2024

Processus de naissance et de mort et quelques applications

Butoyi, Anitha

UB, FS

UNIVERSITE DU BURUNDI
FACULTE DES SCIENCES
DEPARTEMENT DE MATHEMATIQUES
CENTRE DE RECHERCHE EN MATHEMATIQUES ET PHYSIQUE

PROCESSUS DE NAISSANCE ET DE MORT ET POLYNOMES ORTHOGONAUX

Par :

Anitha BUTOYI

Mémoire présenté et défendu publiquement en vue de l'obtention du Diplôme de Master en Mathématiques Fondamentales et Appliquées.
Option : Mathématiques Fondamentales et Appliquées

Sous la direction de :
Pr François NDAYIRAGIJE

Bujumbura, Mai 2024

Les membres du Jury

Professeur Gaspard BANGEREZAKO (**Président**)
Docteur Denis NKURUNZIZA (**Secrétaire**)
Docteur Jean Bosco KAYOYA (**membre**)
Professeur François NDAYIRAGIJE (**Directeur**)

Dédicace

A Dieu le Tout Puissant,
A mon cher mari et à nos enfants,
A mes très chers parents,
A mes frères et sœurs ,
A toute ma famille,
A nos amis et connaissances.

Remerciement

Nous tenons en premier lieu à remercier Dieu de nous avoir donné le courage, la patience et la volonté pour réaliser ce travail.

Nous tenons à exprimer nos remerciements au Directeur de ce mémoire Professeur François NDAYIRAGIJE, pour avoir dirigé ce mémoire, pour ses précieux conseils, son aide, son soutien scientifique et humain qui nous a donné la patience et la volonté d'accomplir ce modeste travail.

Mes vifs remerciements, vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre mémoire en acceptant d'examiner notre travail. Nous adressons aussi nos remerciements à tous les enseignants qui nous ont aidé tout au long de cette formation de Master. Enfin, nous tenons à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Résumé

Dans ce travail, nous présentons l'intérêt des processus de naissance et de mort face aux nombreux problèmes que nous rencontrons en mathématiques appliquées et en physique théorique et mathématiques. On y développe entre autre la notion des polynômes orthogonaux, des processus stochastiques, des processus de Markov à temps discret et continu et à espace d'état discret. Enfin, nous appliquons ces processus de naissance et de mort aux polynômes orthogonaux.

Mots-clés : Processus de naissance, Processus de mort, Processus stochastique, Polynômes orthogonaux.

Abstract

In this work, we present the interest of the processes of birth and death in the face of many problems that we encounter in applied mathematics and in theoretical physics and mathematics. We develop, among other things, the notion of orthogonal polynomials, stochastic processes, discrete and continuous time Markov processes, and discrete state space. Finally, we apply these birth and death processes on orthogonal polynomials.

Keywords : Birth process, Death process, Stochastic process, Orthogonal polynomials.

Table des matières

Liste des tableaux

Table des figures

Avant-propos

Le présent travail de mémoire a été réalisé dans le cadre de l'obtention du diplôme de fin d'études du deuxième cycle des enseignements de l'Université du Burundi. L'objectif de ce mémoire est motivé après avoir vu que beaucoup de phénomènes stochastiques c'est à dire qui viennent par hasard nécessitent une modélisation.

Dans ce travail, nous allons rappelé les notions de base qui nous conduisent aux généralités sur les polynômes orthogonaux. Nous allons aussi rappelé les notions de processus stochastiques, les chaînes et les processus de Markov à temps discret et la notion de processus de naissance et de mort. Enfin, nous allons parlé de processus de naissance et de mort et les polynômes orthogonaux.

INTRODUCTION

Notre travail est basé sur la notion de processus stochastiques dont les applications sont connues en fonctions spéciales notamment en polynômes orthogonaux. Ces polynômes étant les solutions de l'équation différentielle du type hypergéométrique $\sigma(x)y''(x) + \tau(x)y'(x) + \lambda y(x) = 0$, où $\sigma(x)$ est un polynôme de degré $\leqslant 2$, $\tau(x)$ est un polynôme de degré 1 et λ est une constante. Cette équation est utilisée presque partout dans de nombreux problèmes de mathématiques appliquées et de physique théorique et mathématique.

L'idée de travailler sur ce sujet est motivée après avoir vu que beaucoup de phénomènes stochastiques c'est à dire qui viennent par hasard nécessite une modélisation.

Nous allons utiliser les modèles markoviens qui sont donc des cas particuliers de processus markoviens généraux appelées processus de vie (naissance) et de mort et dont ses applications nous aide à modéliser des phénomènes stochastiques.

Le premier chapitre traite des généralités sur les polynômes orthogonaux dans le but d'introduire les chapitres deux et trois. Le deuxième chapitre traite les processus stochastiques qui vont nous donner la notion complète des processus de naissance et de mort et le troisième chapitre parle des processus de naissance et de mort et les polynômes orthogonaux.

Chapitre 1

Théorie générale sur les polynômes orthogonaux

Dans ce chapitre, nous rappelons les théories de base qui nous conduisent aux généralités des polynômes orthogonaux.

Les référence principales de ce chapitre sont :[1, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14].

1.1 Espace de Hilbert

La réference principale de cette section est [16]

Nous allons utiliser des espaces vectoriels H avec produit scalaire

$$\langle .\,,\,. \rangle : H \times H \to \mathbb{R},$$

avec les propriétés suivantes : pour tous f, g, $h \in H$ on a

1. $\langle f, g\rangle = \langle g, f\rangle$, $\quad$ (*symétrie*),

2. $\langle af + bg, h\rangle = a\langle f, h\rangle + b\langle g, h\rangle$ pour tous $a, b \in \mathbb{R}$, $\quad$ (linéarité),

3. $\langle f, f\rangle \geqslant 0$ et $\langle f, f\rangle = 0 \Leftrightarrow f = 0$. $\quad$ (positivité)

On introduit la norme

$$\|f\|_2 = \sqrt{\langle f, f\rangle},$$

et on a **l'inégalité de cauchy-schwartz**

$$|\langle f, g\rangle| \leq \| f \|_2 \| g \|_2.$$

On a aussi

$$\|f + g\|_2^2 = \langle f + g, f + g\rangle = \langle f, f\rangle + \langle g, g\rangle + 2\langle f, g\rangle,$$

$$\|f - g\|_2^2 = \langle f - g, f - g\rangle = \langle f, f\rangle + \langle g, g\rangle - 2\langle f, g\rangle,$$

et donc

$$\| f + g \|_2^2 + \| f - g \|_2^2 = 2(\| f \|_2^2 + \| g \|_2^2)$$

ce qu'on appelle **la règle du parallélogramme**. Cette règle est valide seulement si la norme se déduit d'un produit scalaire.

