverte de la cryptographie. Elles jouent également un rôle important dans les algorithmes de tri par exemple et aussi dans certains algorithmes concernant les graphes que vous étudierez en deuxième année, dans les mélanges de cartes, dans l'étude du Rubik's cube, et bien d'autres domaines encore...

Remarque

Formule de Stirling

Pour n « grand », on peut estimer que :

$$n! \underset{+\infty}{\sim} \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

3 2 2 Nombre d'applications injectives - arrangements sans répétitions

Une source veut envoyer à chacun de ses cinq clients un fichier différent parmi les vingt-six situés dans un certain répertoire.

Combien de possibilités a la source pour envoyer un fichier à chacun des 26 clients ?

On peut modéliser cette situation : il s'agit de compter les applications injectives (pourquoi « injectives » ?) de l'ensemble E des clients vers l'ensemble F des fichiers. Il ne peut y avoir répétition donc il y a 26 choix de fichiers pour le premier client, 25 pour le deuxième, etc. ce qui donne $26 \times 25 \times 24 \times 23 \times 22$ possibilités.

On peut généraliser.

Definition 1 - 2 (Arrangement sans répétition) Une application injective d'un ensemble E de cardinal p dans un ensemble F de cardinal n (avec $p \leq n$) est appelée un arrangement sans répétition de p parmi n .

Un exemple fameux est la recherche de tiercés dans l'ordre (sans ex-æquo...) : pourquoi ?

Nous pouvons formuler cette propriété plus synthétiquement. En effet

$$A_n^p = n(n-1)(n-2)\dots(n-p+1) \times \frac{(n-p)(n-p-1)(n-p-2)\dots \times 2 \times 1}{(n-p)(n-p-1)(n-p-2)\dots \times 2 \times 1} = \frac{n!}{(n-p)!}$$

d'où

Nombre d'arrangements

Le nombre d'arrangements sans répétition de p parmi n est

$$A_n^p = n(n-1)(n-2)\dots(n-p+1) = \frac{n!}{(n-p)!}$$

Théorème 1 - 3

Preuve ?

3 2 3 Combinaisons sans répétition

Il n'y a pas assez de gagnants au loto. Les règles en sont donc modifiées. Il s'agit maintenant de trouver 3 numéros parmi 5. Combien y a-t-il de grilles (combinaisons) possibles ?

Par exemple, on pourrait dire que j'ai cinq manières de choisir le premier numéro, quatre choix pour le deuxième et trois choix pour le troisième, donc il y a $5 \times 4 \times 3$ grilles différentes mais dans ce cas, je compte des 3-listes ordonnées alors que les 3-listes $(1, 2, 3)$ et $(3, 2, 1)$ correspondent à la même grille ou combinaison $\{1, 2, 3\}$.

Posons une petite définition pour clarifier les débats. Donnons en fait un nom à une grille du loto, c'est à dire à un sous-ensemble (une partie) contenant p éléments d'un plus grand ensemble contenant n éléments.

Definition 1 - 3 (Combinaison sans répétition) Soit n et p deux entiers naturels et E un ensemble contenant n éléments. Un sous-ensemble de E contenant p éléments est appelé une combinaison de p éléments de E ou encore une p -combinaison d'éléments de E .

Or ce qui nous intéresse, c'est le nombre de ces combinaisons, donc introduisons une notation :

Definition 1 - 4 (Nombre de combinaisons sans répétitions) Le nombre de p -combinaisons d'un ensemble E contenant n éléments est noté $\binom{p}{|E|} = \binom{p}{n}$ ou encore $C_{|E|}^p = C_n^p$

Revenons à notre mini-loto. Considérons une grille quelconque (i.e. une 3-combinaison de l'ensemble des 5 numéros) : par exemple $\{2, 4, 5\}$. Nous avons vu dans un paragraphe précédent qu'il y a $3!$ façons d'ordonner ces nombres. Finalement, il y a $C_5^3 \times 3!$ suites de 3 nombres ordonnées : c'est le nombre d'arrangements de p parmi n . Or nous en avons comptées $5 \times 4 \times 3$ tout à l'heure. Nous en déduisons finalement que

$$C_5^3 = \frac{A_5^3}{3!} = \frac{5 \times 4 \times 3}{3!}$$

Il est alors possible de généraliser la formule suivante :

Théorème 1 - 4

$$C_n^p = \frac{A_n^p}{p!} = \frac{n(n-1)(n-2)\dots(n-(p-1))}{p!} = \frac{n!}{p!(n-p)!}$$

Preuve ?

3 2 4 Triangle de Pascal - Binôme de Newton

À l'aide des formules précédentes, on prouve facilement (faites-le !) le résultat suivant :

Théorème 1 - 5

$$C_n^p = C_n^{n-p}$$

On établit ensuite, toujours par le calcul, la relation suivante, dite **Relation de Pascal** même si les mathématiciens chinois l'avaient mise en évidence avant lui :

Relation de Pascal

Théorème 1 - 6

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$



Il est possible de démontrer cette formule à l'aide d'un raisonnement ensembliste : pour former des groupes de p éléments dans un ensemble en contenant n , on distingue un élément n_0 quelconque :

- soit le groupe le contient et alors il y a C_{n-1}^{p-1} choix des $p-1$ éléments distincts de n_0 parmi les $n-1$ restant ;
- soit le groupe ne le contient pas et il faut donc choisir p éléments parmi les $n-1$ éléments distincts de n_0 : il y a donc cette fois C_{n-1}^p .

Cette relation permet de démontrer la formule du binôme de Newton (que nous admettrons) :

Formule du binôme

Soit \mathbb{A} un anneau commutatif. Dans $\mathbb{A}[X]$, pour tout entier naturel non nul n :

Théorème 1 - 7

$$(1+X)^n = \sum_{k=0}^n C_n^k X^k$$

et plus généralement :

$$\forall (a, b) \in \mathbb{A}^2, \quad (a+b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$$

RECHERCHES

Recherche 1 - 1 Middle-square (1)

Quel nombre suit 1010 101010 par la méthode du *middle-square*.

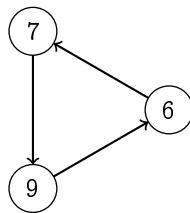
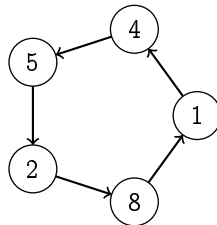
Recherche 1 - 2 Permutations

On peut être amené à rechercher les *cycles* d'une permutations. Il est aisé de les représenter à l'aide de *graphes* orientés...cycliques.

Par exemple, si l'on considère la permutation :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 8 & 3 & 5 & 2 & 9 & 6 & 1 & 7 \end{pmatrix}$$

alors on trouve trois cycles :



1. Décomposez $\pi = (2\ 1\ 3\ 5\ 4\ 7\ 9\ 6\ 8)$ en cycles.

2. La *longueur* d'un cycle est le nombre de ses sommets.

Soit p une permutation. On note $p^k = p \circ p^{k-1}$ avec $p^0 = \text{id}$, la permutation identité.

Soit $p \in \mathfrak{S}_n$. On définit une relation R sur $E = \{1, 2, \dots, n\}$ par :

$$iRj \Leftrightarrow \exists k \in \mathbb{N}^*, p^k(i) = j$$

Est-ce que R est une relation d'équivalence? Quels sont ses classes?

3. L'*ordre* d'une permutation est le plus petit entier naturel strictement positif k tel que $p^k = \text{id}$.

Déterminez l'ordre de π . Est-ce que toute permutation a un ordre fini? Comment le calculer?

Recherche 1 - 3 Middle-square (2)

Observons quelques suites obtenues par ma méthode du middle-square en modifiant la dernière ligne de longVonNeumann ainsi : `return (L,k)`.

```

1  [[500,0], 0),
2  ([501,100,0], 0),
3  ([502,200,0], 0),
4  ([503,300,0], 0),
5  ([504,401,80,40,60], 60),
6  ([505,502,200,0], 0),
  
```

```

7 ([506,603,360,960,160,560], 360),
8 ([507,704,561,472,278,728,998,600,0], 0),
9 ([508,806,963,736,169,856,273,452,430,490,10], 10),
10 ([509,908,446,891,388,54,91,28,78,8,6,3,0], 0),
11 ([510,10], 10),
12 ([511,112,254,451,340,560,360,960,160], 560),
13 ([512,214,579,524,457,884,145,102,40,60], 60),
14 ([513,316,985,22,48,30,90,10], 10),
15 ([514,419,556,913,356,673,292,526,667,488,814,259,708,126,587,456,793,884,145,102,40,60], 60),
16 ([515,522,248,150,250], 250),
17 ([516,625,62,84,5,2,0], 0),
18 ([517,728,998,600,0], 0),
19 ([518,832,222,928,118,392,366,395,602,240,760], 760),
20 ([519,936,609,88,74,47,20,40,60], 60),
21 ([520,40,60], 60),
22 ([521,144,73,32,2,0], 0),
23 ([522,248,150,250], 250),
24 ([523,352,390,210,410,810,610], 210),
25 ([524,457,884,145,102,40,60], 60),
26 ([525,562,584,105,102,40,60], 60),
27 ([526,667,488,814,259,708,126,587,456,793,884,145,102,40,60], 60),
28 ([527,772,598,760], 760),
29 ([528,878,88,74,47,20,40,60], 60),
30 ([529,984,825,62,84,5,2,0], 0),
31 ([530,90,10], 10),
32 ([531,196,841,728,998,600,0], 0),
33 ([532,302,120,440,360,960,160,560], 360),
34 ([533,408,646,731,436,9,8,6,3,0], 0),
35 ([534,515,522,248,150,250], 250),
36 ([535,622,688,334,155,402,160,560,360,960], 160),
37 ([536,729,144,73,32,2,0], 0),
38 ([537,836,889,32,2,0], 0),
39 ([538,944,113,276,617,68,62,84,5,2,0], 0),
40 ([539,52,70,90,10], 10),
41 ([540,160,560,360,960], 160),
42 ([541,268,182,312,734,875,562,584,105,102,40,60], 60),
43 ([542,376,137,876,737,316,985,22,48,30,90,10], 10),
44 ([543,484,425,62,84,5,2,0], 0),
45 ([544,593,164,689,472,278,728,998,600,0], 0),
46 ([545,702,280,840,560,360,960,160], 560),
47 ([546,811,772,598,760], 760),
48 ([547,920,640,960,160,560,360], 960),
49 ([548,30,90,10], 10),
50 ([549,140,960,160,560,360], 960)]

```

Essayons de généraliser. Supposons que l'on veuille engendrer une suite d'entiers X_0, X_1, \dots tous compris entre 0 et un certain entier m . Soit f une fonction *stable* sur $\llbracket 0, m \rrbracket$. On fabrique la suite (X_n) selon le mécanisme $X_{n+1} = f(X_n)$.

1. Est-ce que la méthode *middle-square* vérifie ces conditions? Que vaut m dans le cas précédent?
2. Démontrez que la suite est forcément périodique, c'est-à-dire qu'il existe deux entiers μ et λ tels que les entiers :

$$X_0, X_1, \dots, X_\mu, \dots, X_{\mu+\lambda-1}$$

soient distincts mais que $X_{n+\lambda} = X_n$ dès que $n \geq \mu$ (la période est alors λ). Essayez d'abord de déterminer λ et μ dans les cas observés. Est-ce que vous pouvez trouver les valeurs maxi et mini de μ et λ ?

3. Démontrez qu'il existe un entier $n > 0$ tel que $X_n = X_{2n}$ et que le plus petit n est compris entre μ et $\mu + \lambda$.
4. Déterminez un algorithme qui calcule λ et μ pour toute fonction f et toute graine X_0 utilisant seulement $O(\lambda + \mu)$ étapes et traduisez-le en Python.

```

1 In [62]: trouvePeriode(prendMilieu,511)
2 Out[62]: ([511,112,254,451,340,560,360,960,160], 560, 'mu = 5, lambda = 4')

```

On pourra par exemple tirer parti du type `dict` de Python. Un dictionnaire est un ensemble de couples (clé, valeur) indexé par la clé.

Par exemple :

```

1 In [67]: rep
2 Out[67]: {'A': 45, 'B': 72, 'Z': 789, 'C': 93}
3
4 In [68]: rep = {'A':45, 'B':72 , 'C' :93}
5
6 In [69]: rep['B']
7 Out[69]: 72
8
9 In [70]: rep['Z'] = 789
10
11 In [71]: rep
12 Out[71]: {'A': 45, 'B': 72, 'Z': 789, 'C': 93}

```

Recherche 1 - 4 GCL : disparition de la graine

Dans le premier exemple de GCL du cours, $X_0 = X_4$. Donner un exemple (voire une infinité...) avec $m = 10$ tel que X_0 n'apparaît plus jamais dans la liste.

Pensez partie

Recherche 1 - 5 GCL : cas particuliers

1. Que pensez-vous d'un GCL où $a = 0$ ou $a = 1$?
2. On note $b = a - 1$.

Démontrez que pour tous entiers naturels k et n , et tout entier $a \geq 2$, on a :

$$X_{n+k} = (a^k X_n + (a^k - 1)c/b) \pmod{m}$$

3. Que se passe-t-il si m est une puissance de 2 et si a est pair ?

Recherche 1 - 6 /dev/urandom et /dev/random

Regarder l'aide de `/dev/random` et `/dev/urandom` puis l'aide de `od`.

Que fait `od -An -N 1 -t u2 /dev/random` ?

Créez n fichier texte qui contient 100000 tirages de nombres aléatoires (quel format ?) et tracez l'histogramme correspondant puis animez-le...

Un peu d'aide :

```

1 import os
2 import numpy as np
3 import matplotlib.pyplot as plt
4
5 def tirage_uni(i):
6     fichier_alea = open('fichier_alea.dat', 'w+')
7     os.system('od ?????????? | cat > fichier_alea.dat')
8     tirage = np.array(fichier_alea.read().split()).astype(np.int)
9     plt.hist(tirage, bins=256, histtype='step', normed=True)
10    fichier_alea.close()
11
12
13 for k in range(2,8):
14     fig = plt.figure()
15     tirage_uni(k)
16     plt.savefig('./animHistoRand' + str(k) + '.jpg')
17     plt.close()
18
19 os.system('animate animHistoRand*.jpg') #nécessite imagemagick. On freine avec >

```

Expliquez et complétez les différentes lignes ainsi que les résultats obtenus.

Adaptez ce code pour tester le générateur RANDU utilisé dans les années 1960 :

$$X_{n+1} \equiv (65539 \times X_n) \pmod{2^{31}}$$

avec X_0 impair.

Que pensez-vous de RANDU ? Il a été introduit à l'époque car $65539 = 2^{16} + 3$ ce qui facilite les calculs : pourquoi ?

Calculez m^2 modulo 2^{31} en fonction de m .

Déduisez-en une relation simple entre X_{n+2} , X^{n+1} et X_n .

Si on normalise les X_n en posant $r_n = X_n/2^{31}$ qu'obtient-on ?

Si on change r_n et r_{n+1} de 1%, quel est l'amplitude du changement sur r_{n+2} ? Que peut-on en dire point de vue sécurité ?

Recherche 1 - 7 Comment tricher au Poker en ligne...

...quand les programmeurs ne sont pas des inforMATHiciens :-)

Voici l'implémentation de Random() sur les compilateurs *Borland* :

```

1 long long RandSeed = ##### ;
2
3 unsigned long Random(long max)
4 {
5     long long x ;
6     double i ;
7     unsigned long final ;
8     x = 0xffffffff;
9     x += 1 ;
10
11     RandSeed *= ((long long)134775813);
12     RandSeed += 1 ;
13     RandSeed = RandSeed % x ;
14     i = ((double)RandSeed) / (double)0xffffffff ;
15     final = (long) (max * i) ;
16
17     return (unsigned long)final;
18 }

```

Commentez...

Souvent, la graine utilisée est la valeur renvoyée par l'horloge du système, dans le cas du jeu de Poker de *PlanetPoker*, il s'agissait du nombre de millisecondes écoulées depuis minuit.

L'algo de mélange partait d'un jeu de 52 cartes ordonné : combien de mélanges sont possibles ? Cela fait à peu près combien exprimé en puissance de 2 ?

Mais à chaque mélange, le jeu est ré-ordonné et on choisit une graine selon l'algo Random() : combien de mélanges sont alors en fait possibles ?

Il faut aussi tenir compte du choix de la graine comme temps écoulé depuis minuit. Si on se donne une fourchette de 3 minutes, combien de mélanges sont maintenant possibles ?

Comment allez-vous organiser votre appli de triche sachant que vous jouez au *Texas Hold'em* : vous recevez deux cartes et il y a trois cartes affichées sur la table ?

Recherche 1 - 8 Poker sur Python

Une main au poker est constituée de 5 cartes tirées d'un jeu de 52 cartes. Deux lettres identiques (par exemple XX)

correspondent à deux cartes de même hauteur (par exemple deux dames).

1.
 - i. Combien y a-t-il de mains contenant des carrés (XXXXY) ?
 - ii. des fulls (XXXYY) ?
 - iii. des brelans (XXXYZ) ?
 - iv. des doubles paires (XXYYZ) ?
 - v. des paires (XXYZA) ?
 - vi. des quintes (cinq cartes dont les niveaux se suivent) ?
 - vii. des couleurs (5 cartes de la même couleurs quine soient pas des quintes) ?
 - viii. des quintes flush (des quintes d'une même couleur) ?

| POKER HAND RANKINGS | |
|------------------------|-------------------|
| Royal Flush | 10♥ J♥ Q♥ K♥ A♥ |
| Straight Flush | 4♣ 5♣ 6♣ 7♣ 8♣ |
| Four of a Kind | K♠ K♥ K♣ K♦ 3♠ |
| Full House | 10♥ 10♠ 10♦ A♠ A♣ |
| Flush | 10♠ K♠ 2♠ 6♠ 7♠ |
| Straight | 7♣ 8♠ 9♦ 10♠ J♥ |
| Three of a Kind | 5♠ 5♥ 5♣ J♦ A♦ |
| Two Pair | A♠ A♥ 3♣ 3♠ J♣ |
| One Pair | Q♦ Q♥ 2♥ 8♠ 9♣ |

2. Nous pouvons calculer tout ça avec Python (ou Haskell ou ...)

On rentre d'abord la liste des symboles et la liste des valeurs :

```

1 # Liste des symboles des cartes
2 symboles = ['Trefle', 'Carreau', 'Coeur', 'Pique']
3
4 # Liste des valeurs dans un jeu de 32 cartes
5 vals_32 = ['Sept', 'Huit', 'Neuf', 'Dix', 'Valet', 'Dame', 'Roi', 'As']
6
7 # Ensemble des valeurs dans un jeu de 52 cartes
8 vals_52 = ['Deux', 'Trois', 'Quatre', 'Cinq', 'Six'] + vals_32

```

Il s'agit ensuite de créer l'ensemble de toutes les mains possibles de cinq cartes.

On pourrait faire ça :

```

1 [(valeur, couleur) for valeur in vals_32 for couleur in symboles]

```

mais on va plutôt utiliser des itérateurs : c'est plus rapide.

```

1 from itertools import product
2
3 def jeu(vals):
4     """Ensemble des cartes selon les valeurs choisies (2 -> As ou 7 -> As) """
5     return product(vals, symboles)

```

On lit un peu la doc :

```

1 itertools.product(*iterables, repeat=1)
2
3 Cartesian product of input iterables.
4
5 Equivalent to nested for-loops in a generator expression.
6 For example, product(A, B) returns the same as ((x,y) for x in A for y in B).
7
8 The nested loops cycle like an odometer with the rightmost element advancing
9 on every iteration.
10 This pattern creates a lexicographic ordering so that if the input's iterables
11 are sorted, the product tuples are emitted in sorted order.
12
13 To compute the product of an iterable with itself, specify the number of
14 repetitions with the optional repeat keyword argument.
15 For example, product(A, repeat=4) means the same as product(A, A, A, A).
16

```


17 This function is equivalent to the following code, except that the actual
18 implementation does not build up intermediate results in memory:

```
19
20 def product(*args, repeat=1):
21     # product('ABCD', 'xy') --> Ax Ay Bx By Cx Cy Dx Dy
22     # product(range(2), repeat=3) --> 000 001 010 011 100 101 110 111
23     pools = [tuple(pool) for pool in args] * repeat
24     result = [[]]
25     for pool in pools:
26         result = [x+[y] for x in result for y in pool]
27     for prod in result:
28         yield tuple(prod)
```

Python lorgne un peu sur l'évaluation paresseuse de Haskell.

Pour créer l'ensemble des mains de 5 cartes, on va utiliser combinations :

```
1 itertools.combinations(iterable, r)
2
3 Return r length subsequences of elements from the input iterable.
4
5 Combinations are emitted in lexicographic sort order. So, if the input iterable is sorted,
6 the combination tuples will be produced in sorted order.
7
8 Elements are treated as unique based on their position, not on their value.
9 So if the input elements are unique, there will be no repeat values in each combination.
10
11 The number of items returned is  $n! / r! / (n-r)!$  when  $0 \leq r \leq n$  or zero when  $r > n$ .
```

itertools contient aussi une fonction map qui reprend les caractéristiques du map de Haskell.

Cela nous permet d'avoir les hauteurs d'une main :

```
1 def hauteurs(main):
2     """Liste des hauteurs d'une main"""
3     return map(hauteur,main)
```

Ensuite, il faut compter les sorties identiques. Il existe un outil Python tout fait pour ça : la fonction Counter de la bibliothèque collections. Par exemple :

```
1 In [62]: from collections import Counter
2
3 In [63]: Counter("maman papa")
4 Out[63]: Counter({'a': 4, 'p': 2, 'm': 2, ' ': 1, 'n': 1})
```

Un carré au poker contient donc un 4 dans son ensemble de valeurs (obtenu avec la méthode values en Python).

```
1 def est_carre(main):
2     """Teste si une main contient quatre cartes de même hauteur"""
3     return 4 in Counter(hauteurs(main)).values()
```

On aurait pu faire autrement :

```
1 def est_carre2(main):
2     return {1,4} == set(Counter(hauteurs(main)).values())
3
4 def est_carre3(main):
5     return [1,4] == sorted(list(Counter(hauteurs(main)).values()))
```

Déterminez les fonctions pour les full houses, double-paires, brelans, paires.

```

1 In [79]: compte_mains(est_full,vals_32)
2 Out[79]: 1344
3
4 In [80]: compte_mains(est_brelan,vals_32)
5 Out[80]: 10752
6
7 In [81]: compte_mains(est_double_paire,vals_32)
8 Out[81]: 24192
9
10 In [82]: compte_mains(est_paire,vals_32)
11 Out[82]: 107520

```

Il ne reste plus qu'à s'occuper des couleurs et des quintes en faisant attention aux quintes flush : une couleur ne doit pas être une quinte en même temps et vice-versa. Pour trouver les quintes, on peut utiliser `islice` qui, comme son nom l'indique, coupe un itérateur en tranches :

```

1 itertools.islice(iterable, start, stop[, step])
2
3 Make an iterator that returns selected elements from the iterable.
4 If start is non-zero, then elements from the iterable are skipped until
5 start is reached.
6 Afterward, elements are returned consecutively unless step is set higher
7 than one which results in items being skipped. If stop is None, then
8 iteration continues until the iterator is exhausted, if at all; otherwise,
9 it stops at the specified position. Unlike regular slicing, islice() does
10 not support negative values for start, stop, or step. Can be used to extract
11 related fields from data where the internal structure has been flattened
12 (for example, a multi-line report may list a name field on every third line).
13
14 Equivalent to:
15
16 def islice(iterable, *args):
17     # islice('ABCDEFGH', 2) --> A B
18     # islice('ABCDEFGH', 2, 4) --> C D
19     # islice('ABCDEFGH', 2, None) --> C D E F G
20     # islice('ABCDEFGH', 0, None, 2) --> A C E G
21     s = slice(*args)
22     it = iter(range(s.start or 0, s.stop or sys.maxsize, s.step or 1))
23     nexti = next(it)
24     for i, element in enumerate(iterable):
25         if i == nexti:
26             yield element
27             nexti = next(it)
28
29     If start is None, then iteration starts at zero.
30     If step is None, then the step defaults to one.

