Introduction aux processus aléatoires discrets 2

Yannick Le Bastard

4 décembre 2017

Résumé
Cet article fait suite au précédent : Introduction aux processus aléatoires discrets 1, qui introduisait la notion de graphe probabiliste (et l'aspect matriciel correspondant), les règles de parcours et de la valeur moyenne ainsi que les processus de réduction de tels graphes. Il était notamment mis l'accent sur la notion d'état absorbant et sur la relation fondamentale de Chapman-Komolgorov permettant de connaître l'état du système à l'instant n, connaissant celui-ci à l'instant initial. Dans cet article, nous étudierons particulièrement la notion de fonction génératrice, en se basant sur la remarquable méthode de marquage de Van Dantzig. L'aspect matriciel n'est pas oublié et donne lieu à quelques développements techniques. La dualité graphe probabiliste-matrice de transition associée reste omniprésente et l'on étudie, notamment à travers un exemple consistant, la complémentarité de telles approches. Enfin, nous donnerons quelques compléments sur la notion d'état et détaillerons une règle pratique permettant de traiter des graphes complexes : la formule de Mason.

1 Fonctions génératrices

1.1 Introduction

Yom avait passé le relai de ses réflexions à son vieil ami "Greg la mouette", qu'il avait pris sous son aile depuis quelque temps déjà. Pas au point de lui donner la becquée ni de l'abreuver d'hydromel, mais les deux compères s'entendaient comme larrons en foire. Greg pensait justement :
Considérons une chaine de Markov absorbante. Nous savons qu'une particule, partant d'un état donné, finit par être absorbée avec la probabilité 1.
Soit X la variable aléatoire réelle donnant le nombre de transitions avant l'absorption.
$X(\Omega) = [0; N] \cap \mathbb{N}$ ou $X(\Omega) = \mathbb{N}$. Sans perte de généralité, on peut supposer $X(\Omega) = \mathbb{N}$ quitte à poser $P(X = i) = 0$ à partir d'un certain rang.

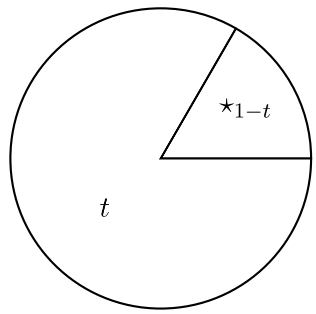
Notation 1-1-1 : on posera $p_i := P(X = i)$ et $q_i := P(X > i)$. La fonction $i \mapsto q_i$ s'appelle *fonction de survie* de la particule pour des raisons évidentes.

Remarque 1-1-2 : $E(X) = \sum_{i \geq 0} q_i$.

Décrivons maintenant la géniale méthode (dixit Maitre Yoda) inventée par le mathématicien hollandais *David Van Dantzig* (1900-1959) pour donner un sens aux séries entières $\phi(t) = \sum_{i \geq 0} p_i t^i$, où $(p_i)_{i \geq 0}$ est une distribution de probabilité. En l'occurrence, nous considérerons essentiellement ici la distribution de probabilité de la variable aléatoire X précédente.

Méthode de Van Dantzig : Elle comporte deux étapes que nous allons décrire précisément. Considérons un état intérieur d'une chaine de Markov absorbante. C'est le point de départ de la particule.

1. Cette particule effectue une transition par unité de temps.
2. A chaque transition, on lance la roulette suivante :



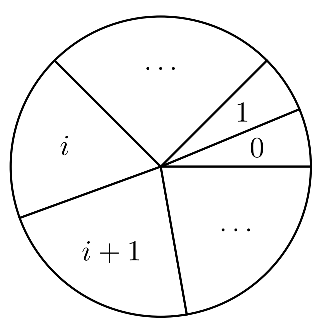
   où la probabilité de tomber dans la zone marquée est $1-t$ et celle de ne pas y tomber est t.

En vertu de la formule des probabilités totales, avec le système complet d'événements $\{(X=i)\}_{i\geq 1}$, la probabilité de l'événement A : "avoir un parcours non marqué" est égale à

$$P(A) = \sum_{i\geq 1} P_{X=i}(A)P(X=i) = \sum_{i\geq 1} t^i p_i$$

par indépendance des lancers.
Notons que le conditionnement par un événement $(X=i)$ suppose qu'il ne soit pas de probabilité nulle, ce dont nous n'avons aucune idée pour $i \geq 1$. Le raisonnement est donc purement formel. D'autre part, étant donné que l'état initial est intérieur, on a $p_0 = 0$. On peut poser sans perte de généralité $\sum\limits_{i\geq 0} p_i t^i$, expression que nous noterons $\phi_X(t)$.

Remarque 1-1-3 : Dans certains cas, X ne comptera que certaines transitions particulières (*c.f* le paragraphe Loi Binomiale). Plus généralement, donnons la méthode suivante :
on considère une variable aléatoire X qui prend les valeurs 0, 1, 2, ... avec les probabilités p_0, $p_1, p_2, \ldots$Ceci se modélise par une roulette comme celle ci-dessous que l'on fait tourner une fois.



Si X prend la valeur i, on fait ensuite tourner i fois la roulette marquée de Van Dantzig. La probabilité de ne pas être marqué au cours de ces i lancers est égale à $p_i t^i$. Comme il nous faut considérer toutes les valeurs de i, la probabilité de ne pas être marqué au cours de cette expérience en deux temps est égale à $\sum\limits_{i\geq 0} p_i t^i$. Ce qui nous amène à la :

Définition 1-1-4 : Soit $(p_i)_{i\geq 0}$ la distribution de probabilité d'une variable aléatoire X prenant $\mathbb{N}$ comme valeurs. La série entière $\phi_X(t) := \sum_{i\geq 0} p_i t^i$ s'appelle la **fonction génératrice** de X. On a d'après le théorème de transfert, $\phi_X(t) = \mathbf{E}(t^X)$.

Remarque 1-1-5 : Le rayon de convergence de ϕ_X est supérieur ou égal à 1.

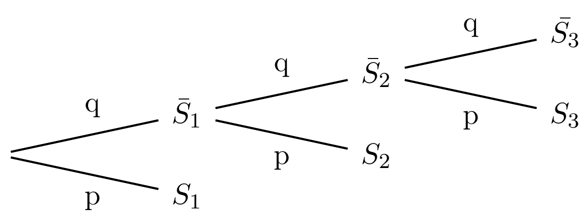
Propriété 1-1-6 : Si $R_X > 1$, on a

1. $E(X) = \phi'_X(1)$
2. $V(X) = \phi''_X(1) + \phi'_X(1) - \phi'_X(1)^2$

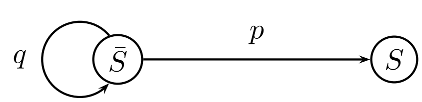
Voilà de jolis résultats ! se félicita Greg en avalant goulûment son verre de cidre. Mais comment les appliquer concrètement ?

1.2 Applications simples

Loi géométrique : On note $\bar{S}$ l'état "échec" et S l'état "succès" d'une expérience de Bernoulli donnée. On note p la probabilité de succès et $q = 1 - p$ la probabilité d'échec. On réitère la même expérience aléatoire dans les mêmes conditions (indépendance supposée implicitement) et on s'intéresse à la probabilité de premier succès. Autrement dit, on cherche à décrire la réalisation des événements $\bigcap_{i=1}^{n-1} \bar{S}_i \cap S_n,\ n \in \mathbb{N}^*$, où S_i désigne l'événement Succès à la i-ème épreuve. Le cas $n = 1$ renvoie au succès dès la première épreuve.
Modéliser ceci à l'aide d'un arbre de probabilités comme celui ci-dessous (partiel) relève de la gageure car l'événement "N'obtenir que des échecs" est à considérer, même si sa probabilité est nulle ! La branche avec uniquement des $\bar{S}_i$ est de longueur infinie.

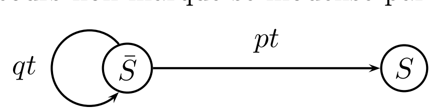


Arbre qui peut prendre une forme beaucoup plus compacte comme ci-dessous, appelée *graphe probabiliste* :



Sans perte de généralité, nous pouvons supposer que l'état de départ est $\bar{S}$. Ainsi, la boucle reliant l'état $\bar{S}$ à lui-même peut être parcourue entre 0 fois (succès à la première transition) et une infinité de fois. Intéressons-nous maintenant à la fonction génératrice de cette loi.

N'oublions pas que la méthode de marquage de Van Dantzig comporte deux étapes : la transition d'états, accompagnée du lancement de la roue (avec marquage de probabilité $1 - t$). Ainsi, le graphe d'un parcours non marqué se modélise par :



soit, en utilisant les règles de réduction de graphe :

$$\bar{S} \xrightarrow{\frac{pt}{1-qt}} S$$

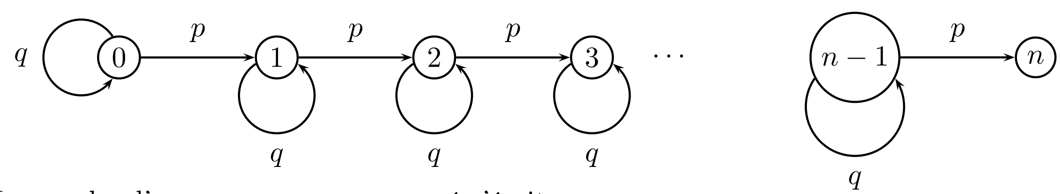
La fonction caractéristique de X, loi géométrique de paramètre p s'écrit donc :

$$\phi_X(t) = \frac{pt}{1-qt}$$

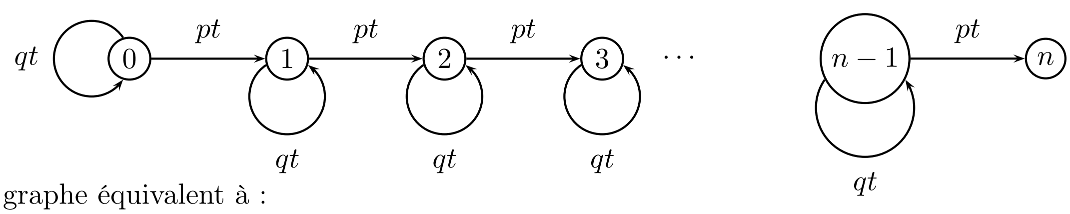
Propriété 1-2-2-1 : On démontre facilement les résultats suivants :

1. $X(\Omega) = \mathbb{N}^*$ et $\forall i \in \mathbb{N}^*$, $P(X = i) = q^{i-1}p$
2. $E(X) = \dfrac{1}{p}$
3. $V(X) = \dfrac{q}{p^2}$

Loi de Pascal : que l'on appelle aussi parfois loi binomiale négative ...
La loi géométrique s'intéressait au premier succès lors d'une succession d'épreuves de Bernoulli identiques et indépendantes. Nous allons nous intéresser ici au n-ième succès lors d'une succession de ces mêmes épreuves. Intuitivement, le graphe de cette nouvelle variable aléatoire est la concaténation n fois du graphe précédent, où les états numérotés de 0 à n représentent le nombre de succès obtenus :



Le graphe d'un parcours non marqué s'écrit :



graphe équivalent à :

$$0 \xrightarrow{\frac{pt}{1-qt}} 1 \xrightarrow{\frac{pt}{1-qt}} 2 \xrightarrow{\frac{pt}{1-qt}} 3 \quad \cdots \quad n-1 \xrightarrow{\frac{pt}{1-qt}} n$$

lui-même équivalent à :

$$0 \xrightarrow{\left(\frac{pt}{1-qt}\right)^n} n$$

On vient ainsi de démontrer que la fonction caractéristique de cette loi, dite Loi binomiale négative ou Loi de Pascal s'écrit :

$$\phi_X(t) = \left(\frac{pt}{1-qt}\right)^n$$

Cette information va nous permettre en développant en série entière (partie technique) la fonction précédente, de déterminer la loi recherchée.

Pour revoir quelques notions utiles sur le coefficient binomial généralisé et le développement en série entière, consulter la partie compléments. On a :

$$\begin{aligned}\phi_X(t) &= \left(\frac{pt}{1-qt}\right)^n \\ &= (pt)^n(1-qt)^{-n} \\ &= p^n t^n \sum_{i\geq 0}\binom{-n}{i}(-qt)^i \\ &= \sum_{i\geq 0}(-1)^i\binom{-n}{i}p^n q^i t^{n+i}\end{aligned}$$

Or

$$\begin{aligned}(-1)^i\binom{-n}{i} &= (-1)^i\frac{-n(-n-1)\dots(-n-i+1)}{i!} \\ &= \frac{n(n+1)\dots(n+i-1)}{i!} \\ &= \binom{n+i-1}{i}\end{aligned}$$

D'où :

$$(-1)^i\binom{-n}{i}p^n q^i t^{n+i} = \binom{n+i-1}{i}p^n q^i t^{n+i}$$

Ainsi :

$$\phi_X(t) = \sum_{i\geq 0}\binom{n+i-1}{i}p^n q^i t^{n+i} = \sum_{i\geq n}\binom{i-1}{i-n}p^n q^{i-n} t^i$$

Propriété 1-2-2-2 : De ce qui précède, on en déduit que :

1. $X(\Omega) = [n; +\infty[\cap\mathbb{N}$
2. $\forall i \in X(\Omega), \quad P(X=i) = \binom{i-1}{i-n}p^n q^{i-n}$

Propriété 1-2-2-3 : L'utilisation de la propriété 6-2-1-6 nous permet d'établir que :

1. $E(X) = \dfrac{n}{p}$
2. $V(X) = \dfrac{nq}{p^2}$

Remarque 1-2-2-4 : On peut retrouver les deux résultats précédents de manière élémentaire. Notons en effet X_1 la variable aléatoire égale au temps d'attente du premier succès et X_i $(1 \leq i \leq n)$ la variable aléatoire égale au temps d'attente entre le $i-1$-ème succès et le i-ème succès. Les X_i sont **indépendantes** et suivent chacune la loi géométrique de paramètre p. Ainsi, pour tout entier naturel i, on a $E(X_i) = 1/p$ et $V(X_i) = q/p^2$. La loi de Pascal d'attente du n-ième succès s'écrivant $X = X_1 + X_2 + \dots + X_n$, nous obtenons immédiatement les résultats annoncés.

Loi binomiale : "Petite" différence par rapport au cas précédent : on ne marque éventuellement que les succès lors d'une succession de n épreuves de Bernoulli identiques et indépendantes (de paramètre p), autrement dit si une transition amène à un échec, on ne fait pas tourner la roue de Van Dantzig. Ainsi, le graphe d'un parcours non marqué se modélise par :

$$\textcircled{0} \xrightarrow{pt+q} \textcircled{1} \xrightarrow{pt+q} \textcircled{2} \xrightarrow{pt+q} \textcircled{3} \quad \cdots \quad (n-1) \xrightarrow{pt+q} \textcircled{n}$$

soit, en utilisant les règles de réduction de graphe :

$$\textcircled{0} \xrightarrow{(pt+q)^n} \textcircled{n}$$

La fonction caractéristique de X, loi binomiale de paramètre p s'écrit donc :

$$\phi_X(t) = (pt+q)^n = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} t^i$$

On retrouve ainsi la loi de probabilité de X :

1. $X(\Omega) = [0; n] \cap \mathbb{N}$
2. $\forall i \in X(\Omega), \quad P(X = i) = \binom{n}{i} p^i q^{n-i}$

Propriété 1-2-2-5 : Toujours grâce à la propriété 1-2-1-6 ou plus classiquement, nous obtenons :

1. $E(X) = np$
2. $V(X) = npq$

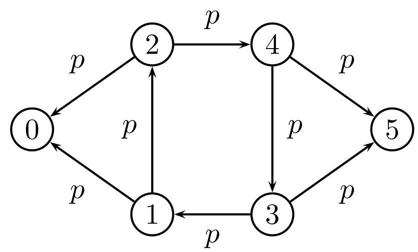
1.3 Lire une fonction génératrice sur un graphe

Nous travaillerons toujours avec des chaines de Markov absorbantes, si bien que les règles de réduction de graphes s'appliquent (*cf* l'article précédent). Traitons deux exemples :

Exemple 1-3-1 : On lance une pièce biaisée. La probabilité de faire Pile est de p. On posera $q = 1 - p$, probabilité de faire face. X est la variable aléatoire réelle égale au nombre d'essais nécessaires pour obtenir pour la première fois une suite de 3 Piles consécutifs.