L'espace vectoriel H est un espace de Hilbert si l'espace est complet pour la norme $\| \,.\, \|_2$. Ça veut dire que chaque suite de Cauchy $(f_n)_{n\in\mathbb{N}}$ en H a une limite $f \in H$:

$$\forall\epsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n,\ m \in \mathbb{N} : n,\ m \geqslant n_0 \Rightarrow \| f_n - f_m \|_2 < \epsilon$$

implique qu'il existe une fonction $f \in H$ avec

$$\lim_{n\to\infty} \| f_n - f \|_2 = 0,$$

(convergence de f_n vers f en norme).

Exemple 1.1. $H = \mathbb{R}^n = \{x = (x_1, x_2, \cdots, x_n) \,|\, x_i \in \mathbb{R},\ 1 \leq i \leq n\}$ *avec*

$$\langle x, y\rangle = \textstyle\sum_{k=1}^n x_k y_k, \qquad \| x \|_2 = \sqrt{\textstyle\sum_{k=1}^n x_k^2}$$

est l'espace euclidien de dimension n.

Exemple 1.2. $H = l^2(\mathbb{N}) = \{x = (x_n)_{n\in\mathbb{N}} \,|\, \sum_{k=0}^\infty x_k^2 < \infty\}$ *avec*

$$\langle x, y\rangle = \textstyle\sum_{k=0}^\infty x_k y_k, \qquad \| x \|_2 = \sqrt{\textstyle\sum_{k=0}^\infty x_k^2},$$

est l'espace de suites qui sont de carré sommable.

Exemple 1.3. $H = L^2(\omega) = \{f : [a, b] \to \mathbb{R} \,|\, \int_a^b |f(x)|^2 \omega(x) dx < \infty\}$, *où* $\omega : [a, b] \to \mathbb{R}_+$

est une fonction de poids, avec

$$\langle f, g\rangle = \textstyle\int_a^b f(x)g(x)\omega(x)dx, \qquad \| f \|_2 = \sqrt{\textstyle\int_a^b |f(x)|^2 \omega(x)dx}.$$

En effet il faut considérer l'espace de classes d'équivalence pour des fonctions mesurables sur [a,b] où $f \equiv g$ quand l'ensemble $\{x \in [a, b] \,|\, f(x) \neq g(x)\}$ à une mesure de Lebesgue zéro.

1.2 Approximation des moindres carrés

Suppose H est espace de Hilbert et $V_n \subset H$ est un sous-espace de dimension n avec une base $\{\varphi_1, \varphi_2, \varphi_3, \cdots, \varphi_n\}$, donc chaque élément $\varphi \in V_n$ est de la forme $\varphi = \sum_{k=1}^{n} a_k \varphi_k$, avec des coefficients $a_1, a_2, \cdots, a_n \in \mathbb{R}$. Nous voulons construire pour $f \in H$ une approximation $f_n \in V_n$ qui est la meilleur approximation en V_n pour la norme $\| \ . \ \|_2$. Le problème est donc de minimiser $\| \ f - f_n \ \|_2$ parmi tous les fonctions $f_n \in V_n$ ou de trouver

$$\min_{a_1, a_2, \cdots, a_n \in \mathbb{R}} \| \ f - \textstyle\sum_{k=1}^{n} a_k \varphi_k \ \|_2$$

On appelle le f_n qui atteint le minimum **une approximation de moindre carrés** de f en V_n. On a

$$\begin{aligned} \| \ f - \sum_{k=1}^{n} a_k \varphi_k \ \|_2^2 &= \langle f - \sum_{k=1}^{n} a_k \varphi_k, f - \sum_{k=1}^{n} a_k \varphi_k \rangle \\ &= \langle f, f \rangle - 2 \sum_{k=1}^{n} a_k \langle f, \varphi_k \rangle + \sum_{k=1}^{n} \sum_{l=1}^{n} a_k a_l \langle \varphi_k, \varphi_l \rangle \end{aligned} \tag{1.1}$$

et les variables $a_1, a_2, \cdots, a_n$ appartiennent d'une manière quadratique. Pour le minimiser, on a les conditions nécessaires

$$\frac{\partial}{\partial_{a_j}} \| \ f - \textstyle\sum_{k=1}^{n} a_k \varphi_k \ \|_2^2 = 0, \qquad 1 \leq j \leq n \ .$$

On observe que

$$\textstyle\sum_{k=1}^{n} \sum_{l=1}^{n} a_k a_l \langle \varphi_k, \varphi_l \rangle = \sum_{k=1}^{n} a_k^2 \langle \varphi_k, \varphi_k \rangle + \sum_{k \neq l}^{n} a_k a_l \langle \varphi_k, \varphi_l \rangle,$$

et donc les conditions nécessaires sont

$$-2\langle f, \varphi_j \rangle + 2a_j \langle \varphi_j, \varphi_j \rangle + 2 \textstyle\sum_{k \neq j}^{n} a_k \langle \varphi_k, \varphi_j \rangle = 0,$$

ou

$$\sum_{k=1}^{n} a_k \langle \varphi_k, \varphi_j \rangle = \langle f, \varphi_j \rangle, \quad 1 \leq j \leq n. \tag{1.2}$$