```

Par exemple :

```

1 def suites(vals):
2     """Ensemble d'ensembles de hauteurs de mains contenant une quinte"""
3     return [set(islice(vals,k,k+5,1)) for k in range(len(vals) - 4)]

```

et vous devez obtenir :

```

1 In [84]: compte_mains(est_quinte_32,vals_32)
2 Out[84]: 4080
3
4 In [85]: compte_mains(est_couleur_32,vals_32)
5 Out[85]: 208
6
7 In [86]: compte_mains(est_flush_32,vals_32)
8 Out[86]: 16

```

Vous pouvez vous amuser à faire quelque chose de semblable (en mieux) en Haskell...

Recherche 1 - 9 Téléphone



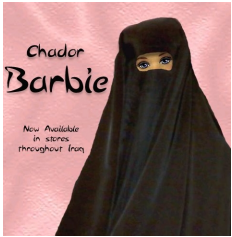
On suppose que les numéros de téléphone sont de la forme $0Z\text{-}YX\text{-}XX\text{-}XX\text{-}XX$ avec X un nombre entier entre 0 et 9, Y un nombre entier entre 2 et 9 et Z un entier entre 1 et 9.

En France, il y a environ 35 millions d'abonnements à un « fixe » et 79 millions à un « mobile » (données du deuxième trimestre 2014).

Est-on tranquille ?

Supposons qu'en Syldavie, il y ait 109 millions d'abonnés à un téléphone fixe et que le Y soit fixé par opérateur et qu'il n'y ait que deux opérateurs. De combien de zones a-t-on besoin ?

Recherche 1 - 10 Barbie



Un étudiant de l'IUT d'informatique de Klow en Syldavie a reçu pour Noël sept Barbie et sept séries de dix robes de couleurs différentes. Il habille chaque Barbie d'une de ces robes.

1. Combien y a-t-il d'habillages possibles selon la couleur ?
2. Combien y a-t-il d'habillages de sorte que chaque Barbie ait une robe de couleur différente des autres ?
3. Combien y a-t-il d'habillages de sorte qu'au moins deux Barbie aient une robe de la même couleur ?
4. Combien y a-t-il d'habillages de sorte qu'exactly deux Barbie aient une robe de la même couleur ?

Recherche 1 - 11 Jus

Lors de la soirée d'intégration de l'IUT de Klow, 9 étudiants ont bu un verre d'eau, 25 un verre de jus de tomate, 14 un verre de jus de betterave, 7 un verre de jus de tomate et un verre d'eau, 4 un verre d'eau et un verre de jus de betterave, 10 un verre de jus de betterave et un verre de jus de tomate. Les étudiants d'Info s'occupaient des boissons et ceux de GEA des entrées or ces derniers ont perdu leur liste et ne savent plus combien d'étudiants ont participé à la soirée. Les étudiants d'Info, connaissant les statistiques sur les consommations de boissons, pourront-ils corriger la gaffe des GEA sachant également que lors de la valse, tous les participants dansaient en couple ?

Recherche 1 - 12 k-uplets

Soit E un ensemble fini non vide de cardinal n .

- a) Combien y a-t-il de parties de E formées de k éléments ?
- b) Combien y a-t-il de k -uplets d'éléments de E ?
- c) Combien y a-t-il de k -uplets d'éléments deux à deux distincts de E ?
- d) Combien y a-t-il de k -uplets d'éléments deux à deux distincts de E , tel que le premier élément est le plus petit et le dernier élément est le plus grand ?
- e) Combien y a-t-il de k -uplets d'éléments de E ordonnés dans l'ordre strictement croissant ?

Déterminez des fonctions Python qui effectuent ces calculs en fonction d'un ensemble donné en paramètre.

Recherche 1 - 13 Arrangements

Comment modéliser le tiercé brillamment organisé par le PMU ? (Qui est E ? Qui est F ? Quelle est l'application de E dans F ?) Combien y a-t-il de tiercés possibles (sans ex-æquo) dans une course de 19 partants ?

Modélisez la recherche du nombre d'arrangements de p parmi n à l'aide d'un tirage de boules dans une urne.

Recherche 1 - 14 L'âge du capitaine

Le capitaine des pompiers de New-York est marié et a quatre enfants dont un mange un yaourt aux fruits tous les matins. Il ne fume pas, aime regarder des films de gladiateurs et réside à l'angle de la 1^{ère} avenue et de la 33^{ème} rue. La caserne se trouve à l'angle de la 9^{ème} avenue et de la 40^{ème} rue. Il s'y rend tous les jours à pied en sifflant « \mathcal{D} du schner Westermald » et sans perdre de temps (i.e. dans le sens des numéros croissants aussi bien pour les rues que pour les avenues). Sachant qu'il a commencé à travailler le jour de ses 18 ans, et sachant qu'il n'est jamais passé deux fois par le même chemin, qu'il boîte légèrement de la jambe droite après avoir participé au championnat d'Écosse

de lancement d'enclume, qu'il est sourd de l'oreille gauche depuis qu'il a plongé dans un lac gelé pour sauver son petit frère John qui était en train de se noyer après que la glace sur laquelle il patinait craqua, quel est l'âge (maximum) du capitaine ?

Recherche 1 - 15 Mélange de cartes : épisode 1



On dispose d'un jeu de 52 cartes et on le mélange en *queue d'aronde* : on divise le jeu en deux parties de 26 cartes et on intercale les cartes de la moitié gauche dans la moitié droite sans modifier l'ordre de chaque partie bien sûr. Est-ce que quatre mélanges en queue d'aronde successifs permettent de fournir un ordre aléatoire des cartes ?

Il faudrait préciser d'abord cette notion : combien y a-t-il d'ordres possibles des cartes ?

Existe-t-il alors des ordres qui ne pourront être obtenus par notre succession de mélanges ? Pour répondre à cette question, il faudrait calculer le nombre d'ordres qu'on peut obtenir par la succession de quatre mélanges... Imaginez que vous disposiez de 26 cartes bleues et 26 rouges. Vous les mélangez avec la méthode de la queue d'aronde : pouvez-vous reconstituer le jeu initial ? Combien y a-t-il alors d'ordres possibles après le premier mélange ?

Recherche 1 - 16 Mélange de carte : le mélange de ASF Software

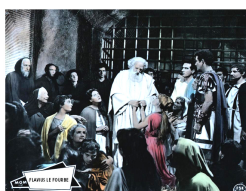
Afin de convaincre les utilisateurs que leur algorithme de mélange était juste, la société ASF Software, qui produit les logiciels utilisés par de nombreux sites de jeu, avait publié cet algorithme :

```

1  procedure TDeck.Shuffle;
2  var
3      ctr: Byte;
4      tmp: Byte;
5
6      random_number: Byte;
7  begin
8      { Fill the deck with unique cards }
9      for ctr := 1 to 52 do
10         Card[ctr] := ctr;
11
12     { Generate a new seed based on the system clock }
13     randomize;
14
15     { Randomly rearrange each card }
16     for ctr := 1 to 52 do begin
17         random_number := random(51)+1;
18         tmp := card[random_number];
19         card[random_number] := card[ctr];
20         card[ctr] := tmp;
21     end;
22
23     CurrentCard := 1;
24     JustShuffled := True;
25 end;
```

1. Considérez un jeu de 3 cartes. Dressez l'arbre de tous les mélanges possibles en suivant cet algorithme. Que remarquez-vous ?
2. Proposez un algorithme qui corrige ce problème. Dressez l'arbre correspondant pour un jeu de trois cartes.

Recherche 1 - 17 Flavius Josèphe



Flavius Josèphe (37 - 100) ou plutôt Yossef ben Matityahou HaCohen est un historien romain d'origine juive et de langue grecque et qui ne devait pas être trop mauvais en mathématiques si on en croit la sinistre anecdote suivante. Lors de la première guerre judéo-romaine, Yossef fut piégé dans une grotte avec 39 autres de ses compagnons en juillet 67. Ne voulant pas devenir esclaves, ils mirent au point un algorithme d'auto-destruction : il s'agissait de se mettre en cercle et de se numéroter de 1 à 40. Chaque septième devait être tué jusqu'à ce qu'il n'en reste plus qu'un qui devait alors se suicider. Ce dernier fut Yossef lui-même...et il ne se suicida pas ! Après

deux ans de prison, il fut libéré, il entra au service des romains comme interprète et acquit la citoyenneté romaine deux ans plus tard. Quel numéro portait Yossef? Créez un programme Python (ou Haskell..) qui donne l'ordre des exécutions quelque soit le nombre de prisonniers et le nombre sinistre.

Dans une autre situation (pour ne pas vous donner la réponse quand même...), on doit obtenir avec 20 compagnons et 6 comme nombre sinistre suite :

```

1 La 1e victime est le no 6
2 La 2e victime est le no 12
3 La 3e victime est le no 18
4 La 4e victime est le no 4
5 La 5e victime est le no 11
6 La 6e victime est le no 19
7 La 7e victime est le no 7
8 La 8e victime est le no 15
9 La 9e victime est le no 3
10 La 10e victime est le no 14
11 La 11e victime est le no 5
12 La 12e victime est le no 17
13 La 13e victime est le no 10
14 La 14e victime est le no 8
15 La 15e victime est le no 2
16 La 16e victime est le no 9
17 La 17e victime est le no 16
18 La 18e victime est le no 13
19 La 19e victime est le no 1
20 Out[62]: 'Et le survivant est le no20'
```

Peut-on prévoir qui sera le survivant sans dresser de liste ?

Recherche 1 - 18 7 manières de calculer un coefficient binomial avec Python

Déterminez au moins 5 méthodes pour calculer $\binom{p}{n}$ et testez les différents temps obtenus avec `%timeit`.

Dressez ensuite un tableau où on retrouve le coefficient $\binom{j}{i}$ à la ligne i et la colonne j .

Transformez ce tableau en remplaçant les nombres pairs par des espaces et les nombres impairs par des % par exemple : que voyez-vous ?

Espace probabilisé discret

INCROYABLE!
 RECALÉ AU BAC,
 IL JOUE SES NOTES
 AU LOTO ET GAGNE
 37 MILLIONS!



« Le nom seul de calcul des probabilités est un paradoxe : la probabilité, opposée à la certitude, c'est ce qu'on ne sait pas, et comment peut-on calculer ce que l'on ne connaît pas? »

Henri POINCARÉ - La Science et l'hypothèse (1908)

« [La théorie des probabilités] donne les aperçus les plus sûrs qui puissent nous guider dans nos jugements. [...] Il n'est point de science plus digne de nos méditations, et qu'il soit plus utile de faire entrer dans le système de l'instruction publique. »

Pierre Simon de LAPLACE - Essai philosophique sur les probabilités (rédigé de 1795 à 1825)

« The true logic of this world lies in the calculus of probabilities »

James C. MAXWELL - The Scientific Letters and Papers of James C. Maxwell, Vol. 1, 1846-1862

« We cannot predict whether a given photon will arrive at A or B. All we can predict is that out of 100 photons that come down, an average of 4 will be reflected by the front surface. Does this mean that physics, a science of great exactitude, has been reduced to calculating only the probability of an event, and not predicting exactly what will happen? Yes. That's a retreat, but that's the way it is : Nature permits us to calculate only probabilities. Yet science has not collapsed. »

Richard P. FEYNMAN - The strange theory of light and matter (1985)

1 Probabilités ?

Le calcul des probabilités est très important en informatique et dans toute science en général. Il nous permet de résoudre des problèmes tout en étant confrontés à l'incertitude ou à des systèmes très complexes.

Contrairement aux clichés, la théorie des probabilités, en tant que discipline mathématique, est dotée de toute la rigueur et de la précision possibles. Elle puise ses sources dans tous les champs des mathématiques. Il n'est dès lors pas étonnant de constater que les prix mathématiques les plus prestigieux sont attribués depuis quelque temps à des travaux liés aux probabilités.

En informatique, leur usage est proportionnel à la complexité des problèmes abordés.

Un informaticien dirait qu'en probabilité le modèle est totalement spécifié, permet et impose la plus grande rigueur suite à un raisonnement déductif.

Il ne faut pas confondre les probabilités avec la statistique qui complète un modèle inconnu à l'aide d'observations (raisonnement inductif) et fait plutôt appel à l'intuition qui n'est pas la principale qualité d'un ordinateur...

Enfin, s'il est tout à fait possible de calculer des probabilités sans faire de statistique, le contraire est quasiment impossible c'est pourquoi nous aborderons la statistique inférentielle après avoir étudié les probabilités.

2 Avant la formalisation

Essayons, grâce à notre petite expérience probabiliste acquise au lycée, de dégager les grandes lignes de ce que pourrait être une théorie formalisée, nécessaire au « dialogue » avec une machine. Lors de l'étude des probabilités, on est amené à effectuer une expérience aléatoire, c'est-à-dire un processus dont on ne peut prévoir à l'avance le résultat qui peut donner des résultats différents même si l'on répète l'expérience dans des conditions qui semblent identiques.

Nous allons considérer par la suite le lancement d'un dé cubique équilibré et tout et tout.

On doit d'abord connaître l'ensemble des résultats possibles : c'est l'univers que nous noterons par la suite Ω .

Par exemple, si je lance une fois le dé, $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$.

Si je lance deux fois de suite le dé et que je note dans l'ordre le numéro des faces, j'obtiens dans ce cas $\Omega_2 = \{11, 12, 13, \dots, 66\}$

Si je lance les dés jusqu'à obtenir deux 6 à la suite, le nombre de lancers n'est maintenant pas fixé à l'avance et l'univers est a priori infini mais en demeurant discret : chaque événement peut en effet être « numéroté ». Plus rigoureusement, Ω_3 peut être mis en bijection avec \mathbb{N} : $\Omega_3 = \{1366, 3546213466, 55236221466, \dots\}$.

Il faut maintenant étudier la notion d'« événement ^a ». Disons dans un premier temps qu'il s'agit d'une assertion vérifiable relativement au résultat d'une expérience aléatoire. Il est donc « associé » à un sous-ensemble (une partie) de Ω . Un élément de Ω sera associé à un événement élémentaire.

Reprenons le deuxième exemple. Voici quelques événements :

- « Obtenir deux fois 3 » : $E_a = \{33\}$. C'est un événement élémentaire ;
- « Ne pas obtenir deux fois 3 » : $E_b = \Omega_2 \setminus E_a \subseteq \Omega_2$;
- « Obtenir deux faces dont la somme des numéros fait 11 » : $E_c = \{56, 65\} \subseteq \Omega_2$;
- « Obtenir deux faces dont le numéro est inférieur à 7 » : $E_d = \Omega_2$. Il s'agit donc de l'événement certain ;
- « Obtenir deux faces dont la somme des numéros fait 37 » : $E_e = \emptyset$. Il s'agit cette fois de l'événement impossible.

On comprend que l'univers lui-même et l'ensemble vide sont des événements. On voudrait que la réunion et l'intersection d'événements soit encore un événement et que le complémentaire de la partie associée à un événement définisse encore un événement : les mathématiciens (français...)

a. Depuis 1990, l'Académie française préconise de mettre un accent grave au deuxième e d'événement qui doit donc s'écrire événement...

disent alors que l'ensemble de tous ces événements forment une **tribu**. Dans le reste du monde, on parle plutôt de σ -algèbre de Boole : les ensembles concernés vérifient en effet certaines propriétés *algébriques*.

Il reste enfin à définir un moyen de **mesurer** les probabilités : c'est une sorte de degré de certitude correspondant à une échelle graduée de 0 à 1, 0 correspondant au degré de l'évènement impossible et 1 au degré de l'évènement certain. Il est important de comprendre que ce degré de certitude est donné **a priori** et permet donc de prévoir l'issue de l'expérience avant de la réaliser.

Il faudrait que cette mesure assure une cohérence entre les événements et les degrés.

Par exemple, si $A \subseteq B$, on aimerait que $\Pr(A) \leq \Pr(B)$. On aimerait également que l'union de deux parties de l'univers se « transforme » en somme de leur degré de certitude, mais on « sent » bien que cela ne pourra être vrai que si les événements sont totalement **incompatibles** : $A \cap B = \emptyset \rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$. On retrouve ici cette notion algébrique de « transport » des opérations : la notion d'*isomorphisme*.

Considérons pour illustrer ces dernières notions le premier exemple du lancer unique d'un dé. Notre univers Ω_1 contient 6 éléments et prenons comme tribu naturellement associée les $2^6 = 64$ parties de Ω_1 qui forment l'ensemble $P(\Omega_1)$.

Imposons pour notre modèle que la mesure du degré de certitude ne dépende pas des numéros des faces mais uniquement du cardinal de la partie considérée : chacun des six événements élémentaires aura donc la même probabilité : $\Pr(\{1\}) = \Pr(\{2\}) = \dots = \Pr(\{6\}) = \frac{1}{6}$.

La probabilité d'obtenir un numéro pair sera donc, sachant que l'évènement est associé à $A = \{2, 4, 6\}$: $\Pr(A) = \Pr(\{2\}) + \Pr(\{4\}) + \Pr(\{6\}) = \frac{1}{2}$.

Mais si le dé est truqué ? Si l'univers est infini ? S'il n'est pas dénombrable ?

Il est temps de définir clairement et rigoureusement les notions afin de ne pas dire de bêtises...

3

Espace probabilisable - Espace probabilisé

Dans ce chapitre, Ω est un ensemble dénombrable. Nous le considérerons parfois fini.

Definition 2 - 1 (Tribu) Une tribu T_Ω définie sur un univers Ω est une partie de $P(\Omega)$ qui vérifie :

1. $\Omega \in T_\Omega$;
2. $A \in T_\Omega \rightarrow C_\Omega A \in T_\Omega$;
3. $\forall I \subseteq \mathbb{N}, (A_i)_{i \in I} \in T_\Omega^{|I|} \rightarrow \bigcup_{i \in I} A_i \in T_\Omega$.

Attention ! Une tribu est une partie de $P(\Omega)$: c'est donc un ensemble de sous-ensembles de Ω !

Definition 2 - 2 (Espace probabilisable) Soit T_Ω une tribu définie sur l'univers Ω . On dit que (Ω, T_Ω) est un **espace probabilisable** (ou **espace d'évènements**). Les éléments de la tribu sont appelés **évènements**.

Attention ! Pour une certaine expérience aléatoire, il peut y avoir de nombreux espaces probabilisables différents.

Par exemple, en reprenant un des exemples introduits dans la section précédente :

- $(\Omega_1, \{\emptyset, \Omega_1\})$ est un espace probabilisable : on ne pourra s'intéresser qu'aux évènements « obtenir un numéro entre 1 et 6 » et « ne pas obtenir de numéro entre 1 et 6 » ce qui est un peu sommaire ;
- $(\Omega_1, \{\emptyset, A, C_\Omega A, \Omega_1\})$ avec A associé à « obtenir un 6 » ;
- $(\Omega_1, P(\Omega_1))$ est ce qu'on peut faire de plus précis.

Vérifiez qu'il s'agit bien d'espaces probabilisables !

Definition 2 - 3 (Axiomes de КОЛМОГОРОВ) Soit (Ω, T) un **espace probabilisable**. On appelle **mesure de probabilité** (ou **probabilité tout court**) une application \Pr définie sur T vérifiant :

1. $\forall A \in T, \Pr(A) \in [0, 1]$;
2. $\Pr(\Omega) = 1$;

3. $\forall (A, B) \in T \times T, A \cap B = \emptyset \rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$

On dit alors que (Ω, T, \Pr) est un espace probabilisé.



Андрей Николаевич
КОЛМОГОРОВ
(1903 - 1987)

Vous pouvez vérifier que l'espace probabilisable $(\Omega, P(\Omega))$ avec $\Omega = \{1, 2, 3, 4, 5, 6\}$ muni de l'application :

$$\Pr: \begin{matrix} P(\Omega) & \rightarrow & [0, 1] \\ A & \mapsto & \frac{|A|}{6} \end{matrix}$$

définit un espace probabilisé.

Pour l'axiome 3, on se souviendra de la formule du crible (cf théorème ?? page ??).

Comme Ω est dénombrable, il existe une certaine famille I d'entiers telle que $\Omega = \bigcup_{i \in I} \{\omega_i\}$ et pour tout évènement A de T , il existe une famille $J \subseteq I$ telle que $A = \bigcup_{j \in J} \{\omega_j\}$.

Alors, par application itérée du troisième axiome :

$$\Pr(A) = \Pr\left(\bigcup_{j \in J} \{\omega_j\}\right) = \sum_{j \in J} \Pr(\{\omega_j\})$$

Ainsi la donnée des valeurs prises par \Pr sur les évènements élémentaires suffit à caractériser \Pr sur T .

Dans le cas où Ω est fini et si tous les évènements élémentaires ont la même probabilité (on dit qu'il y a équiprobabilité), alors :

$$\Pr(\omega_1) = \Pr(\omega_2) = \dots = \Pr(\omega_{|\Omega|}) = \frac{1}{|\Omega|}$$

Dans ce cas, \Pr est appelée probabilité uniforme sur Ω .



Attention ! Toute probabilité n'est pas uniforme... Votre portable sonne : quelle est la probabilité que cela soit Monica B. qui vous fasse une déclaration d'amour ? La tribu contient deux évènements en plus de l'évènement certain et de l'évènement impossible : « c'est Monica » et « ce n'est pas Monica ». Le rapport entre cas favorables et cas possibles est bien $\frac{1}{2}$ et pourtant...

Vous noterez enfin que lorsque l'univers est infini dénombrable, il ne peut pas y avoir équiprobabilité : Pourquoi?...

Principales propriétés

Soit (Ω, T, \Pr) un espace probabilisé. La probabilité \Pr vérifie :

1. $\Pr(C_\Omega A) = 1 - \Pr(A)$;
2. $\Pr(\emptyset) = 0$;
3. $\forall (A, B) \in T \times T, A \subset B \implies \Pr(A) \leq \Pr(B)$;
4. $\forall (A, B) \in T \times T, \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$ (cas particulier de la formule du crible de POINCARÉ).

Théorème 2 - 1

Dans la suite, quand il n'y aura pas ambiguïté, on notera $C_\Omega A$ par \bar{A} .

Démonstration?...

Definition 2 - 4 (Système complet d'évènements) Une famille A_1, A_2, \dots, A_n d'évènements non vides de T forme un système complet d'évènements si, et seulement si :

- $\forall i \neq j, A_i \cap A_j = \emptyset$ (évènements incompatibles deux à deux) ;
- $\bigcup_{i=1}^n A_i = \Omega$.

On dit aussi que la famille (A_i) forme une partition de Ω .

Théorème 2 - 2

Théorème des probabilités totales (version 1)

Soit $(A_i)_{1 \leq i \leq n}$ un système complet d'évènements, alors :

$$\forall E \in T, \Pr(E) = \sum_{i=1}^n \Pr(E \cap A_i)$$

Preuve ?

4 Probabilités conditionnelles

Dans toute cette section on travaille dans l'espace probabilisé (Ω, T, \Pr) .

4.1 Un exemple pour comprendre

Considérons l'expérience simplissime consistant à lancer deux fois un dé à six faces. L'univers Ω est donc constitué de l'ensemble des couples (i, j) , avec i et j appartenant à l'ensemble $[[1, 6]]$: il y a donc 36 éléments dans Ω . Intéressons nous à la somme des deux chiffres et soit A l'évènement « le total fait neuf ». On prendra $T = P(\Omega)$.

$$A = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$$

donc

$$\Pr(A) = \frac{4}{36} = \frac{1}{9}$$

Soit B l'évènement : « on obtient 3 au premier lancer », alors

$$B = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\}$$

donc

$$\Pr(B) = \frac{6}{36} = \frac{1}{6}$$

Jusqu'ici, tout roule. Posons-nous maintenant le problème suivant : sachant que B est réalisé, i.e. que l'on obtient 3 au premier lancer, quelle est la probabilité que le total fasse 9 ? On ne considère plus tous les éléments de Ω . Il semble alors nécessaire de définir un nouvel univers (notre nouvel ensemble des possibles) :

$$\Omega' = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\} = \Omega \cap B$$

Puisque l'univers change, la probabilité aussi. L'évènement A' « le total fait neuf » dans ce nouveau modèle s'écrit

$$A' = \{(3, 6)\} = A \cap B$$

et donc

$$\Pr'(A') = \frac{1}{6} = \frac{\text{nombre d'éléments de } A \cap B}{\text{nombre d'éléments de } B} = \frac{\Pr(A \cap B)}{\Pr(B)}$$

4.2 Définition

Cela nous incite à poser la définition suivante :

Definition 2 - 1 (Probabilité conditionnelle) Soient A et B deux évènements d'un espace probabilisé muni d'une probabilité \Pr , avec $\Pr(B) \neq 0$. La probabilité de A sachant B est définie par

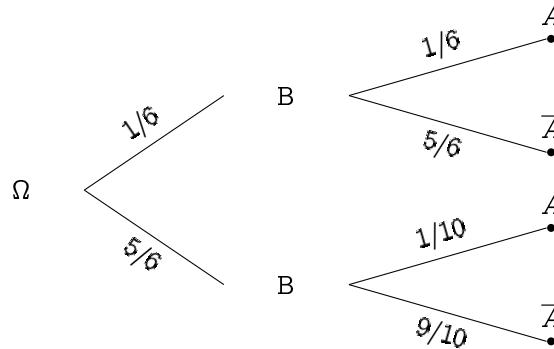
$$\Pr(A | B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

On parle de probabilité conditionnelle. Mais est-ce bien une probabilité ? C'est-à-dire est-ce que $(\Omega, T, \Pr(\cdot | B))$ est un espace probabilisé ? D'après la définition 2 - 3 page 25, il y a trois conditions à vérifier : faites-le !