1. Dessiner un graphe probabiliste modélisant l'expérience précédente. En déduire le graphe d'un parcours non marqué.
2. En réduisant le graphe précédent, donner l'expression de la fonction génératrice de X.
3. Généraliser au cas de n Piles consécutifs.

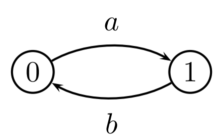
Exemple 1-3-2 : On reprend l'exemple des taulières de l'article précédent, dont nous redonnons le graphe probabiliste ci-dessous (p=1/2) :



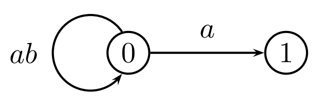
1. Donner le graphe d'un parcours non marqué et le réduire.
2. En déduire la fonction caractéristique de X, variable donnant le temps d'absorption.
3. Vérifier que le temps moyen d'absorption est égal à 2.

Solutions : Nous nous appuierons sur les méthodes de réduction de graphes étudiées au chapitre 5.

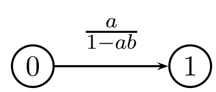
Exercice 1 : Donnons d'abord un résultat préliminaire très utile. Nous laissons sa démonstration à la sagacité du lecteur.
Le graphe :



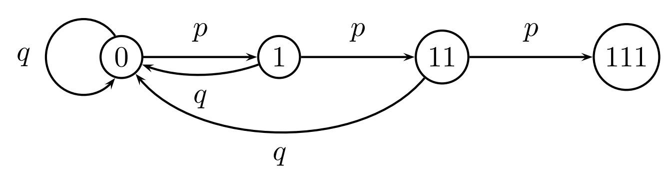
équivaut à :



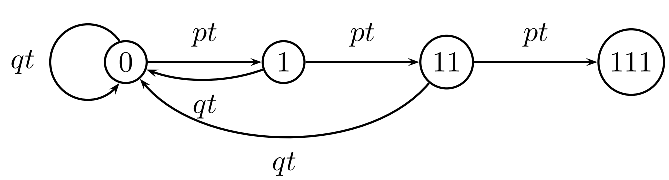
équivaut à :



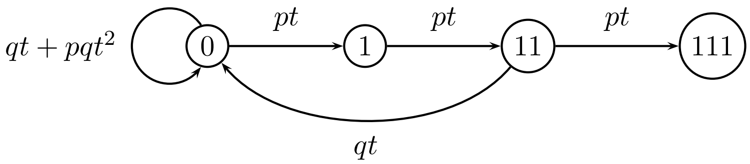
Question 1 :
Nous poserons 0 pour Face et 1 pour Pile. Sans perte de généralité, on peut supposer que l'état initial est 0. Le graphe probabiliste associé à notre expérience aléatoire prend alors la forme :



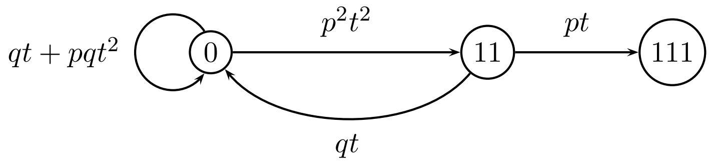
Question 2 : Il suffit de multiplier chacune des probabilités indiquées par t, la probabilité de ne pas être marqué en faisant tourner la roue de Van Dantzig.



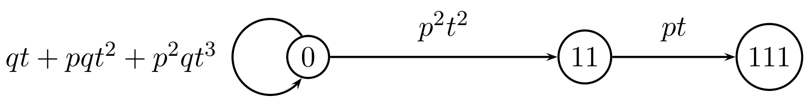
équivaut à :



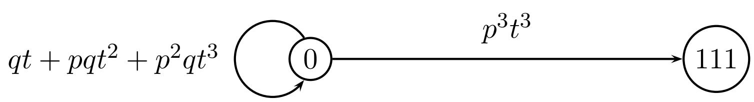
lui-même équivalent à :



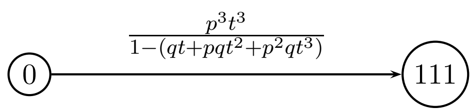
qui équivaut à :



équivaut à :



équivaut à :



De plus, $qt + pqt^2 + p^2qt^3 = \dfrac{qt(1-p^3t^3)}{1-pt}$, d'où :

$$\phi_X(t) = \frac{(pt)^3(1-pt)}{1-pt-qt(1-(pt)^3)} = \frac{(pt)^3(1-pt)}{1-t+qt(pt)^3}$$

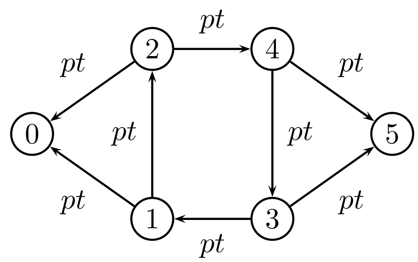
soit :

$$\phi_X(t) = \frac{(1-pt)p^3t^3}{1-t+qp^3t^4}$$

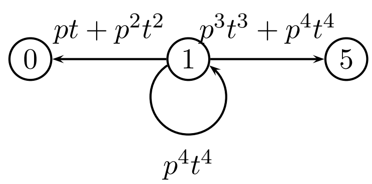
Question 3 :
On laisse le lecteur se persuader que $\phi_X(t) = \dfrac{(1-pt)p^nt^n}{1-t+qp^nt^{n+1}}$.

Exercice 2 : il y a deux états absorbants cette fois-ci.

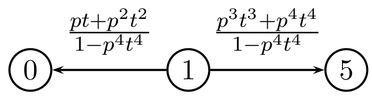
Question 1 :



équivaut à :



lui-même équivalent à :



Question 2 :
On en déduit que $\phi_X(t) = \dfrac{pt + (pt)^2 + (pt)^3 + (pt)^4}{1 - (pt)^4}$.
Mais $pt + (pt)^2 + (pt)^3 + (pt)^4 = \dfrac{pt(1 - (pt)^4)}{1 - pt}$, d'où $\phi_X(t) = \dfrac{pt}{1 - pt}$.

Question 3 :
Il s'agit de calculer $E(X) = \phi'_X(1)$.
Or $\phi'_X(t) = \dfrac{p}{(1 - pt)^2}$, d'où $E(X) = \dfrac{p}{(1 - p)^2} = 2 \quad (p = 1/2)$.

1.4 Fonctions génératrices et indépendance

Les résultats énoncés ici sont redoutables en efficacité. Par analogie avec le logarithme népérien qui transforme des produits en sommes, nous allons exprimer les fonctions génératrices de sommes de variables aléatoires discrètes indépendantes comme produit de leurs fonctions génératrices.

Théorème 1-4-1 : Soit $(X_i)_{1 \leq i \leq n}$ une suite de variables aléatoires réelles discrètes identiquement distribuées et indépendantes. Soit $S_n = \displaystyle\sum_{i=1}^{n} X_i$. Alors :

1. X_n et S_{n-1} sont indépendantes,
2. $\phi_{S_n} = \phi_{X_1} \dots \phi_{X_n}$

Démonstration : On supposera $X_i(\Omega) = \mathbb{N}$.

1. Résulte du résultat classique et extrêmement utile suivant : si les X_i $(1 \leq i \leq n)$ sont mutuellement indépendantes, alors pour tout $k \in [1; n - 1]$ et pour toutes fonctions boréliennes $f : \mathbb{R}^k \to \mathbb{R}$ et $g : \mathbb{R}^{n-k} \to \mathbb{R}$, $f(X_1, \dots, X_k)$ et $g(X_{k+1}, \dots, X_n)$ sont indépendantes.

2. Au vu du résultat précédent, il suffit de prouver que si X_1 et X_2 sont deux variables aléatoires indépendantes, alors $\phi_{X_1+X_2} = \phi_{X_1}\phi_{X_2}$.
   On rappelle le produit de deux séries absolument convergentes : Si $a := \sum_{n\geq 0} a_n$ et $b := \sum_{n\geq 0} b_n$ sont deux séries absolument convergentes, on définit la série $c := \sum_{n\geq 0} c_n$ par $c_n := \displaystyle\sum_{k=0}^{n} a_k b_{n-k}$. Alors c converge absolument et $c = ab$.
   Sur ce, pour tout $t < R_{X_1}$, on a :

$$\phi_{X_1}(t)\phi_{X_2}(t) = \sum_{i\geq 0}\sum_{k=0}^{i} P(X_1 = k)P(X_2 = i-k)t^i$$

   Par indépendance de X_1 et X_2, $P(X_1 = k)P(X_2 = i-k) = P(X_1 = k \cap X_2 = i-k)$.
   Posons $X = X_1 + X_2$.
   L'événement $(X = i)$ est la réunion disjointe $\bigcup_{k=0}^{i}(X_1 = k \ \cap \ X_2 = i-k)$.
   D'où $P(X = i) = \displaystyle\sum_{k=0}^{i} P(X_1 = k \cap X_2 = i-k) = \sum_{k=0}^{i} P(X_1 = k)P(X_2 = i-k)$.
   Ainsi, $\phi_X(t) = \phi_{X_1}(t)\phi_{X_2}(t)$.

Remarque 1-4-2 : Nous pouvons donner une démonstration alternative au point 2 du théorème précédent, basée sur la méthode de marquage de Van Dantzig. Nous avons vu à la remarque 1-1-3 l'interprétation probabiliste de la fonction génératrice d'une variable aléatoire X prenant tous les entiers naturels i comme valeurs avec les probabilités p_i $(i \geq 0)$:
on tourne une roue donnant les nombres 0, 1, 2, ... avec les probabilités respectives p_i. Si l'on tombe sur i, on tourne i fois la roue marquée de Van Dantzig. La fonction génératrice de X donnant la probabilité d'un parcours non marqué s'écrit $\phi_X(t) = \displaystyle\sum_{i\geq 0} p_i t^i$.
Nous effectuons n épreuves indépendantes de la sorte. La k-ième expérience n'est pas marquée avec la probabilité $\phi_{X_k}(t)$. Si l'on pose $X = X_1 + \cdots + X_n$, nous obtenons alors :

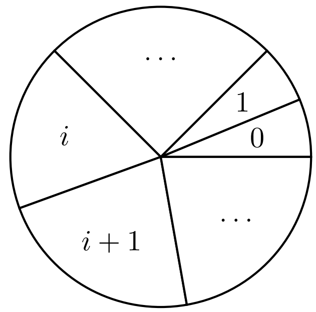
$$\phi_X(t) = \prod_{k=1}^{n} \phi_{X_k}(t)$$

La loi de Pascal se déduit ainsi de la loi géométrique de manière naturelle ! Ce que confirment leurs graphes probabilistes comme vu auparavant.

1.5 Exercices d'application :

Un peu de calcul ...

1. Soient X et Y deux variables aléatoires indépendantes suivant respectivement les lois de Poisson de paramètre λ et μ. Prouver que $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.
2. Considérons une population constituée de n bactéries de même type. On dit qu'elle est dans l'état n.
   On suppose qu'à la génération 0, il y a un seul individu. On fait alors tourner la roue suivante qui indique en combien d'individus k (avec la probabilité p_k) se divise la bactérie initiale.



Nous obtenons la génération 1.
Pour chacune des k bactéries de la génération 1 on fait tourner la roulette précédente et on obtient la génération 2. Le processus se poursuit ainsi. On l'appelle **processus de branchement**.

(a) Soit Z_n la variable aléatoire égale au nombre de bactéries présentes à la génération n. On note ϕ_n la fonction génératrice de Z_n. Justifier que $\phi_0 \equiv 1$ et que $\phi_1(t) = t$. On posera $\phi = \phi_1$.

(b) Déterminer une relation de récurrence entre Z_{n+1} et Z_n à l'aide de ϕ.

(c) En déduire une expression de ϕ_n en fonction de ϕ.

(d) On pose $\mu = E(Z_1)$ et $\sigma^2 = V(Z_1)$. Exprimer $E(Z_n)$ et $V(Z_n)$ en fonction de μ et de σ.

(e) Application : Un insecte éphémère (mourant lorsque la génération suivante nait) donne naissance à 0, 1, 2 ou 3 insectes avec les probabilités respectives 0,05 ; 0,45 ; 0,3 ; 0,2. Calculer la probabilité d'extinction de la lignée de cet insecte.

2 Aspect matriciel

Comme nous l'avons vu dans l'article précédent avec le théorème de Chapman-Komolgorov, l'étude des puissances successives de la matrice de transition nous permet de retrouver la loi de X, variable aléatoire donnant le temps d'absorption, et donc sa fonction génératrice, songea Greg. Rappelons son corollaire le plus usité :
Notons $\vec{p}_0$ l'état initial du système et $\vec{p}_n$ son état après n transitions, tous deux écrits sous la forme d'un vecteur ligne de longueur Card(S). Alors

$$\boxed{\vec{p}_n = \vec{p}_0(\mathcal{P})^n}$$

Maintenant que j'ai fait connaissance avec Monsieur Van Dantzig, l'idée est de traduire matriciellement les graphes de parcours non marqués. Allez, je mets les mains dans le cambouis pour vérifier quelques exemples ...

2.1 Loi géométrique

Soit p le paramètre de la loi géométrique. La matrice de transition de l'état 0 (échec) à l'état 1 (succès) s'écrit :

$$(\mathcal{P}) = \begin{pmatrix} q & p \\ 0 & 1 \end{pmatrix}$$

donc celle d'un parcours non marqué est (en conservant la même notation) :

$$(\mathcal{P}) = \begin{pmatrix} qt & pt \\ 0 & 1 \end{pmatrix}$$

Une récurrence immédiate nous permet de démontrer que pour tout entier naturel n non nul :

$$(\mathcal{P})^n = \begin{pmatrix} (qt)^n & \sum_{i=0}^{n-1}(qt)^i pt \\ 0 & 1 \end{pmatrix}$$

soit :

$$(\mathcal{P})^n = \begin{pmatrix} (qt)^n & \sum_{i=1}^{n} q^{i-1}pt^i \\ 0 & 1 \end{pmatrix}$$

On suppose $|qt| < 1$, de sorte que $\lim\limits_{n\to+\infty}(qt)^n = 0$. Ceci assure aussi la convergence de $\sum\limits_{i=1}^{\infty} q^{i-1}pt^i$. Ainsi :

$$(\mathcal{P})^\infty = \begin{pmatrix} 0 & \sum_{i=1}^{\infty} q^{i-1}pt^i \\ 0 & 1 \end{pmatrix}$$

On retrouve ainsi la fonction génératrice de la loi géométrique : $\phi_X(t) = \sum\limits_{i=1}^{\infty} q^{i-1}pt^i$, avec en prime la loi : pas besoin de développer en série entière !