C'est un système linéaire pour les coefficients $a_1, a_{2 \ldots}, a_n$ qu'on peut écrire sous forme matricielle

$$\begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_2, \varphi_1 \rangle & \ldots & \langle \varphi_n, \varphi_1 \rangle \\ \langle \varphi_1, \varphi_2 \rangle & \langle \varphi_2, \varphi_2 \rangle & \ldots & \langle \varphi_n, \varphi_2 \rangle \\ \vdots & \vdots & \ldots & \vdots \\ \langle \varphi_1, \varphi_n \rangle & \langle \varphi_2, \varphi_n \rangle & \ldots & \langle \varphi_n, \varphi_n \rangle \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \langle f, \varphi_2 \rangle \\ \vdots \\ \langle f, \varphi_n \rangle \end{pmatrix}, \tag{1.3}$$

ce qu'on appelle **le système normale** pour l'approximation des moindre carrés. La matrice

$$G_n = \begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_2, \varphi_1 \rangle & \ldots & \langle \varphi_n, \varphi_1 \rangle \\ \langle \varphi_1, \varphi_2 \rangle & \langle \varphi_2, \varphi_2 \rangle & \ldots & \langle \varphi_n, \varphi_2 \rangle \\ \vdots & \vdots & \ldots & \vdots \\ \langle \varphi_1, \varphi_n \rangle & \langle \varphi_2, \varphi_n \rangle & \ldots & \langle \varphi_n, \varphi_n \rangle \end{pmatrix} \tag{1.4}$$

où $(G_n)_{ij} = \langle \varphi_j, \varphi_i \rangle$ est une matrice symétrique qu'on appelle **la matrice de Gram** . Cette matrice est définie positive : pour chaque $x = (x_1, x_n, \cdots, x_1) \in \mathbb{R}^n$ on a $xG_nx^T \geq 0$ et $xG_nx^T = 0$ si et seulement si $x = (0, 0, \cdots, 0)$. En effet, on a

$$xG_nx^T = \sum_{k=1}^{n}\sum_{j=1}^{n} x_k \langle \varphi_k, \varphi_j \rangle x_j$$

$$= \langle \sum_{k=1}^{n} x_k\varphi_k, \sum_{j=1}^{n} x_j\varphi_j \rangle$$

$$= \| \sum_{k=1}^{n} x_k\varphi_k \|_2^2 \geq 0.$$

On observe que $xG_nx^T = 0$ est possible si et seulement $\sum_{k=1}^{n} x_k\varphi_k = 0$ et comme $\varphi_1, \cdots, \varphi_n$ sont linéairement indépendantes, cela est possible si et seulement si

$x_1 = x_2 = \cdots = x_n$. Donc $det\, G_n > 0$ et la matrice de Gram est non-singulière et la solution $(a_1, \cdots, a_n)$ du système (1.3) existe et est unique.

Théorème 1.1. *Supposons $f \in H$ et $V_n \subset H$ est un espace vectoriel de dimension n avec base $\{\varphi_1, \cdots, \varphi_n\}$. L'approximation $f_n \in V_n$ des moindres carrés est donnée par $f_n = \sum_{k=1}^{n} a_k\varphi_k$, où $(a_1, \cdots, a_n)$ est la solution unique du système normal* (1.3)*. L'erreur $f - f_n$ est orthogonale sur l'espace V_n*

$$\langle f - f_n, g \rangle = 0, \qquad \forall g \in V_n, \tag{1.5}$$

et sa norme satisfait

$$\| f - f_n \|_2^2 = \| f \|_2^2 - \| f_n \|_2^2 . \tag{1.6}$$

On dit que f_n est la projection orthogonale de f dans l'espace V_n et (1.6) *est la formule de Pythagore pour les approximations des moindres carrés.*

Démonstration.

Nous avons déjà trouver que $(a_1, a_2, \cdots, a_n)$ est la solution unique de (1.3)*.*

Pour montrer que $\langle f - f_n, g \rangle = 0$ quand $g \in V_n$, on observe que pour $g = \varphi_j$ on a

$$\langle f - f_n, \varphi_j \rangle = \langle f, \varphi_j \rangle - \textstyle\sum_{k=1}^{n} a_k \langle \varphi_k, \varphi_j \rangle = 0$$

grâce à (1.2)*. Par linéarité on a donc*

$$\langle f - f_n, \textstyle\sum_{k=1}^{n} c_j\varphi_j \rangle = 0,$$

pour tous $c_1, c_2, \cdots, c_n \in \mathbb{R}$ et alors $\langle f - f_n, g \rangle = 0$ pour toute $g \in V_n$. Finalement on a

$\| f - f_n \|_2^2 = \langle f - f_n, f - f_n \rangle = \langle f, f - f_n \rangle - \langle f_n, f - f_n \rangle = \langle f, f - f_n \rangle$, par ce que $f_n \in V_n$ et grâce à (1.5)*.*

En continuant, on a

$\| f - f_n \|_2^2 = \langle f, f\rangle - \langle f, f_n\rangle = \langle f, f\rangle - \langle f - f_n, f_n\rangle - \langle f_n, f_n\rangle,$

et comme $\langle f - f_n, f_n\rangle = 0$ *on trouve (1.6)*

1.3 Équation différentielle du type hypergéométrique et les polynômes orthogonaux très classique d'une variable continue

Les référence principales de cette section sont [1, 2, 3, 6, 9, 10]

1.3.1 Équation du type hypergéométrique

On va considérer des équations différentielles de la forme

$$\sigma(x)y''(x) + \tau(x)y'(x) + \lambda y(x) = 0, \tag{1.7}$$

où σ est un polynômes de degré $\leqslant 2$, τ est un polynômes de degré 1 et λ est une constante. Nous sommes intéressés aux solutions polynomiales. Si y est un polynôme de degré n, on peut comparer les coefficients de x^n dans (1.7) pour trouver $n(n-1)\frac{\sigma''}{2} + n\tau' + \lambda = 0$, et cela donne que la constante λ doit être égale à

$$\lambda_n = -n(n-1)\frac{\sigma''}{2} - n\tau'. \tag{1.8}$$

Nous introduisons la fonction w comme solution de

$$(\sigma w)' = \tau w, \tag{1.9}$$

avec des conditions aux limites $(\sigma w)(a) = (\sigma w)(b) = 0$ pour des points $a, b \in \mathbb{R} \cup \{\pm\infty\}$ à déterminer. Cette fonction w est un facteur intégrable et on peut écrire (1.7) comme