4 3 Arbre

Reprenons l'exemple de la section 2.4.1 page précédente en répondant aux mêmes questions à l'aide d'un arbre *pondéré*, c'est à dire un arbre dont chaque branche est marquée de la probabilité (du *poids*) correspondant.

Alors la somme des probabilités de chaque « ramification » est égale à 1.



À l'aide de la formule des probabilités totales, il est aisé d'obtenir les résultats suivant :

1. Calcul de $\Pr(A)$

B et \bar{B} forment une partition. D'après la formule des probabilités totales,

$$\Pr(A) = \Pr(B \cap A) + \Pr(\bar{B} \cap A) = \Pr(B) \times \Pr(A | B) + \Pr(\bar{B}) \times \Pr(A | \bar{B})$$

$$\Pr(A) = \frac{1}{6} \times \frac{1}{6} + \frac{5}{6} \times \frac{1}{10} = \frac{1}{9}$$

2. Calcul de $\Pr(\bar{A})$

B et \bar{B} forment une partition. D'après la formule des probabilités totales,

$$\Pr(\bar{A}) = \Pr(B \cap \bar{A}) + \Pr(\bar{B} \cap \bar{A}) = \Pr(B) \times \Pr(\bar{A} | B) + \Pr(\bar{B}) \times \Pr(\bar{A} | \bar{B})$$

$$\Pr(\bar{A}) = \frac{1}{6} \times \frac{5}{6} + \frac{5}{6} \times \frac{9}{10} = \frac{8}{9}$$

3. Probabilités conditionnelles

$$\Pr(B | A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{\frac{1}{36}}{\frac{1}{9}} = \frac{1}{4}$$

$$\Pr(B | \bar{A}) = \frac{\Pr(B \cap \bar{A})}{\Pr(\bar{A})} = \frac{\frac{5}{36}}{\frac{8}{9}} = \frac{5}{32}$$

$$\Pr(\bar{B} | A) = \frac{\Pr(\bar{B} \cap A)}{\Pr(A)} = \frac{\frac{1}{12}}{\frac{1}{9}} = \frac{3}{4}$$

$$\Pr(\bar{B} | \bar{A}) = \frac{\Pr(\bar{B} \cap \bar{A})}{\Pr(\bar{A})} = \frac{\frac{3}{4}}{\frac{8}{9}} = \frac{27}{32}$$

4 4 Formule de BAYES

Cette formule correspond au « retournement » d'un arbre.

Deuxième formule de Bayes (probabilité des causes)

Soit (Ω, T, \Pr) un espace probabilisé et $(A_i)_{i \in I \subseteq \mathbb{N}}$ un système complet d'évènements.

Théorème 2 - 3

$$\forall B \in T, \quad \Pr(A_k | B) = \frac{\Pr(B | A_k) \times \Pr(A_k)}{\sum_{i \in I} \Pr(B | A_i) \times \Pr(A_i)}$$

Preuve ?

On utilisera souvent ce théorème dans le cas particulier de deux évènements :

Première formule de Bayes

Théorème 2 - 4

$$\Pr(A | B) = \frac{\Pr(B | A) \times \Pr(A)}{\Pr(B)}$$

On aurait pu écrire :

$$\Pr(\text{cause} \mid \text{effet}) = \frac{\Pr(\text{effet} \mid \text{cause}) \times \Pr(\text{cause})}{\Pr(\text{effet})}$$

Statistiquement, on a accès aux effets connaissant les causes (c'est le déterminisme...) mais dans la réalité, on subit les effets et on voudrait en connaître la cause.

Le cas le plus célèbre d'utilisation de la *statistique bayésienne* est le test de grossesse. Les médecins « disposent » d'un échantillon de femmes enceintes ou non (et donc de $\Pr(E)$ et $\Pr(\bar{E})$) et d'un test de grossesse. Ils font subir le test à tout l'échantillon et disposent donc de statistiques sur $\Pr(T \mid E)$, $\Pr(\bar{T} \mid \bar{E})$, $\Pr(T \mid \bar{E})$ et $\Pr(\bar{T} \mid E)$ mais ce qui intéresse le médecin et surtout la femme qui utilise le test c'est $\Pr(E \mid T)$ et $\Pr(\bar{E} \mid \bar{T})$...

4 5 Indépendance

La notions d'évènements indépendants est l'une des difficultés du calcul des probabilités. Comme souvent, cette notion purement mathématique renvoie, par son appellation, à une notion intuitive utilisée dans le langage courant. Il faut bien garder en mémoire que le mot du vocabulaire courant est souvent un « faux ami ». ^b Donnons tout d'abord sa définition.

Definition 2 - 2 (Évènements indépendants) Les évènements A et B sont dits indépendants si $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$

Ainsi, dans notre exemple de dés, on a $\Pr(A) \times \Pr(B) = \frac{1}{6} \times \frac{6}{36} = \frac{1}{6} = \Pr(A \cap B)$, donc le lancer du 2^{ème} dé est indépendant du premier, ce qui est rassurant.

Cette notion est purement abstraite et ne renvoie qu'à des propriétés mathématiques dont la principale est :

Indépendance et probabilités conditionnelles

Théorème 2 - 5

Soient A et B deux évènements tels que $\Pr(B) \neq 0$. Les évènements A et B sont indépendants si et seulement si $\Pr(A \mid B) = \Pr(A)$

On peut faire dire ce que l'on veut à des probabilités selon le modèle choisi.

Supposons que sur un groupe de 100 personnes, 20 portent des sous-vêtements en polystyrène expansé (pour avoir chaud l'hiver), 50 se grattent la tête avec l'index gauche et 10 font les deux à la fois. On met ces 100 personnes dans une boîte et on en tire une au hasard. Vérifiez que les évènements « la personne tirée porte des sous-vêtements en polystyrène expansé » et « la personne tirée se gratte la tête avec l'index gauche » sont indépendants.

Étudiez le même problème en considérant cette fois-ci que 15 personnes se grattent la tête avec l'index gauche (pourquoi pas). ^c

Veillez à ne pas confondre évènements *indépendants* et évènements *incompatibles*.

On peut montrer d'ailleurs que deux évènements incompatibles de probabilité non nulle ne sont jamais indépendants.

En effet, A et B sont incompatibles si, et seulement si, $A \cap B = \emptyset$, donc $\Pr(A \cap B) = \Pr(\emptyset) = 0$. Si $\Pr(A) \neq 0$ et $\Pr(B) \neq 0$, on a forcément $\Pr(A \cap B) \neq \Pr(A) \times \Pr(B)$.

La seule idée à retenir est que, si A et B sont indépendants, avoir observé la réalisation de A ne modifie pas la probabilité d'une éventuelle réalisation de B .

^b. On aurait aussi bien parler de *szjwrtrpfgklance*, mais cela aurait été plus difficile à prononcer.

^c. De façon générale, la définition probabiliste de l'indépendance est plus large que la notion intuitive.

Ainsi, en supposant que la Française des Jeux n'utilise pas de boules truquées, on peut considérer que deux tirages successifs du loto sont indépendants.

Indépendance mutuelle

Soit $(A_i)_{i \in I \subseteq \mathbb{N}}$ une suite d'événements de T .

1. On dit que ces événements sont **indépendants deux à deux** si, et seulement si :

$$\forall (i, j) \in I^2, i \neq j, \Pr(A_i \cap A_j) = \Pr(A_i) \times \Pr(A_j)$$

Théorème 2 - 6

2. On dit que ces événements sont **mutuellement indépendants** si, et seulement si, pour toute partie J de I on a :

$$\Pr\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \Pr(A_j)$$

5 Variables aléatoires réelles finies

On travaillera *en général* dans l'espace probabilisé (Ω, T, \Pr) avec $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ un univers dénombrable, voire le plus souvent fini de cardinal n .

5 1 Préambule : le retour du Duc de Toscane

Lors de l'étude du problème du Duc de Toscane, on ne s'intéressait pas aux événements eux-mêmes (les numéros des faces des dés) mais plutôt à la valeur de la somme des faces.

Au lieu de s'intéresser à 6^3 événements équiprobables, on a en fait considéré 16 événements non équiprobables (somme = 3 ou 4 ou 5 ou ... ou 18).

5 2 Définition

L'ensemble $P(\mathbb{R})$ est bien une tribu mais il pose quelques problèmes lors d'une étude plus approfondie des probabilités. Dans \mathbb{R} , on utilisera donc une tribu particulière qui correspond en fait à la tribu à laquelle on pense « naturellement », celle créée à partir des intervalles :

Definition 2 - 1 (Tribu borélienne) On note $\mathcal{B}(\mathbb{R})$ l'ensemble des parties engendrées (par intersections et réunions dénombrables) par les intervalles de \mathbb{R} .

Vous vérifierez qu'il s'agit d'une tribu.

Definition 2 - 2 (Variable aléatoire réelle) On appelle *variable aléatoire réelle finie* toute application de Ω dans \mathbb{R} vérifiant :

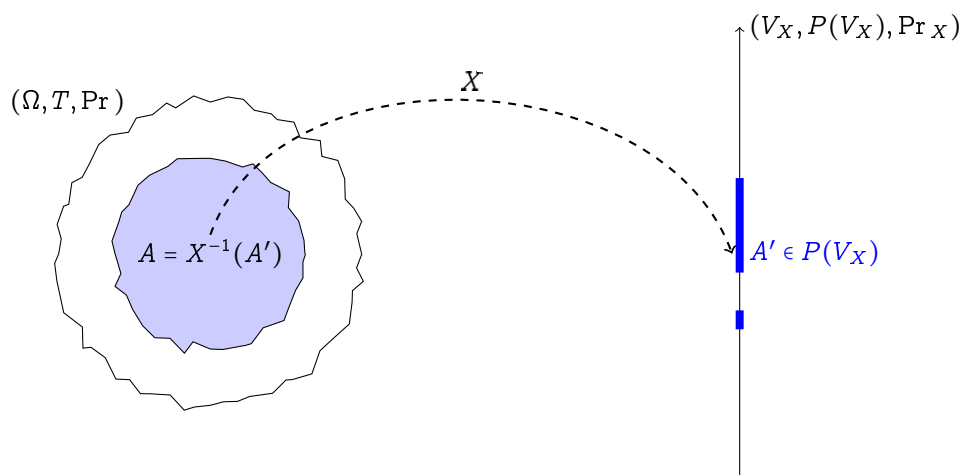
$$\forall I \subseteq \mathbb{R}, X^{-1}(I) \in T$$

Cette définition nous assure que $X^{-1}(I)$ est un événement dont on peut calculer la probabilité. Une v.a. est réelle car X est à valeurs dans \mathbb{R} et finie car l'univers est fini donc l'ensemble image de X aussi.

Danger

Vous noterez qu'une variable aléatoire réelle (v.a.r. pour les intimes) est en fait une fonction!...

Nous noterons $V_X = X(\Omega)$ l'ensemble des valeurs prises par la v.a. X .



Soit x_1, \dots, x_k les différentes valeurs prises par la fonction X .

On note $\{X = x_i\}$ l'événement « la variable aléatoire prend la valeur x_i ». Il se note rigoureusement $\{\omega \in \Omega \mid X(\omega) = x_i\}$, ce qui se lit « l'ensemble des ω tels que $X(\omega) = x_i$ » ou encore $X^{-1}(\{x_i\})$.

La v.a.r. X induit une nouvelle probabilité (ou plutôt loi de probabilité comme nous allons le voir) \Pr_X sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Par exemple, $\Pr_X(\{2\}) = \Pr(\{\omega \in \Omega \mid X(\omega) = 2\})$.

La définition précédente n'est pas très parlante, il faut essentiellement retenir que l'on va s'intéresser à des événements décrits par des nombres, les variables aléatoires nous permettent de décrire, et donc de noter, facilement des événements à l'aide de nombres.

Si V_X est réduit à un singleton, $V_X = \{\alpha\}$, on dit que X est une variable certaine, on note alors $X = \alpha$.

5 3 Loi de probabilité

Definition 2 - 3 (Loi de probabilité d'une v.a.r.) Soit X une variable aléatoire réelle sur Ω . On appelle loi de probabilité de X la fonction \Pr_X de $\mathcal{B}(\mathbb{R})$ dans $[0, 1]$ définie par

$$\Pr_X: \begin{array}{l} \mathcal{B}(\mathbb{R}) \rightarrow [0, 1] \\ B \mapsto \Pr(\{X^{-1}(B)\}) = \Pr(\{\omega \in \Omega \mid X(\omega) \in B\}) = \Pr(X(\Omega) \in B) = \Pr_X(B) \end{array}$$

Vérifiez que \Pr_X vérifie les axiomes de Kolmogorov donc que \Pr_X est bien une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Dans la plupart des cas, on étudiera $\Pr(X = x)$, c'est-à-dire $\Pr(\{\omega \in \Omega \mid X(\omega) = x\})$.

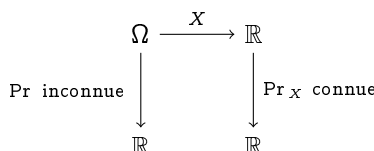
Si $x \notin X(\Omega)$, alors $\{X = x\} = \emptyset$ et donc $\Pr(\{X = x\}) = 0$.

Définir la loi de probabilité d'une expérience aléatoire reviendra donc à :

- déterminer toutes les valeurs possibles x_1, \dots, x_n prises par X ;
- calculer les probabilités p_1, \dots, p_n des événements correspondants ;
- regrouper les résultats dans un tableau du type :

| | | | | |
|-----------------------------------------------|-------|-------|-----|-------|
| Valeurs prises par X | x_1 | x_2 | ... | x_n |
| Probabilité correspondante $\Pr(\{X = x_i\})$ | p_1 | p_2 | ... | p_n |

Vous n'oublierez pas de vérifier que $p_1 + \dots + p_n = 1$ d'après la loi des probabilités totales. En général, on se trouve dans la situation suivante :



À retenir

Dans la vraie vie, on connaît les valeurs prises par la v.a.r. mais on ne connaît pas bien Ω , voire pas du tout. On réfléchit donc sur les lois et non pas sur des ensembles comme au lycée.

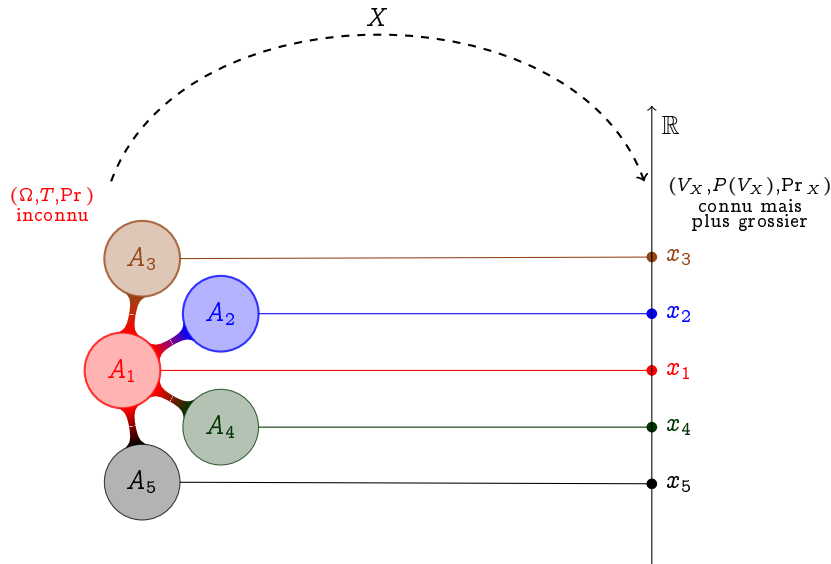
Par exemple, on peut vouloir réfléchir au nombre de plantages d'un réseau. On estime statistiquement les probabilités de tomber 0, 1, 2, etc. fois en panne dans une journée. On ne sait pas trop ce qu'est Ω mais on peut parler de la loi de X .

En fait, une variable aléatoire « transporte » les calculs de probabilités d'un univers inconnu vers des valeurs réelles connues.

La v.a.r. a pour effet de remplacer la tribu T par une tribu $X^{-1}(P(V_X))$ en général plus petite : elle condense l'information sur l'univers en une information plus grossière (numérique) mais c'est déjà bien car en général on ne connaît de toute manière pas ou peu l'univers de départ !...

On retiendra donc que le « praticien » travaille la plupart du temps avec des lois de probabilités et non pas des ensembles : cela simplifie son travail car il peut travailler sur des espaces complexes (l'univers de départ est souvent trop complexe voire inconnu) sans les connaître.

L'information donnée par la loi est souvent plus grossière : par exemple, on lance deux dés et on regarde la somme des deux faces. La tribu de départ contient 2^{36} évènements alors que la tribu de la loi n'en contient que 2^{11} (il y a 11 sommes possibles).



À partir de la loi, on induit un système complet d'évènements sur Ω : la connaissance de la loi nous donne des renseignements sur l'univers mais souvent de manière grossière (par « paquets »).

Danger

Deux v.a.r. peuvent avoir la même loi sans être égales ! On lance n fois une pièce non truquée. Considérez X le nombre de pile tombés et $n - X$ le nombre de face. Ces deux v.a.r. suivent la même loi binomiale de mêmes paramètres n et $1/2$ et pourtant elles ne sont pas égales (ni même égales presque partout ce qui signifie sauf sur un évènement de probabilité nulle mais ceci est une autre histoire).

5 4 Variables aléatoires indépendantes

La notation $\Pr(X = x_i, Y = y_j)$ signifie usuellement $\Pr(\{X = x_i\} \cap \{Y = y_j\}) = \Pr(X = x_i \text{ et } Y = y_j)$.

On peut donc définir un vecteur aléatoire $Z = (X, Y)$ tel que

$$\Pr(Z = (x_i, y_j)) = \Pr(X = x_i, Y = y_j)$$

On peut se représenter l'espace produit à l'aide d'une représentation cartésienne : on parle alors d'évènements « parallèles » à des « axes » (les Ω_i), ces « axes » étant « orthogonaux ». Un système complet d'évènements est alors une « base ». Tous ces ponts de vocabulaire reliant plusieurs théories est un plaisir évident pour le mathématicien qui pourra faciliter la vie de l'informaticien mais vous pouvez laisser cela de côté dans un premier temps...

Definition 2 - 4 (Couple de variables indépendantes) Les v.a.r. X et Y sont dites indépendantes si, et seulement si, pour toute partie $(A, B) \in V_X \times V_Y$,

$$\Pr(X \in A, Y \in B) = \Pr(X \in A) \cdot \Pr(Y \in B)$$

Definition 2 - 5 (Variables aléatoires mutuellement indépendantes) Si $X_1, X_2, X_3, \dots, X_k$ sont k v.a. définies sur Ω , on dit qu'elles sont mutuellement indépendantes si, et seulement si, quels que soit $(I_1, I_2, \dots, I_k) \in V_{X_1} \times V_{X_2} \times \dots \times V_{X_k}$ on a :

$$\Pr(X_1 \in I_1, X_2 \in I_2, \dots, X_k \in I_k) = \Pr(X_1 \in I_1) \times \Pr(X_2 \in I_2) \times \dots \times \Pr(X_k \in I_k)$$

Attention ! L'indépendance deux à deux de variables aléatoires n'entraîne pas leur indépendance mutuelle, comme nous le verrons en TD.

5 5 Fonctions de répartition

Definition 2 - 6 (Fonction de répartition) Soit X une v.a.r. On appelle fonction de répartition de X la fonction F de \mathbb{R} dans \mathbb{R} définie par :

$$F_X(x) = \Pr(X \leq x) = \Pr_X([\!-\infty, x])$$

On note souvent cette fonction F_X .
Voici quelques propriétés que vous démontrerez :

Théorème 2 - 7

Propriétés de F_X

1. $\forall x \in \mathbb{R}, F_X(x) \in [0, 1]$
2. F_X est croissante (sens large).
3. F_X est continue à droite en tout point.
4. $\lim_{x \rightarrow -\infty} F_X(x) = 0$
5. $\lim_{x \rightarrow +\infty} F_X(x) = 1$
6. $\Pr(a < X \leq b) = \Pr_X(]a, b]) = F_X(b) - F_X(a)$
7. $F_X(a) = \Pr(X \leq a) = \Pr_X([\!-\infty, a]) = 1 - \Pr(X > a)$
8. $\Pr(X = a) = \lim_{x \rightarrow a^+} F_X(x) - \lim_{x \rightarrow a^-} F_X(x) = F_X(a) - \lim_{x \rightarrow a^-} F_X(x)$. Par conséquent en tout point a où F_X est continue on a $\Pr(X = a) = 0$.

5 6 Espérance mathématique

Definition 2 - 7 (Espérance mathématique) On appelle espérance mathématique de la variable aléatoire discrète X le nombre noté $\mathbb{E}(X)$ défini par

$$\mathbb{E}(X) = x_1 \cdot \Pr_X(x_1) + x_2 \cdot \Pr_X(x_2) + \dots + x_n \cdot \Pr_X(x_n) = \sum_{i=1}^n x_i \cdot \Pr_X(x_i)$$

Exemple : on lance un dé honnête. On définit la variable aléatoire X qui prend la valeur 2 si le numéro du dé est pair et 1 sinon.

Notons ω_i l'événement « le numéro de la face est i », alors $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ et

| | | | | | | |
|-----------------|-----|-----|-----|-----|-----|-----|
| i | 1 | 2 | 3 | 4 | 5 | 6 |
| $\Pr(\omega_i)$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |
| $X(\omega_i)$ | 1 | 2 | 1 | 2 | 1 | 2 |

D'après le théorème précédent, on a $\mathbb{E}(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 1 \times \frac{1}{6} + 2 \times \frac{1}{6} = \frac{3}{2}$
Nous aurions pu procéder autrement pour définir la loi de probabilité :

| | | |
|------------------------------|-----|-----|
| Valeurs x_i prises par X | 1 | 2 |
| $\Pr_X(x_i)$ | 1/2 | 1/2 |

alors, d'après la définition, $\mathbb{E}(X) = 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}$

Espérance de $f \circ X$

Soit X une v.a.r. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $f \circ X$ soit une v.a.r. Alors

Théorème 2 - 8

$$\mathbb{E}(f \circ X) = \sum_{n \in \mathbb{N}} f(x_n) \cdot \Pr(X = x_n) = \sum_{n \in \mathbb{N}} f(x_n) \cdot \Pr_X(x_n)$$

Cette propriété sera admise dans un premier temps.