2.2 Loi de Pascal

Soient p et n les paramètres de la loi de Pascal. La matrice de transition d'un parcours non marqué $(\mathcal{P}) \in \mathcal{M}_{n+1}(\mathbb{R})$ s'écrit :

$$(\mathcal{P}) = \begin{pmatrix} V_n & U \\ 0 & 1 \end{pmatrix}$$

avec $V_n = \begin{pmatrix} qt & pt & 0 & \cdots & 0 \\ 0 & qt & pt & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & pt \\ 0 & 0 & 0 & \cdots & qt \end{pmatrix}$ et $U = pte_n$, où e_n est le n-ième vecteur de la base canonique de $\mathbb{R}^n$.

On en déduit que pour tout entier naturel m :

$$(\mathcal{P})^m = \begin{pmatrix} V_n^m & \sum_{i=0}^{m-1} V_n^i U \\ 0 & 1 \end{pmatrix}$$

avec la convention usuelle $V_n^0 = I_n$.
Puisque e_1^T est la distribution initiale, et sous réserve que $\lim\limits_{m\to+\infty} V_n^m = 0$, l'élément $(\mathcal{P})^\infty_{1n}$ nous permet de retrouver la fonction génératrice associée à la loi de X.

Remarquons que $V_n = qtI_n + ptW_n$, où $W_n \in \mathcal{M}_n(\mathbb{R})$ s'écrit : $W_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$

Un exercice classique de premier cycle (très bon a refaire) nous assure que la matrice W_n est nilpotente d'ordre n. De plus, pour tout $1 \leq i \leq n-2$, on a :

$$W_n^i = \begin{pmatrix} 0 & \overbrace{\cdots}^{i-1} & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \quad \textbf{et} \quad W_n^{n-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

Or, I_n et W_n commutent, d'où : $V_n^m = (ptW_n + qtI_n)^m = \sum_{i=0}^{m} \binom{m}{i} p^i W_n^i q^{m-i} t^m$.
De la nilpotence de W_n, on en tire que pour tout entier $m \geq n-1$:

$$V_n^m = \sum_{i=0}^{n-1} \binom{m}{i} p^i q^{m-i} W_n^i t^m$$

Mais pour $i = 0, \dots, n-1$ **on a** $W_n^i U = pte_{n-i}$ (1), $n-i$-ème vecteur de la base canonique de $\mathbb{R}^n$. Dans l'idée de réduction de graphe faisant "apparaître" la fonction caractéristique de X, on fait tendre m vers $+\infty$.
Considérons donc formellement la somme :

$$\sum_{i \geq 0} V_n^i U = \sum_{i \geq 0} \sum_{j=0}^{i} \binom{i}{j} p^j q^{i-j} W_n^j U t^i$$

D'après (1), si $j \leq n-1$, $W_n^j U = pte_{n-j}$ et $W_n^j = 0$ si $j \geq n$. La somme précédente s'écrit donc :

$$\sum_{i \geq 0} V_n^i U = \sum_{i \geq 0} \sum_{j=0}^{n-1} \binom{i}{j} p^j q^{i-j} pe_{n-j} t^{i+1}$$

Seul le terme colinéaire à e_1 nous intéresse, ce qui revient à choisir $j = n-1$. On est donc ramené à étudier :

$$S := \sum_{i \geq 0} \binom{i}{n-1} p^{n-1} q^{i-n+1} t^{i+1} p = \sum_{i \geq 0} \binom{i}{n-1} p^n q^{i-n+1} t^{i+1}$$

Comparons cette somme avec celle obtenue en développant en série entière $\phi_X(t)$ dans le paragraphe précédent :

$$\phi_X(t) = \sum_{i \geq 0} \binom{n+i-1}{i} p^n q^i t^{n+i} = \sum_{i \geq n} \binom{i-1}{i-n} p^n q^{i-n} t^i$$

Remarquons que pour $i = 0, \dots, n-2$ on a $\binom{i}{n-1} = 0$; ainsi

$$S = \sum_{i \geq n-1} \binom{i}{n-1} p^n q^{i-n+1} t^{i+1} = \sum_{i \geq n} \binom{i-1}{n-1} p^n q^{i-n} t^i$$

Or

$$\binom{i-1}{n-1} = \frac{(i-1)!}{(n-1)!(i-n)!} = \binom{i-1}{i-n}$$

D'où l'égalité voulue !

Remarque : Il n'y avait ici qu'un seul état absorbant. L'élément $(\mathcal{P})^{\infty}_{1,n}$ nous donnait alors l'expression $\phi_X(t)$ écrite sous la forme d'une série. On a donc accès directement à la loi de X au prix de quelques calculs. Quid s'il y a plusieurs états absorbants ? Quelle forme peut alors prendre la matrice de transition $(\mathcal{P})$ judicieusement écrite ? Nous en reparlerons dans la sous-section **Généralisation partielle**.

Donnons pour finir un script version fréquentiste simulant la loi de Pascal :

```
def reussite(p):
    alea=random()
    if alea<=p:
        succes=1
    else:
        succes=0
    return succes

def Pascal(n):
    S,nb_essais=0,0
    while S<n:
        S=S+reussite(p)
        nb_essais=nb_essais+1
    return nb_essais

def frequence(N):
    S=0
    for i in range(N):
        S=S+Pascal(n)
    return S/N

#Programme principal
from random import *
p=float(input("Probabilite de succes par experience ? "))
n=int(input("Nombre de succes souhaites ? "))
N=int(input("Nombre d'experiences ? "))
print("Nombre moyen d'essais avant d'obtenir",n,"succes",frequence(N))
```

2.3 Loi binomiale

Le graphe probabiliste est presque le même que celui associé à la Loi de Pascal, mais on se limite au calcul de $(\mathcal{P})^n$:

La matrice de transition d'un parcours non marqué $(\mathcal{P}) \in \mathcal{M}_{n+1}(\mathbb{R})$ s'écrit :

$$(\mathcal{P}) = \begin{pmatrix} V_n & U \\ 0 & 1 \end{pmatrix}$$

avec $V_n = \begin{pmatrix} q & pt & 0 & \cdots & 0 \\ 0 & q & pt & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & pt \\ 0 & 0 & 0 & \cdots & q \end{pmatrix}$ et $U = pte_n$, où e_n est le n-ième vecteur de la base canonique de $\mathbb{R}^n$.

Comme nous n'avons pas nécessairement absorption au bout de n transitions (seul cas favorable), nous observons cette fois-ci toute la première ligne de $(\mathcal{P})^n$. En omettant le terme t^k, nous avons $P(X = k) = (\mathcal{P})^n_{1,k}$, où X compte le nombre de succès au cours de n transitions.

Effectuons quelques calculs :

La probabilité d'avoir n succès est égale au coefficient de $\sum_{i=0}^{n-1} V_n^i U$ colinéaire à e_1, soit :

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\binom{i}{n-1}p^n q^{i-n+1} = p^n$$

La probabilité d'avoir $k < n$ succès est égale (exercice) à $e_1^T V_n^n e_{k+1}$.
Or :

$$V_n^n = \sum_{j=0}^{n-1}\binom{n}{j}p^j q^{n-j} W_n^j$$

D'où pour tout entier $0 \leq k \leq n-2$:

$$e_1^T V_n^n e_{k+1} = \sum_{j=0}^{n-1}\binom{n}{j}p^j q^{n-j} e_1^T W_n^j e_{k+1} = \binom{n}{k+1-1}p^{k+1-1}q^{n-(k+1-1)}$$

Soit pour $0 \leq k \leq n-1$:

$$P(X = k) = \binom{n}{k}p^k q^{n-k}$$

égalité restant vraie pour $k = n$, d'où : $\phi_X(t) = \sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k} t^k$.

2.4 Généralisation partielle

Dans ce paragraphe, nous ne considérons pas des parcours non marqués, mais les matrices de transition telles quelles. Il semble que toute chaine de Markov absorbante puisse être mise sous la forme matricielle :

$$(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & I_r \end{pmatrix}$$

où V est une matrice carrée $(n-r)\times(n-r)$ telle que $\lim_{n\to+\infty} V^n = 0$, U une matrice $(n-r)\times r$ telle que $\lim_{n\to+\infty}\sum_{k=0}^{n-1} V^k U$ existe et soit finie (pour une norme matricielle donnée quelconque).
La matrice identité I_r **regroupe** les r états absorbants de la chaine de Markov.
De plus, la matrice $I_{n-r} - V$ n'est pas singulière *i.e* $I_{n-r} - V$ est inversible. Ceci résulte du résultat classique suivant :

Théorème 2-4-1 : Soit $A \in \mathcal{M}_n(\mathbb{C})$. Les propriétés suivantes sont équivalentes :

1. $\rho(A) < 1$ (où $\rho(A)$ désigne la plus grande en module des valeurs propres de A),
2. $\lim_{k\to+\infty} A^k = 0$,
3. $I_n - A$ est inversible et $\sum_{k\geq 0} A^k = (I_n - A)^{-1}$.

Souvenons-nous que la matrice de transition $(\mathcal{P})$ d'un graphe probabiliste est une matrice stochastique (la somme des coefficients de ses lignes est égale à 1). Si c'est celle d'une chaine de Markov absorbante, on peut regrouper les états absorbants comme annoncé juste avant. Ainsi, $(\mathcal{P})$ s'écrit :

$$(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & I_r \end{pmatrix}$$

Choisissons une norme $\|.\|$ sur $\mathcal{M}_n(\mathbb{R})$ telle que $\|V\| < 1$ (comme V est stochastique, on peut prendre par exemple $\|V\| = \max\limits_{1 \leq i,j \leq n-r} v_{i,j}$).

En appliquant le théorème 2-4-1, la matrice $I_{n-r} - V$ n'est pas singulière et son inverse est égale à $\sum\limits_{n \geq 0} V^n$. D'où $\lim\limits_{n \to +\infty} V^n = 0$.

De plus, $\forall n \in \mathbb{N}^*$, $\sum\limits_{k=0}^{n-1} V^k U = \left(\sum\limits_{k=0}^{n-1} V^k\right) U$, d'où l'existence (et l'unicité) de $\sum\limits_{k \geq 0} V^k U = (I_{n-r} - V)^{-1} U$.

Ceci justifie a posteriori tous les calculs effectués précédemment.

Mais quid de la fonction génératrice ? C'est bien là notre sujet d'étude, se reprit Greg en piquant un far ... dans son frigo.

Première remarque : mon état initial, en numérotant bien les sommets de mon graphe afin d'avoir la décomposition par blocs annoncée peut s'écrire sous la forme $\vec{p_0} = (p_1, p_2, \ldots, p_{n-r}, 0, \ldots, 0)$ comme je pars d'un état intérieur. Ce qui me permet d'identifier ce vecteur-ligne de $\mathbb{R}^n$ à un vecteur-ligne de $\mathbb{R}^{n-r}$: $(p_1, p_2, \ldots, p_{n-r})$ que je noterai toujours $\vec{p_0}$;

Deuxième remarque, la matrice de transition $(\mathcal{P}_t)$ d'un parcours non marqué s'écrit donc :

$$(\mathcal{P}_t) = \begin{pmatrix} tV & tU \\ 0 & I_r \end{pmatrix}$$

si bien que :

$$(\mathcal{P}_t)^\infty = \begin{pmatrix} 0 & t(I_{n-r} - tV)^{-1} U \\ 0 & I_r \end{pmatrix}$$

Chapman & Komolgorov me permettent alors de conclure que :

$$\boxed{\phi_X(t) = \sum_{j=1}^{r} \left(\sum_{i=1}^{n-r} p_i [t(I_{n-r} - tV)^{-1} U]_{i,j} \right)}$$

Si i est l'état de départ, la loi initiale est simplement $(0, \ldots, 1, \ldots, 0)$ où le 1 est situé en i-ème position et l'on obtient l'expression plus sympathique : $\phi_X(t) = \sum\limits_{j=1}^{r} [t(I_{n-r} - tV)^{-1} U]_{i,j}$

2.5 Exercices d'application

1. Est-il plus probable en lançant autant de fois que nécessaire une pièce équilibrée d'obtenir pour la première fois (Face,Face,Face) ou (Pile,Pile,Face) ?
   (a) Dessiner un graphe probabiliste modélisant la situation,
   (b) Écrire la matrice de transition associée sous une forme adaptée et calculer si possible ses puissances successives ou utiliser ce qui précède,
   (c) Conclure.

2. Deux compartiments A et B contiennent à eux deux, N particules numérotées de 1 à N. À chaque instant, on choisit un nombre de 1 à N, avec la probabilité 1/N. L'ensemble des états est $S = \{0, 1, \ldots, N\}$. Le processus est dit être dans l'état j si le compartiment A contient j particules. Si le processus est dans l'état 0 (resp. dans l'état N), alors la probabilité de passer dans l'état 1 (resp. dans l'état $N-1$) est égale à 1.
(a) Dessiner un graphe probabiliste modélisant la situation,
(b) Écrire la matrice de transition associée,
(c) On suppose qu'à l'instant initial, le compartiment A contient les N particules. Soit X_n la variable aléatoire égale au nombre de particules présentes dans le compartiment A à chaque instant. Décrire le processus (X_n).

3 Une histoire de rencontre (suite)

Rappelons l'histoire de nos scarabées proposée à la sagacité du lecteur à titre d'exercice-défi dans l'article précédent.

On étudie la première rencontre entre deux scarabées situés symétriquement sur un polygone à 2^p $(p \geq 2)$ côtés. Le jeu se déroule ainsi :

- À l'instant initial, deux scarabées sont situés symétriquement par rapport à O, centre d'un polygone régulier à 2^p côtés.
- Chaque seconde on lance une pièce pour chacun des scarabées. Si la pièce tombe sur pile le scarabée concerné tourne dans le sens des aiguilles d'une montre, sinon il tourne dans le sens inverse des aiguilles d'une montre.

3.1 Approche fréquentiste

Question 1 : Écrire un programme en Python (éventuellement sous la forme d'une fonction) qui modélise la marche des 2 scarabées sur un carré jusqu'à temps qu'ils se rencontrent pour la première fois. Vous utiliserez la notion de congruence via la commande modulo % pour construire la condition de (non) rencontre des scarabées : Tant que ceux-ci ne sont pas sur le même sommet leur promenade continue.

Question 2 : Modifier le programme précédent pour tester le script précédent sur N marches (N choisi par l'utilisateur) et évaluer le temps moyen de rencontre des scarabées. Vous le testerez avec N =10000.

Nous répondons maintenant à la question posée initialement. Considérons donc un polygone régulier à 2^p $(p \geq 2)$ côtés dont les sommets sont numérotés de 0 à $2^p - 1$.

Question 3 : Le scarabée 1 est situé au sommet numéroté 0. Le scarabée 2 est le symétrique du scarabée 1 par rapport à O, centre du polygone. Quel est le numéro du sommet sur lequel il est situé ?

Question 4 : Écrire une fonction Python marche(p) qui prend comme paramètre p, qui modélise la marche aléatoire des 2 scarabées sur le polygone régulier à 2^p côtés et qui renvoie comme valeur le temps de première rencontre.

Question 5 : En vous servant de la fonction précédemment créée, écrire un programme qui demande à l'utilisateur de saisir N, nombre de marches à effectuer par les scarabées, et qui renvoie le temps moyen de rencontre sur ces N essais. Vous testerez votre programme plusieurs fois avec N=100 000 pour compléter le tableau suivant :

Valeur de p	Nombre de sommets	Temps moyen T_p de rencontre
2		
3		
4		
5		
6		

Question 6 : Subodorer une relation fonctionnelle entre p et T, c'est-à-dire trouver l'expression d'une fonction f telle que $T_p = f(p)$.