$$\Big(\sigma(x)w(x)y'(x)\Big)' + \lambda w(x)y(x) = 0. \tag{1.10}$$

On appelle équation (1.9) l'équation de Pearson, après le statisticien Karl Pearson. L'équation différentielle (1.10) est auto-adjointe : si on prend le produit scalaire

$$\langle f, g\rangle = \textstyle\int_a^b f(x)g(x)dx,$$

En intégrant par parties on obtient

$$\langle \sigma w f, g\rangle' = -\langle \sigma w f', g'\rangle,$$

et on a toujours

$$\langle \sigma w f', g'\rangle = \langle f', \sigma w g'\rangle.$$

Une autre intégration par partie donne

$$\langle f', \sigma w g'\rangle = -\langle f, (\sigma w g')'\rangle,$$

et on conclut que

$$\langle (\sigma w f')', g\rangle = \langle f, (\sigma w g')'\rangle,$$

ou

$$\int_a^b \Big(\sigma(x)w(x)f'(x)\Big)' g(x)dx = \int_a^b f(x)\Big(\sigma(x)w(x)g(x)'\Big)' dx. \tag{1.11}$$

1.3.2 Polynômes orthogonaux très classique d'une variable continue

De l'équation (1.7), nous allons considérer plusieurs cas.

1.3.2.1 Polynômes d'Hermite

Nous prenons $deg\, \sigma = 0$ et $\sigma(x) = 1$, et $\tau(x) = Ax + B$.

De ce fait, la solution de (1.9) est

$log\, w(x) = \frac{Ax^2}{2} + Bx + C$, ce qui donne

$$w(x) = e^{\frac{Ax^2}{2}+Bx+C}.$$

Nous prenons aussi $A = -2$, $B = 0$ et, la solution $\omega(x)$ pour ce cas est

$$w(x) = e^{-x^2},\ a = -\infty < x < b = \infty.$$

La solution polynomiale de (1.7) après la substitution des valeurs de λ calculées à l'aide de(1.8), de τ et de σ, nous obtenons

$$y''(x) - 2xy'(x) + 2ny(x) = 0$$

qui est **le polynôme d'Hermite** H_n de degré n , avec

$$a = -\infty,\ b = +\infty, w(x) = e^{-x}$$

Si on prend $f = H_n$, $g = H_m$ dans (1.11), on trouve

$$\int_a^b \left(e^{-x^2} H_n'(x)\right)' H_m(x)dx = \int_a^b H_n(x)\left(e^{-x^2} H_m'(x)\right)' dx,$$

et en utilisant l'équation différentielle auto-adjointe

$$\left(e^{-x^2} H_n'(x)\right)' + 2ne^{-x^2} H_n(x) = 0,$$

on a

$-2n \int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}dx = -2m \int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}dx,$

et ainsi

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}dx = 0, n \neq m,$$

ce qui donne l'orthogonalité du polynôme d'Hermite.

1.3.2.2 Polynômes de Laguerre

Le polynôme σ est de degré 1 . On prend $\sigma = x$ et $\tau(x) = Ax + B$.

La solution de (1.9) après les calculs est $log\,[xw(x)] = Ax + B[log\,|x|] + C$, ce qui donne

$$xw(x) = |x|^B e^{ax+B}.$$

Nous prenons $A = -1, B = \alpha + 1$ et la solution ω pour ce cas est

$$w(x) = x^\alpha e^{-x}, 0 < x < \infty, \alpha > -1.$$

La solution polynomiale de (1.7) après la substitution des valeurs de λ calculées à l'aide de (1.8), de τ et de σ, nous obtenons

$$xy''(x) + (\alpha + 1 - x)y'(x) + ny(x) = 0,$$

qui est **le polynôme de Laguerre** $L_n^{(\alpha)}$ de degré n et de paramètre α.

Par la même procédé que pour le cas précédent, on arrive aussi à l'orthogonalité de la forme

$$-n \int_0^\infty x^\alpha e^{-x} L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)dx = -m \int_0^\infty x^\alpha e^{-x} L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)dx,$$

et ainsi

$$\int_0^\infty x^\alpha e^{-x} L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)dx = 0, \quad n \neq m.$$

ce qui donne l'orthogonalité du polynôme de Laguerre, avec :

$$a = 0, b = +\infty, w(x) = e^{-x}x^\alpha, \alpha > -1.$$

1.3.2.3 Polynômes de Jacobi

Le polynôme σ est de degré 2 , $\tau(x) = Ax + B$ et $\sigma(x) = 1 - x^2$(est le seul cas intéressant quand σ a deux zéros simples et réels). Pour certaines constantes α, β, la solution de (1.9) est

$$1 - x^2 w(x) = (1 - x)^{\alpha+1}(1 + x)^{\beta+1} \ .$$

Prenons

$$\tau(x) = -x(\alpha + \beta + 2) - \alpha + \beta,$$
$$w(x) = (1 - x)^{\alpha}(1 + x)^{\beta},$$

et on obtient la solution des polynômes de Jacobi $P_n^{(\alpha,\beta)}(x)$ avec paramètres $\alpha, \beta > -1$, $a = -1$, $b = +1$.

Les polynômes de Legendre sont des cas particulier avec $\alpha = \beta = 0$.