5 7 Variance

La variance mesure l'« écart » par rapport à l'espérance en faisant la somme des carrés des « distances » entre chaque valeur prise par la variable aléatoire et l'espérance pondérées par la probabilité correspondante, ce qui donne :

$$\text{Definition 2 - 8 (Variance)} \quad \mathbb{V}(X) = \sum_{i=1}^n (x_i - \mathbb{E}(X))^2 \cdot \Pr(X = x_i) = \sum_{i=1}^n (x_i - \mathbb{E}(X))^2 \cdot \Pr_X(x_i)$$

On a choisi d'utiliser les carrés de manière arbitraire pour ne pas avoir de problèmes de signes ; on aurait pu choisir une autre méthode mais celle-ci a l'avantage de rappeler la distance euclidienne bien connue. La variance est en ce sens homogène au carré d'une distance. On va donc définir une distance proprement dite en en prenant la racine carrée : c'est ce qu'on appelle l'écart-type.

$$\text{Definition 2 - 9 (Écart-type)} \quad \sigma(X) = \sqrt{\mathbb{V}(X)}$$

Vous pouvez obtenir espérance, variance et écart-type très simplement à l'aide des modules statistiques de vos calculatrices. Il suffit de rentrer les valeurs prises par la variable aléatoire en liste 1, les probabilités correspondantes en liste 2.

5 8 Linéarité de l'espérance

À partir de variables aléatoires existantes, on peut en créer de nouvelles. Avec des notations usuelles on obtient

- $aX + b : \omega_i \mapsto aX(\omega_i) + b$ avec a et b des réels.
- $X + Y : \omega_i \mapsto X(\omega_i) + Y(\omega_i)$

On peut alors démontrer les propriétés suivantes :

Linéarité de l'espérance

Théorème 2 - 9

$$\begin{aligned} \text{— } \mathbb{E}(aX + b) &= \sum_{i=1}^n (ax_i + b) \cdot \Pr_X(x_i) = a\mathbb{E}(X) + b \\ \text{— } \mathbb{E}(X + Y) &= \mathbb{E}(X) + \mathbb{E}(Y) \end{aligned}$$

Les démonstrations sont similaires. Notons $Y = aX + b$.

Pour tout $\omega_i \in \Omega$ on a $Y(\omega_i) = aX(\omega_i) + b$ et donc $y_i = ax_i + b$ avec les notations usuelles donc $\{\omega \in \Omega | Y(\omega) = y_i\} = \{\omega \in \Omega | aX(\omega) + b = ax_i + b\} = \{\omega \in \Omega | X(\omega) = x_i\}$

$$\text{Alors } \mathbb{E}(Y) = \sum_{i=1}^n y_i \Pr(Y = y_i) = \sum_{i=1}^n (ax_i + b) \Pr(X = x_i) = a \sum_{i=1}^n x_i \Pr(X = x_i) + b \sum_{i=1}^n \Pr(X = x_i)$$

On en déduit que $\mathbb{E}(Y) = a\mathbb{E}(X) + b$.

5 9 Théorème de König-Huygens

Johann KÖNIG (1712-1757) fut un mathématicien allemand, élève de Jean BERNOULLI et Christian HUYGENS (1629-1695) était lui néerlandais.

Théorème de König-Huygens

Théorème 2 - 10

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

À démontrer en TD...

5 10 V.a.r. centrée réduite

Definition 2 - 10 Soit X une v.a.r. qui admet une variance non nulle. On appelle v.a.r. centrée réduite associée à X la v.a.r. :

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$$

Cette définition en est-elle une ? Que valent $\mathbb{E}(X^*)$ et $\mathbb{V}(X^*)$?

5 11 Fonction indicatrice d'un évènement

Soit A un évènement aléatoire appartenant à T_Ω . Définissons la variable aléatoire X :

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

On note cette variable aléatoire $X = 1_A$ et on l'appelle fonction indicatrice de A .

Interprétation d'une probabilité en terme d'espérance d'un v.a.

Théorème 2 - 11

$$\mathbb{E}(1_A) = \Pr(A)$$

Preuve ?

6 Quelques lois discrètes classiques

6 1 Loi uniforme (discrète)

C'est la loi qui correspond au tirage équiprobable d'un jeton blanc dont on note le numéro dans une urne qui contient n jetons blancs numérotés.

Definition 2 - 1 (Loi uniforme) On dit que $X : \Omega \rightarrow \llbracket 1, n \rrbracket$ suit une loi uniforme sur $\llbracket 1, n \rrbracket$ et on note $X \sim U(\llbracket 1, n \rrbracket)$ si :

$$X(\Omega) = \llbracket 1, n \rrbracket \quad \text{et} \quad \forall k \in \llbracket 1, n \rrbracket, \Pr(X = k) = \frac{1}{n}$$

Calculons l'espérance et la variance de $X \sim U(\llbracket 1, n \rrbracket)$.

$$\mathbb{E}(X) = \sum_{k=1}^n k \Pr(X = k) = \sum_{k=1}^n k \frac{1}{n} = \frac{1}{n} \sum_{k=1}^n k = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$$

De manière analogue, on obtient que $\mathbb{V}(X) = \frac{n^2-1}{12}$ sachant que $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.

6 2 Loi de Bernoulli : to be or not to be



On place des jetons blancs et des noirs dans une urne. On suppose que toutes les conditions sont réunies pour effectuer le tirage d'un jeton dans le cadre d'équiprobabilité et que la proportion de jetons blancs est p . Que dire de la variable aléatoire qui vaut 1 si le jeton tiré est blanc et 0 s'il est noir ?

Definition 2 - 2 (Loi de Bernoulli) On dit que la v.a.r. finie X suit la loi de Bernoulli de paramètre p si :

- $p \in]0; 1[$, $X(\Omega) = \{0, 1\}$;
- $\Pr(X = 1) = p$.

On note alors $X \sim B(1, p)$

On dit aussi que X est une variable aléatoire de Bernoulli. Sa loi de probabilité est simple à décrire :

| | | |
|----------------|-----|---|
| x_i | 0 | 1 |
| $\Pr(X = x_i)$ | 1-p | p |

Il est simple de déterminer son espérance et sa variance :

- $\mathbb{E}(X) = 0 \times \Pr(X = 0) + 1 \times \Pr(X = 1) = p$;
- $\mathbb{V}(X) = (0 - p)^2 \times \Pr(X = 0) + (1 - p)^2 \times \Pr(X = 1) = p^2(1 - p) + (1 - p)^2 p = p(1 - p)$.

6 3 Loi binomiale

Dans les mêmes conditions que précédemment, on effectue n tirages d'un jeton avec remise. À chaque tirage on associe une v.a.r. indépendante des autres qui suit la loi de Bernoulli de paramètre p . Les issues élémentaires ω de ces n tirages sont des « mots » de n lettres, chaque lettre étant un B (pour blanc) ou un \bar{B} . On définit alors la variable aléatoire X à valeurs dans $\{0, 1, 2, \dots, n\}$, donnant le nombre de jetons blancs de ces issues élémentaires, c'est à dire que X est la somme des v.a.r. finies élémentaires suivant la loi de Bernoulli de manière indépendante.

Definition 2 - 3 (Loi binomiale) On dit que la v.a.r. finie X suit la loi binomiale de paramètres n et p si $n \in \mathbb{N} \setminus \{0\}$, $p \in]0; 1[$, $X(\Omega) = \llbracket 1, n \rrbracket$ et si l'expérience est une suite de n expériences de Bernoulli de paramètre p menées indépendamment.

On note alors $X \sim B(n, p)$

On s'intéresse maintenant à l'événement $(X = k)$, c'est à dire l'ensemble des mots de n lettres écrits avec k B et $n - k$ \bar{B} . Notons B_j l'événement « obtenir un jeton blanc au $j^{\text{ème}}$ tirage ». Chaque tirage d'une boule étant indépendant des autres, on a

$$\Pr(\omega) = \Pr(B_1) \times \dots \times \Pr(B_k) \times \Pr(\bar{B}_{k+1}) \times \dots \times \Pr(\bar{B}_n) = p^k (1 - p)^{n-k}$$

comme l'illustre cet arbre :

Il reste à déterminer combien il y a de tels mots. Ce sont les anagrammes (permutations distinctes) du mot de n lettres $BB\dots B\bar{B}\bar{B}\dots\bar{B}$. Il y en a $n!$ qu'il faut diviser par les $k!$ permutations des B et les $(n - k)!$ permutations des \bar{B} , c'est-à-dire C_n^k . Finalement, on obtient :

Théorème 2 - 12

Pour toute v.a.r. finie $X \sim B(n, p)$, on a, pour tout $k \in X(\Omega)$,

$$\Pr(X = k) = C_n^k p^k (1 - p)^{n-k}$$

Vous démontrerez que $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1 - p)$.

6 4 Loi hypergéométrique

Modèle : on considère une urne contenant N boules ($N \in \mathbb{N}^*$), la proportion de boules blanches dans l'urne étant p . On tire simultanément n boules (les boules sont choisies au hasard) et X est la v.a. qui indique le nombre de boules blanches tirées. On dit que X suit la loi hypergéométrique de paramètres N , n et p notée $H(N, n, p)$.

On retiendra :

- X peut prendre toute valeur entière de l'intervalle $\{\text{Max}\{0, n - N(1 - p)\}, \text{Min}\{n, Np\}\}$
- $k \in V_X$, $\Pr(\{X = k\}) = \frac{C_{Np}^k C_{N(1-p)}^{n-k}}{C_N^n}$
- $\mathbb{E}(X) = np$, $\mathbb{V}(X) = np(1 - p) \frac{N - n}{N - 1}$

6 5 Loi géométrique

Modèle : on répète indéfiniment une épreuve de Bernoulli dont la probabilité de succès est p . Les répétitions étant mutuellement indépendantes, X désigne le rang de l'essai donnant pour la première fois un succès. On dit que X suit la loi géométrique de paramètre p et on écrit $X \sim G(p)$. On retiendra :

$$\begin{aligned} V_X &= \mathbb{N}^* \\ k \in \mathbb{N}^*, \Pr(\{X = k\}) &= p(1 - p)^{k-1} \\ \mathbb{E}(X) &= \frac{1}{p}, \mathbb{V}(X) = \frac{1 - p}{p^2} \end{aligned}$$

La loi géométrique modélise donc le temps d'attente jusqu'au premier succès.



Problème du collectionneur

Voici une application classique de la loi géométrique : en Syldavie, le petit Ivan collectionne les images du chocolat Wonka. Il y a n images différentes. Lorsqu'il aura obtenu les n images, il aura droit de visiter les usines Wonka. À chaque fois qu'il achète une tablette, la probabilité d'obtenir une image quelconque est $1/n$. Soit T_n le temps d'attente (en nombre de tablettes achetées) pour obtenir les n images. Calculez $\mathbb{E}(T_n)$.

Soit X_i la variable aléatoire correspondant au temps d'attente pour passer de $i-1$ images différentes à i images différentes. Quand on est dans l'état $i-1$, on dispose de $i-1$ images différentes. Si on tombe sur l'une des $i-1$ images dont on dispose (avec la probabilité $\frac{i-1}{n}$), on reste dans l'état $i-1$, sinon, on passe à l'état i (avec la probabilité $\frac{n-(i-1)}{n}$).

Ainsi, $X_i \sim G(p_i)$ avec $p_i = \frac{n-(i-1)}{n}$. On en déduit que $\mathbb{E}(X_i) = \frac{1}{p_i} = \frac{n}{n+1-i}$.

Or, $T_n = X_1 + X_2 + \dots + X_n$, donc $\mathbb{E}(T_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n)$:

$$\mathbb{E}(X_n) = \sum_{i=1}^n \frac{1}{p_i} = n \sum_{i=1}^n \frac{1}{n+1-i} = n \left(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{2} + 1 \right)$$

Par exemple, avec $n = 10$:

Avec Caml

```

1 let liste_entiers = fun min max ->
2   let rec listentier = fun m acc ->
3     if m > max then
4       acc
5     else
6       listentier (m + 1) (m :: acc)
7   in listentier min [];;
8
9 let somme_r = List.fold_left (+.) (0.) ;;
10
11 somme_r ((List.map (fun x -> 1. /. (int_of_float x)) (liste_entiers 1 10)))
12
13 - : float = 29.2896825396825378

```

Avec Haskell :

```

1 *Main> 10 * sum [1 / k | k <- [1..10]]
2 29.289682539682538

```

Avec Scala :

```

1 scala> 10 * (for (x <- List.range(1,11)) yield 1. / x).sum
2 res1: Double = 29.289682539682538

```

Avec Python

```

1 In [1]: 10 * sum(1 / k for k in range(1,11))
2 Out[1]: 29.289682539682538

```

Il faut en moyenne acheter 30 tablettes pour obtenir les dix images.

6 6 Loi de Pascal

Modèle : on répète indéfiniment une épreuve E de Bernoulli de probabilité de succès p .

Les répétitions étant mutuellement indépendantes, X désigne le rang de l'essai amenant le $r^{\text{ème}}$ succès. On dit que X suit la loi de Pascal de paramètres r et p et on écrit $X \sim P_r(r, p)$. On retiendra (démonstration en TD) :

$$\begin{aligned}
 V_X &= \{r, r+1, r+2, r+3, \dots\} \\
 k \in V_X, \Pr(\{X = k\}) &= C_{k-1}^{k-r} p^r (1-p)^{k-r} \\
 \mathbb{E}(x) &= \frac{r}{p}, \mathbb{V}(X) = \frac{r(1-p)}{p^2}
 \end{aligned}$$

Si $r = 1$ on retrouve la loi géométrique.

6 7 Loi binomiale négative

Modèle : on répète indéfiniment une épreuve E de Bernoulli de probabilité de succès p . Les répétitions étant mutuellement indépendantes, X désigne le nombre d'échecs précédant le $r^{ème}$ succès. On dit que X suit la loi binomiale négative de paramètres r et p et on écrit $X \rightsquigarrow J(r, p)$. On retiendra (démonstration en TD) :

$$V_X = \mathbb{N}$$

$$k \in \mathbb{N}, \Pr(\{X = k\}) = C_{k+r-1}^k p^r (1-p)^k$$

$$\mathbb{E}(x) = \frac{r(1-p)}{p}, \mathbb{V}(X) = \frac{r(1-p)}{p^2}$$

6 8 Loi de Poisson

La variable aléatoire X suit la loi de Poisson de paramètre λ si X peut prendre toute valeur entière k avec la probabilité $\Pr(\{X = k\}) = \frac{\lambda^k e^{-\lambda}}{k!}$. On retiendra aussi : $\mathbb{E}(X) = \lambda$ et $\mathbb{V}(X) = \lambda$.

6 8 1 Que modélise la loi de Poisson ?

Considérons une probabilité définie par :

$$p_j(\pi_n, n) = \begin{cases} C_n^j (\pi_n)^j (1 - \pi_n)^{n-j} & \text{si } j \leq n \\ 0 & \text{si } j > n \end{cases}$$

Ainsi, cette suite de probabilités est une extension sur \mathbb{N} tout entier de la loi binomiale $B(n, \pi_n)$. *Supposons* que la suite (π_n) tende vers 0 quand n tend vers l'infini de telle sorte que

$$\lim_{n \rightarrow +\infty} n\pi_n = \lambda \in \mathbb{R}_+^*$$

Il faut également se souvenir que :

$$\forall t \in \mathbb{R}, \lim_{n \rightarrow +\infty} \left(1 + \frac{t}{n}\right)^n = e^t$$

Alors on montre (au tableau pendant l'amphi...) que :

$$\lim_{n \rightarrow +\infty} p_j(n, \pi_n) = \frac{\lambda^j e^{-\lambda}}{j!}$$

Ainsi, la loi de Poisson modélise bien le nombre d'apparitions de phénomènes rares (le nombre de pannes sur un systèmes GNU, le nombre de sessions sans panne d'un système Windows, le nombre d'étudiants d'info travaillant leur cours, etc.) dans une suite infinie d'évènements ($\pi_n \sim \frac{\lambda}{n}$ est « petit » et n est « grand »).



Siméon-Denis POISSON
(1781 - 1840)

7 Somme de v.a.r. indépendantes - Stabilité

7 1 Convolution

Le produit de convolution intervient dans de nombreux domaines, particulièrement en probabilités et en théorie du signal. On l'utilise sous sa forme continue (à l'aide d'une intégrale qui est une somme...continue) ou discrète (à l'aide d'une somme...discrète).

On peut l'exprimer de manière universelle avec un symbole « \int » bien pratique car souvent les propriétés des sommes discrètes et continues sont les mêmes (par exemple, l'espérance d'une loi continue ou discrète peut s'exprimer de la même manière).

$$f(x) \otimes g(x) = \int f(t) \cdot g(x - t) = \int f(x - t) \cdot g(t)$$

signifie que cet opérateur \otimes signifie à la fois :

$$f(x) \otimes g(x) = \int_{\mathbb{R}} f(t) \cdot g(x - t) dt = \int_{\mathbb{R}} f(x - t) \cdot g(t) dt$$

et

$$f(n) \otimes g(n) = \sum_{k \in \mathbb{Z}} f(k) \cdot g(n - k) = \sum_{k \in \mathbb{Z}} f(n - k) \cdot g(k)$$

En particulier, avec l'écriture usuelle des suites :

$$u_n \otimes v_n = \sum_{k \in \mathbb{Z}} u_k \cdot v_{n-k} = \sum_{k \in \mathbb{Z}} u_{n-k} \cdot v_k$$

7 2 Somme de v.a.r. indépendantes

Somme de v.a.r. indépendantes
Si X et Y sont deux v.a.r. indépendantes

Théorème 2 - 13

$$\Pr(X + Y = k) = \sum_{i \in \mathbb{Z}} \Pr(X = i) \cdot \Pr(Y = k - i) = \sum_{i \in \mathbb{Z}} \Pr(X = k - i) \cdot \Pr(Y = i) = \Pr(X = k) \otimes \Pr(Y = k) =$$

Ceci reste à démontrer...en TD of course!

7 3 Applications

7 3 1 Loi binomiale comme somme de lois de Bernoulli

Commencez par considérer la somme de 1 variable aléatoire indépendante de Bernoulli de même paramètre p... puis généralisez par récurrence au cas de n variables aléatoires de Bernoulli de même paramètre p.

Cela nous donnera également un moyen d'obtenir l'espérance de la loi binomiale B(n, p).

Remarque

Une loi de Bernoulli de paramètre p est une loi binomiale de paramètre (1, p)...

7 3 2 Stabilité de la loi binomiale

Montrez que si X ~ B(n, p) et Y ~ B(m, p) sont deux v.a.r. indépendantes alors X + Y ~ B(n + m, p).

7 3 3 Stabilité de la loi de Poisson

$$\left. \begin{array}{l} X \sim P(\lambda) \\ Y \sim P(\mu) \\ X \text{ et } Y \text{ indépendantes} \end{array} \right\} \Rightarrow Z = X + Y \sim P(\lambda + \mu)$$

Démonstration : cherchons la loi de probabilité de Z. Comme les v.a. sont indépendantes

$$\begin{aligned} \Pr(Z = k) &= \sum_{i=1}^k \Pr(X = i) \cdot \Pr(Y = k - i) \\ \Pr(Z = k) &= \sum_{i=1}^k \frac{\lambda^i e^{-\lambda}}{i!} \cdot \frac{\mu^{k-i} e^{-\mu}}{(k-i)!} = e^{-(\lambda+\mu)} \sum_{i=1}^k \frac{1}{k!} \frac{k!}{i!(k-i)!} \lambda^i \mu^{k-i} \\ \Pr(Z = k) &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=1}^k C_k^i \lambda^i \mu^{k-i} = \frac{(\lambda + \mu)^k}{k!} \cdot e^{-(\lambda+\mu)} \end{aligned}$$

Preuve de $\lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n}\right)^n = e^x$

Montrez que $\lim_{u \rightarrow 0} \frac{\ln(1+u)}{u} = 1$ en reconnaissant un taux de variation. Alors

Aparté

$$\begin{aligned} (1 + x/n)^n &= e^{n \ln(1+x/n)} \\ &= e^{n \frac{\ln(1+x/n)}{x/n} \cdot x/n} \\ &= e^{x \cdot \frac{\ln(1+x/n)}{x/n}} \end{aligned}$$

Conclusion ?

RECHERCHES

Recherche 2 - 1 Formule du crible de Poincaré : cas de trois évènements

Déterminez $\Pr(A \cup B \cup C)$ en fonction des probabilités de $A, B, C, A \cap B, A \cap C, B \cap C$ et $A \cap B \cap C$.

Recherche 2 - 2 Un peu de technique

Soit (Ω, T, \Pr) un espace probabilisé. Déterminez n tel que, pour tous évènements A et B de T ,

$$\Pr(A \cup B) + \Pr(A \cup \bar{B}) + \Pr(\bar{A} \cup B) + \Pr(\bar{A} \cup \bar{B}) = n$$

Déterminez également les signes de $\Pr(A \cap B) - \Pr(A) + \Pr(\bar{B})$ et de $\Pr(A \cap B \cap C) - 1 + \Pr(\bar{A}) + \Pr(\bar{B}) + \Pr(\bar{C})$.

Recherche 2 - 3 Définition d'une probabilité

Soit $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ un ensemble d'éléments tous distincts. On suppose que, pour tout entier naturel non nul k , $\Pr(\{\omega_k\}) = \frac{\lambda}{37^k}$ avec λ un réel.

Existe-il une valeur de λ telle que $(\Omega, P(\Omega), \Pr)$ soit un espace probabilisé ?

Recherche 2 - 4 Paradoxe des anniversaires et collisions dans un réseau

Quelle est la probabilité qu'au moins deux personnes dans cette salle aient leur anniversaire le même jour (on négligera les années bissextiles...)?

Essayons maintenant de généraliser ce problème pour étudier les allocations dynamiques d'adresses dans un réseau. On suppose que les allocations d'adresses pour chaque nœud d'un réseau se font indépendamment des autres nœud : il y a donc un risque de collision car des nœuds peuvent choisir la même adresse.

Par exemple, le protocole IPv4 permet aux nœuds de choisir une adresse codée sur 32 bits.

Supposons donc qu'un réseau contienne n nœuds et que leurs adresses puissent être codées sur a bits.

Notons p la probabilité qu'il y ait au moins une collision.

Calculez $1 - p$ en fonction de a et n puis le logarithme népérien de ce nombre.

Supposons que n soit « petit » devant 2^a .

Vous vous rappelez du résultat trouvé l'an passé lors de l'étude des polynômes de TAYLOR :

$$\ln(1 + u) \approx u \quad \text{pour } u \text{ « petit »}$$

Utilisez-le pour déterminer une expression de p sous la forme $1 - e^{f(n,a)}$.

Déduisez-en une estimation du nombre maximum de nœuds avec $a = 32$ et $p = 2^{-20}$.

Déterminez une estimation de la taille des adresses en bits si le réseau contient 2^{32} nœuds.

Recherche 2 - 5 Duc de Toscane

Cosme II de Médicis (Florence 1590-1621), Duc de Toscane, fut le protecteur de l'illustre Gallilée (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.

Observons nous-même sans perdre trop d'argent mais avec seulement 2 dés pour aller plus vite :

```

1 from numpy.random import randint
2 from collections import Counter
3 sObs = [sum( randint(1,7,size = 2) ) for k in range(144000)]
4 cObs = Counter(sObs)

```

On obtient le dictionnaire des effectifs observés :

```

1 Counter({7: 23996, 8: 20029, 6: 19934, 9: 15942, 5: 15767, 4: 12152,
2         10: 12030, 11: 8011, 3: 8003, 12: 4095, 2: 4041})

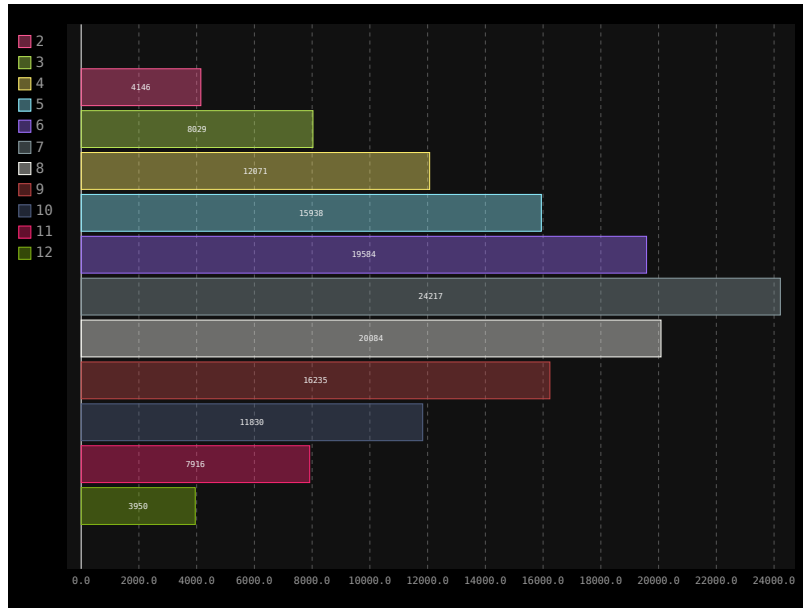
```


Pour un rendu graphique d'une grande qualité, on utilise le module pygal de Python :

```
1 from pygal import *
2 toscane = HorizontalBar()
3 for v in cObs:
4     toscane.add(str(v), cObs[v])
5 toscane.render_to_file("toscano.svg")
```

Ensuite, depuis ipython on peut visualiser sur firefox par exemple avec la commande :

```
1 In [9]: !firefox ./toscano.svg
```



Remplissez le tableau suivant :

| | | | | | | | | | | | |
|--------|---|---|---|---|---|---|---|---|----|----|----|
| Sortie | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Proba | | | | | | | | | | | |

et celui-ci :

| | | | | | | | | | | | |
|---------|------|------|---|---|---|---|---|---|----|----|----|
| Sortie | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Observé | 4041 | 8003 | | | | | | | | | |
| Attendu | 4000 | | | | | | | | | | |

On ne tombe jamais sur ce qui était demandé : combien y a-t-il de suites de 144 000 lancers de 2 dés possibles ? Est-ce que les dés lancés par Python sont pipés ?