Solution : Détaillons les solutions des questions précédentes de ce problème qui a été donné en projet informatique à des élèves d'une classe de première S en 2012.

Question 1 : Modélisation d'une expérience

```
def rencontre():
    tempsRencontre=0
    position1,position2=0,2
    while (position2-position1)%4!=0:
        alea1,alea2=randint(1,2),randint(1,2)
        if alea1==1:
            position1=position1+1
        else:
            position1=position1-1
        if alea2==1:
            position2=position2+1
        else:
            position2=position2-1
        tempsRencontre=tempsRencontre+1
    return tempsRencontre
```

Remarque : La condition (position2−position1)%4!=0 signifie que les valeurs de position1 et de position2 de chacun des scarabées, qui peuvent devenir négatives au cours de leur évolution dans l'algorithme, correspondent à des sommets du carré différents. On peut comprendre cette condition en visualisant le mouvement des scarabées avec un jeu de Pile ou Face. Comme en trigonométrie usuelle où l'on travaille modulo 2π, à un même sommet du carré on associe une infinité de valeurs qui diffèrent toutes d'un multiple de 4.

Question 2 : Temps moyen de rencontre des deux scarabées

```
from random import randint
N=int(input("Combien d'experiences a effectuer ? "))
tempsTotal=0
for i in range(N):
    tempsTotal=tempsTotal+rencontre()
print("Temps moyen de premiere rencontre :",tempsTotal/N)
```

Question 3 : Le second scarabée est situé initialement au sommet $2^p/2 = 2^{p-1}$.

Question 4 : Le script est presque identique à celui de la question 2 mais on travaille cette fois-ci modulo 2^p.

```
def marche(p):
    tempsRencontre=0
    position1,position2=0,2**(p-1)
    while (position2-position1)%2**p!=0:
        alea1,alea2=randint(1,2),randint(1,2)
        if alea1==1:
            position1=position1+1
        else:
            position1=position1-1
        if alea2==1:
            position2=position2+1
        else:
            position2=position2-1
        tempsRencontre=tempsRencontre+1
    return tempsRencontre
```

Question 5 : Calcul du temps moyen de rencontre

```
from random import randint
p=int(input("p = ? "))
N=int(input("Combien d'experiences a effectuer ? "))
tempsTotal=0
for i in range(N):
    tempsTotal=tempsTotal+marche(p)
print("Temps moyen de premiere rencontre :",tempsTotal/N)
```

Question 6 : Avec N = 10 000, on trouve des valeurs du type (car fluctuation d'échantillonnage !) :

Valeur de p	Nombre de sommets	Temps moyen T_p de rencontre
2	4	1.99
3	8	8.005
4	16	32.016
5	32	127.48
6	64	514.475

On peut remarquer que $1 = 2 \times 2 - 3$ et que $T_2 \approx 2^1$; que $3 = 2 \times 3 - 3$ et que $T_3 \approx 2^3$; que $5 = 2 \times 4 - 3$ et que $T_4 \approx 2^5$, etc. On peut donc supposer que $T_p = 2^{2p-3}$.

Nous allons maintenant prouver rigoureusement ce que notre intuition TICE nous a permis de subodorer, mais en changeant de point de vue.

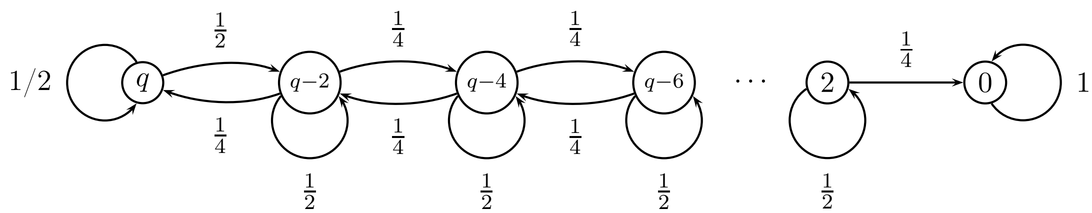
3.2 Première modélisation

Définissons les états pour commencer : on s'intéresse au nombre de segments du chemin le plus court joignant les deux scarabées le long du polygone à un instant donné. Initialement ce nombre est 2^{p-1}, à chaque seconde il peut :

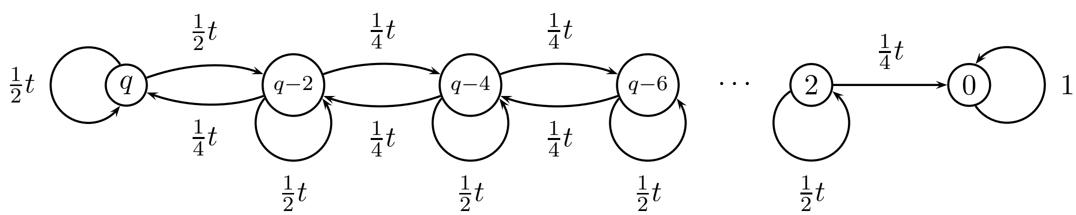
- rester constant si les deux pièces donnent le même résultat (probabilité $\frac{1}{2}$) ;
- diminuer de 2 si le résultat des lancers font que les scarabées se rapprochent (probabilité $\frac{1}{4}$)

– augmenter de 2 si le résultat des lancers font que les scarabées s'éloignent (probabilité $\frac{1}{4}$)

sauf lorsque les scarabées sont diamétralement opposés ou au même point. Dans le premier cas le nombre de segments séparant les scarabées reste constant ou diminue de 2 de façon équiprobable. L'état 0 est par définition absorbant. Il y a ainsi (petit exercice sur les suites arithmétiques) $2^{p-2}+1$ états : 0, 2, 4, …, $q := 2^{p-1}$.
Le but du problème est toujours d'étudier la variable aléatoire T_p donnant le temps avant la première rencontre des deux scarabées.
La remarque précédente conduit au graphe suivant :



Le graphe d'un parcours non marqué par la roulette de Van Dantzig est donc :



Le graphe semble difficile à réduire ! Il est tentant d'utiliser l'approche matricielle pour en déduire $\lim\limits_{n\to+\infty}(\mathcal{P})^n$, où $(\mathcal{P})$ est la matrice de transition d'un parcours non marqué. A titre d'exemple, faisons-le pour $p=3$ (nous avons $2^3=8$ côtés, soit un octogone).

$$(\mathcal{P}) = \begin{pmatrix} t/2 & t/2 & 0 \\ t/4 & t/2 & t/4 \\ 0 & 0 & 1 \end{pmatrix}$$

Posons $V = \begin{pmatrix} t/2 & t/2 \\ t/4 & t/2 \end{pmatrix}$ et $U = \begin{pmatrix} 0 \\ t/4 \end{pmatrix}$, de sorte que $(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & 1 \end{pmatrix}$. XCas donne :

$$V^n = \begin{pmatrix} \frac{\sqrt{2}\cdot t\left(\frac{t}{-2\sqrt{2}+4}\right)^n}{t\cdot 2\sqrt{2}} + \frac{\sqrt{2}\cdot t\left(\frac{t}{2\sqrt{2}+4}\right)^n}{t\cdot 2\sqrt{2}} & \frac{2\cdot t\left(\frac{t}{-2\sqrt{2}+4}\right)^n}{t\cdot 2\sqrt{2}} - \frac{2\cdot t\left(\frac{t}{2\sqrt{2}+4}\right)^n}{t\cdot 2\sqrt{2}} \\ \frac{\left(\frac{t}{-2\sqrt{2}+4}\right)^n}{2\sqrt{2}} - \frac{\left(\frac{t}{2\sqrt{2}+4}\right)^n}{2\sqrt{2}} & \frac{\sqrt{2}\left(\frac{t}{-2\sqrt{2}+4}\right)^n}{2\sqrt{2}} + \frac{\sqrt{2}\left(\frac{t}{2\sqrt{2}+4}\right)^n}{2\sqrt{2}} \end{pmatrix}$$

si bien que :

$$V^nU = \left(\frac{t\left(\frac{2\cdot\frac{1}{2}\cdot t\left(\frac{t}{-2\sqrt{2}+4}\right)^n}{\sqrt{2}\cdot t} - \frac{2\cdot\frac{1}{2}\cdot t\left(\frac{t}{2\sqrt{2}+4}\right)^n}{\sqrt{2}\cdot t}\right)}{4}, \frac{t\left(\frac{\sqrt{2}\cdot\frac{1}{2}\left(\frac{t}{-2\sqrt{2}+4}\right)^n}{\sqrt{2}} + \frac{\sqrt{2}\cdot\frac{1}{2}\left(\frac{t}{2\sqrt{2}+4}\right)^n}{\sqrt{2}}\right)}{4} \right)^T$$

Intéressons-nous uniquement au premier terme de V^nU. En effet, $(\mathcal{P})^n = \begin{pmatrix} V^n & \sum_{i=0}^{n-1} V^iU \\ 0 & 1 \end{pmatrix}$.

Posons $\alpha(t) = \dfrac{t}{-2\sqrt{2}+4}$ et $\beta(t) = \dfrac{t}{2\sqrt{2}+4}$, de sorte que :

$$\sum_{i=0}^{n-1} V^iU = \frac{t}{4\sqrt{2}}\left[\frac{\alpha(t)^n-1}{\alpha(t)-1} - \frac{\beta(t)^n-1}{\beta(t)-1}\right]$$

On a $4+2\sqrt{2} > 4-2\sqrt{2} > 1$, donc pour $|t| < 4-2\sqrt{2}$, $\lim_{n\to+\infty} \alpha(t)^n = \lim_{n\to+\infty} \beta(t)^n = 0$.
Ainsi,

$$\phi_X(t) = \sum_{i=0}^{\infty} V^i U = \frac{t}{4\sqrt{2}} \left[\frac{-1}{\alpha(t)-1} + \frac{1}{\beta(t)-1} \right] = \sum_{n\geq 0} (\alpha(t)^n - \beta(t)^n) \frac{t}{4\sqrt{2}}$$

Soit :

$$\phi_X(t) = \sum_{n\geq 0} \left[\frac{1}{(4-2\sqrt{2})^n} - \frac{1}{(4+2\sqrt{2})^n} \right] \frac{t^{n+1}}{4\sqrt{2}}$$

Le calcul de $\phi'_X(1)$ et de $\phi''_X(1)$ à l'aide d'un logiciel de calcul formel nous permet de conclure que $E(X) = 8$ et $V(X) = 40$.
Il semble délicat de généraliser, la difficulté venant du calcul de V^n ; dans le cas général, V est en effet tridiagonale *presque* symétrique :

$$V = \begin{pmatrix} a & \mathbf{a} & 0 & 0 & \dots & \dots & 0 \\ b & a & b & 0 & & & \vdots \\ 0 & b & a & b & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & a & b \\ 0 & \dots & \dots & \dots & 0 & b & a \end{pmatrix} \text{ avec } b = a/2.$$

C'est ce **presque** qui nous ennuie. Il y a cependant une petite astuce basée sur la nature particulière du graphe pour s'en sortir. Nous y reviendrons.
Pour le moment, reformulons-donc le problème (méthode due à Jean-Michel Dardié) :

3.3 Seconde modélisation

Question 1 : La position du deuxième scarabée est repérée dans un repère dont le premier est l'origine et l'unité vaut 2 côtés du polygone, vérifier que la situation peut être décrite à l'aide d'une marche aléatoire sur $\mathbb{Z}$ dont vous préciserez les probabilités de transitions. Formuler dans ce contexte la variable aléatoire T_k correspondant à un polygone à $4k$ côtés.

Question 2 : On généralise la question en supposant que la position à l'instant 0 n'est pas nécessairement k mais un entier i compris entre 0 et $2k$, on note toujours T_k la variable aléatoire donnant le temps de première arrivée en 0 ou en $2k$. Pour $n \in \mathbb{N}$ et $r \in \mathbb{Z}$, exprimer la probabilité $P_{X_0=i}(X_n = r)$ en conditionnant par rapport à la variable aléatoire X_1.
Justifier que $P_{X_0=i}(X_n = r) = P_{X_1=i}(X_{n+1} = r)$, en déduire que pour tous entiers naturels n et i, $0 < i < 2k$,

$$P_{X_1=i}(T_k = n) = P_{X_0=i}(T_k = n-1)$$

en déduire une relation sur les $P_{X_0=i}(T_k = n)$.

Question 3 : On fixe k pour simplifier les notations et on note ϕ_i la fonction définie sur $[0,1]$ pour tout entier i compris entre 0 et $2k$ par :

$$\phi_i(x) = \sum_{n=0}^{\infty} P_{X_0=i}(T_k = n)\, x^n$$

en particulier $\Phi_k = \phi_k$ est la fonction génératrice des probabilités de la variable aléatoire T_k du temps de première rencontre entre les deux scarabées sur un polygone à $4k$ côtés. En utilisant la question 2 montrer que pour $x \neq 0$ les $\phi_i(x)$ sont solution d'une équation de récurrence linéaire d'ordre 2, préciser ϕ_0 et ϕ_{2k} puis déterminer Φ_k.

Question 4 : Déterminer l'espérance mathématique et la variance de T_k.

Solution : C'est parti !

Question 1 : Dans ce repère initialement le (deuxième) scarabée est à la position k, si on note X_n sa position à l'instant n, alors X_{n+1} prend l'une des valeurs X_n+1, X_n ou bien X_n-1 et pour tout $i \in \mathbb{Z}$

$$\begin{cases} P_{X_n=i}(X_{n+1}=i+1) = \frac{1}{4} \\ P_{X_n=i}(X_{n+1}=i) = \frac{1}{2} \\ P_{X_n=i}(X_{n+1}=i-1) = \frac{1}{4} \end{cases}$$

La variable T_k dans ce contexte est le temps de première arrivée en 0 ou en $2k$.

Question 2 : D'après la question 1, on a :
$P_{X_0=i}(X_n=r) = P_{X_1=i+1}(X_n=r) \cdot P_{X_0=i}(X_1=i+1) +$
$P_{X_1=i}(X_n=r) \cdot P_{X_0=i}(X_1=i) + P_{X_1=i-1}(X_n=r) \cdot P_{X_0=i}(X_1=i-1)$
$= \frac{1}{4}P_{X_1=i+1}(X_n=r) + \frac{1}{2}P_{X_1=i}(X_n=r) + \frac{1}{4}P_{X_1=i-1}(X_n=r)$

D'autre part si on connaît la position à un instant donné, les positions antérieures n'interviennent plus dans les probabilités des positions postérieures, la différence entre les deux écritures se réduit donc à un décalage d'indice. La deuxième relation vient de ce que l'événement : la première arrivée en 0 ou $2k$ est au temps n sachant que $X_1 = i$ signifie que sous l'hypothèse $X_1 = 1$ on a X_n égal à 0 ou $2k$ et pour $j < n$, $0 < X_j < 2k$. Il suffit ensuite d'utiliser la relation précédente.
Finalement on obtient :

$$\begin{aligned} P_{X_0=i}(T_k=n) &= \frac{1}{4}P_{X_1=i+1}(T_k=n) + \frac{1}{2}P_{X_1=i}(T_k=n) + \frac{1}{4}P_{X_1=i-1}(T_k=n) \\ &= \frac{1}{4}P_{X_0=i+1}(T_k=n-1) + \frac{1}{2}P_{X_0=i}(T_k=n-1) + \frac{1}{4}P_{X_0=i-1}(T_k=n-1) \end{aligned}$$

Question 3 : D'après la question 2, on a pour $n > 0$

$$P_{X_0=i}(T_k=n)\,x^n = \frac{1}{4}P_{X_0=i+1}(T_k=n-1)\,x^n + \frac{1}{2}P_{X_0=i}(T_k=n-1)\,x^n + \frac{1}{4}P_{X_0=i-1}(T_k=n-1)\,x^n$$

l'événement $T_k = 0$ étant certain si i vaut 0 ou $2k$ et impossible si non. Comme toutes ces séries sont absolument convergentes, il vient :

$$\phi_i(x) = \frac{x}{4}\phi_{i+1}(x) + \frac{x}{2}\phi_i(x) + \frac{x}{4}\phi_{i-1}(x)$$

on a donc :

$$\begin{cases} \frac{x}{4}\phi_{i+1}(x) - \left(1 - \frac{x}{2}\right)\phi_i(x) + \frac{x}{4}\phi_{i-1}(x) \text{ pour } 0 < i < 2k \\ \phi_0(x) = \phi_{2k}(x) = 1 \end{cases}$$

et donc :

$$\Phi_k(x) = \frac{2}{R^k(x) + R^{-k}(x)} \text{ avec } R(x) = \frac{1 - \frac{x}{2} - \sqrt{1-x}}{\frac{x}{2}}$$

On a en particulier $\Phi_1(x) = \dfrac{\frac{x}{2}}{1-\frac{x}{2}}$ qui confirme que la loi du temps de rencontre sur un carré est géométrique de probabilité $\frac{1}{2}$.