En résume, on obtient un tableau de valeurs pour ces quatre types de polynômes qui est de la forme

Tableau 1.1 – Polynômes orthogonaux classiques

	Hermite	Laguerre	Jacobi	Legendre
	$H_n(x)$	$L_n^{(\alpha)}(x)$	$P_n^{(\alpha,\beta)}(x)$	$P_n(x)$
interval	(-∞, ∞)	[0, ∞)	[-1,1]	[-1,1]
$\omega(x)$	e^{-x^2}	$x^\alpha e^{-x}$	$(1-x)^\alpha(1+x)^\beta$	1
$\sigma(x)$	1	x	$1-x^2$	$1-x^2$
$\tau(x)$	$-2x$	$-x+\alpha+1$	$-x(\alpha+\beta+2)-\alpha+\beta$	$-2x$
λ_n	$2n$	n	$n(n+\alpha+\beta+1)$	$n(n+1)$
remarques		$\alpha > -1$	$\alpha, \beta > -1$	$\alpha = \beta = 0$

1.4 Polynômes orthogonaux classiques d'une variable discrète

1.4.1 Introduction

Soit la formule de Rodrigues

$$y_n(x) = \frac{B_n}{\rho(x)} \bigtriangledown^n [\rho_n(x)], \tag{1.12}$$

avec

$$\begin{aligned} \Delta y_n(x) &= -\lambda_n \frac{B_n}{\rho_1(x)} \bigtriangledown^{n-1} [\rho_n(x)] \\ &= -\lambda_n \frac{B_n}{B_{n-1}}(x)\overline{y}_{n-1}(x), \end{aligned} \tag{1.13}$$

Nous écrivons l'équation aux différences $\Delta(\sigma\rho) = \tau\rho$ sous la forme

$$\frac{\rho(x+1)}{\rho(x)} = \frac{\sigma(x) + \tau(x)}{\sigma(x+1)}. \tag{1.14}$$

La solution d'une équation aux différences

$\frac{\rho(x+1)}{\rho(x)} = f(x)$ possède la propriété suivante

Si $\rho_1(x)$ et $\rho_2(x)$ sont des solutions des équations

$$\frac{\rho_1(x+1)}{\rho_1(x)} = f_1(x),$$
$$\frac{\rho_2(x+1)}{\rho_2(x)} = f_2(x),$$

alors les solutions de l'équation

$$\frac{\rho(x+1)}{\rho(x)} = f(x),$$

avec $f(x) = f_1(x)f_2(x)$ est

$$\rho(x) = \rho_1(x)\rho_2(x),$$

et avec $f(x) = \frac{f_1(x)}{f_2(x)}$ est

$$\rho(x) = \frac{\rho_1(x)}{\rho_2(x)}.$$

Puisque le membre de droite de (1.14) est une fonction rationnelle, sa solution peut être exprimé en termes de solutions des équations aux différences

$$\frac{\rho(x+1)}{\rho(x)} = \gamma + x, \tag{1.15}$$

$$\frac{\rho(x+1)}{\rho(x)} = \gamma - x, \tag{1.16}$$

$$\frac{\rho(x+1)}{\rho(x)} = \gamma, \tag{1.17}$$

où γ est une constante. De

$$\gamma + x = \frac{\Gamma(\gamma + x + 1)}{\Gamma(\gamma + x)},$$

une solution particulière de (1.15) est de la forme $\gamma + x$.

Similairement en utilisant

$$\gamma - x = \frac{\Gamma(\gamma - x + 1)}{\Gamma(\gamma - x)} = \frac{1}{\Gamma[(\gamma + 1) - (x + 1)]} : \frac{1}{\Gamma(\gamma - x)}$$

on obtient une solution particulière de(1.16)

$$\rho(x) = \frac{1}{\Gamma(\gamma + 1 - x)},$$

et une solution particulière de (1.17) est

$$\rho(x) = \gamma^x.$$

Trouvons maintenant les solutions de (1.14) correspondant aux différents degrés de $\sigma(x)$.

1.4.2 Polynôme de Hahn de première espèce et de Chebyshev

Soit $\sigma(x) = x(\gamma_1 - x)$, $\sigma(x) + \tau(x) = (x + \gamma_2)(\gamma_3 - x)$.

Ici $(\gamma_1, \gamma_2, \gamma_3)$ sont des constantes.

Dans ce cas (1.14) a la forme

$$\frac{\rho(x+1)}{\rho(x)} = \frac{(x+\beta+1)(N-1-x)}{(x+1)(N+\alpha-1-x)}. \tag{1.18}$$

La solution de cette équation est

$$\rho(x) = \frac{\Gamma(N+\alpha-x)\Gamma(x+\beta+1)}{\Gamma(x+1)\Gamma(N-1)},\ \alpha > -1,\ \beta > -1. \tag{1.19}$$

Les polynômes $y_n(x)$ obtenus par la formule de Rodrigues (1.12) lorsque $B_n = \frac{(-1)^n}{n!}$, de fonction de poids $\rho(x)$ définie par (1.19), sont appelés **les polynômes de Hahn de première espèce** et sont notés $h_n^{(\alpha,\beta)}(x, N)$.

On peut aussi utiliser la notation $h_n^{(\alpha,\beta)}(x)$ quant N est fixé.

Un cas spécial important des polynômes de Hahn, appelés polynômes de CHEBYSHEV d'une variable discrète, notés par

$t_n(x) = h_n^{(0,0)}(x, N)$,

apparaissent quand $\rho(x) = 1$.

1.4.3 Polynôme de Hahn de seconde espèce

Soit $\sigma(x) = (x + \gamma_1)$, $\sigma(x) + \tau(x) = (\gamma_2 - x)(\gamma_3 - x)$, $\gamma_1 > -1, \gamma_2 > N - 2, \gamma_3 = N - 1$.

Dans ce cas, la solution de (1.14) est

$$\rho(x) = \frac{1}{\Gamma(x+1)\Gamma(x+\mu+1)\Gamma(N+\nu-x)\Gamma(N-x)}, \quad \mu > -1, \nu > -1, \tag{1.20}$$

ici $\mu = \gamma_1$, $\nu = \gamma_2 - N + 1$.

Les polynômes $y_n(x)$ obtenus par la formule de Rodrigues avec $B_n = \frac{1}{n!}$, lorsque $\rho(x)$ est défini par (1.20), sont aussi appelés **polynômes de Hahn de seconde espèce** et sont notés $\tilde{h}_n^{(\mu,\nu)}(x, N)$.

1.4.4 Polynôme de Meixner, Kravchuk et Charlier

Soit $\sigma(x) = x$, nous considérons trois cas :

$$\sigma(x) = \begin{cases} \mu(\gamma + x), \\ \mu(\gamma - x), \\ \mu. \end{cases} \tag{1.21}$$

On a alors la solution :

$$\rho(x) = \begin{cases} C\frac{\mu^x \Gamma(\gamma+x)}{\Gamma(x+1)} \\ C\frac{\mu^x}{\Gamma(x+1)\Gamma(\gamma+1-x)} \\ C\frac{\mu^x}{\Gamma(x+1)}. \end{cases} \tag{1.22}$$

Si nous prenons $C = \frac{1}{\Gamma(\gamma)}$, nous obtenons la distribution de Pascal à partir de la théorie des probabilités,

$\rho(x) = \frac{\mu^x(\gamma)_x}{\Gamma(x+1)}$.