Gailleurs bien faitz en piperie
 Pour ruer les ninars au loing
 A l'asault tost sans suerie
 Que les mignons ne soient au gaing
 Farciz d'un plumbis a coing
 Qui griffe au gard le duc
 Et de la dure si tres loing
 Dont l'ambourex luy rompt le suc.

in *Ballade en Jargon* - Attribué à François VILLON - 1489

Une première idée est de mesurer la distance de la distribution observée avec la distribution attendue.

Notons O_i les valeurs observées et $A_i = 144\,000 \times p_i$ les valeurs attendues.

Comment la mesurer ? On peut utiliser la distance euclidienne habituelle :

$$D^2 = (O_2 - A_2)^2 + (O_3 - A_3)^2 + \dots + (O_{12} - A_{12})^2$$

Mais par exemple, 7 apparaît beaucoup plus que 2 mais dans le calcul précédent, mais les distances ont le même poids dans le calcul précédent. On va donc pondérer les écarts par leurs effectifs théoriques :

$$D^2 = \frac{(O_2 - A_2)^2}{A_2} + \frac{(O_3 - A_3)^2}{A_3} + \dots + \frac{(O_{12} - A_{12})^2}{A_{12}}$$

Un peu de calcul :

```

1 sTheo = [x + y for x in range(1,7) for y in range(1,7)]
2 cTheo = {x: v*len(sObs)/len(sTheo) for x,v in Counter(sTheo).items()}
3 khi2Details = [(x,((cObs[x] - cTheo[x])**2) /cTheo[x]) for x in cTheo]
4 khi2 = sum([(cObs[x] - cTheo[x])**2) /cTheo[x] for x in cTheo])

```

Alors :

```

1 In [35]: khi2Details
2 Out[35]:
3 [(2, 0.42025),
4  (3, 0.001125),
5  (4, 1.9253333333333333),
6  (5, 3.3930625),
7  (6, 0.2178),
8  (7, 0.0006666666666666666),
9  (8, 0.04205),
10 (9, 0.21025),
11 (10, 0.075),
12 (11, 0.015125),
13 (12, 2.25625)]
14
15 In [36]: khi2
16 Out[36]: 8.556912500000001

```

Comment interpréter ce 8.55 ?

Ceci est une autre histoire : il nous faudra étudier un peu plus avant les probabilités à densité et la loi du χ^2

On peut généraliser à l'étude de la somme d'un nombre quelconque de dés :

```

1 from collections import Counter
2 from itertools import product
3
4 nObs = 144000
5
6 # séries d'observations de séries de lancers de 3 dés
7 sObs = lambda n : [sum( randint(1,7,size = n) ) for k in range(nObs)]
8 sTheo = lambda n : [sum(t) for t in product(range(1,7), repeat = n)]
9
10 # dictionnaire des fréquences de ces séries
11 # dico.items() renvoie la liste des couples (clé,valeur) de dico
12 def pObs(n) :
13     s = sObs(n)
14     return {x: 100*v/len(s) for x,v in Counter(s).items()}
15
16 def pTheo(n) :
17     s = sTheo(n)
18     return {x: 100*v/len(s) for x,v in Counter(s).items()}

```

Recherche 2 - 6 Autre problème historique

Le Chevalier de Méré soutient à Pascal que les deux jeux suivants sont favorables au joueur : obtenir au moins un 6 en lançant 4 fois de suite un dé, et obtenir au moins un double 6 en lançant 24 fois de suite 2 dés.

Vous simulerez ces expériences à l'aide de votre logiciel préféré et vous démontrerez les résultats observés.

Recherche 2 - 7 Simulations

1. On lance trois fois de suite une pièce de monnaie. On compte combien de fois on obtient pile : simuler informatiquement et démontrez dans le cas général.

2. Faites de même avec l'expérience aléatoire suivante : on dispose de trois urnes, la première contenant 7 boules blanches et 4 noires, la seconde 5 blanches et 2 noires et la troisième 6 blanches et 3 noires. On tire une boule dans chaque urne et on note le nombre de boules blanches obtenues.
3. On tire successivement et avec remise quatre boules dans une urne qui contient 7 boules blanches et 3 roses. On compte le nombre de tirages contenant exactement deux boules blanches - au moins une boule blanche.
4. On lance dix fois une pièce de monnaie et on s'intéresse au nombre maximal de résultats consécutifs égaux.

Recherche 2 - 8 Test aléatoire

On considère deux polynômes écrits différemment, par exemple l'un sous forme factorisée, l'autre sous forme développée : comment faire pour vérifier que les deux polynômes sont bien égaux ?

Quelle est la complexité en nombre de calculs arithmétiques basiques de votre méthode ?

Imaginons maintenant que les polynômes soient de degré d . On choisit un nombre entier n dans l'intervalle $[0, 100d]$ selon une distribution uniforme et on calcule l'image de ce nombre par chacun des polynômes : comment exploiter le résultat ? Évaluer sa « sûreté » ? Complexité ?

Que pensez-vous de l'issue de l'algorithme par rapport aux issues habituelles ? Pouvez-vous prouver que l'algorithme est correct ?

Que se passe-t-il si on répète cette expérience ? Soyez précis(e) dans la description de votre tirage.

Recherche 2 - 9 Indépendance mutuelle

On lance un dé deux fois. On considère les événements :

- A : « Le premier lancer a donné un numéro pair » ;
- B : « Le second lancer a donné un numéro impair » ;
- C : « La somme des deux numéros est un chiffre pair ».

Les événements sont-ils deux à deux indépendants ? Mutuellement indépendants ?

Recherche 2 - 10 Roulette

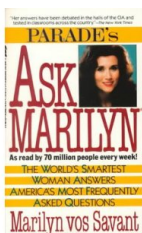
Le rouge est sorti dix fois de suite à la roulette : quelle est la probabilité qu'il sorte une onzième fois ?

On se placera dans le cas très improbable que la roulette n'est pas truquée et que le casino n'est pas dirigé par un mafieux ou un ex-entraîneur d'une équipe de France et ex-ministre.

Recherche 2 - 11 Formule de B. et Monica B.

En Syldavie, 70% des habitants sont de sexe masculin et 60% des hommes et 25% des femmes ont une photo de Monica B. dans leur portefeuille. Un extra-terrestre enlève au hasard une personne syldave et découvre une photo de Monica B. dans son portefeuille : quelle est la probabilité que la personne enlevée soit une femme ?

Recherche 2 - 12 Marylin et les chèvres



Dans le numéro du 9 septembre 1990 de *Parade*, Marilyn VOS SAVANT connue pour avoir un QI de 186, soit l'un des plus élevés au monde, répondit au courrier suivant : *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors ?*

Craig. F. Whitaker Columbia, MD

Cette lettre décrit en fait le déroulement d'un jeu télévisé des années 1970, *Let's make a deal*, animé par Carol MERRILL et Monty Hall : c'est en référence à ce dernier qu'on donna son nom à ce problème célèbre (plus tard, un haut représentant de la culture française reprit ce jeu : ce fut Lagaf et son *Bigdil...*).

Marilyn répondit qu'il valait mieux pour le candidat changer de porte quand le présentateur l'y invitait. Qu'en pensez-vous ?

Vous pourrez également proposer une simulation informatique.



Recherche 2 - 13 Formule de Bayes et générateur biaisé

On dispose de trois générateurs pseudo-aléatoires. On sait que l'un des trois a été mal programmé. On demande à chacun de générer un bit. Le générateur biaisé sort un 1 avec une probabilité de $2/3$. Deux générateurs ont sorti un 1 et le troisième un 0. Le problème est qu'on ne sait pas qui est le générateur biaisé. Quelle est la probabilité que le générateur biaisé soit le premier ?

Quelle quantité d'information a-t-on gagnée ?

Recherche 2 - 14 Duel

Alexandre et Évariste s'affrontent en duel au pistolet suivant les règles suivantes :

- Ils tirent chacun leur tour. Le premier qui atteint la cible a gagné...
- Lorsqu'il tire, Alexandre atteint la cible avec la probabilité a ($0 < a < 1$) et il la rate avec la probabilité $\bar{a} = 1 - a$.
- Lorsqu'il tire, Évariste atteint la cible avec la probabilité b ($0 < b < 1$) et il la rate avec la probabilité $\bar{b} = 1 - b$.
- Alexandre tire le premier. Ainsi, Alexandre (resp. Évariste) n'effectue que des tirs de rang impair (resp. pair).

On considère, pour tout entier $n \geq 1$, les événements A_{2n-1} : « Alexandre gagne à l'issue du tir numéro $2n - 1$ », B_{2n} : « Évariste gagne à l'issue du tir numéro $2n$ ».



1. Calculez, en fonction de a et b , les probabilités des événements A_1 , B_2 et A_3 . Plus généralement, calculez $\mathbb{P}(A_{2n-1})$ et $\mathbb{P}(B_{2n})$.
2. Pour $n \geq 1$, on note C_n (resp. D_n) l'évènement : « Alexandre (resp. Évariste) gagne à un tir dont le numéro est entre 1 et $2n - 1$ (resp. entre 2 et $2n$) ». Calculer $\mathbb{P}(C_n)$ et $\mathbb{P}(D_n)$.
3. Calculer $\alpha = \lim_{n \rightarrow +\infty} \mathbb{P}(C_n)$ et $\beta = \lim_{n \rightarrow +\infty} \mathbb{P}(D_n)$. Vérifier que $\alpha + \beta = 1$. Comment interpréter ce résultat ?

Recherche 2 - 15 QCM

Dites si les variables aléatoires X définies ci-dessous suivent une loi binomiale ? Donnez quand c'est possible $X(\Omega)$, les paramètres de la loi ainsi que l'espérance.

1. Dans une classe on tire au sort et sans remise 5 élèves. X est le nombre d'élèves travaillant pour les services secrets bordures dans le lot tiré au sort.
2. Dans un sac de 20 billes contenant 7 noires et 13 blanches, on tire avec remise 3 d'entre elles, X étant le nombre de billes noires obtenues.
3. On lance 10 dés, X est le nombre de « 5 » obtenus.
4. Un circuit comprend 32 lampes en série. Pour chacune d'elle, la probabilité qu'elle fonctionne est de $3/100$. X est le nombre de lampes qui s'allument lorsqu'on appuie sur l'interrupteur.
5. Même question avec cette fois des lampes en parallèle.

Recherche 2 - 16 Lois usuelles

Dans chacune des expériences aléatoires ci-dessous, reconnaître une loi usuelle pour la variable aléatoire X . Donner alors son espérance.

1. On choisit une carte au hasard dans un jeu de 32 cartes numérotées. X = « hauteur de la carte choisie »
2. Un lion mange à chaque repas une gazelle (avec probabilité $2/3$) ou un zèbre (avec probabilité $1/3$). On suppose que la composition d'un repas est indépendante de celle des autres. X = « nombre de zèbres mangés sur 10 repas consécutifs ».
3. Dans un enclos se trouvent 15 gazelles et 30 zèbres. Un lion mange à chaque repas une gazelle ou un zèbre, pris parmi les animaux de l'enclos. X = « nombre de zèbres mangés sur 10 repas consécutifs ».
4. On tire une à une les cartes d'un jeu de 32 cartes, jusqu'à obtenir le valet de pique. X = « nombre de tirage effectuées ».
5. On lance un dé à 6 faces. X = « chiffre obtenu ».
6. Monica a dans sa poche 7 allumettes et 12 pièces de monnaie. Il met sa main dans sa poche et en sort 5 objets. X = « nombre de pièces obtenues ».

7. Monica fait passer une audition pour trouver des partenaires pour son prochain film. 3000 candidats passent l'audition. Chaque candidat réussit son entretien avec probabilité $2/3$. $X =$ « nombre de candidats sélectionnés par Monica ».
8. James dort chez son amie Monica qui lui a prêté son appartement. Il dispose d'un trousseau de 7 clés, et ne sait pas laquelle ouvre la porte de l'appartement. Il les essaye une par une, en mettant de côté chaque clé essayée. $X =$ « nombre d'essais nécessaires pour trouver la bonne clé ».

Recherche 2 - 17 Loi hypergéométrique

Justifiez la valeur donnée dans le cours de $\Pr(\{X = k\})$ lorsque $X \sim H(N, n, p)$.

Recherche 2 - 18 Probabilités, « experts » et journalistes

Le 3 juin 2011, est paru dans Libération (http://www.liberation.fr/politiques/2011/06/03/accident-nucleaire-une-certitude-statistique_740208) un article signé par Bernard LAPONCHE, physicien nucléaire et expert en politique de l'énergie et de Benjamin DESSUS, ingénieur et économiste. Du solide quoi!

la probabilité d'occurrence d'un accident majeur sur ces parcs serait donc de 50% pour la France et de plus de 100% pour l'Union européenne.

Bon, évidemment, quand un mathématicien lit qu'une probabilité dépasse 100%, cela le laisse perplexe... Étienne GHYS, grand mathématicien français qui a lu l'article, a relevé cette bourde dans « Image des mathématiques » le lendemain.

L'article commence par estimer la probabilité d'un accident majeur par réacteur nucléaire et par année de fonctionnement. Selon l'article, le parc mondial actuel de réacteurs cumule 14000 réacteurs-ans (environ 450 réacteurs pendant 31 ans). Pendant cette période, il y a eu quatre accidents majeurs, ce qui mène à une probabilité d'accident majeur d'environ 0,0003 par an pour chaque réacteur.

Les auteurs en « déduisent » donc que la probabilité d'un accident majeur en France (avec ses 58 réacteurs) pendant les trente prochaines années serait de 58 fois 30 fois 0,0003, donc d'environ 50%. Quant à la probabilité d'un accident en Europe (143 réacteurs) dans les trente prochaines années, elle « est » de 143 fois 30 fois 0,0003, « donc » d'environ 129%. Comment une probabilité pourrait-elle dépasser 100%? Plus sûr que la certitude? Les auteurs ont sans doute mauvaise conscience d'écrire que la probabilité est de 129% alors ils se contentent d'écrire qu'elle est de « plus de 100% »... Ils concluent « La réalité, c'est que le risque d'accident majeur en Europe n'est pas très improbable, mais au contraire une certitude statistique. »

Quand on voit leur CV...

Bernard LAPONCHE : polytechnicien, docteur ès sciences et en économie de l'énergie, il a été ingénieur au Commissariat à l'énergie atomique (CEA), responsable syndical à la CFDT, puis directeur général de l'Agence Française pour la Maîtrise de l'Énergie (AFME) dans les années 80. Co-fondateur avec Florence Rosenstiel et directeur du bureau d'étude ICE (International Conseil Energie) de 1988 à 1998, il a été conseiller pour l'énergie et la sûreté nucléaire auprès de la Ministre de l'aménagement du territoire et de l'environnement Dominique Voynet en 1998 et 1999.

Benjamin DESSUS, ingénieur et économiste, a débuté sa carrière aux laboratoires de Marcoussis dans le domaine l'électronique quantique et des lasers avant de rejoindre les Etudes et Recherches d'Electricité de France pour y monter un laboratoire de métrologie optique. En 1982, à la création de l'AFME (Agence Française de la Maîtrise de l'Énergie, devenue ADEME), il prend la direction des services techniques de cette agence. Il la quitte en 1987 pour rejoindre le CNRS où il assumera jusqu'en 2001 la direction de plusieurs programmes interdisciplinaires de recherche (PIRSEM, Ecotech, ECODEV) consacrés aux problèmes d'énergie et d'environnement. Parallèlement, il a contribué à la formulation de la stratégie climat du Fonds pour l'Environnement Mondial (FEM) - dont il a fait partie de 1991 à 1994 du Scientific and Technical Advisory Panel (alors présidé par Robert Watson) - puis présidé de 1994 à 2003 le conseil scientifique et technique du Fonds Français pour l'Environnement Mondial (FFEM). Spécialiste reconnu des questions énergétiques et en particulier du nucléaire (rapport Charpin-Dessus-Pellat relatif à l'Étude économique prospective de la filière électrique nucléaire en juillet 2000).

Bon, pouvez-vous corriger les calculs de nos experts?

Si l'on garde cette probabilité de 0,0003 par an et par réacteur de tomber en panne (mais on peut vraiment en discuter), quelle est la probabilité qu'il n'y ait pas d'accident dans les 30 prochaines années sur les 143 réacteurs européens. On fera beaucoup d'hypothèses : on suppose que la loi suivie est une loi binomiale, qu'il y a indépendance entre les réacteurs et les années.

Qu'on calculé les « experts » en fait?

Prenons un autre exemple moins discutable : quelle est la probabilité d'avoir au moins une fille dans une famille de quatre enfants ? Nos experts auraient dit : « proba d'avoir une fille 0,5 donc $4 \times 0,5 = 2$ donc la probabilité vaut 2 ». Qu'en pensez-vous ?

Autre exemple : Google a le monopole de la sécurité des centrales nucléaires. Son programme commet une erreur 1 fois sur 100. Il y a 450 réacteurs dans le monde : quelle est la probabilité qu'il n'y ait aucun accident. Quelle est l'espérance du nombre d'accidents ?

po

Recherche 2 - 19 Espérance géométrique

Nous avons prouvé au module précédent que :

$$\forall q \in \mathbb{R}, |q| < 1 \rightarrow S = \sum_{n \geq 0} q^n = \frac{1}{1-q}$$

Pour $|q| < 1$, notons $T = \sum_{n \geq 1} nq^{n-1}$ en supposant que cette écriture a un sens.

Calculer $(1-q)T$ en fonction de S et déduisez-en T puis l'espérance d'une loi géométrique.

Recherche 2 - 20 Attentes

Combien faut-il en moyenne de lancers de dés pour obtenir les six faces ?

Combien faut-il en moyenne de tirages d'un entier choisi au hasard entre 0 et 9 pour les obtenir tous ?

Recherche 2 - 21 Roulette



Une roulette contient 36 cases numérotées de 1 à 36 dont 18 sont rouges et 18 sont noires, plus une case numérotée 0 de couleur verte.

Un joueur qui mise sur la couleur rouge ou noire gagne deux fois sa mise si la couleur choisie sort.

Un joueur qui mise sur un numéro de 1 à 36 gagne 36 fois sa mise si le numéro sort.

Il est interdit de miser sur le zéro.

1. Un joueur mise $a \text{ €}$ sur une couleur. Soit C la variable aléatoire correspondant au gain associé. Trouvez la loi de C puis calculez $\mathbb{E}(C)$ et $\sigma(C)$.
2. Un joueur mise $a \text{ €}$ sur un numéro. Soit N la variable aléatoire correspondant au gain associé. Trouvez la loi de N puis calculez $\mathbb{E}(N)$ et $\sigma(N)$.
3. Vaut-il mieux miser sur une couleur ou un numéro ?

Recherche 2 - 22 Dé truqué

On considère 6 dés, cinq étant équilibrés. Le dernier est pipé de manière à ce que lorsque l'on lance ce dé, chacun des chiffres apparaît avec une probabilité proportionnelle à ce chiffre.

- a) Donner la loi, l'espérance et la variance de la variable aléatoire égale au chiffre donné par le dé truqué lorsqu'on le lance.

On effectue n tirages successifs et indépendants d'un dé parmi les six.

- b) Quelle est la loi suivie par la variable aléatoire égale au nombre de fois où on a tiré le dé truqué ? Combien de tirages doit-on effectuer pour que la probabilité d'avoir obtenu le dé truqué parmi ceux tirés soit supérieure ou égale à $1/2$?

On effectue n tirages ($n \leq 6$) successifs sans remise d'un dé parmi les six.

- c) Quelle est la loi suivie par la variable aléatoire égale au nombre de fois où est tiré le dé truqué ? Combien de tirages doit-on effectuer pour que la probabilité d'avoir obtenu le dé truqué parmi ceux tirés soit supérieure ou égale à $1/2$?

Recherche 2 - 23 « C'est sans danger... »



Afin d'effectuer des économies dans le budget de l'État, le premier ministre syldave a supprimé la faculté de médecine. Dans le but de promouvoir les liens sociaux, les dentistes sont maintenant choisis parmi les anciens boxeurs aveugles et parkinsoniens. Ils arrachent les dents de leurs patients au hasard.

Suite à quelques plaintes, une étude statistique a été menée dans le cabinet d'un de ces nouveaux dentistes. Les syldaves venant le consulter ont toujours une seule dent de malade parmi les trente-deux qu'ils possèdent encore avant l'intervention des tenailles ou des poings, c'est selon (une étude précédente a en effet montré que les syldaves ne retournent jamais chez leur dentiste après une première consultation).

L'étude a été effectuée sur n patients. On note X la variable aléatoire égale au nombre de dents malades extraites à bon escient.

1. Déterminez la loi de probabilité de la variable aléatoire X . Calculez la probabilité pour qu'aucune dent malade n'ait été extraite.
2. Combien de patients doit-il traiter pour extraire au moins une dent malade avec une probabilité supérieure à 0,9 ?
3. Le dernier patient est prêt à se sacrifier pour la gloire des statistiques syldaves et se laisse arracher les dents une à une tant que la dent malade n'a pas été extraite. On note Y le nombre de dents saines que ce vaillant patriote voit tomber de ses mâchoires endolories.

Calculez la probabilité pour qu'il reparte complètement édenté, puis $\mathbb{E}(Y)$ et $\sigma(Y)$.

Recherche 2 - 24 Indépendance de v.a.

L'expérience aléatoire est le lancer de deux dés parfaits. X_1 est la variable de (Bernoulli) qui prend la valeur 1 si le premier dé donne un résultat pair 0 sinon, X_2 est la variable (de Bernoulli) qui prend la valeur 1 si le deuxième dé donne un résultat pair et 0 sinon et X_3 est la variable (de Bernoulli) qui prend la valeur 1 si la somme des numéros tirés est paire et 0 sinon.

Les variables aléatoires sont-elles mutuellement indépendantes? Deux à deux indépendantes?

Recherche 2 - 25 Inégalité de Jensen

Supposons que l'on choisisse la longueur d'un côté d'un carré selon une distribution uniforme dans l'intervalle entier $[1,99]$.

Quelle est l'aire moyenne du carré obtenu ?

Notons X la var égale à la longueur du côté tirée aléatoirement.

On cherche donc $\mathbb{E}(X^2)$.

Calculez $\mathbb{E}(X)$ puis $(\mathbb{E}(X))^2$.

Pour le calcul de $\mathbb{E}(X^2)$, on se souviendra que $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$.

A-t-on $(\mathbb{E}(X))^2 = \mathbb{E}(X^2)$?

Introduire la var $Y = (X - \mathbb{E}(X))^2$ et calculez son espérance par linéarité. Conclure.

En fait, il s'agit d'un cas particulier de l'inégalité de JENSEN qui permet d'obtenir que

$$\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$$

pour toute fonction f convexe. Pour le démontrer, on peut utiliser un développement de TAYLOR à l'ordre 2 et exploiter le fait que f'' est à valeurs positives car f est convexe, mais je m'arrêterai là...