Question 4 : Calcul de la dérivée Φ'_k,

$$\Phi'_k = -2k\frac{R'}{R}\cdot\frac{R^k-R^{-k}}{\left(R^k+R^{-k}\right)^2} = -2k\frac{R'\left(R-R^{-1}\right)}{R}\cdot\frac{S_k}{\left(R^k+R^{-k}\right)^2} = k\frac{\Phi'_1}{{\Phi_1}^2}\cdot S_k{\Phi_k}^2$$

où si $k=2p+1$, $S_k = 1+2\sum_{i=1}^{p}\frac{1}{\Phi_{2i}}$ et si $k=2p$, $S_k = 2\sum_{i=0}^{p-1}\frac{1}{\Phi_{2i+1}}$. Comme $\Phi_i(1)=1$ pour tout i , on a dans tous les cas $S_k(1)=k$. De plus $\Phi'_1(1)=E(T_1)=2$, il vient

$$E(T_k) = \Phi'_k(1) = 2k^2$$

Calcul de la dérivée seconde Φ''_k :

$$\Phi''_k = k\frac{\Phi''_1\Phi_1-{\Phi'_1}^2}{{\Phi_1}^3}\cdot S_k{\Phi_k}^2 + k\frac{\Phi'_1}{{\Phi_1}^2}\cdot\left(S'_k{\Phi_k}^2+2S_k\Phi'_k\Phi_k\right)$$

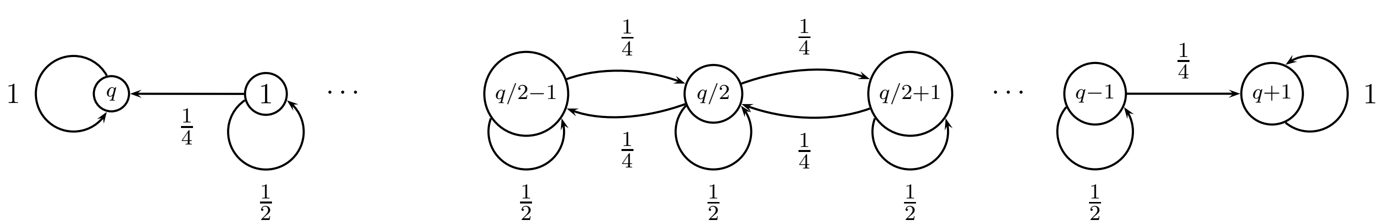
la seule fonction nouvelle est S'_k. En observant que $S'_k(1)$ est à un facteur près la somme des carrés des entiers pairs ou impairs suivant le cas, on obtient $S'_k(1) = -\frac{2}{3}k\left(k^2-1\right)$, d'où :

$$\Phi''_k(1) = 4k^4 + \frac{8}{3}k^2\left(k^2-1\right)$$
$$V(T_k) = \Phi''_k(1) - \Phi'_k(1) + \Phi'_k(1)^2 = \frac{8}{3}k^4 - \frac{2}{3}k^2$$

Dans le cas des polygones à 2^p cotés : $E(T_p) = 2^{2p-3}$ et $V(T_p) = \frac{1}{3}\left(2^{4p-5}-2^{2p-3}\right)$.

3.4 Première modélisation (suite)

Greg avançant tel Hercule Poirot dans son enquête eut un éclair de génie :
L'état initial $q = 2^{p-1}$ du graphe probabiliste est particulier. C'est lui qui casse la symétrie observée tout au long du graphe ! Matriciellement, cela se traduit par le *tridiagonal presque symétrique*. L'idée est donc, comme nous l'avons déjà vu en exercice de prolongement au concours C 2012, de déplier le graphe pour obtenir un graphe équivalent concernant la variable aléatoire X : temps d'absorption, mais avec deux états absorbants. Re-numérotons les sommets efficacement en regroupant les états absorbants : $q,\ 1,\ 2,\ \dots,\ ,q/2,\ \dots,\ q-1,\ q+1$



"Bizarre … Bizarre ! Il me semble que nous retombons sur la modélisation précédente. Continuons notre enquête, mais matriciellement", marmonna Greg en mâchouillant un caramel au beurre salé.
Ainsi, la matrice de transition d'un parcours non marqué s'écrit $(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & I_2 \end{pmatrix}$ avec :

$$V = \begin{pmatrix} a & b & 0 & 0 & \dots & \dots & 0 \\ b & a & b & 0 & & & \vdots \\ 0 & b & a & b & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & a & b \\ 0 & \dots & \dots & \dots & 0 & b & a \end{pmatrix} \text{ où } a = t/2,\ b = t/4 \text{ et } U = \begin{pmatrix} b & 0 \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & b \end{pmatrix}$$

Purée de nouilles et de brocolis : V est symétrique réelle, donc diagonalisable dans une base orthonormée de vecteurs propres ! Dans le cas présent, on peut même préciser tout ce petit monde.

Problème 3-4-1 : L'algèbre linéaire en secours ...

1. On désigne par $(D_n)_{n\geq 0}$ la suite définie par : $D_0 = 1,\ \ D_1 = a$ et pour tout $n \geq 2,\ D_n = \det(V)$. Exprimer D_n à l'aide d'une relation de récurrence liant $D_n,\ D_{n-1}$ et D_{n-2}. En déduire D_n pour tout $n \geq 0$.
2. Prouver que le système $(E_1)\ :\ VX = \lambda X \quad (X \neq 0)$, où $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ se ramène à la résolution d'une suite récurrente linéaire d'ordre 2 dont l'équation caractéristique peut s'écrire :
$$(E_2)\ :\ r^2 + \left(\frac{a-\lambda}{b}\right) r + 1 = 0$$
3. Prouver que (E_1) n'a pas de solutions si $\lambda \notin]0; 2a[$.
4. On suppose donc $\lambda \in]0; 2a[$. Prouver que pour tout entier naturel $1 \leq j \leq n,\ x_j$ est de la forme $2\mathrm{i}A\sin(j\theta_m)$, avec $\theta_m = \dfrac{m\pi}{n+1},\ m = 1, 2, \dots, n$ ($m = 0$ exclus).
5. En déduire que les valeurs propres de V sont au nombre de n (ainsi on retrouve le fait que V est diagonalisable) et s'écrivent : $\lambda_m = a\,(1 + \cos(\theta_m))$.
6. Notons $\mathcal{B}$ la base canonique de $\mathbb{R}^n$ et $\mathcal{B}'$ la base $\{X_1; \dots X_n\}$, où $X_j = \begin{pmatrix} \sin(\theta_j) \\ \sin(2\theta_j) \\ \vdots \\ \sin(n\theta_j) \end{pmatrix}$.
En remarquant que $\theta_j = j\theta_1 = j\frac{\pi}{n+1}$, donner le terme générique p_{ij} de P.
7. Notons $D = \mathrm{diag}\,(a(1 + \cos(\theta_i)))$. Exprimer V^m en fonction de $m,\ P,\ D$, et de P^{-1}.
8. Démontrer que $\|X_k\|^2 = \frac{n+1}{2}$ indépendamment de k.
9. Calculer $P^T P$ et en déduire l'expression de P^{-1} en fonction de P.
10. Démontrer que $\lim\limits_{n\to+\infty} V^m = 0$.
11. Déterminer la i-ème ligne de la matrice $V^m U$ et calculer la somme $S_i^{(m)}(t)$ de ses (deux) composantes.
12. Justifier que l'état de départ peut s'identifier à la matrice ligne $e_{2^{p-2}}^T$ de $\mathbb{R}^{2^{p-1}-1}$ et que $\phi_X(t) = \sum\limits_{m\geq 0} S_{2^{p-2}}^{(m)}(t)$.

13. En déduire que $\phi_X(t) = \sum_{m\geq 0} \alpha_m t^{m+1}$, où :

$$\alpha_m = \frac{1}{2^{p-1}} \sum_{j=0}^{2^{p-2}-1} \frac{(1+\cos((2j+1)\theta_1))^m}{2^m}(-1)^j \sin((2j+1)\theta_1)$$

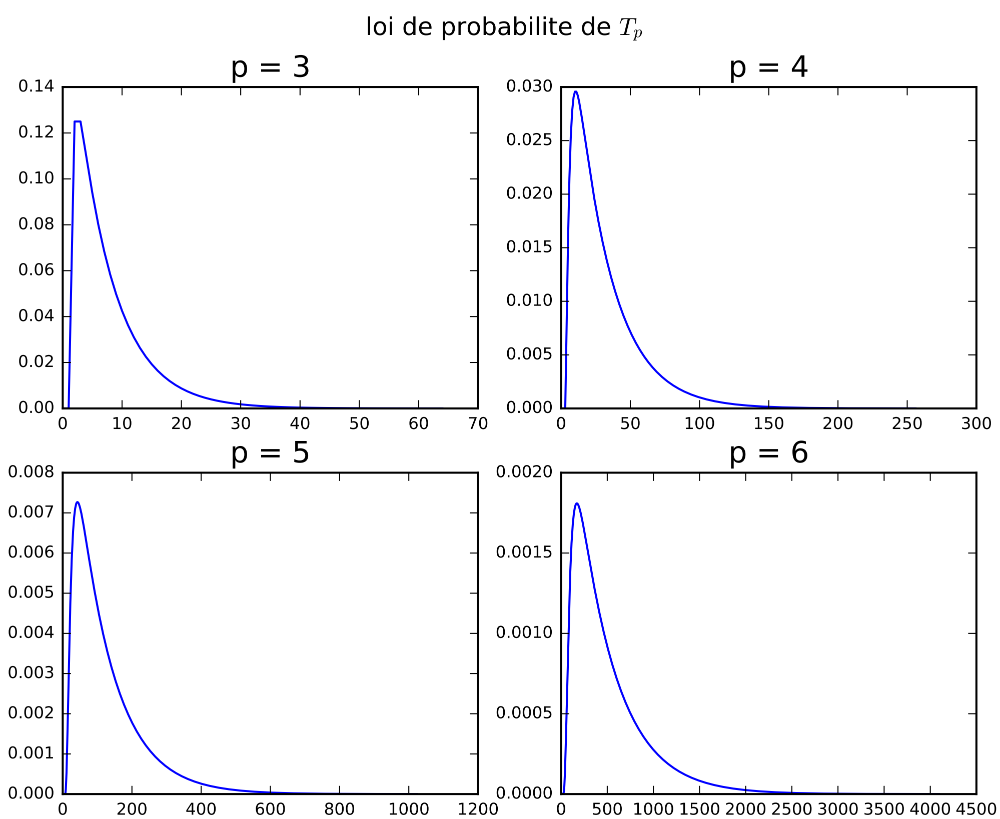
.

14. Application au cas $p=3$: retrouver le résultat du paragraphe précédent :

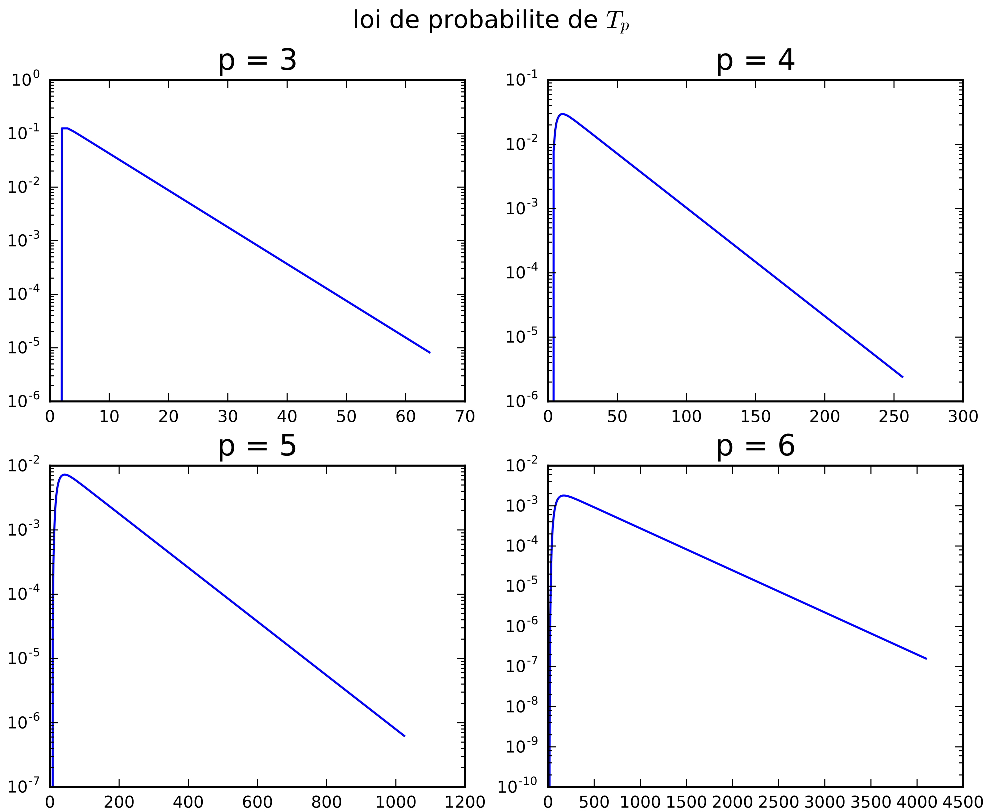
$$\phi_X(t) = \sum_{n\geq 0} \left[\frac{1}{(4-2\sqrt{2})^n} - \frac{1}{(4+2\sqrt{2})^n}\right] \frac{t^{n+1}}{4\sqrt{2}}$$

15. Donner un équivalent de α_m au voisinage de l'infini.

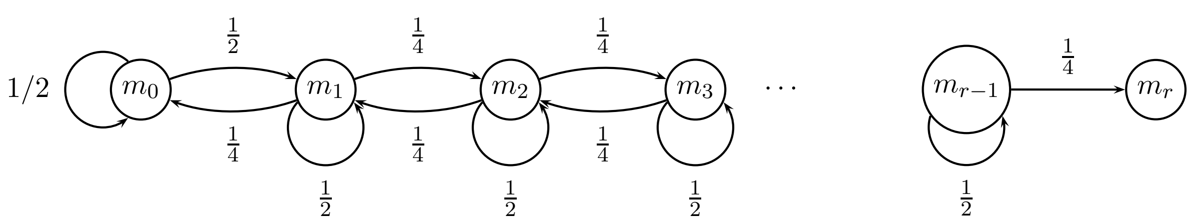
Conclusion : En regardant bien un graphe probabiliste, nous constatons qu'il n'est nul besoin de se ruer vers les matrices ! La fonction génératrice n'est jamais bien loin, du moment que l'on puisse symétriser ou simplifier le problème à l'aide du graphe. En outre, ceci incite à choisir (et c'est ici la difficulté) une modélisation du problème élégante, quand cela est possible. Bien des calculs lourds sont ainsi évités. L'approche matricielle a cependant pour intérêt de donner directement la loi de la variable aléatoire réelle considérée. Comme cela nécessite le calcul des puissances successives (ou de l'inverse) d'une matrice, les techniques visant à réduire la difficulté de cette opération (diagonalisation ou trigonalisation) sont mises en œuvre. Pour autant, selon le cas considéré, les calculs peuvent s'avérer lourds et compliqués.
Ci-joint les graphes (nous laissons le script écrit à l'aide du module numpy de Python à la sagacité du lecteur) donnant les distributions de probabilités de la loi "Premier temps de rencontre sur un polygone à 2^p côtés pour $p = 3, 4, 5, 6$:



La décroissance semble exponentielle, ce qui se vérifie en utilisant une échelle logarithmique :



Remarquons que la seconde règle des parcours nous permet malgré tout de calculer rapidement le temps moyen d'absorption :
Notons $m_0, m_1, \ldots, m_{2^{p-2}}$ les temps moyens d'absorption à partir de l'état $q = 2^{p-1}$ jusqu'à l'état 0, comme figuré dans le graphe probabiliste (non symétrisé) modélisant le problème que nous rappelons, en mettant à l'intérieur des disques chaque temps moyen d'absorption. Posons $r = 2^{p-2}$ pour simplifier.