Prenant $B_n = \mu^{-n}$, les polynômes correspondants sont **les polynômes de Meixner** $m_n^{(\gamma,\mu)(x)}$.

$\gamma = n, \quad \mu = \frac{p}{q}, \quad p + q = 1\,(p > 0,\, q > 0), \quad C = q^N N!$.

Si nous prenons $B_n = (-1)^n \frac{q^n}{n!}$, les polynômes correspondants sont **les polynômes de Kravchuk** $k_n^{(p)}(x, N)$.

Prenant $a = 0, b = +\infty, C = e^{-\mu}$, nous avons la distribution de Poisson

$$S(x_i) = \frac{e^{-\mu}\mu^i}{i!}$$

Les polynômes orthogonaux d'une variable discrète correspondant, avec $B_n = \mu^{-n}$, sont appelés polynômes de Charlier $c_n^{(\mu)}(x)$.

De (1.13) nous obtenons les formules aux différences des Polynômes pour **Hahn**, **Meixner**, **Kravchuk** et **Charlier** :

$\Delta h_n^{(\alpha,\beta)}(x, N) = (\alpha + \beta + n + 1) h_{n-1}^{(\alpha+1,\beta+1)}(x, N - 1);$

$\Delta \tilde{h}_n^{(\mu,\nu)}(x, N) = -(\mu + \nu + 2N - n - 1) \tilde{h}_{n-1}^{(\mu,\nu)}(x, N - 1);$

$\Delta k_n^{(p)}(x, N) = k_{n-1}^{(p)}(x, N - 1),$

$\Delta m_n^{(\gamma,\mu)}(x) = -\frac{n(1-\mu)}{\mu} m_{n-1}^{(\gamma+1,\mu)}(x);$

$\Delta c_n^{(\mu)}(x) = -\frac{n}{\mu} c_{n-1}^{(\mu)}(x).$

En résumé, on a le tableau suivant :

Tableau 1.2 – Polynômes orthogonaux d'une variable discrete

	Meixner	Charlier	Kravuchuk	Hahn première espèce	Hahn de seconde espèce
	$m_n^{(\gamma,\mu)}(x)$	$c_n^{(\mu)}(x)$	$k_n^{(p)}(x,N)$	$h_n^{(\alpha,\beta)}(x,\,N)$	$\tilde{h}_n^{(\mu,\nu)}(x,\,N)$
$\sigma(x)$	x	x	x	$x(\gamma_1 - x)$	$x(x+\gamma_1)$
$\sigma(x)+\tau(x)$	$\mu(\gamma+x)$	μ	$\mu(\gamma-x)$	$(\gamma_2+x)(\gamma_3-x)$	$(\gamma_2-x)(\gamma_3-x)$
B_n	μ^{-n}	μ^{-n}	$\frac{(-1)^n q^n}{n!}$	$\frac{(-1)^n}{n!}$	$\frac{1}{n!}$
ρ_n	$C\frac{\mu^x(\gamma)_x}{\Gamma(x+1)}$	$C\frac{\mu^x}{\Gamma(x+1)}$	$C\frac{\mu^x}{\Gamma(x+1)\Gamma(\gamma-x+1)}$	$\frac{\Gamma(N+\alpha-x)\Gamma(\beta+x+1)}{\Gamma(N-x)\Gamma(x+1)}$	$\frac{1}{\Gamma(x+1)\Gamma(x+\mu+1)\Gamma(N+\nu-x)\Gamma(N-x)}$

Remarque :

Pour les polynômes de Kravchuk, $\mu = \frac{p}{q}$, $p+q=1$.

Pour les polynômes de Hahn première espèce, $\gamma_1 = N+\alpha$, $\gamma_2 = \beta+1$, $\gamma_3 = N-1$.

Pour les polynômes de Hahn deuxième espèce espèce,$\mu = \gamma_1$, $\nu = \gamma_2 - N + 1$, $\gamma_1 > -1, \gamma_2 > N-2$, $\gamma_3 = N-1$.

1.5 Schéma d'Askey pour les polynômes orthogonaux

Les polynômes orthogonaux très classiques d'Hermite, Laguerre et Jacobi et les polynômes orthogonaux discrets classiques de Charlier et Meixner se retrouvent sur les trois niveaux inférieurs du schéma d'Asky des polynômes hypergéométriques.