Recherche 2 - 26 Théorème de König-Huygens

Démontrez ce théorème...

Recherche 2 - 27 Tirages sans remise avec condition d'arrêt

Une urne contient n boules. On y effectue des tirages sans remise.

1. Quelle est la probabilité d'obtenir p boules données en k tirages sans remise ?



Johan JENSEN
(1859 - 1925)

2. L'urne contient p boules rouges et $n-p$ boules blanches. On note T la variable aléatoire égale au nombre de tirages nécessaires à l'obtention des p boules rouges.
- Déterminer $T(\Omega)$ puis $\Pr(T \leq k)$ pour $k \in T(\Omega)$. Vérifier que cette probabilité s'exprime en fonction de $\frac{C_k^p}{C_n^p}$.
 - Déterminer la loi de T .
 - On rappelle que, pour tout $r \in \mathbb{N}$, pour tout $q \in \{0, \dots, r\}$: $\sum_{k=q}^r C_k^q = C_{r+1}^{q+1}$. Calculer $\mathbb{E}(T)$.

Recherche 2 - 28 Ordinateur syldave

Piotr a récupéré un vieil ordinateur de bord muni d'un processeur 4 bits sur un avion bordure abattu par les forces du peuple syldave. Il crée un compilateur qui pour l'instant ne renvoie qu'un quartet (une chaîne de 4 bits) au hasard, chaque bit valant 0 avec une probabilité p . Soit X la variable aléatoire égale au quartet renvoyé exprimé en base 10.

- Établissez la loi de probabilité de X .
- Calculez son espérance et sa variance.
- Calculez la probabilité d'obtenir un nombre pair.
- Calculez la probabilité que le quartet soit un palindrome.

Recherche 2 - 29 Produit de variables aléatoires indépendantes

Soit X_1 et X_2 deux v.a.r. discrètes finies prenant leurs valeurs respectives dans I_1 et I_2 . On rappelle que X_1 et X_2 sont indépendantes si, et seulement si, pour tout couple (x_1, x_2) de $I_1 \times I_2$:

$$\Pr(\{X_1 = x_1\} \cap \{X_2 = x_2\}) = \Pr(X_1 = x_1) \times \Pr(X_2 = x_2)$$

Montrez alors le théorème suivant :

Théorème 2 - 14

Espérance d'un couple de v.a.r. indépendantes

Si X et Y sont indépendantes alors $\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y)$

Recherche 2 - 30 Opérations sur la variance

Comment écrire $\mathbb{V}(aX)$, $\mathbb{V}(X+b)$ et $\mathbb{V}(X+Y)$ en fonction de $\mathbb{V}(X)$ et $\mathbb{V}(Y)$?

Recherche 2 - 31 Approximation de la loi binomiale par la loi de Poisson

On considère une suite de v.a.r. X_n ayant une loi de probabilité binomiale de paramètres n et λ/n . Montrez que :

$$\lim_{n \rightarrow +\infty} \Pr_{X_n}(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

On dit alors que la suite de variables aléatoires X_n converge en loi (c.a.d. $\lim_{n \rightarrow +\infty} \Pr_{X_n} = \Pr_X$: la loi de X_n converge vers la loi de X). On étudie en fait un cas particulier de convergence de *suite de fonctions...* vers une v.a.r. X de loi de Poisson de paramètre λ .

Déduisez-en une approximation de $C_n^k p^k (1-p)^{n-k}$ lorsque $n \gg 1$.

Application : la fusion amoindrit les risques



Deux sociétés assurent des biens très chers mais dont le risque d'accident est très faible : vaut-il mieux fusionner ou bien les risques s'additionnent-ils ?

Prenons le cas d'une société d'assurance sicilienne ayant pour client la société des croisières Costa.

- Supposons que la société Costa possède 500 navires dont la valeur est de 5 millions de Schpziřšw. L'assureur rembourse la perte totale d'un navire, événement dont la probabilité est estimée à 0,001 pour une année. On suppose que les risques de pertes des navires sont indépendants. Donner la loi de X égale au nombre de navires perdus au cours de l'année.
- L'assureur rembourse son client une fois par an. Combien doit-il avoir dans son coffre pour qu'il puisse rembourser son client avec une probabilité de 0,999. On utilisera l'approximation du préambule et on utilisera une table de la loi de Poisson.

3. Une autre compagnie assure les 500 navires de la White Star dans les mêmes conditions. Les deux compagnies décident de fusionner suite au suicide du président de la première compagnie dont on a retrouvé le corps percé de 52 coups de couteaux. La compagnie d'assurance prend donc en charge à présent 1000 navires. Refaire les calculs de la question précédente dans ces conditions. Commentaires ?

Recherche 2 - 32 Opération sur une v.a.r.

Soit X une v.a.r. telle que $X(\Omega) = \mathbb{N}^*$ et $\Pr(\{X = n\}) = \frac{\lambda}{n!}$.

1. Déterminez λ .
2. Calculer $\mathbb{E}(X)$.
3. Calculer $\mathbb{E}\left(\frac{1}{1+X}\right)$

Recherche 2 - 33 Espérance et variance de la loi de Poisson

Démontrer les formules donnant l'espérance et la variance de la loi de Poisson.

On se souviendra de la série exponentielle évoquée au module précédent :

$$\forall x \in \mathbb{R}, \quad \sum_{k \geq 0} \frac{x^k}{k!} = e^x$$

Recherche 2 - 34 Les étudiants et le poisson



Une étude sérieuse vient de prouver que le nombre d'étudiants arrivant en salle F22 entre 12h30 et 12h35 les jeudis suit une loi de Poisson de paramètre $\lambda = 3$.

1. Quel est, en moyenne, le nombre d'étudiants qui arrivent entre 12h30 et 12h35 ?
2. Calculer la probabilité, qu'entre 12h30 et 12h35 aucun étudiant ne se présente en F22.
3. Calculer la probabilité, qu'entre 12h30 et 12h35 trois étudiants se présentent en F22.
4. Calculer la probabilité, qu'entre 12h30 et 12h35 plus de 4 étudiants se présentent en F22.
5. Calculer la probabilité, qu'entre 12h30 et 12h35 au moins un étudiant se présente en F22.

Recherche 2 - 35 MURD3RS



Les américains ont un goût prononcé pour les tueurs en série. Le profil psychologique de certains de ces héros maléfiques correspond à une personne logique, froide, méthodique, dangereuse, qui connaît bien son métier, qui travaille suffisamment lentement pour bien faire son boulot et suffisamment rapidement pour être efficace. Certains sont des sociopathes, d'autres sont très à l'aise en société : on retrouve exactement les qualités recherchées chez un programmeur... Bref, si on apprenait plus de Haskell à l'école, les tueurs en série disparaîtraient.

Un policier canadien, Kim ROSSMO, s'est lui aussi passionné pour les mathématiques et la programmation et devint en 1995 le premier policier canadien à obtenir un doctorat en criminologie (https://en.wikipedia.org/wiki/Kim_Rossmo).

Il a publié un article (http://www.popcenter.org/library/crimeprevention/volume_04/10-Rossmo.pdf) sur l'analyse géographique des crimes et proposa une formule pour localiser le lieu d'habitation d'un tueur en série, ou plutôt pour calculer le lieu le plus probable :

$$p_{i,j} = k \sum_{n=1}^{(\text{total crimes})} \left[\underbrace{\frac{\phi_{ij}}{(|X_i - x_n| + |Y_j - y_n|)^f}}_{1^{\text{st}} \text{ term}} + \underbrace{\frac{(1 - \phi_{ij})(B^{g-f})}{(2B - |X_i - x_n| - |Y_j - y_n|)^g}}_{2^{\text{nd}} \text{ term}} \right]$$

avec

$$\phi_{ij} = \begin{cases} 1, & \text{if } (|X_i - x_n| + |Y_j - y_n|) > B \\ 0, & \text{else} \end{cases}$$

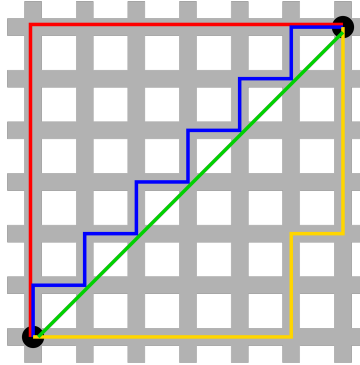
Bon, plus synthétiquement, ça peut s'écrire :

$$P(x) = \sum_{\text{lieu du crime } c} \frac{\varphi}{d(x, c)^f} + \frac{(1 - \varphi)B^{g-f}}{(2B - d(x, c))^g}$$

...euh...qui est qui là-dedans ?

On discrétise un plan, de ville de préférence, à l'aide d'une grille. On calcule alors la probabilité que le tueur habite dans la case x .

Comme on travaille dans une ville, la distance d sera la distance du taxi de Manhattan :



ou ça si vous préférez



B est le rayon de la « Buffer zone » (les tueurs en série ont des buffers comme emacs) ou zone tampon : c'est un « cercle » (au sens des taxis) autour du lieu du crime. La fonction φ est la fonction caractéristique de cette zone tampon : elle vaut 1 si la cellule est dans la zone et 0 sinon. f et g sont des paramètres arbitraires qui permettent de s'adapter aux données des crimes précédents.

On cherche à présent à dresser une carte des probabilités. On va utiliser la fonction `pcolormesh` de `matplotlib.pyplot`.

Vous pouvez regarder l'aide. Ce qui va motiver notre construction, c'est la nature des arguments de cette fonction :

```

1 def dessine_rossmo( lieux, vrai, B, f, g, taille_x, taille_y) :
2     t_x = np.arange(taille_x)
3     t_y = np.arange(taille_y)
4     X, Y = np.meshgrid(t_x, t_y)
5     Z = np.vectorize( proba_rossmo(lieux, B, f, g) )(X, Y)
6     g = cg(lieux)
7     plt.pcolormesh( X, Y, Z, cmap = cm.jet)# , shading='gouraud')
8     plt.plot(vrai[0], vrai[1], 'ks', mew = 7)
9     plt.plot(g[0],g[1], 'ko', mew = 7)
10    plt.colorbar()
11    plt.show()

```

Ici, `lieux` est l'ensemble des lieux des crimes, `vrai` est le véritable domicile du tueur, `taille_x` et `taille_y` sont les dimensions de la grille.

La fonction `cg` calcule les coordonnées du centre de gravité des lieux des crimes pour comparer avec le candidat de la méthode ROSSMO.

`colorbar` permet d'avoir une échelle de couleur pour lire la carte.

Il va falloir bien comprendre la construction de `Z`...À vous de menez l'enquête informatique.

ROSSMO fournit des données pour deux tueurs en série célèbres :

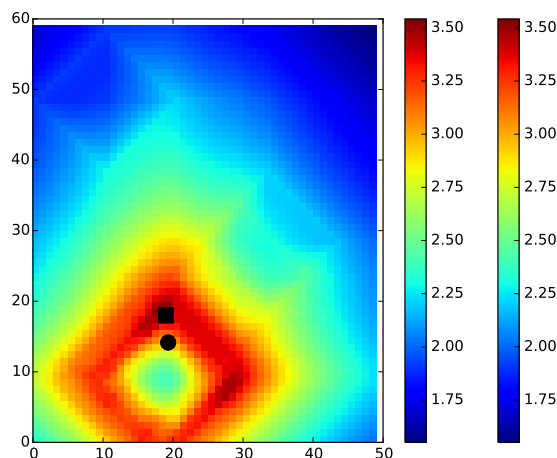
— le vampire de Sacramento (https://en.wikipedia.org/wiki/Richard_Chase)

```
1 vampire = {(3, 17), (15, 3), (19, 27), (21, 22), (25, 18)}
2 vrai_vampire = ([19],[17])
```

— L'étrangleur de Boston (https://en.wikipedia.org/wiki/Albert_DeSalvo)

```
1 boston = [[10, 48], [13, 8], [15, 11], [17, 8], [18, 7],
2           [18, 9], [19, 4], [19, 8], [20, 9], [20, 10], [20, 11],
3           [29, 23], [33, 28]]
4
5 vrai_boston = ([19],[18])
```

Vous obtiendrez des cartes qui ressemblent à ça :



Recherche 2 - 36 Algorithme probabiliste : amélioration du tri rapide

Rappelez le principe du tri rapide (quicksort).

Vous pouvez vous le remémorer ici :

<http://www.youtube.com/watch?v=ywWBy6J5gz8>

Quelle peut être sa complexité ? Y a-t-il des cas pires que d'autres ?

On peut introduire une dose de hasard de deux façons : on choisit le pivot au hasard ou on prend toujours le premier élément de la liste mais on permute aléatoirement les autres éléments de la liste.

On distingue deux classes d'algorithmes probabilistes (nous ne rentrerons pas dans les détails des classes de complexité, des machines de TURING probabilistes, etc.) :

- les algorithmes de Monte-Carlo : l'algorithme peut se tromper mais on connaît la probabilité qu'il donne la bonne réponse : on peut donc diminuer cette probabilité en augmentant le nombre d'exécution.
- les algorithmes de Las Vegas : ils donnent toujours la bonne réponse mais le temps d'exécution est une variable aléatoire.

On peut passer de Monte-Carlo à Las Vegas à la condition de pouvoir déterminer si l'issue de Monte-Carlo est correcte : comment ?

Revenons au cas du tri rapide.

Notons $\{x_1, x_2, \dots, x_n\}$ la suite à trier et y_1, y_2, \dots, y_n la suite triée. Soit $i < j$. On considère la var X_{ij} qui vaut 1 si y_i et y_j ont été comparés et 0 sinon.

Soit X le nombre total de comparaisons, alors

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}$$

Que vaut $\mathbb{E}(X)$?

Que vaut $\mathbb{E}(X_{ij})$ compte-tenu de la nature de la var X_{ij} ?

Maintenant, y_i et y_j sont comparés si, et seulement si, l'un des deux est choisi comme pivot de la sous-suite $y_i \dots y_j$.

En effet, si aucun des deux n'est pivot de cette suite, ils seront dans des sous-listes séparées et ne seront jamais comparés.

Si, l'un des deux est choisi comme pivot de la sous-suite $y_i \dots y_j$, ils seront forcément comparés.

Écrivez plus formellement cette preuve à l'aide de la logique des propositions.

Quelle est donc la probabilité que $X_{ij} = 1$?

En déduire $\mathbb{E}(X)$ en fonction de $H_n = \sum_{k=1}^n \frac{1}{k}$.

Ce terme est une somme partielle de la série harmonique qui est équivalente à $\ln(n)$.

Concluez.

Recherche 2 - 37 π par la méthode de Monte-Carlo

Imaginez un moyen de calculer une valeur approchée de π avec un algorithme probabiliste.

Recherche 2 - 38 Détection d'erreur par répétition

On considère une source binaire d'information constituée d'une suite de 1 et de 0 indépendants et équiprobables. La source émet à destination d'un récepteur à travers un canal de transmission bruyant entraînant des erreurs éventuelles en réception.

Soit p la probabilité, commune, qu'un bit émis soit détecté comme son complémentaire. On suppose que les erreurs commises sur chaque bit sont indépendantes mutuellement.

Pour tenter de lutter contre ce bruit, on utilise le protocole simple suivant :

- à l'émission, on répète une fois chaque bit ;
- à la réception, la réception de (b, b) entraîne la détection du symbole b ; la réception de (b, \bar{b}) donne lieu à une demande de retransmission d'un nouveau couple de bits identiques. On continue tant qu'un couple (b, \bar{b}) est reçu.

1. En quoi consiste l'erreur résiduelle d'une telle stratégie ? Donnez des exemples de chaînes transmises aboutissant à une erreur sur le bit détecté.
2. Montrez que $\Pr(\text{erreur}) = \Pr(\text{erreur} \mid 0 \text{ émis}) = \frac{p^2}{1-2p+2p^2}$.
3. Soit N la v.a.r. donnant le nombre de symboles émis par l'émetteur pour un symbole utile émis. Étudiez la loi de N et donnez son espérance et sa variance. Vous pourrez montrer que la loi de $N/2$ est géométrique. Vous tracerez l'allure des courbes de la loi, de l'espérance et de la variance en fonction de p . Commentez. Étudiez le cas particulier où $p \ll 1$.

Recherche 2 - 39 Loi faible des grands nombres

Cet exercice est en fait du cours et représente même un résultat plus que primordial...

Souvenez-vous : il y a bien longtemps, vous étiez en classe de seconde et vous découvriez les probabilités. Vous lanciez un dé et notiez la fréquence d'obtention de tel ou tel résultat. Déjà fort perspicace, vous aviez remarqué que plus vous lanciez le dé, plus la fréquence d'obtention de 4, par exemple, tendait vers $1/6$ qui, vous l'avez appris depuis, est la probabilité d'obtenir 4 en lançant un dé équilibré.

Pour résumer, plus vous répétez une expérience aléatoires dans des conditions identiques et avec indépendance des résultats, plus la fréquence observée de succès tend vers la probabilité de succès : c'est ce qu'on appelle la loi faible des grands nombres et nous allons essayer de prouver ce résultat dans le cas particulier de la loi binomiale.

Inégalité de Markov

Soit X une variable aléatoire suivant une loi de probabilité \Pr_X et ne prenant que des valeurs positives (rangées comme d'habitude dans l'ordre croissant) x_1, x_2, \dots, x_n .

Soit a un nombre (qui sera notre seuil d'erreur fréquence/probabilité) strictement positif fixé.

Le monde se sépare en deux catégories : les x_i strictement inférieurs à a et ceux qui lui sont supérieurs.

Supposons par exemple que $x_1 \leq x_2 \leq \dots \leq x_{k-1} < a \leq x_k \leq \dots \leq x_n$

1. Rappelez la définition de l'espérance $\mathbb{E}(X)$.
2. Montrez que $\mathbb{E}(X) \geq a \sum_{i=k}^n \Pr_X(x_i)$.
3. Déduisez-en l'inégalité suivante $\Pr_X([a, +\infty[) \leq \frac{\mathbb{E}(X)}{a}$.
4. Généralisez afin d'obtenir l'inégalité de MARKOV : $\Pr_X([a, +\infty[) \leq \frac{\mathbb{E}(X^p)}{a^p}$ pour tout entier naturel non nul p .

5. Déduisez-en cette autre formulation : $\Pr_X([0, a[) \geq 1 - \frac{\mathbb{E}(X^p)}{a^p}$

6. Cas particulier de la loi binomiale

On suppose que X suit la loi binomiale $B(n, p)$. Donnez une majoration de $\Pr\left(\frac{X}{n} \geq a\right)$ et une minoration de $\Pr\left(\frac{X}{n} < a\right)$ à l'aide de l'inégalité de MARKOV.

Inégalité de Bienaymé-Tchebychev

1. Variance

La variance d'une variable aléatoire est la quantité qui mesure la dispersion de X autour de sa moyenne, c'est à dire son espérance. Plus les valeurs sont dispersées, plus sa variance augmente. On attend donc que si X est une variable aléatoire constante, sa variance sera nulle.

On aura remarqué que la variance est « homogène » à X^2 .



2. Appliquez l'inégalité de Markov astucieusement pour obtenir l'inégalité de Bienaymé-Tchebychev

$$\Pr(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}$$

Majorez alors la probabilité qu'une v.a.r. s'éloigne de $10\sigma(X)$ de son espérance.

3. Une application de ces inégalités

Les inégalités de nos amis franco-russes permettent d'avoir une estimation de certaines probabilités sans qu'on connaisse la loi de probabilité. il faut toutefois être conscient que la probabilité exacte peut être assez éloignée de la borne proposée.

Par exemple, le nombre de caleçons molletonnés fabriqués dans une usine syldave en une semaine est une variable aléatoire d'espérance 50 et de variance 25.

Estimez, grâce à l'inégalité de Markov, la probabilité que la production de la semaine à venir dépasse 75 caleçons molletonnés.

Estimez, grâce à l'inégalité de Bienaymé-Tchebychev, la probabilité que la production de la semaine à venir soit strictement comprise entre 40 et 60 caleçons molletonnés.

Loi faible des grands nombres dans le cas de la loi binomiale

1. Soit λ un réel strictement positif et X une variable aléatoire quelconque. Montrez que $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$ puis que $\mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X)$.

2. Considérons maintenant le cas où X obéit à la loi binomiale $B(n, p)$.

Intéressons-nous à la probabilité d'obtenir k succès au cours des n répétitions de l'épreuve.

Le rapport k/n est alors le nombre relatif (la fréquence) de succès.

Nous allons essayer de confirmer notre étude expérimentale, à savoir que, plus n est grand, plus le rapport k/n est proche de la probabilité p de succès à chacune des épreuves.

On a démontré que la variance de X vérifie $\mathbb{V}(X) = np(1 - p)$

Montrez que

$$\Pr\left(\left|\frac{X}{n} - p\right| > a\right) \leq \frac{p(1-p)}{na^2} \quad \text{puis que} \quad \lim_{n \rightarrow +\infty} \Pr\left(\left|\frac{X}{n} - p\right| \leq a\right) = 1$$

Qu'en déduisez-vous ?

Recherche 2 - 40 Norme et Espace \mathcal{L}^1

Nous étudierons l'an prochain la notion d'espace vectoriel. Disons pour faire (très) court que c'est un espace qui est stable par combinaisons linéaires : si j'additionne deux vecteurs ou si je multiplie un vecteur par un scalaire, je reste dans l'espace et il existe un vecteur nul.

Une norme est une application définie sur un espace vectoriel E , à valeurs dans \mathbb{R}_+ , notée le plus souvent $\|\cdot\|$ et vérifiant :

- 1. $\forall x \in E, \|x\| = 0 \Leftrightarrow x = 0$;
- 2. $\forall x \in E, \forall \lambda \in \mathbb{R}, \|\lambda \cdot x\| = |\lambda| \times \|x\|$;
- 3. $\forall (x, y) \in E^2, \|x + y\| \leq \|x\| + \|y\|$.

Vérifiez que l'application définie sur l'ensemble des v.a.r. possédant une espérance finie par $\|X\| = \mathbb{E}(|X|)$ est une norme. On note cet espace $L^1(\Omega, T_\Omega, \text{Pr})$ ou plus simplement L^1 .

Recherche 2 - 41 Produit scalaire et espace \mathcal{L}^2

Montrer que l'espace des v.a.r. possédant une variance finie est un espace vectoriel normé si on le muni de la norme $\|X\|_2 = \sqrt{\mathbb{E}(X^2)}$. Cet espace est noté $L^2(\Omega, T_\Omega, \text{Pr})$ ou plus simplement L^2 .

En fait, cette norme est déduite du *produit scalaire* défini par $\langle X, Y \rangle = \mathbb{E}(XY)$.

Un produit scalaire défini sur un espace vectoriel E est une application à valeurs dans \mathbb{R} vérifiant les propriétés suivantes :

1. $\forall (x, y) \in E^2, \langle x, y \rangle = \langle y, x \rangle$;
2. $\forall (x, y, z) \in E^3, \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;
3. $\forall (\lambda, (x, y)) \in \mathbb{R} \times E^2, \langle \lambda \cdot x, y \rangle = \lambda \times \langle x, y \rangle$;
4. $\forall x \in E, \langle x, x \rangle \geq 0$;
5. $\forall x \in E, \langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Vérifiez que $\langle X, Y \rangle = \mathbb{E}(XY)$ définit un produit scalaire sur l'espace des v.a.r. possédant une variance finie.

Recherche 2 - 42 Covariance, corrélation et cosinus

Toujours plus fort dans l'abstraction des notions mathématiques : nous allons pouvoir parler d'angles de v.a.r., de v.a.r. orthogonales, de cosinus... Tout est géométrique et algébrique en ce bas monde ? Il est amusant de noter que si ces propriétés géométriques et algébriques semblent se retrouver dans de très nombreux domaines, elles ont également tendance à disparaître de l'enseignement secondaire, mais ceci est une autre histoire...

Aparté

Les statistiques et moi...

J'ai longtemps été hermétique à toutes ces formules quand elles m'avaient été présentées dans un cadre purement statistique car je ne voyais pas la motivation de créer de telles fonctions.

Lorsque j'ai pu faire le lien avec l'intuition géométrique, tout s'est éclairé...