On en déduit en utilisant la seconde règle des parcours que :

$$\begin{cases} m_0 = 1 + \frac{1}{2}m_0 + \frac{1}{2}m_1, \quad \text{soit } m_0 = 2 + m_1 \\ m_i = 1 + \frac{1}{2}m_i + \frac{1}{4}(m_{i+1} + m_{i-1}), \quad \text{soit } m_{i-1} - 2m_i + m_{i+1} = -4 \quad (1 \leq i \leq r-1) \\ m_r = 0 \quad \text{(condition au bord).} \end{cases}$$

L'équation caractéristique de la suite récurrente linéaire associée à la suite récurrente affine d'ordre 2 : $m_{i-1} - 2m_i + m_{i+1} = -4 \quad (1 \leq i \leq 2^{p-2} - 1)$ s'écrit $x^2 - 2x + 1 = 0$ qui a pour unique solution $x = 1$. Ainsi, les solutions de l'équation linéaire homogène associée sont de la forme $\tilde{u}_n = (An + B)1^n = An + b$. D'autre part, on remarque aisément que $w_n = -2n^2$ est

une solution particulière de notre équation initiale.
Ainsi, pour tout entier naturel n tel que $1 \leq n \leq 2^{p-2} - 1$, on a : $m_n = An + B - 2n^2$.
Ruse classique : posons $m_{-1} = m_1$, de sorte que l'équation $m_{n-1} - 2m_n + m_{n+1} = -4$ reste valable pour $n = 0$. Ainsi, $m_0 = B$. Or $m_0 = 2 + m_1 = A + B$, d'où $A = 0$.
De plus, $m_{2^{p-2}} = 0$; on aimerait donc bien écrire que $B - 2.2^{2p-4} = 0$, ce qui donnerait la valeur de B. Mais la relation $m_{n-1} - 2m_n + m_{n+1} = -4$ est vraie pour $0 \leq n \leq 2^{p-2} - 1$
Ceci dit, en choisissant $n = 2^{p-2} - 1$ et en utilisant le fait que $m_{2^{p-2}} = 0$, on obtient que $m_{2^{p-2}-2} - 2m_{2^{p-2}-1} = -4$, soit $B - 2(2^{p-2} - 2)^2 - 2[B - 2(2^{p-2} - 1)^2] = -4$, qui conduit après un bref calcul à $B = 2^{2p-3}$ i.e $m_0 = 2^{2p-3}$.
Résultat que nous avions d'abord subodoré puis prouvé lors de la seconde modélisation.

4 Plus sur les chaines de Markov

4.1 Classification des états

Considérons une chaine de Markov, absorbante ou non. On note toujours $(\mathcal{P})$ sa matrice de transition. Nous allons répartir les états en classes à l'aide de $(\mathcal{P})$.

Définition 4-1-1 : On dit que l'état j est *accessible* à partir de l'état i s'il existe un entier naturel n tel que $p_{ij}^{(n)} > 0$. On note $i \rightsquigarrow j$.

Propriété 4-1-2 : La relation d'accessibilité entre états est réflexive ($i.e$ $i \rightsquigarrow i$) et transitive ($i.e$ si $i \rightsquigarrow j$ et $j \rightsquigarrow k$, alors $i \rightsquigarrow k$).

Propriété 4-1-3 : Soient i et j deux états. Les propositions suivantes sont équivalentes :

1. L'état j est accessible à partir de l'état i, soit $i \rightsquigarrow j$.
2. Le processus, partant de i, passe par j avec une probabilité strictement positive.

La propriété 4-1-2 ne dit cependant pas que la relation d'accessibilité est symétrique. Par exemple, si la chaine de Markov est absorbante, tout état du bord est accessible, mais une fois atteint, il n'y a plus de retour possible.

Définition 4-1-4 : On dit que deux états i et j *communiquent* si $i \rightsquigarrow j$ et $j \rightsquigarrow i$. On note $i \leftrightsquigarrow j$.

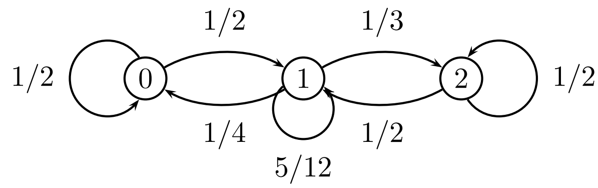
Propriété 4-1-5 : La relation de communication entre états est une relation d'équivalence. On notera C_i la classe d'équivalence de l'état i.
Nous en déduisons que **l'ensemble de états S est partitionné en classes** (non vides et disjointes), dites *classes indécomposables*, dont on peut trouver un système de représentants.

Remarque 4-1-6 : À l'intérieur de chaque classe, tous les états communiquent. En particulier, tout état communique avec lui-même.
Si C_1 et C_2 sont deux classes distinctes, on peut éventuellement relier un état de C_1 à un état de C_2, mais le retour n'est pas possible.
Certaines classes peuvent ne comporter qu'un seul élément, par exemple :

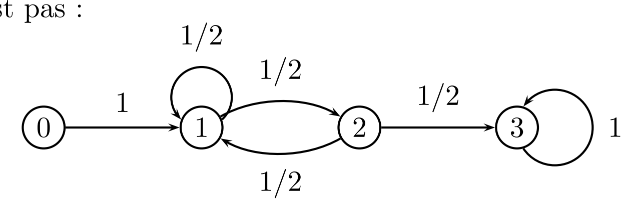
– un *état de non-retour* i : $p_{i,i}^{(0)} = 1$, $p_{i,i}^{(n)} = 0$ pour $n \geq 1$,
– un état absorbant i : $p_{i,i}^{(0)} = 1$, $p_{i,i}^{(n)} = 1$ pour $n \geq 1$.

Définition 4-1-7 : Une chaine de Markov qui ne possède qu'une seule classe d'équivalence (*i.e* tous les états communiquent) est dite *irréductible*.

Exemple 4-1-8 : Le graphe suivant est irréductible : tous les états communiquent.



Alors que celui-ci ne l'est pas :



Les états 0, 1 et 2 ne sont visités qu'un nombre fini de fois ; l'état 0 est un état de non-retour, l'état 3 est absorbant. En tout, il y a trois classes : $\{0\}$, $\{1,2\}$ et $\{3\}$.
Remarquons que la matrice de transition de ce graphe s'écrit :

$$(\mathcal{P}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

L'état de non-retour 0 est caractérisé par une première colonne nulle.

4.2 Temps d'atteintes et de retour

Greg prit sa plus belle voix d'outre-tombe et déclama devant son miroir comme jadis Flaubert le faisait afin de peaufiner ses textes :
Dans ce paragraphe, nous allons préciser la notion de temps d'atteinte, ce qui nous permettra de généraliser la méthode et les concepts employés lors de la seconde modélisation du problème de première rencontre des deux scarabées. Nous retrouverons également la géniale seconde règle de la valeur moyenne vue dans l'article précédent. Et comme je suis généreux, nous disserterons également sur la probabilité de retour d'un état à lui-même, et sur le nombre moyen de tels retours, tout ceci de la manière la plus élémentaire qui soit. Pas de tribu, pas de vocabulaire cauchemardesque : états récurrents, transients ... bien que ces derniers ne soient plus maintenant ... hors d'atteinte !

Définition 4-2-1 : Pour tout état j, on appelle *temps d'atteinte* de la chaine $(X_n)_{n\geq 0}$ dans l'état j à partir de l'instant 1, l'entier (éventuellement infini pour le moment) :

$$T_j := \inf\{n \geq 1; X_n = j\}$$

Remarquons que $(T_j = n) = (X_1 \neq j) \cap (X_2 \neq j) \cap \cdots \cap (X_{n-1} \neq j) \cap (X_n = j)$. Cet événement ne dépend donc que de $X_1, \ldots, X_n$. Il n'y a pas d'hypothèse sur l'état j (intérieur, absorbant, etc.). Cependant, si l'état j est absorbant, on a vu que $T_j < \infty$ (le bord est atteint avec la probabilité 1).

Notation 4-2-2 : Posons $a_{i,j}^{(n)} := P(T_j = n|X_0 = i)$ pour tout entier $n \geq 1$, que l'on note aussi $P^i(T_j = n)$ en prenant comme loi initiale e_i^T.
Ainsi, $a_{i,j}^{(n)}$ est la probabilité pour que le processus, partant de l'état i, atteigne l'état j **pour la première fois**, à l'instant n. Pour tout couple d'états (i,j), on pose conventionnellement $a_{i,j}^{(0)} := 0$.

Théorème 4-2-3 : Pour tout entier $n \geq 1$, on a :

$$p_{i,j}^{(n)} = \sum_{k=0}^{n} a_{i,j}^{(k)} p_{j,j}^{(n-k)}$$

avec la convention $p_{i,j}^{(0)} = \delta_{i,j}$.

Démonstration (abrégée) : Le processus passe de l'état i à l'état j en n étapes s'il passe de i à j pour la première fois en k étapes $(0 \leq k \leq n)$ et s'il passe ensuite de j à j en les $(n-k)$ étapes suivantes. Ces chemins, sont, pour des k distincts, disjoints, et la probabilité pour un chemin fixé est égale à $a_{i,j}^{(k)} p_{j,j}^{(n-k)}$.

Théorème et définition 4-2-4 : L'espérance mathématique de T_j par rapport à la loi P^i est notée $M_{i,j}$. C'est le **temps moyen d'atteinte** de j à partir de i.
La quantité $M_{i,i}$ est appelée **temps de retour moyen** dans i.
On a :

$$M_{i,j} = 1 + \sum_{k \neq j} p_{i,k} M_{k,j}$$

Remarque 4-2-5 : Posons $M = (M_{i,j})$, U la matrice dont tous les coefficients sont égaux à 1, et $\Delta = \mathrm{Diag}(M_{i,i})$. Matriciellement, l'égalité précédente se réécrit :

$$M = U + (\mathcal{P})(M - \Delta)$$

On en déduit que : $(I - (\mathcal{P}))M = U - (\mathcal{P})\Delta$.
Attention, si la chaine de Markov est absorbante, la matrice $I - \mathcal{P}$ est singulière.

Continuons notre étude des chaines de Markov absorbantes, songea Greg. Cette écriture par blocs que peut prendre la matrice de transition $(\mathcal{P})$, je sens qu'il est possible de l'exploiter encore plus ! Allez, hop ! un verre de Guillevic[1] et tout ceci va apparaître au grand jour ! On a[2] :

$$(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & I_r \end{pmatrix}$$

où V est une matrice carrée $(n-r) \times (n-r)$ telle que $I_{n-r} - V$ soit inversible, d'inverse $\sum_{k \geq 0} V^k$ et U une matrice $(n-r) \times r$. La matrice identité I_r **regroupe** les r états absorbants de la chaine de Markov. Si bien que, en passant à la limite sur les puissances successives de $(\mathcal{P})$:

$$(\mathcal{P})^{\infty} = \begin{pmatrix} 0 & (I_{n-r} - V)^{-1} U \\ 0 & I_r \end{pmatrix}$$

Notons :

1. La Guillevic est une variété réputée de pomme à cidre
2. section Généralisation partielle

- $I = \{1, 2, \dots, n-r\}$ l'ensemble des états intérieurs,
- m_i le nombre moyen de sauts jusqu'à l'absorption, partant de l'état i,
- $t_{i,j}$ le nombre moyen de passage en j en partant de i.

Théorème et définition 4-2-6 : Pour tout $(i,j) \in I^2$, on a :

$$\boxed{t_{i,j} = \delta_{i,j} + \sum_{k \in I} p_{i,k} t_{k,j}}$$

soit sous forme matricielle si l'on pose $T = (t_{i,j})$:

$$T = I + VT$$

relation de laquelle on tire (puisque $I - V$ n'est pas singulière) :

$$\boxed{T = (I - V)^{-1}}$$

La matrice T s'appelle la *matrice fondamentale* de la chaine de Markov absorbante.

Corollaire 4-2-7 : Soit $i \in I$.

$$\boxed{m_i = \sum_{k \in I} t_{i,k}}$$

Autrement dit, m_i est égale à la somme des éléments de la $i-$ème ligne de la matrice T.

Démonstration : Soit $i \in I$ l'état de départ. Notons D_i la durée de séjour en I et X_i la variable aléatoire réelle "nombre de pas avant absorption". On a clairement $D_i = X_i$, donc $E(D_i) = m_i$. Mais par définition des $t_{i,j}$, on a $E(D_i) = \sum_{k \in I} t_{i,k}$. Comme quoi, changer de point de vue est encore une fois bénéfique !

Théorème 4-2-8 : Notons B le bord de la chaine de Markov absorbante. Soit $(i,j) \in I \times B$. Notons $b_{i,j}$ la probabilité d'être absorbé en j en partant de i. Alors :

$$\boxed{b_{i,j} = p_{i,j} + \sum_{k \in I} p_{i,k} b_{k,j}}$$

soit sous forme matricielle, si l'on pose $B = (b_{i,j})$:

$$B = U + VB$$

qui se réécrit aisément :

$$\boxed{B = (I - V)^{-1} U \quad i.e \quad B = TU}$$

4.3 Comportement asymptotique

Définition 4-3-1 : La distribution de probabilité $\vec{p} = (p_0, p_1, \dots, p_n)$ est dite *stationnaire* si $\vec{p}(\mathcal{P}) = \vec{p}$.

Définition 4-3-2 : Une matrice de transition $(\mathcal{P})$ est dite *régulière* s'il existe un entier naturel m tel que $(\mathcal{P})^m$ ait tous ses éléments strictement positifs.

Remarque 4-3-3 : Ceci implique entre autres que tous les états communiquent *i.e* que la chaine de Markov associée est irréductible. En particulier, il n'y a aucun état absorbant.