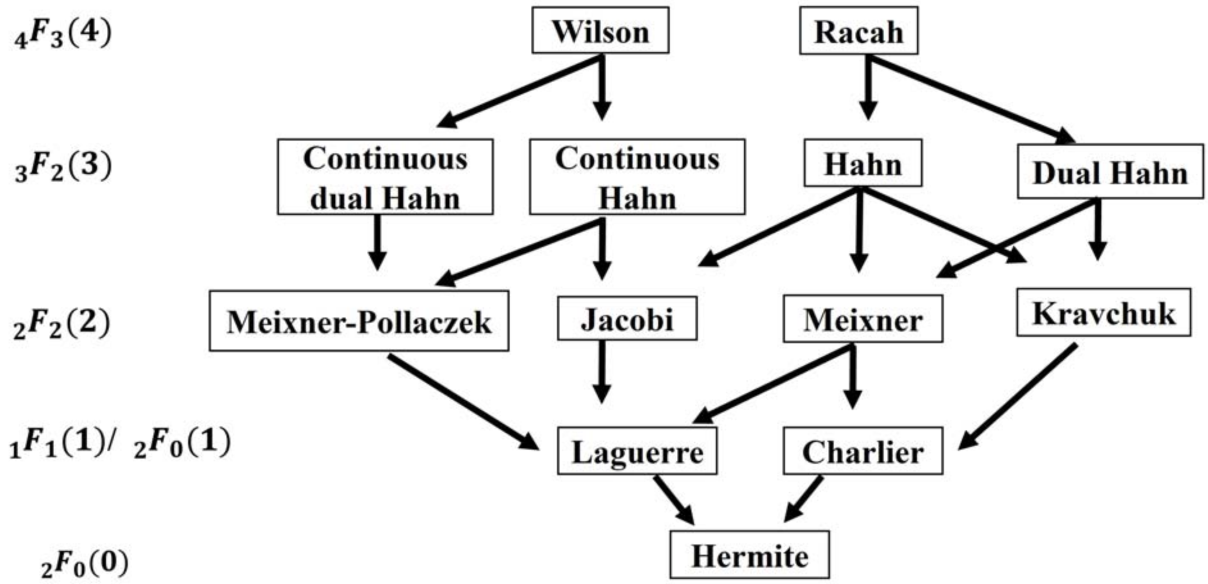

FIGURE 1.1 – Schéma d'Askey pour les polynômes orthogonaux

où $\quad {}_rF_s \left[\begin{matrix} a_1, a_2, \cdots, a_r \\ \\ b_1, b_2, \cdots, b_s \end{matrix} ; z \right] = \sum_{k=0}^{\infty} \frac{(a_1)_k, (a_2)_k \cdots, (a_r)_k}{(b_1)_k, (b_2)_k \cdots, (b_s)_k} \frac{z^k}{k!},$

z est "le point d'évaluation", r et s sont respectivement le nombre de paramètres "hauts" et "bas" de la série.

Les paramètres "hauts" $a_1, a_2, \cdots, a_r$, "bas" $b_1, b_2, \cdots, b_s$ et la variable z est réel ou complexe.

$(a)_k$ est le symbole de pochhammer :

$$(a)_k = \begin{cases} 1,\ si\, k = 0 \\ a(a+1)(a+2)\cdots(a+k-1) = \frac{(a+k-1)!}{(a-1)!},\ si\, k = 1, 2, 3, \cdots. \end{cases}$$

$$\begin{aligned} (a)_k &= (1)_k \\ &= 1(1+1)(1+2)\cdots(1+k-1) \\ &= 1.2.3\cdots k \\ &= k!. \end{aligned}$$

Les polynômes de kravchuk sont des polynômes orthogonaux sur un nombre fini d'entiers et correspondent aux polynômes de Meixner avec le choix spécial $\beta = -N$ du paramètre. Les polynômes de Meixner -Pollaczek correspondent à une équation de Pearson avec un opérateur aux différences des translatons purement imaginaires. Les polynômes de Hahn sont des polynômes orthogonaux discrets sur un nombre fini d'entiers. Les polynômes "Dual Hahn ", les polynômes Hahn continus et les polynômes"Dual Hanh" continus, les polynômes de Wilson et de Racah ont une équation de Pearson avec un opérateur aux différences divisé. Les polynômes de Racah et les polynômes "Dual Hahn " sont des polynômes orthogonaux discrets sur un réseau quadratique, alors que les polynômes de Wilson et les polynômes de Hahn continus et les "Dual Hahn " continus sont orthogonaux sur un intervalle infini.

1.6 Orthogonalité de Gram-Schmidt

L'espace de Hilbert $H = L^2\ (\mu)$, avec μ une mesure positive sur $\mathbb{R}$ (ou sur un ensemble $S \subset \mathbb{R}$) contient les fonctions μ-mesurables f sur S avec

$$\int |f(x)|^2 d\mu(x) < \infty.$$

Quand μ est absolument continue avec la dérivée de Radon-Nikodym $\frac{d\mu}{dx} = w(x)$, on retrouve l'espace $L^2(w)$. Le produit scalaire est

$$\langle f, g\rangle = \int f(x)g(x)d\mu(x).$$

Nous voulons trouver les polynômes orthogonaux pour ce produit scalaire. On peut le faire avec **la méthode de Gram-Schmidt** : les fonctions $\{1, x, x^2, ..., x^n\}$ constituent une base pour l'espace V_{n+1} des polynômes de degré $\leqslant n$. Le polynôme orthonormé de degré 0 est

$$p_0(x) = \frac{1}{\|1\|_2}.$$

Supposons que nous avons déjà construit les polynômes orthonormés $p_0, p_1, ..., p_k$, alors le polynôme

$$P_{k+1}(x) = x^{k+1} - \sum_{j=0}^{k} \langle x^{k+1}, p_j\rangle p_j(x).$$

est un polynôme de degré k+1 avec

$$\langle P_{k+1}, p_m\rangle = \langle x^{k+1}, p_m\rangle - \sum_{j=0}^{k} \langle x^{k+1}, p_j\rangle \langle p_j, p_m\rangle$$

et l'orthogonalité $\langle p_j, p_m\rangle = \delta_{j,m}$ pour $j, m \leqslant k$ donne

$$\langle P_{k+1}, p_m\rangle = \langle x^{k+1}, p_m\rangle - \langle x^{k+1}, p_m\rangle = 0,\ m \leqslant k,$$

donc P_{k+1} est orthogonal sur $p_0, p_1, ..., p_k$. Le polynôme

$$p_{k+1} = \frac{P_{k+1}(x)}{\|P_{k+1}\|_2}$$

est donc le polynôme orthonormé de degré k + 1. On dit que P_n est le polynôme orthogonal monique, et

$$\gamma_n = \frac{1}{\|P_n\|_2} > 0,$$

donc le polynôme orthonormé est $p_n(x) = \gamma_n P_n(x) = \gamma_n x^n + \ldots$. Une méthode plus directe utilise les moments de la mesure μ

$$m_n = \int x^n d\mu(x).$$

On a

$$P_n(x) = \frac{1}{\triangle_{n-1}} det \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_n \\ m_1 & m_2 & m_3 & \cdots & m_{n+1} \\ m_2 & m_3 & m_4 & \cdots & m_{n+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-1} \\ 1 & x & x^2 & \cdots & x^n \end{pmatrix} \tag{1.23}$$

où $\triangle_n$ sont les déterminants

$$\triangle_n = det \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_n \\ m_1 & m_2 & m_3 & \cdots & m_{n+1} \\ m_2 & m_3 & m_4 & \cdots & m_{n+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-1} \\ m_n & m_{n+1} & m_{n+2} & \cdots & m_{2n} \end{pmatrix}.$$