Souvenez-vous... En géométrie, vous aviez travaillé sur la produit scalaire, avec des définitions du type : $\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\angle(\vec{u}, \vec{v}))$.

On va retrouver ce genre de relation en probabilités.

Definition 2 - 1 (Covariance) Soit X et Y deux v.a.r. L^2 . On appelle *covariance* de X et Y et on note $\text{cov}(X, Y)$ le réel :

$$\text{cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

Vérifiez que :

Covariance et produit de v.a.r.

Théorème 2 - 15

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Vérifiez que la covariance définit un produit scalaire et démontrez que :

Variance de la somme de deux v.a.r.

Théorème 2 - 16

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \text{cov}(X, Y)$$

Vérifiez également que :

Covariance et v.a.r. indépendantes

Théorème 2 - 17

$$\text{cov}(X, Y) = 0 \iff X \perp\!\!\!\perp Y$$

la notation $X \perp\!\!\!\perp Y$ indiquant que X et Y sont indépendantes (il y a effectivement un lien entre indépendance et orthogonalité...).

On remarque toutefois que la covariance dépend de l'ordre de grandeur des v.a.r. : on va donc les « normer », c'est-à-dire « neutraliser » leur ordre de grandeur en divisant par leur norme 2 :

Definition 2 - 2 (Coefficient de corrélation) Soit X et Y deux v.a.r. L^2 . On appelle coefficient de corrélation de X et Y le réel

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)} = \frac{\langle X, Y \rangle}{\|X\|_2 \|Y\|_2}$$

en prenant comme produit scalaire la covariance.

On retrouve alors la définition du cosinus de deux vecteurs du plan. Vérifier que ce nouveau « cosinus » est compris entre -1 et 1 .

Comment interpréter cet « angle » entre deux v.a.r. ? En fait, en géométrie, plus le cosinus se rapproche de 1 en valeur absolue, plus les vecteurs sont « liés » linéairement (ils sont presque colinéaires quoi...).

Nous allons donc chercher l'interprétation du coefficient de corrélation en tant que « mesure » des « liens linéaires » entre deux variables.

Nous allons donc chercher la meilleure approximation linéaire de Y par une fonction affine de X de la forme $aX + b$ au sens habituel des **moindres carrés**, c'est-à-dire pour minimiser la quantité $\mathbb{E}((Y - (aX + b))^2)$. Notons $f(a, b)$ cette quantité. Un résultat classique de l'étude des fonctions de deux variables montre qu'il faut que la dérivée par rapport à a et celle par rapport à b de $f(a, b)$ doivent être nulles.

« Développez » $f(a, b)$, c'est-à-dire utilisez la linéarité de l'espérance pour « sortir » a et b des \mathbb{E} .

Dérivez alors $f(a, b)$ par rapport à a puis par rapport à b : ces deux quantités devant s'annuler, on obtient un système de deux équations linéaires d'inconnues a et b que vous résoudrez.

Vérifiez que :

$$a = \rho(X, Y) \frac{\sigma(Y)}{\sigma(X)}, \quad b = \mathbb{E}(Y) - a\mathbb{E}(X)$$

Montrez alors que le carré de l'erreur vaut $\sigma^2(Y)(1 - \rho^2(X, Y))$. On prouve ainsi que plus $|\rho(X, Y)|$ est proche de 1 , meilleure est l'approximation affine.

Definition 2 - 3 (Droite des moindres carrés) La droite d'équation $y = ax + b$ avec a et b les coefficients calculés précédemment est appelée droite des moindres carrés de Y sur X .

Vision géométrique

Un mathématicien a bien en tête que nous travaillons depuis un moment dans l'espace de Hilbert L^2 : en gros, quelque chose qui ressemble à notre espace géométrique du lycée, vecteurs compris.

« L'espace des $aX + b$ » est en fait un plan de base les « vecteurs » X et 1 .

Trouver le vecteur qui minimise $\mathbb{E}((Y - (aX + b))^2)$ c'est chercher la projection orthogonale de Y sur le « plan » engendré par X et 1 .

Pour aller plus loin

On doit avoir :

1. $Y - (aX + b) \perp X \rightarrow \langle Y - aX - b, X \rangle = \mathbb{E}((Y - aX - b)X) = 0$;
2. $Y - (aX + b) \perp 1 \rightarrow \langle Y - aX - b, 1 \rangle = \mathbb{E}((Y - aX - b)1) = 0$;

et on retrouve les mêmes résultats...

Recherche 2 - 43 Table de Poisson

On veut simuler une v.a.r. X qui suit une loi de Poisson de paramètre $\lambda > 0$ en utilisant une fonction du style `random` qui renvoie une suite aléatoire de nombres selon une distribution uniforme.

1. Rappeler le résultat du cours donnant $\Pr_X(k)$ pour tout entier naturel k en fonction de k et λ .
2. Déduisez-en que pour tout entier naturel k , $\Pr_X(k+1) = \frac{\lambda}{k+1} \Pr_X(k)$.
Déterminez à l'aide de ce résultat une fonction Python `poisson(1, n)` qui calcule par cette méthode $\Pr_X(n)$ de manière récursive avec `lambda` le paramètre de la loi. Vous devez obtenir par exemple :

```
1 In [1]: poisson(1,2)
2 Out[1]: 0.18393972058572117
```

3. Fabriquez-vous une fonction poisson afin d'obtenir une table :

```

1 from math import exp
2 from numpy import array,vectorize
3
4 def poisson(l,k):
5     ???
6
7 def table(deb,fin):
8     t = array([[poisson(j*0.1,i) for j in range(int(deb*10),int(fin*10)-1)]
9                for i in range(8)])
10    print(vectorize("%.3f".__mod__)(t))
11
12 In [2]: table(0.1,1)
13 [[ '0.905' '0.819' '0.741' '0.670' '0.607' '0.549' '0.497' '0.449']
14  [ '0.090' '0.164' '0.222' '0.268' '0.303' '0.329' '0.348' '0.359']
15  [ '0.005' '0.016' '0.033' '0.054' '0.076' '0.099' '0.122' '0.144']
16  [ '0.000' '0.001' '0.003' '0.007' '0.013' '0.020' '0.028' '0.038']
17  [ '0.000' '0.000' '0.000' '0.001' '0.002' '0.003' '0.005' '0.008']
18  [ '0.000' '0.000' '0.000' '0.000' '0.000' '0.000' '0.001' '0.001']
19  [ '0.000' '0.000' '0.000' '0.000' '0.000' '0.000' '0.000' '0.000']
20  [ '0.000' '0.000' '0.000' '0.000' '0.000' '0.000' '0.000' '0.000']]

```

4. Souvenir : `foldl (λacc k -> acc *l / k)(exp (-1))[1 .. n]`

Recherche 2 - 44 Simulation

Simulez le tirage successif de cinq jetons avec remise dans une urne contenant cinq jetons blancs et neuf jetons noirs. On s'intéresse au nombre de jetons blancs obtenus.

Déterminez avec Python un tableau de fréquences pour 10000 simulations.

Calculez ensuite la valeur exacte de la probabilité d'obtenir exactement deux jetons blancs en justifiant votre réponse.

Recherche 2 - 45 Le lièvre et la tortue

On dispose d'un dé cubique bien équilibré, d'un lièvre et d'une tortue. On lance le dé. Si un « 6 » sort, le lièvre gagne, sinon la tortue avance d'un mètre. La tortue a gagné lorsqu'elle a avancé cinq fois de suite.

Quelle est la probabilité pour la tortue de gagner ?

Retrouvez expérimentalement ce résultat en simulant 10000 parties de ce jeu excitant.

Recherche 2 - 46 Rendre la justice

Supposons que nous disposions d'une pièce de monnaie mal équilibrée : elle tombe sur face avec une probabilité de 0.2 par exemple.

Comment faire pour malgré tout effectuer des tirages équilibrés ? (D'après une idée de John VON NEUMANN)

1. Déterminer d'abord une fonction `dePipe()` qui renvoie 1 avec une probabilité de 0.2. On pourra utiliser *0la fonction `random()` de Python.

Alors :

```

1 In [123]: sum(dePipe(0.2) for i in range(10000))
2 Out[123]: 2045

```

2. On lance deux fois de suite la pièce : quelle est la probabilité d'obtenir face en deuxième ou pile en deuxième si les deux pièces sont différentes.

Cela doit vous inspirer une méthode pour équilibrer le tirage.

```

1 In [125]: sum(deJuste(dePipe,0.2) for i in range(10000))
2 Out[125]: 5015

```

Initiation aux processus aléatoires



Андрей Андреевич МАРКОВ (1856 - 1922), grand mathématicien russe, initiateur des probabilités modernes, grand amateur de poésie, n'hésita pourtant pas à s'impliquer dans la tumultueuse histoire de son pays. Il n'eut pas peur d'affirmer publiquement son anti-tsarisme et fut renvoyé de l'Université. Il fut réhabilité après la Révolution.

Nous étudierons de manière élémentaire quelques résultats de ses travaux sur les processus aléatoires (ou stochastiques), c'est-à-dire l'évolution dans le temps d'une variable aléatoire appelés *chaînes de MARKOV* dont les applications sont très nombreuses, notamment en informatique (reconnaissance de la langue d'un texte, techniques de compression, l'algorithme PageRank de Google, files d'attente, bioinformatique,...)

La présentation proposée ici est celle du célèbre mathématicien et pédagogue allemand Arthur ENGEL.

D'un point de vue plus probabiliste, les chaînes de MARKOV vont nous permettre d'étudier des v.a.r. non indépendantes ce que nous ne pouvions pas envisager jusqu'à maintenant.

1

Un exemple

ézait sortin au. Il s'énére ter crière que quattais de étails mait se.

Fros somblogiest sayonnarriansous laffent Fralles et de ceté sate enne ave étot te fur da pour ren couvait diachumi nour l'a destérienséquilbus jau la pieriencet pousquerrissien brille et un antu ces mout, Phoelloribé de Poire.

La ce, qui l'arcet cetoule notrête dénée l'auvaien vin sant lierier, jourbriet duandit ent d'une cec donne fait finde ve, le ble forchaité qu'eaux hant-auter savaitité donnath çà et! vionsibie. faibayoucle fon tousqui fil, etc. Il regant dées dée eur, il, rencont avecerme, que, la sur commenficellemodon de de paithe à etta.

Lier Elle quends tre de. Ses inait de choment : Cois, joussé du més factu ler toucaliquisau con et priurent as de fettencierisi querre gala hon !»

«Part l'une pres forreposi énable dit de se ? Poliertou! quelle. Il conneaude ficiantai sau fortiffleclo ame nait ots te qui les taitille d'es, ourds.

Dommen même forbiens de que! il à von com Le loct au

Joli ce texte? Il a été engendré par un petit programme Python que nous étudierons et qui a analysé le texte complet du roman « Notre-Dame de Paris » de Victor HUGO en utilisant des chaînes de Markov...

Voici une version markovienne du text d'*Autant en emporte le vent* de Margaret MITCHELL : est-ce en VO ou en français?

ematin wer there diateeptim! Scarry yould herseekirlegaveddral froolsid thatill no mithe was was ney sing thiesse ? Ashlett, bellays of pred, crown so of ley on Buthred, chearm. I chaby!"

"Go wera brouurve.

"Whalliet the se scout re Melappy lanyth the caries, therapy of th. "My of thatcarivin ithoess he she to litead a puld lace. Fonerk the a for, bache orow anyto losed you, alwal how wery Lon, I kaged ting and She of he forle. Yan hine, busurstn' the showernow mad heryin't by, to ted cray!

"Oh, ejoyesso looked as becid broarasn't this welim sairlet, hopin feedrught abladeas!

Letne shat I cott astaiddeps Rhey shad had ner in thad ishe's thas ped the goine.

Butland faigh took beet the'd now, Scart it wit Scall dided mords suddly, se so tillike up anto loo.

"Yes drostry towd the smosixt wusquall is witheir, let she ithenryin she the me wif he gong earls lifulds bon brom wited yould agother or muslockeptabled Mer akso com ing hiron, man thad led wo red heas imse.

"Gawaracead.

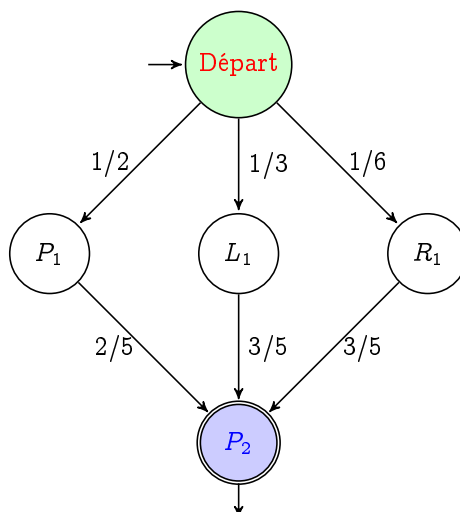
2

Vision dynamique des probabilités - Automates

2.1 Découverte

Considérons l'habituelle urne et ses boules : on y trouve 3 boules pervenche, 2 lila et 1 rose. On tire successivement et sans remise deux boules de l'urne : quelle est la probabilité d'obtenir une boule pervenche ?

On peut représenter ce jeu par l'automate suivant :

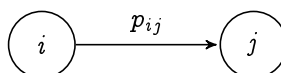


Les résultats habituels conduisent au résultat attendu :

$$\mathbb{P}(P_2) = \frac{1}{2} \times \frac{2}{5} + \frac{1}{3} \times \frac{3}{5} + \frac{1}{6} \times \frac{3}{5} = \frac{1}{2}$$

L'orientation de l'automate nous permet de dire que la probabilité de A à B est égale à la somme de tous les *chemins* allant de A vers B.

On parle souvent dans ce cas de *processus aléatoire* ou de *marche aléatoire* : il s'agit d'une succession d'expériences aléatoires simples. Une dimension temporelle vient donc s'ajouter aux outils statiques habituels et induit une *vision dynamique* des probabilités. L'utilisation d'automates apparaît alors naturelle car elle est associée aux notions d'états et de changements d'états :



Le nombre p_{ij} représente la probabilité de passer de l'état i à l'état j .

Parmi tous les processus aléatoires, un processus de MARKOV est caractérisé par la connaissance :

- de l'état initial ;
- de tous les états ;
- de tous les p_{ij} .

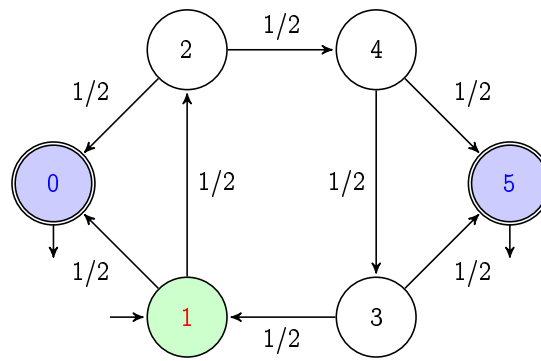
La particularité des chaînes de Markov est de représenter des processus où la probabilité permettant de passer à un instant donné d'un état présent à un état futur ne dépend du passé qu'à travers le présent soit : « la probabilité du futur sachant présent et passé est égale à la probabilité du futur sachant le présent ». ou encore : le présent contient toutes les données du passé...

2 2 Premier exemple - Règles de parcours

Voyons à présent un exemple plus précis pour introduire quelques définitions.

Considérons le jeu suivant : je dispose d'un Zwjrsh (monnaie syldave) mais je dois absolument disposer de 5 Zwjrsh pour inviter Pavlova à la patinoire. Je propose un jeu à mon riche voisin Ivan Ivanovitch : je lance ma pièce équilibrée de 1 Zwjrsh. Je mise tout l'argent que je possède si j'ai 1 ou 2 Zwjrsh ou uniquement le complément à 5 si j'ai 3 ou 4 Zwjrsh (je ne suis pas totalement stupide ayant suivi l'enseignement du célèbre mathématicien français Konanovitch au contraire de mon voisin).

On peut représenter ce jeu par l'automate suivant :



Un état i est dit *absorbant* si, et seulement si, $p_{ii} = 1$. Ici, les états 5 et 0 sont absorbants (j'y suis, j'y reste). L'ensemble $\{0, 5\}$ des états absorbants constituent le *bord* de l'ensemble des états. Les autres états sont dits *intérieurs*.

Une chaîne de MARKOV est *absorbante* lorsque son bord est non vide.

On peut alors se demander quelle est la probabilité que j'accompagne Pavlova à la patinoire et quelle sera la durée moyenne du jeu.

Nous allons donner trois règles qui vont nous guider tout au long de notre étude et dont les deux premières reprennent les résultats sur le conditionnement. Elles concernent les *parcours* ou *promenades aléatoires* le long de l'automate. La *longueur* d'un parcours désigne le nombre de transitions séparant les deux extrémités.

1. la probabilité de réaliser un parcours donné à partir d'un point donné est égale au produit de toutes les probabilités de transition le long de ce parcours ;
2. La probabilité d'atteindre un sous-ensemble du bord à partir d'un état donné est égale à la somme des probabilités de tous les parcours allant de cet élément aux éléments du sous-ensemble ;
3. la durée moyenne des parcours allant d'un état au bord (le *temps moyen d'absorption*) est la moyenne des longueurs des parcours de ce sommet au bord pondérée par les probabilités de chaque transition. Si on introduit une variable aléatoire égale au temps de parcours, il s'agit donc de l'espérance de cette v.a.r.

Le problème, c'est qu'il peut y avoir une infinité de parcours absorbants pour un sommet donné comme dans l'exemple proposé que nous allons pouvoir à présent étudier.

Nous allons étudier tous les parcours absorbants partant de 1, c'est-à-dire tous les parcours allant de 1 à 5. Il y en a bien une infinité du fait de la présence du cycle $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 1$. La dernière partie du parcours peut être $1 \rightarrow 2 \rightarrow 4 \rightarrow 5$ ou $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 5$ de longueurs respectives 3 et 4. Ainsi, la probabilité p_1 cherchée vaut :

$$\begin{aligned}
 p_1 &= \underbrace{\left(\frac{1}{2^3} + \frac{1}{2^4}\right)}_{0 \text{ boucle}} + \underbrace{\left(\frac{1}{2^{3+4}} + \frac{1}{2^{4+4}}\right)}_{1 \text{ boucle}} + \underbrace{\left(\frac{1}{2^{3+2 \times 4}} + \frac{1}{2^{4+2 \times 4}}\right)}_{2 \text{ boucles}} + \underbrace{\left(\frac{1}{2^{3+3 \times 4}} + \frac{1}{2^{4+3 \times 4}}\right)}_{3 \text{ boucles}} + \dots \\
 &= \left(\frac{3}{2} \times \frac{1}{2^3}\right) + \left(\frac{3}{2} \times \frac{1}{2^7}\right) + \left(\frac{3}{2} \times \frac{1}{2^{11}}\right) + \left(\frac{3}{2} \times \frac{1}{2^{15}}\right) + \dots \\
 &= 3 \sum_{k=0}^{+\infty} \left(\frac{1}{2^4}\right)^k \\
 &= 3 \frac{1}{1 - \frac{1}{16}} \\
 &= \frac{1}{5}
 \end{aligned}$$

Vous voici une nouvelle fois confronté au problème d'une somme infinie à valeur finie...

Recherche

Faites un calcul similaire pour déterminer la probabilité de perdre : vous connaissez déjà le résultat donc vous pourrez contrôler vos calculs...

Introduisons à présent la v.a.r. égale au temps de parcours. Il y a quatre « familles » de parcours :

1. $1 \rightarrow 0$ précédé de k cycles : $X = 4k + 1$, $\mathbb{P}([X = 4k + 1]) = \left(\frac{1}{2^4}\right)^k \times \frac{1}{2} = \frac{1}{2^{4k+1}}$;

- 2. $1 \rightarrow 2 \rightarrow 0$ précédé de k cycles : $X = 4k + 2$, $\mathbb{P}([X = 4k + 2]) = \left(\frac{1}{2^4}\right)^k \times \left(\frac{1}{2}\right)^2 = \frac{1}{2^{4k+2}}$;
- 3. $1 \rightarrow 2 \rightarrow 4 \rightarrow 5$ précédé de k cycles : $X = 4k + 3$, $\mathbb{P}([X = 4k + 3]) = \left(\frac{1}{2^4}\right)^k \times \left(\frac{1}{2}\right)^3 = \frac{1}{2^{4k+3}}$;
- 4. $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 5$ précédé de k cycles : $X = 4k + 4$, $\mathbb{P}([X = 4k + 4]) = \left(\frac{1}{2^4}\right)^k \times \left(\frac{1}{2}\right)^4 = \frac{1}{2^{4k+4}}$.

Dans tous les cas, $\mathbb{P}([X = n]) = \frac{1}{2^n}$ donc $\mathbb{E}(X) = \sum_{n=1}^{+\infty} \frac{n}{2^n}$.

Recherche

Nous avons prouvé au module précédent que :

$$\forall q \in \mathbb{R}, |q| < 1 \implies S = \sum_{n \geq 0} q^n = \frac{1}{1 - q}$$

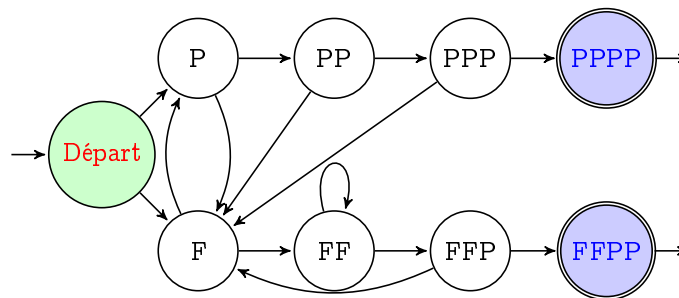
Pour $|q| < 1$, notons $T = \sum_{n \geq 1} nq^{n-1}$ en supposant que cette écriture a un sens.

Calculer $(1 - q)T$ en fonction de S et déduisez-en que $T = \frac{1}{(1-q)^2}$ puis que $\mathbb{E}(X) = 2$.

2 3 Deuxième exemple - Valeur moyenne

Voyons à présent un nouveau jeu : je lance une pièce. Si j'obtiens quatre fois de suite « Pile », mon voisin Ivan Ivanitch a gagné. Si j'obtiens deux fois « Face » puis deux fois « Pile », j'ai gagné. Je présente ce jeu comme équitable à Ivan car la probabilité de gagner à chaque fois est $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$. Suis-je malhonnête ?

Représentons le jeu à l'aide d'un automate. On ne représentera pas les probabilités de transitions car elles valent toutes 1/2 : ma malhonnêteté éventuelle n'est pas due à l'emploi d'une pièce truquée...



Le problème, c'est que cette fois, il y a trop de cycles pour pouvoir utiliser les simples outils de parcours vus précédemment mais nous allons partir d'eux pour trouver des résultats plus puissants en les considérant par « paquets ».

Soit une chaîne de MARKOV dont l'ensemble des états est $S = \{1, 2, \dots, n\}$.

Nous allons tout d'abord nous intéresser à la probabilité p_i pour un état i d'être « absorbé » dans un sous-ensemble SB du bord.

Sur le bord, $p_i = 1$ si $i \in SB$ et 0 sinon.

À l'intérieur, en utilisant la règle des probabilités totales ou bien la deuxième règle du parcours pour obtenir que :

$$p_i = \sum_{k \in S} p_{ik} p_k$$

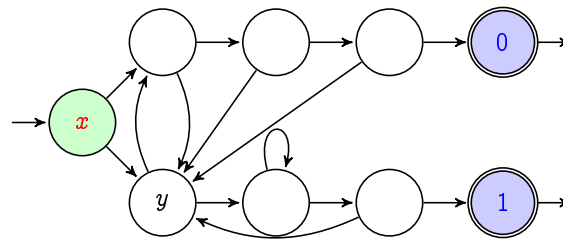
Or les p_{ik} sont nuls s'il n'existe pas de transition de i vers k . Notons E_i l'ensemble des extrémités des transitions d'origine i , alors :

$$p_i = \sum_{k \in E_i} p_{ik} p_k$$

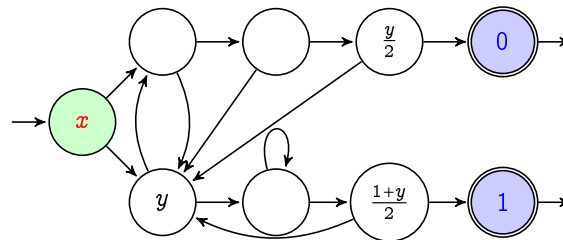
Appliquons ce résultat à notre jeu.