Théorème 4-3-4 : Soit $(\mathcal{P})$ une matrice régulière.

1. $(\mathcal{P})^{\infty} := \lim\limits_{n \to +\infty} (\mathcal{P})^n$ existe, et est une matrice stochastique dont toutes les lignes sont identiques,
2. Soit $\vec{p} = (p_1, p_2, \ldots, p_n)$ une telle ligne. Toutes les composantes de ce vecteur sont strictement positives. C'est de plus l'unique distribution stationnaire de $(\mathcal{P})$,
3. Le temps moyen de retour à l'état i est égal à $M_{i,i} = \dfrac{1}{p_i}$.

Remarque 4-3-5 : Comme pour la convergence simple et la convergence uniforme d'une suite de fonctions en analyse où l'inversion des quantificateurs est cruciale, si $(\mathcal{P})$ est une matrice régulière, alors il existe un entier m tel que pour tout état i, on atteint l'état j en exactement m transitions. Il est maintenant temps de relier la notion de chaine de Markov irréductible à la notion de matrice de transition régulière ...

Définition 4-3-6 : Soit S une chaine de Markov irréductible. Sa matrice de transition $(\mathcal{P})$ est, elle-aussi, dite *irréductible.*

Définition 4-3-7 : Soit $(\mathcal{P})$ une matrice irréductible. Soit $d = \inf\{n \in \mathbb{N}; p_{i,i}^n > 0\}$. d est le PGCD de tous les temps de retour possibles à l'état i.

1. on l'appelle d la *période* de l'état i,
2. $(\mathcal{P})$ est dite *apériodique* si $d = 1$,

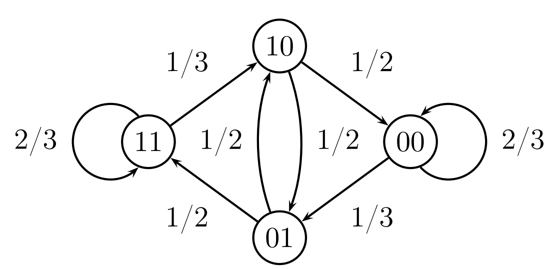
Théorème 4-3-8 : Soit une chaine de Markov irréductible. Alors tous les états i ont la même période.

Théorème 4-3-9 : Une matrice $(\mathcal{P})$ est régulière si elle est irréductible et apériodique.

Corollaire 4-3-10 : De ce qui précède, on en déduit que :

1. $\lim\limits_{n \to +\infty} p_{i,j}^n = p_j$, indépendamment de l'état initial i,
2. Pour tout entier naturel n, $\vec{p}(\mathcal{P})^n = \vec{p}$,

Les exemples d'application fourmillent dans les sujets de baccalauréat "spécialité mathématiques" du bac ES, aussi nous renvoyons le lecteur intéressé vers les annales disponibles sur le site de l'APMEP : `https://www.apmep.fr/`.
Revenons, juste pour le plaisir, sur le concours C 2017 dont nous redonnons ci-dessous le graphe probabiliste, s'esclaffa Greg dans un grand rire psychédélique :



Il est clair que le graphe G de la chaine de Markov est irréductible.
Nous avions re-numéroté les états 11, 10, 01, 00 respectivement par 1, 2, 3, 4. La matrice de transition s'écrivait alors :

$$(\mathcal{P}) = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

Recherchons donc un état stationnaire $\vec{p}$. D'après le théorème 4-3-4, un tel état existe, il est unique et il vérifie : $\vec{p} = \vec{p}(\mathcal{P})^\infty$ (dont l'existence est assurée).
Il nous faut donc déterminer $(\mathcal{P})^\infty$. C'est très limite comme question !
Repoussons-la avec notre meilleure alliée ... la force ! Non, XCas. Nous obtenons sans sourciller :

$$(\mathcal{P})^\infty = \begin{pmatrix} 0,3 & 0,2 & 0,2 & 0,3 \\ 0,3 & 0,2 & 0,2 & 0,3 \\ 0,3 & 0,2 & 0,2 & 0,3 \\ 0,3 & 0,2 & 0,2 & 0,3 \end{pmatrix}$$

La distribution stationnaire apparait immédiatement : $\vec{p} = (0,3; 0,2; 0,2; 0,3)$.

4.4 La formule de Mason

Nous ne nous étendrons pas sur les fondements théoriques de la notion de déterminant d'un endomorphisme, ni sur l'étude détaillée du groupe des permutations. Nous renvoyons le lecteur intéressé aux ouvrages classiques de premier cycle, par exemple [1] ou [2], qui sont d'excellentes références.

Notons S_n le groupe des permutations de $\{1, 2, \dots, n\}$.
Soit f un endomorphisme de $E = \mathbb{R}^n$ muni d'une base quelconque $\mathcal{B}$ et A la matrice de f dans cette base. On appelle déterminant de A dans la base $\mathcal{B}$ le réel $\det_{\mathcal{B}}(A) := \sum_{\sigma \in S_n} \epsilon_\sigma \prod_{i=1}^{n} a_{i,\sigma(i)}$, où ϵ_σ désigne la *signature* de la permutation σ.
Le déterminant est invariant par changement de base. On peut donc parler sans ambiguïté du déterminant de l'endomorphisme f.

Théorème 4-4-1 : Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice vérifiant $\det(A) \neq 0$. Alors A est inversible et on a :

$$A^{-1} = \frac{1}{\det(A)} \mathrm{com}(A)^T$$

Le théorème que nous allons énoncer maintenant est très utilisé en automatique et même en chimie. On peut consulter par exemple :
`http://www.lassc.ulg.ac.be/webCheng00/SYST011/Dyna03_fTransfert.pdf`
ou :
`http://public.iutenligne.net/automatique-et-automatismes-industriels/verbeken/cours_au_mv/chapitre5/chap52.html`

ou encore :
`https://cel.archives-ouvertes.fr/cel-00156394/file/Chapitre5.pdf`

Définition 4-4-2 : Soit S l'ensemble des états d'une chaine de Markov absorbante et G son graphe associé. On appelle *déterminant du graphe* G le réel défini par :

$$\Delta = 1 - \sum L_i + \sum L_i L_j - \sum L_i L_j L_k + \dots$$

La première sommation porte sur toutes les boucles, la seconde sur toutes les paires de boucles disjointes, la troisième sur tous les triplets de boucles disjointes, etc.
Remarquons que $\Delta = \det(I - V)$ dans la décomposition par blocs de la matrice de transition $(\mathcal{P})$ du graphe G.

Théorème 4-4-3 (Formule de Mason) : $T_{i,j}$ (avec $T = (I - V)^{-1}$) est égal à :

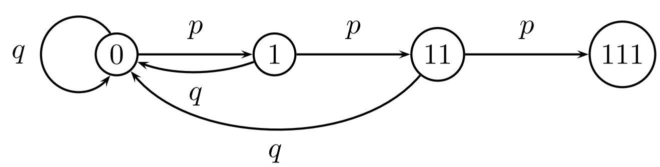
$$\frac{1}{\Delta} \sum_k P_k \Delta_k$$

où P_k est la probabilité de transition d'un **chemin direct** de i vers j et Δ_k le **cofacteur** du chemin k, $i.e$ le déterminant des éléments de G non touchés par le chemin k.

Remarque 4-4-4 : il ne s'agit rien d'autre que de la formule de Cramer ! Nous pouvons en déduire la méthode d'application qui suit :

1. calculer le déterminant Δ du graphe complet,
2. déterminer les K chemins directs reliant l'état i à l'état j,
3. calculer la probabilité de transition P_k de chacun des K chemins directs,
4. calculer pour chaque chaine directe le déterminant Δ_k du graphe obtenu en supprimant tous les nœuds de la k-ième chaine directe,
5. calculer enfin $T_{i,j}$ par la formule de Mason.

Exemple 4-4-5 : Détaillons le raisonnement précédent sur un exemple un peu "tête de nœud", comme l'exemple 1-3-1 dont nous redonnons le graphe probabiliste :



En re-numérotant les états 0, 1, 11, 111 par 1, 2, 3, 4, nous allons calculer $T_{1,i}$ pour $i = 1, 2, 3, 4$.

1. Le graphe comporte trois boucles : $0 \to 0$, $0 \to 1 \to 0$, $0 \to 1 \to 11 \to 0$. Il n'y a donc aucune boucle disjointe. On obtient donc :

$$\Delta = 1 - q - pq - p^2 q = 1 - q(1 + p + p^2)$$

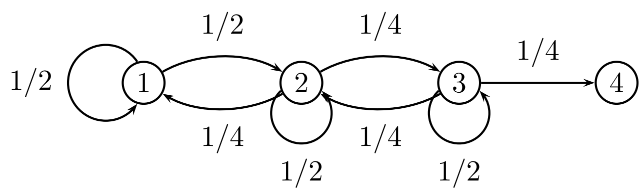
soit $\Delta = 1 - q\frac{1-p^3}{1-p} = p^3$
2. Il n'y a qu'un seul chemin direct menant de l'état 1 à l'état 2, et il a pour probabilité p ; celui (unique aussi) menant de l'état 1 à l'état 3 est de probabilité p^2.

3. Chacun de ces trois chemins directs touchent chacune des boucles donc $\Delta_1 = \Delta_2 = \Delta_3 = 1$,
4. On en déduit d'après la formule de Mason que :

$$T_{1,2} = \frac{p}{p^3} = \frac{1}{p^2},\ T_{1,3} = \frac{p^2}{p^3} = \frac{1}{p}$$

Remarquons que $T_{1,1} = \frac{1}{p^3}$ car il n'existe qu'un seul chemin menant du nœud 1 à lui-même, et donc $p_1 = 1$, et en le supprimant on voit que $\Delta_0 = 1$. Le temps moyen d'absorption, partant de l'état 1 est donc égal à : $m_1 = \displaystyle\sum_{j=1}^{3} T_{1,j} = \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} = \frac{1}{q}\left[\left(\frac{1}{p}\right)^3 - 1\right]$.

Exemple pédagogique 4-4-6 : Le retour des scarabées !
Étudions, à l'aide de la formule de Mason le temps moyen de première rencontre de nos deux scarabées sur un polygone à $2^4 = 16$ côtés. Selon les résultats établis précédemment, on doit trouver $2^{2\times 4-3} = 32$. Dans ce qui suit, il y a une grossière erreur. À vous de la trouver !
Le graphe probabiliste modélisant le problème est le suivant :



Ce graphe comporte un certain nombre de boucles :
- $1 \to 1,\ 2 \to 2,\ 3 \to 3$, toutes de probabilité $\dfrac{1}{2}$,
- $2 \to 3 \to 2$, de probabilité $\left(\dfrac{1}{4}\right)^2$,
- $1 \to 2 \to 1$, de probabilité $\dfrac{1}{4} \times \dfrac{1}{2}$,
- $1 \to 2 \to 3 \to 2 \to 1$, de probabilité $\dfrac{1}{2}\left(\dfrac{1}{4}\right)^3$,

On en déduit immédiatement le déterminant du graphe :

$$\Delta = 1 - \left(3 \times \frac{1}{2} + \left(\frac{1}{4}\right)^2 + \frac{1}{4} \times \frac{1}{2} + \frac{1}{2}\left(\frac{1}{4}\right)^3)\right) + \left(3 \times \left(\frac{1}{2}\right)^2 + \frac{1}{4} \times \left(\frac{1}{2}\right)^2 + \frac{1}{2} \times \left(\frac{1}{4}\right)^2\right) - \left(\frac{1}{2}\right)^3$$

soit :

$$\Delta = \frac{3}{128}$$

Il y a un seul chemin menant de l'état 1 à lui-même : $1 \to 1$. Comme vu avant, $t_{1,1} = \dfrac{\Delta_1}{\Delta}$. En enlevant le nœud 1, il reste trois boucles :
- $2 \to 2$ et $3 \to 3$, disjointes et de probabilité $\frac{1}{2}$,
- $2 \to 3 \to 2$, de probabilité $\frac{1}{16}$.

Ainsi, $\Delta_1 = 1 - \left(2 \times \dfrac{1}{2} + \dfrac{1}{16}\right) + \left(\dfrac{1}{2}\right)^2 = \dfrac{3}{16}$.
On en déduit $t_{1,1} = \dfrac{128}{3} \times \dfrac{3}{16} = 8$.
De même, on laisse le soin au lecteur de vérifier que $t_{1,2} = \dfrac{128}{3} \times \dfrac{1}{4} = \dfrac{32}{3}$ et que $t_{1,3} =$

$\frac{128}{3} \times \frac{1}{8} = \frac{16}{3}$.
Il vient que le temps moyen d'absorption, en partant de l'état 1, est égal à :

$$m_1 = \sum_{j=1}^{3} t_{1,j} = 24$$

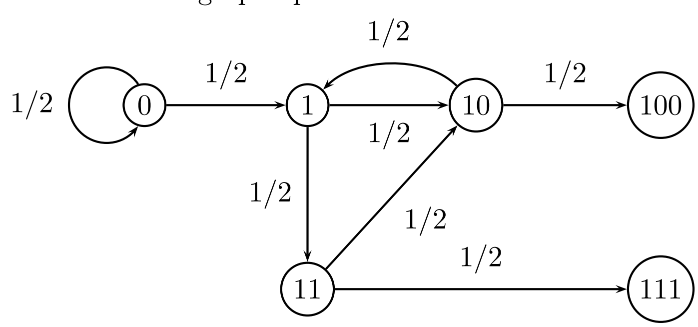
Mais alors, où est l'erreur ? ? ? (exercice)

4.5 Un dernier exemple d'application

Reprenons, à une légère modification près l'exercice 1 de la section 2-5.
On lance une pièce équilibrée jusqu'à l'obtention de (Face,Face,Face) ou de (Face,Pile,Pile).
Nous allons nous intéresser, en variant les points de vue :

1. à la loi de probabilité de ce jeu,
2. à la probabilité de chacun des événement considérés,
3. au temps moyen de jeu.

Notons 0 pour Pile et 1 pour Face. Les états absorbants sont donc 111 et 100. Comme tous les deux débutent par 1 (Face), on peut, sans perte de généralité, supposer que l'état de départ est 0, ce qui nous conduit au graphe probabiliste suivant :



Afin d'étudier la loi de probabilité de la variable aléatoire X : nombre de pas avant l'absorption, nous pouvons considérer indifféremment :
- la matrice de transition du graphe et ses puissances successives, de manière à appliquer la relation de Chapman-Komolgorov ou de manière équivalente
- déterminer sa fonction génératrice ϕ_X

Notons que la seconde approche peut être assouplie en réduisant préalablement le graphe, mais il nous faudra en contrepartie effectuer un développement en série entière, ce qui n'est pas nécessaire avec l'approche matricielle, à condition bien entendu, que le calcul de puissances successives de la matrice de transition soit aisé.

Approche 1 : Tout matriciel !
La matrice de transition du graphe peut s'écrire (avec $p = 1/2$) :

$$(\mathcal{P}) = \begin{pmatrix} p & p & 0 & 0 & 0 & 0 \\ 0 & 0 & p & p & 0 & 0 \\ 0 & p & 0 & 0 & p & 0 \\ 0 & 0 & p & 0 & 0 & p \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

soit :

$$(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & I_2 \end{pmatrix}$$

où $V = \begin{pmatrix} p & p & 0 & 0 \\ 0 & 0 & p & p \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \end{pmatrix}$ et $U = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ p & 0 \\ 0 & p \end{pmatrix}$.

X-Cas donne une expression immonde pour V^n. Même en essayant de diagonaliser V. Il semble tentant de tenter l'approche "fonctions génératrices".

Approche 2 : Fonctions génératrices.
La matrice de transition d'un parcours **non marqué** peut s'écrire (avec $p = 1/2$) :

$$(\mathcal{P}) = \begin{pmatrix} pt & pt & 0 & 0 & 0 & 0 \\ 0 & 0 & pt & pt & 0 & 0 \\ 0 & pt & 0 & 0 & pt & 0 \\ 0 & 0 & pt & 0 & 0 & pt \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

soit :

$$(\mathcal{P}) = \begin{pmatrix} V & U \\ 0 & I_2 \end{pmatrix}$$

où $V = \begin{pmatrix} pt & pt & 0 & 0 \\ 0 & 0 & pt & pt \\ 0 & pt & 0 & 0 \\ 0 & 0 & pt & 0 \end{pmatrix}$ et $U = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ pt & 0 \\ 0 & pt \end{pmatrix}$.

La loi initiale est e_1^T. D'où, en utilisant les résultats de la section Généralisation Partielle :

$$\phi_X(t) = \sum_{j=1}^{2} [(I_4 - V)^{-1} U]_{1,j}$$

Greg éclata de rire en repensant à ses longs exercices de calculs d'inverses de matrice au cours de sa tendre enfance mathématique. Comme je te X-Cas ça maintenant !!!
On obtient aisément la première ligne de $[(I_4 - V)^{-1} U]$:

$$\left(\frac{t(2 \cdot t^3 + 4 \cdot t^2)}{2(t^4 - 4 \cdot t^2 - 8 \cdot t + 16)}, \frac{2 \cdot t^3}{t^4 - 4 \cdot t^2 - 8 \cdot t + 16} \right)$$

Cette dernière nous permet deux options :

1. Calculer directement les probabilités d'être absorbé en 100 et en 111,
2. Donner l'expression explicite de $\phi_X(t)$, puis en calculant $\phi'_X(1)$, d'en déduire le temps moyen d'absorption.

On a en remplaçant t par 1 :
- $P(\text{absorption en } 100) = [(I_4 - V)^{-1} U]_{1,1} = 3/5$
- $P(\text{absorption en } 111) = [(I_4 - V)^{-1} U]_{1,2} = 2/5$

puis en sommant les deux colonnes :

$$\phi_X(t) = \frac{t^4 + 4t^3}{t^4 - 4t^2 - 8t + 16}$$

En utilisant la propriété 1-1-6, on trouve à l'aide d'un logiciel de calcul formel que $E(X) = 28/5$ et que $V(X) = 188/25$.
Utilisons encore X-Cas pour déterminer la matrice fondamentale T de notre chaine de Markov :

$$T = \begin{pmatrix} 2 & \frac{8}{5} & \frac{6}{5} & \frac{4}{5} \\ 0 & \frac{8}{5} & \frac{6}{5} & \frac{4}{5} \\ 0 & \frac{4}{5} & \frac{8}{5} & \frac{2}{5} \\ 0 & \frac{2}{5} & \frac{4}{5} & \frac{6}{5} \end{pmatrix}$$

On a en utilisant le corollaire 4-2-7, que le temps moyen d'absorption, en partant de l'état 0, est égal à $2 + \frac{8}{5} + \frac{6}{5} + \frac{4}{5} = \frac{28}{5}$. Tout colle bien !

Mais pour "gagner" la loi de X, il convient de développer en série entière $\phi_X(t)$.
Remarquons que (merci X-Cas) :

$$\phi_X(t) = 1 + \frac{6}{t-2} - 2\frac{t^2 + 6t + 8}{t^3 + 2t^2 - 8}$$

L'équation $t^3 + 2t^2 - 8 = 0$ a trois solutions : une solution réelle : $\alpha \approx 1,5098$ et deux solutions complexes conjuguées : $\beta, \bar{\beta} \approx -1,7549 \pm 1,4897i$.
La fonction rationnelle $F(t) := \dfrac{t^2 + 6t + 8}{t^3 + 2t^2 - 8}$, considérée comme une fonction de la variable complexe, s'écrit sous la forme : $F(t) = \dfrac{A}{t - \alpha} + \dfrac{B}{t - \beta} + \dfrac{C}{t - \bar{\beta}}$.
Classiquement :

- $A = \lim_{t \to \alpha} (t - \alpha) F(t) = \dfrac{\alpha^2 + 6\alpha + 8}{\alpha^2 - 2\alpha \Re(\beta) + |\beta|^2}$,
- $B = \lim_{t \to \beta} (t - \beta) F(t) = \dfrac{\beta^2 + 6\beta + 8}{(\beta - \alpha)(\beta - \bar{\beta})}$,
- $C = \bar{B}$.

Rappelons que pour $a \in \mathbb{C}^*$ et pour $|t| < |a|$, on a :

$$\sum_{n \geq 0} \frac{1}{a^{n+1}} t^n = \frac{1}{a - t}$$

On a :

$$\phi_X(t) = 1 - 3\frac{1}{1 - t/2} + 2A\frac{1}{\alpha - t} + 2B\frac{1}{\beta - t} + 2\bar{B}\frac{1}{\bar{\beta} - t}$$

donc pour $|t| < \min(2, \alpha, |\beta|) = \alpha$:

$$\phi_X(t) = 1 - 3\sum_{n \geq 0} \left(\frac{t}{2}\right)^n + 2A\sum_{n \geq 0} \frac{1}{\alpha^{n+1}} t^n + 2B\sum_{n \geq 0} \frac{1}{\beta^{n+1}} t^n + 2\bar{B}\sum_{n \geq 0} \frac{1}{\bar{\beta}^{n+1}} t^n$$

soit :

$$\phi_X(t) = 1 + \sum_{n \geq 0} \left(\frac{-3}{2^n} + \frac{2A}{\alpha^{n+1}} + 4\Re\left(\frac{\bar{B}\beta^{n+1}}{|\beta|^{2(n+1)}}\right)\right) t^n$$

En particulier, on vérifie que $P(X = 0) = P(X = 1) = P(X = 2) = 0$, ce qu'indiquait le graphe probabiliste.

La loi de probabilité de X est donc définie par :

$$\begin{cases} X(\Omega) = [3; +\infty[\cap\mathbb{N} \\ \forall n \in X(\Omega),\ P(X = n) = \dfrac{-3}{2^n} + \dfrac{2A}{\alpha^{n+1}} + \dfrac{4}{|\beta|^{2(n+1)}}\Re(\bar{B}\beta^{n+1}) \end{cases}$$

Remarque : Nous pouvions également nous servir de la formule de Mason afin de déterminer le temps moyen d'absorption.
Posons $p = 1/2$ et re-numérotons les sommets 0, 1, 10, 11, 100, 111 respectivement par 1, 2, 3, 4, 5, 6.
Les boucles du graphe sont :

- S_1 : $1 \to 1$, de probabilité p,
- S_2 : $2 \to 3 \to 2$, de probabilité p^2,
- S_3 : $2 \to 4 \to 3 \to 2$, de probabilité p^3

Remarquons que le premier chemin S_1 est disjoint des deux autres, qui eux ne le sont pas entre eux.
On en déduit que le déterminant du graphe est égal à : $\Delta = 1-(p+p^2+p^3)+(p^3+p^4) = 5/16$.
Calculons $t_{1,j}$ pour $j = 1, 2, 3, 4$. Ainsi, nous en déduirons le temps moyen d'absorption partant de l'état 1 : $m_1 = \displaystyle\sum_{j=1}^{4} t_{1,j}$.
Commençons par remarquer que pour tout état i, il existe un unique chemin allant de i à i. Ainsi, $t_{i,i} = \dfrac{\Delta_i}{\Delta}$. On a donc $t_{1,1} = \dfrac{\Delta_1}{\Delta}$.

- Calcul de $t_{1,1}$: $S\backslash\{1\}$ n'a aucune paire de boucles disjointes, donc $\Delta_1 = 1-(p^2+p^3) = \dfrac{5}{8}$.
  On en déduit que $t_{1,1} = \dfrac{5}{8} \times \dfrac{16}{5} = 2$.
- Calcul de $t_{1,2}$: il y a un seul chemin direct menant de 1 à 2 : $1 \to 2$, qui est de probabilité p. En supprimant les noeuds 1 et 2 de S, il n'y a plus de boucles donc $\Delta_2 = 1$. D'où $t_{1,2} = \dfrac{1}{2} \times \dfrac{16}{5} = \dfrac{8}{5}$.
- Calcul de $t_{1,3}$: il y a deux chemins directs menant de 1 à 3 : $1 \to 2 \to 3$, de probabilité p^2, et $1 \to 2 \to 4 \to 3$, de probabilité p^3. Comme précédemment, les cofacteurs associés sont égaux à 1, de sorte que $t_{1,3} = \left(\dfrac{1}{4} + \dfrac{1}{8}\right) \times \dfrac{16}{5} = \dfrac{6}{5}$.
- Calcul de $t_{1,4}$: on laisse le lecteur vérifier que $t_{1,4} = \dfrac{4}{5}$.

On retrouve bien la première ligne de $T = (I - V)^{-1}$. Ainsi, $m_1 = \dfrac{28}{5}$.

Greg se frotta les mains en vidant une bouteille d'hydromel dans le bac de tri situé en bas de sa demeure.
"Je crois que nous avons pas mal tourné autour du pot jusqu'à l'ivresse", songeait-il. Une étoile filante traversa alors le ciel de novembre en laissant une traînée argentée persistante qui tardait à s'effacer.
"Je crois qu'il est temps de passer le flambeau ... Et pour diluer ces réflexions, rien de tel que de se diriger vers les bords d'eau. Vers Bordeaux où mon ami "Lolo" a beaucoup à dire avec sa faconde légendaire. Les sièges de ce bar Salsa, proche de la place de la Victoire murmurent encore son nom en rougissant. Encore plus qu'à Nevers ..., au bien nommé "l'Agricole". Allez, il est temps de brancher ! Pour la postérité ...
Au prochain article, nous étudierons en détail les Processus de branchement, étudiant notamment la persistance d'une lignée.

5 Compléments

5.1 Coefficient binomial généralisé

Définition : Soit n un entier naturel et a un nombre complexe. La *factorielle montante* est définie par : $(a)_n := \begin{cases} 1 & \text{si } n = 0, \\ a(a+1)\dots(a+n) & \text{si } n \geq 1 \end{cases}$

Définition : Le *coefficient binomial* $\binom{a}{n}$ (où n un entier naturel et a un nombre complexe) est défini par :

$$\binom{a}{n} := \frac{a(a-1)\dots(a-n+1)}{n!} = (-1)^n \frac{(-a)_n}{n!}$$

5.2 Séries entières

Lemme d'Abel : Considérons une série entière $\sum_{n \geq 0} a_n z^n$. On suppose qu'il existe $z_0 \in \mathbb{C}$ tel que $(a_n z_0^n)_{n \geq 0}$ soit bornée. Alors pour tout $z \in \mathbb{C}$ tel que $|z| < |z_0|$, $\sum_{n \geq 0} a_n z^n$ converge absolument.

Théorème et définition : Soit $\sum_{n \geq 0} a_n z^n$ une série entière. Alors il existe $R \in [0; +\infty]$ vérifiant les conditions équivalentes suivantes :

1. $R = \sup\{|z| \ ; \ \sum_{n \geq 0} a_n z^n \ \text{ converge absolument}\}$
2. $R = \sup\{|z| \ ; \ \sum_{n \geq 0} a_n z^n \ \text{ converge}\}$
3. $R = \sup\{|z| \ ; \ (a_n z^n) \ \text{ tend vers } 0\}$
4. $R = \sup\{|z| \ ; \ (a_n z^n) \ \text{ est bornée}\}$

R s'appelle le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$.

Calcul du rayon de convergence : Le théorème précédent donne à lui seul des moyens pratiques de calcul du rayon de convergence. La place des suites géométriques dans une échelle de comparaison à l'infini est également fort utile :

$$(ln(n))^\alpha << n^\beta << t^n << e^{\gamma n}, \ \beta, \gamma > 0, t > 1$$

Citons néanmoins deux théorèmes classiques de calcul du rayon de convergence :

Formule de Hadamard : Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence R non nul. Alors $R = \dfrac{1}{\overline{\lim}_{n \to +\infty} |a_n|^{1/n}}$.

Formule de Cauchy : Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence R. On suppose qu'à partir d'un certain rang $a_n \neq 0$, alors : $\dfrac{1}{R} = \overline{\lim}_{n \to +\infty} \dfrac{|a_{n+1}|}{|a_n|}$.

Quelques résultats utiles : Soient $\sum_{n\geq 0} a_n z^n$ et $\sum_{n\geq 0} b_n z^n$ deux séries entières de rayons de convergences respectifs R_1 et R_2. Alors :

1. Le rayon de convergence R de $\sum_{n\geq 0}(a_n + b_n)z^n$ vérifie $R \geq \min\{R_1; R_2\}$, avec égalité si $R_1 \neq R_2$.
2. Le rayon de convergence R de $\left(\sum_{n\geq 0} a_n z^n\right)\left(\sum_{n\geq 0} b_n z^n\right)$ vérifie $R \geq \min\{R_1; R_2\}$, avec égalité si $R_1 \neq R_2$.
3. La fonction $f(z) = \sum_{n\geq 0} a_n z^n$ est indéfiniment dérivable à l'intérieur du disque ouvert de convergence D_R et pour tout entier naturel p, on a :

$$f^{(p)}(z) = \sum_{n\geq p} n(n-1)\dots(n-p+1)z^{n-p}$$

Remarquons qu'une série entière et ses dérivées ont même rayon de convergence.

Développement en série entière : Après avoir défini le principe, nous donnerons un formulaire succinct de quelques développements classiques. Notons qu'il s'agit là de la démarche inverse du calcul explicite de l'expression d'une série de fonctions. On a par exemple, pour $|t| < 1$: $\sum_{k\geq 0} t^k = \frac{1}{1-t}$.

Inversement, le développement en série entière de la fonction définie sur $]-1;1[$ par $f(t) = \frac{1}{1-t}$ est $f(t) = \sum_{k\geq 0} t^k$.

Définition : On dit qu'une fonction $f : \mathbb{C} \to \mathbb{C}$, définie sur un voisinage de $z_0 \in \mathbb{C}$, est *développable en série entière* en z_0, si et seulement s'il existe une série entière $\sum_{n\geq 0} a_n z^n$, de rayon de convergence $R > 0$ telle que :

$$\exists \alpha \in]0; R[\ ; \ z \in B(z_0, \alpha) \Rightarrow f(z) = \sum_{n\geq 0} a_n (z - z_0)^n$$

Remarque : Par translation, on peut supposer $z_0 = 0$.
Ce développement est unique, et est précisément égal à la série de Mac-Laurin de f :

$$\sum \frac{1}{n!} f^{(n)}(0) z^n$$

Nous ne nous étendrons pas sur les conditions nécessaires pour une fonction d'être développable en série entière. [1] le fait très bien dans son tome 4.
Nous considérons par ailleurs très souvent des fonctions de la variable réelle. Le tableau des développements usuels peut s'écrire …

comme dans le lien ci-dessous !
`http://www.panamaths.net/Documents/Formulaires/FORMU_DSEUSUELS.pdf`

6 Bibliographie

1. Ramis-Deschamps-Odoux, Cours de mathématiques spéciales, Tome 1, Masson (1993)
2. J. Calais, Théorie des groupes, PUF (1984)
3. A. Engel, Processus aléatoires pour les débutants, L-Cassini (2011)
4. D. Foata, A. Fuchs, Processus stochastiques, Dunod (2004)