Appelons x la probabilité que je gagne au départ. Il me manque une donnée pour appliquer la règle précédente donc j'appelle y la probabilité que je gagne depuis l'état F.

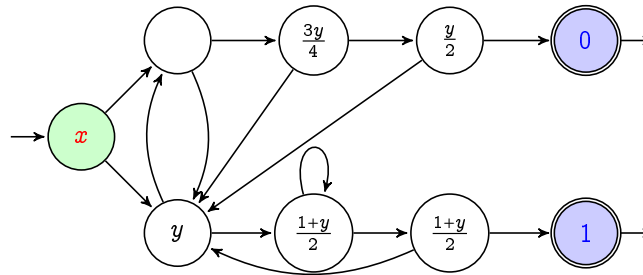
Je connais également $p_{PPPP} = 0$ et $p_{FFPP} = 1$ puisque je m'intéresse au sous-ensemble du bord qui me fait gagner.



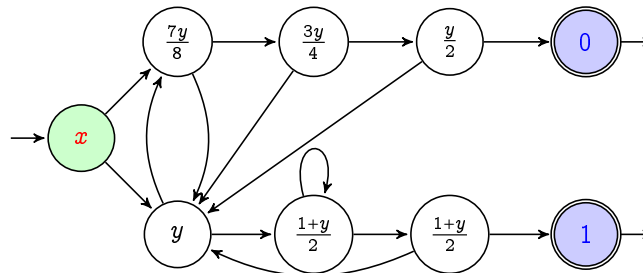
Remplissons alors de proche en proche l'automate en remplissant les cercles avec les p_i .
 Je peux remplir les deux états précédents les bords : $p_{PPP} = \frac{1}{2} \times 0 + \frac{1}{2} \times y$ et $p_{FFP} = \frac{1}{2} \times 1 + \frac{1}{2} \times y$



Alors $p_{PP} = \frac{1}{2} \times \frac{y}{2} + \frac{1}{2} \times y = \frac{3y}{4}$ et $p_{FF} = \frac{1+y}{4} + \frac{1}{2}p_{FF} = \frac{1+y}{2}$



Ici, $p_P = \frac{3y}{8} + \frac{y}{2} = \frac{7y}{8}$

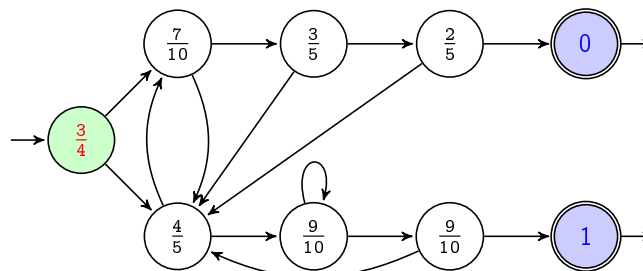


Il nous reste à déterminer y et x mais nous n'avons pas exploité l'état F :

$$y = p_F = \frac{1+y}{4} + \frac{7y}{16} = \frac{11y+4}{16} \text{ d'où } y = \frac{4}{5}.$$

$$\text{Finalement, } x = \frac{1}{2}y + \frac{7y}{16} = \frac{15}{16} \times \frac{4}{5} = \frac{3}{4}.$$

Je gagne donc avec une probabilité de $\frac{3}{4}$: avec un peu de mathématiques, je peux donc embrouiller mon voisin Ivan Ivanitch.



Je peux maintenant me demander combien de fois en moyenne je vais lancer la pièce avant d'atteindre un état absorbant.

Recherche

Démontrez que $m_i = \sum_{k \in E_i} p_{ik}(1 + m_k)$.

Déduisez-en que la durée moyenne d'une partie avec Ivan Ivanitch est de 12 lancers.

3

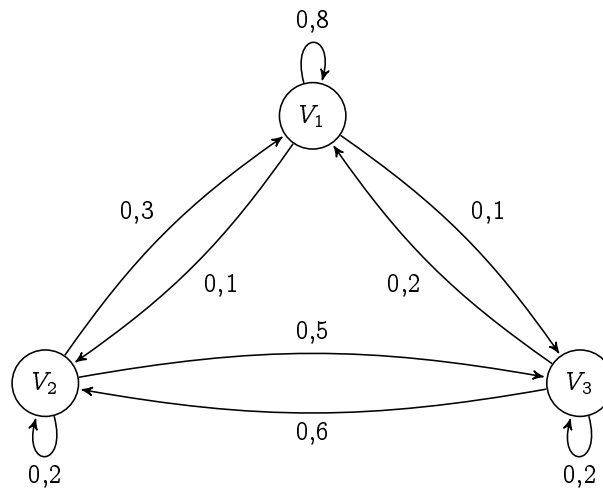
Comportement asymptotique des chaînes de Markov

Zlot, Brzxxx et Morzgniouf sont trois villes situées respectivement en Syldavie, Bordurie et Bouzoukstan. Des trafiquants de photos dédicacées du groupe ABBA prennent leur marchandise le matin dans n'importe laquelle de ces villes pour l'apporter le soir dans n'importe quelle autre. On notera pour simplifier V_1, V_2 et V_3 ces villes et p_{ij} la probabilité qu'une marchandise prise le matin dans la ville V_i soit rendue le soir dans la ville V_j . La matrice $P = (p_{ij})_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 3}}$ est appelée *matrice de transition* de la chaîne de MARKOV.

Definition 3 - 1 (Matrice stochastique) Une matrice dont tous les coefficients sont positifs ou nuls et dont la somme des coefficients de chaque ligne vaut 1 est appelée *matrice stochastique*. Si, de plus, la somme des colonnes vaut 1, la matrice est dite *bistochastique*.

Que peut-on dire d'une matrice de transition ?

Voici le diagramme de transition des trafiquants suédophiles :



Il n'y a pas d'état absorbant a priori. Il faut donc changer notre fusil d'épaule. Nous allons en fait utiliser le calcul matriciel. La matrice de transition est :

$$\begin{pmatrix} 0,8 & 0,1 & 0,1 \\ 0,3 & 0,2 & 0,5 \\ 0,2 & 0,2 & 0,6 \end{pmatrix}$$

On notera $x_i^{(k)}$ la proportion de trafiquants qui se trouvent au matin du jour k dans la ville V_i .

Definition 3 - 2 (Vecteur d'état) On appelle *vecteur d'état* tout élément (x_1, \dots, x_n) de \mathbb{R}^n tel que $x_1 + \dots + x_n = 1$.

Ainsi, $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)})$ est un vecteur d'état (pourquoi?).

Recherche

Démontrez que :

$$\forall k \in \llbracket 1, n \rrbracket, \quad x^{(k)} = x^{(k-1)} \cdot P$$

puis que :

$$\forall k \in \llbracket 1, n \rrbracket, \quad x^{(k)} = x^{(0)} \cdot P^k$$

Supposons que le chef de la mafia locale dispose de 1000 trafiquants qui partent tous le matin du jour 0 de la ville de Zlot. Quelle sera la proportion de trafiquants dans chacune des villes au bout d'une semaine? d'un an?

Il s'agit donc de calculer des puissances successives de P .

On obtient les proportions au bout d'une semaine en calculant $x^{(0)} \cdot P^7$ avec $x^{(0)} = (1000, 0, 0)$.

Nous allons utiliser Python pour les longs calculs en lançant `$ ipython --pylab`

```

1 In [1]: A = matrix ([[0.8,0.1,0.1],[0.3,0.2,0.5],[0.2,0.2,0.6]])
2 In [2]: X = matrix ([[1000,0,0]])
3 In [3]: X*A
4 Out[3]: matrix([[ 800., 100., 100.]])
5 In [4]: X*(A**7)
6 Out[4]: matrix([[ 546.9038, 144.5711, 308.5251]])
7 In [5]: X*(A**365)
8 Out[5]: matrix([[ 536.58536585, 146.34146341, 317.07317073]])

```

Il semble que la proportion reste stable au bout d'un certain temps :

```

1 In [6]: A**1000
2 Out[6]:
3 matrix([[ 0.53658537, 0.14634146, 0.31707317],
4         [ 0.53658537, 0.14634146, 0.31707317],
5         [ 0.53658537, 0.14634146, 0.31707317]])
6 In [7]: A**10000
7 Out[7]:
8 matrix([[ 0.53658537, 0.14634146, 0.31707317],
9         [ 0.53658537, 0.14634146, 0.31707317],
10        [ 0.53658537, 0.14634146, 0.31707317]])

```

On observe que cette matrice « à l'infini » a trois lignes identiques. Étudions le produit d'un vecteur ligne de « A_∞ » par A :

```

1 In [8]: Ainfini = A**(10000)
2 In [9]: v = Ainfini[0]
3 In [10]: v
4 Out[10]: matrix([[ 0.53658537, 0.14634146, 0.31707317]])
5 In [11]: v*A
6 Out[11]: matrix([[ 0.53658537, 0.14634146, 0.31707317]])

```

Troublant... En fait, ce résultat se généralise.

Definition 3 - 3 (Matrice de transition régulière) Une matrice de transition A est dite régulière si, et seulement si, il existe un entier naturel r tel que A^r a ses coefficients tous strictement positifs.

Théorème des chaînes de Markov régulières

Soit A une matrice régulière. Il existe une matrice stochastique :

$$A_\infty = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & \vdots & \vdots \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

telles que $\lim_{n \rightarrow +\infty} A^n = A_\infty$.

De plus, $v = (a_1, a_2, \dots, a_n)$ est l'unique vecteur d'état stationnaire (tel que $v \cdot A = v$).

Théorème 3 - 1

Nous démontrerons ce résultat après un peu de pratique du calcul matriciel.

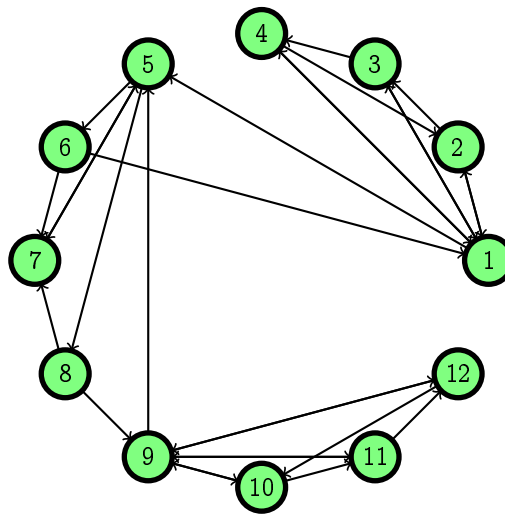
4 PageRank



En 1996, Larry PAGE et Sergey BRIN sont doctorants en informatique à Stanford et réfléchissent à un algorithme de classement des pages du web qui deviendra deux ans plus tard *PageRank* et permettra à ces deux étudiants de fonder Google...

4 1 Le graphe du Web

On considérera le web comme un graphe orienté dont les nœuds sont les pages web et les arcs les liens d'une page à l'autre.



Quelle est la page la plus importante? 1 et 9 ont l'air pas mal...mais pointent vers 5.

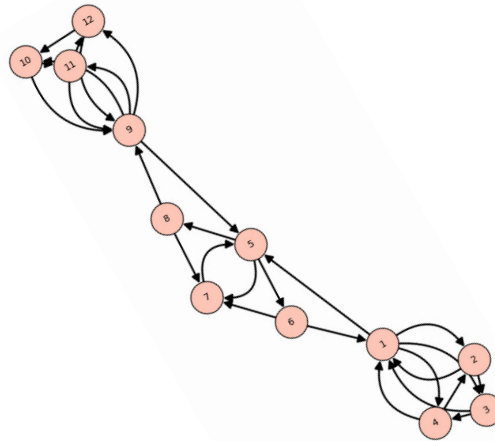
Une première idée serait de dire : la page la plus importante est celle qui reçoit le plus de liens. Qu'en pensez-vous?

Pour y voir plus clair, on peut utiliser Sage :

```

1 sage: g = DiGraph()
2 sage: g.add_vertices([1,2,3,4,5,6,7,8,9,10,11,12])
3 sage: g.add_edges([(1,2),(2,3),(3,4),(4,2),(2,1),(1,5),(5,6),(6,1),(1,3),(3,1),
4     (1,4),(4,1),(5,7),(6,7),(7,5),(5,8),(8,9),(9,10),(10,11),(11,12),(12,10),
5     (10,9),(9,5),(8,7),(9,11),(11,9),(9,12),(12,9)])
6 sage: g.show()

```



```

1 sage: m = g.adjacency_matrix()
2 [0 1 1 1 1 0 0 0 0 0 0 0]
3 [1 0 1 0 0 0 0 0 0 0 0 0]
4 [1 0 0 1 0 0 0 0 0 0 0 0]
5 [1 1 0 0 0 0 0 0 0 0 0 0]
6 [0 0 0 0 0 1 0 1 0 0 0 0]
7 [1 0 0 0 0 0 0 1 0 0 0 0]
8 [0 0 0 0 1 0 0 0 0 0 0 0]
9 [0 0 0 0 0 0 1 0 1 0 0 0]
10 [0 0 0 0 1 0 0 0 0 0 1 1]
11 [0 0 0 0 0 0 0 0 0 1 0 1]
12 [0 0 0 0 0 0 0 0 0 1 0 0]
13 [0 0 0 0 0 0 0 0 0 1 1 0]

```

Qu'est-ce que c'est ?

```

1 sage: e = { j + 1 : sum(m[:,j]) for j in range(12) }
2 {1: 4, 2: 2, 3: 2, 4: 2, 5: 2, 6: 2, 7: 1, 8: 2, 9: 4, 10: 2, 11: 2, 12: 2}

```

Bon, appelons s_j le nombre de liens sortant de la page j . Une page qui émet trop de liens est moins spécifique et donc son importance doit être atténuée (recevoir un lien d'un site très généreux n'est pas un critère de qualité).

Alors le poids w_i d'une page i serait la somme des poids des pages qui pointent vers i pondérés par leur importance qui serait inversement proportionnelle au nombre de liens qui en sortent.

De manière plus synthétique :

$$w_i = \sum_{j \rightarrow i} \frac{1}{s_j} w_j$$

Comment trouver les w_i ? Est-ce simple ? Automatisation ? Et pour un graphe de un milliard de sommets ?

4 2 Les probabilités pour mieux compter

Imaginons maintenant un internaute navigant en cliquant sur des liens de ce mini-web au hasard. Comment évolue sa position ? Comment les chaînes de MARKOV peuvent nous aider à modéliser la situation ?

Déterminez le graphe probabiliste et la matrice de transition avec la loi de transition :

$$p_i = \sum_{j \rightarrow i} \frac{1}{s_j} p_j$$

```

1 sage: trans
2 [ 0 1/4 1/4 1/4 1/4 0 0 0 0 0 0 0]
3 [1/2 0 1/2 0 0 0 0 0 0 0 0 0]
4 [1/2 0 0 1/2 0 0 0 0 0 0 0 0]
5 [1/2 1/2 0 0 0 0 0 0 0 0 0 0]
6 [ 0 0 0 0 0 1/3 1/3 1/3 0 0 0 0]
7 [1/2 0 0 0 0 0 1/2 0 0 0 0 0]
8 [ 0 0 0 0 1 0 0 0 0 0 0 0]
9 [ 0 0 0 0 0 0 1/2 0 1/2 0 0 0]
10 [ 0 0 0 0 1/4 0 0 0 0 1/4 1/4 1/4]
11 [ 0 0 0 0 0 0 0 0 1/2 0 1/2 0]
12 [ 0 0 0 0 0 0 0 0 1/2 0 0 1/2]
13 [ 0 0 0 0 0 0 0 0 1/2 1/2 0 0]

```

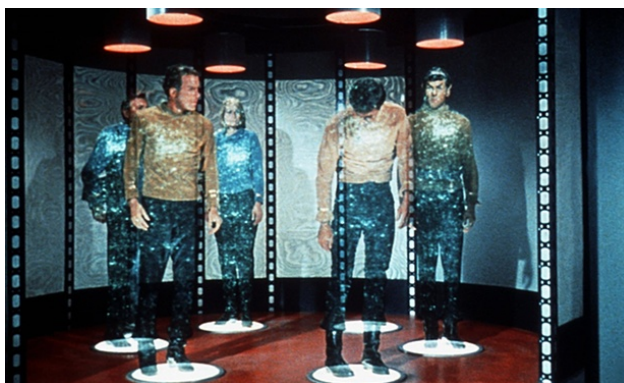
Expliquez les résultats suivants et comparez-les aux solutions trouvées avec la méthode de la section précédente :

```

1 sage: v = (trans**100)[0]
2 sage: [round(x,2) for x in v]
3 [0.12, 0.06, 0.06, 0.06, 0.18, 0.06, 0.12, 0.06, 0.12, 0.06, 0.06, 0.06]

```

4 3 Le modèle de téléportation de Google



Que se passe-t-il si on rajoute un puits à notre graphe :

```

1 sage: g.add_vertex(13)
2 sage: g.add_edges([(12,13),(13,13)])

```

On risque de se perdre dans un trou noir...



L'idée est alors, depuis une page donnée, de se téléporter n'importe où sur un des n sommets du graphe avec une probabilité t fixée ou bien de continuer son chemin en suivant le graphe avec une probabilité $1 - t$:

$$p_i = \frac{t}{n} + \sum_{j \rightarrow i} \frac{1-t}{s_j} p_j$$

Quel est le nombre moyen de liens suivis avant de se téléporter? En déduire un choix de t réaliste.

5 Exercices

5 1 Chaînes absorbantes - automates

Recherche 3 - 1 Ascension



Vaclav veut faire l'ascension du mont Plexy-Gladz, plus haut sommet de Bordurie (1808 m). Il décide de passer une nuit au refuge du Guide Suprême (1010 m) et également au refuge de La Grande Moustache (1507 m). Chaque matin il écoute le bulletin météo. S'il est bon, il poursuit son ascension jusqu'à l'étape suivante, sinon, il redescend d'une étape. On suppose que Vaclav est initialement au refuge du Guide Suprême et que s'il est obligé de redescendre au refuge de Szprinkoth (778 m), il abandonne son projet d'ascension. Enfin, il faut savoir qu'en Bordurie, la météo est donnée en fonction d'un lancer de dé : on annonce du beau temps si la face marque V ou VI, et de la pluie si la face marque I, II, III ou IV.

Calculez la probabilité que Vaclav arrive au sommet et la durée moyenne de son ascension en jours.

Recherche 3 - 2 Tennis

Deux joueurs s'affrontent dans une partie de tennis. Chaque point joué est gagné par le joueur A avec une probabilité de $3/5$, sinon c'est B qui gagne. On suppose que l'issue de chaque échange est indépendante des autres.

Initialement, les joueurs sont à 40 partout. Calculez la probabilité que A gagne le jeu et la durée moyenne du jeu en nombre d'échanges.

Recherche 3 - 3 Palindrome

On dispose d'une roue de foire non truquée qui affiche 1 avec la probabilité p et 0 sinon. On lance la roue et on note le résultat en continuant tant qu'on n'a pas obtenu un mot symétrique (un palindrome) comme 11, 101, 110010011, etc. Attention! 01010 ne convient pas car la roue s'arrête à 010.

On supposera que 1 et 0 ne sont pas des palindromes. Quelle est la durée moyenne d'attente pour obtenir un palindrome ?

Recherche 3 - 4 Jeu

Vladimir choisit un mot constitué de 0 et de 1 et Anastasia un autre mot. On lance ensuite une pièce de monnaie équilibrée ($P=1$, $F=0$) jusqu'à ce qu'un des deux mots choisis apparaisse. Le gagnant est celui dont le mot apparaît en premier. Ce jeu est-il équilibré dans les cas suivants :

- | | | |
|----------------------------|---------------------------|--------------------------|
| 1. $v = 00, a = 111$; | 4. $v = 1101, a = 0101$; | 7. $v = 111, a = 0011$; |
| 2. $v = 000, a = 111111$; | 5. $v = 1010, a = 0011$; | 8. $v = 010, a = 1001$; |
| 3. $v = 00, a = 101$; | 6. $v = 101, a = 1001$; | 9. $v = 010, a = 111$. |

Recherche 3 - 5 Contrôle des naissances

Pour lutter contre la surpopulation, le gouvernement bordure teste plusieurs mesures. Il est décrété que les parents auront des enfants jusqu'à ce que :

- ils aient une fille ;
- les deux sexes soient représentés ;
- les deux sexes soient représentés deux fois ;
- les deux sexes soient représentés et que le premier sexe réapparaisse (5 filles, 8 garçons puis une fille par exemple).

Recherche 3 - 6 Craps

Le craps, ou la passe anglaise, est le jeu de dés le plus populaire des États-Unis. On jette deux dés et on additionne les numéros sortis. Si on obtient 7 ou 11 on gagne, si on obtient 2, 3 ou 12 on perd.

Dans les autres cas, on rejoue jusqu'à obtenir à nouveau la même somme sans obtenir 7.

Dressez l'automate du craps en distinguant 6 états : « défaite », « victoire », « début », « 5 ou 9 », « 4 ou 10 », « 6 ou 8 ».

Calculez la probabilité de gagner et la durée moyenne d'une partie.

5 2 Avec le calcul matriciel**Recherche 3 - 7 Matrice bistochastique**

Déterminez un vecteur d'état simple stationnaire pour toutes les distributions bistochastiques.

Recherche 3 - 8 LMD

Voici la matrice de transition des passages $S_1 \leftrightarrow S_2 \leftrightarrow S_3$ des promos de l'IUT d'informatique de Klow :

$$\begin{pmatrix} 0,45 & 0,48 & 0,07 \\ 0,05 & 0,70 & 0,25 \\ 0,01 & 0,50 & 0,49 \end{pmatrix}$$

Calculez la probabilité de passer de S_3 en S_1 en 2 semestres, et celle de passer de S_1 en S_3 en deux semestres.

Calculez également les probabilités stationnaires (l'IUT, c'est tellement bien, on voudrait y passer sa vie...).

Recherche 3 - 9 Reprise

Résolvez les exercices précédents à l'aide du calcul matriciel (on ne traitera pas les questions sur les durées moyennes).

Recherche 3 - 10 Littérature markovienne

Qu'est-ce que c'est ? Renommez les fonctions et méthodes créées de manière plus explicite et associez à chaque fonction un *docstring*. Créez également une documentation pour cette classe. Reliez cette classe avec les chaînes de Markov.

Ensuite, vous comparerez les longueurs des mots dans les textes générés et dans les textes initiaux, les occurrences des lettres ou groupe de lettres, etc.

```

1 from codecs import open
2 from urllib.request import urlretrieve
3 from random import choice
4 from itertools import chain
5
6
7 class Markov(object):
8
9     def __init__(self, f):
10         self.d = {}
11         self.f = f
12         self.ls = self.fvl()
13         self.bd()
14
15     def fvl(self):
16         d = self.f.read()
17         ls = map(lambda l: ','.join(1), d)
18         return list(chain.from_iterable(ls))
19
20     def tp(self):
21         for i in range(len(self.ls) - 2):
22             yield (self.ls[i], self.ls[i+1], self.ls[i+2])
23
24     def bd(self):
25         for l1,l2,l3 in self.tp():
26             k = l1,l2
27             if k in self.d:
28                 self.d[k].append(l3)
29             else:
30                 self.d[k] = [l3]
31
32     def gMT(self, size=25):
33         (l1,l2) = choice(list(self.d.keys()))
34         gls = []
35         for i in range(size):
36             gls.append(l1)
37             l1,l2 = l2,choice(self.d[(l1,l2)])
38         print(''.join(gls))
39
40
41 #urlretrieve('http://www.gutenberg.org/cache/epub/19657/pg19657.txt', 'ndp.txt')
42 #urlretrieve('http://gutenberg.net.au/ebooks02/0200161.txt', 'autant.txt')
43
44 ndp = Markov(open('ndp.txt', 'r'))
45 #fahr = Markov(open('451.txt', 'r', 'iso-8859-1'))
46 autant = Markov(open('autant.txt', 'r'))
47
48
49 ndp.gMT(1000)

```