En effet, on trouve

$$\langle P_n, x^k \rangle = \int P_n(x) x^k d\mu(x) = \frac{1}{\triangle_{n-1}} det \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_n \\ m_1 & m_2 & m_3 & \cdots & m_{n+1} \\ m_2 & m_3 & m_4 & \cdots & m_{n+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-1} \\ m_k & m_{k+1} & m_{k+2} & \cdots & m_{k+n} \end{pmatrix},$$

et quand $k < n$ la dernière ligne est la même que la ligne k + 1, donc

$$\int P_n(x) x^k d\mu(x) = 0.$$

Quand k = n on a

$$\int P_n(x)x^k d\mu(x) = \frac{\triangle_n}{\triangle_{n-1}},$$

donc le polynôme orthonormé est

$$p_n(x) = \frac{\sqrt{\triangle_{n-1}}}{\sqrt{\triangle_n}} P_n(x) = \frac{1}{\sqrt{\triangle_{n-1}\triangle_n}} det \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_n \\ m_1 & m_2 & m_3 & \cdots & m_{n+1} \\ m_2 & m_3 & m_4 & \cdots & m_{n+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ m_{n-1} & m_n & m_{n+1} & \cdots & m_{2n-1} \\ 1 & x & x^2 & \cdots & x^n \end{pmatrix}.$$

on a donc

$$\gamma_n = \frac{\sqrt{\triangle_{n-1}}}{\sqrt{\triangle_n}}.$$

1.7 Relation de récurrence pour les polynômes orthogonaux

Les polynômes orthogonaux classiques satisfont une équation différentielle de deuxième ordre. Cela est seulement vrai pour les polynômes orthogonaux classiques. Mais tous les polynômes orthogonaux généraux satisfont une relation de récurrence de deuxième ordre : l'orthogonalité est équivalente à une relation de récurrence de deuxième ordre.

Théorème 1.2. *Une suite $(p_n)_{n\in\mathbb{N}}$ de polynômes orthonormés sur $\mathbb{R}$*

$$\int_{\mathbb{R}} p_n(x)p_m(x)d\mu(x) = \delta_{n,m}, \quad n, m \in \mathbb{N},$$

avec une mesure positive μ, toujours satisfait

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_n p_n(x) + a_n p_{n-1}(x), \quad n \geqslant 0, \tag{1.24}$$

avec des valeurs initiales $p_{-1} = 0$ et $p_0 = \frac{1}{\sqrt{m_0}}$ où $m_0 = \int d\mu(x)$. On a les formules

$$a_n = \int_{\mathbb{R}} xp_n(x)p_{n-1}(x)d\mu(x) = \frac{\gamma_{n-1}}{\gamma_n} > 0, \tag{1.25}$$

$$b_n = \int_{\mathbb{R}} xp_n^2(x)d\mu(x). \tag{1.26}$$

Démonstration.

Le polynôme $xp_n(x)$ est un polynôme de degré $n+1$ et on peut l'écrire dans la base $\{p_k(x) \mid 0 \leqslant k \leqslant n+1\}$ comme

$$xp_n(x) = \sum_{k=0}^{n+1} a_{k,n} p_k(x)$$

avec les coefficients de Fourier

$$a_{k,n} = \int_{\mathbb{R}} xp_n(x)p_k(x)d\mu(x) = \int_{\mathbb{R}} p_n(x)xp_k(x)d\mu(x).$$

Le polynôme $xp_k(x)$ est de degré k+1 et l'orthogonalité donne que $a_{k,n} = 0$ quand k + 1 $< n$. Alors on a que trois termes

$$xp_n(x) = a_{n+1,n}p_{n+1}(x) + a_{n,n}p_n(x) + a_{n-1,n}p_{n-1}(x),$$

avec

$$a_{n,n} = \int_{\mathbb{R}} xp_n^2(x)d\mu(x),$$

ce qui correspond avec b_n en (1.26), et

$$a_{n-1,n} = \int_{\mathbb{R}} xp_n(x)p_{n-1}(x)d\mu(x),$$

ce qui corresponde avec a_n en (1.25).

$a_{n+1,n} = a_{n+1}$. En comparant les coefficients de x^{n+1} en (1.24) on trouve $\gamma_n = a_{n+1}\gamma_{n+1}$, cela donne $a_{n+1} = \frac{\gamma_n}{\gamma_{n+1}}$ et cette quantité est positive.

La relation de récurrence pour les polynômes orthogonaux moniques $P_n(x) = \frac{p_n(x)}{\gamma_n}$ est

$$P_{n+1} = (x - b_n)P_n(x) - a_n^2 P_{n-1}(x), \quad n \geqslant 0, \tag{1.27}$$

avec les valeurs initiales $P_{-1} = 0$ et $P_0 = 1$.

1.8 Zéros des polynômes orthogonaux

Théorème 1.3. *Supposons que la mesure μ est supportée sur l'intervalle [a,b] $\subset \mathbb{R}$. Les zéros du polynôme orthogonal p_n de degré n sont simples, réels et dans l'intervalle]a,b[.*

Démonstration.

Supposons que p_n change de signe m < n fois dans l'intervalle]a,b[aux points $a < x_1 < x_2 < ... < x_m < b$. Si on prend $q_m(x) = (x - x_1)(x - x_2)...(x - x_m)$, alors le polynôme $p_n q_m$ ne change pas de signe dans [a,b] et donc

$$\int_a^b p_n(x) q_m(x) d\mu(x) \neq 0.$$

Mais comme q_m est un polynôme de degré $m < n$, l'orthogonalité donne

$$\int_a^b p_n(x) q_m(x) d\mu(x) = 0.$$

Cette contradiction implique que $m \geqslant n$. Mais le nombre de changements de signe de p_n est $\leqslant n$, donc on trouve $m = n$. On conclut que p_n a n zéros simples dans [a,b].□

On peut aussi montrer que p_n a n zéros réels et simples à partir de la relation de récurrence (1.24). C'est une conséquence de l'entrelacement des zéros de p_n et p_{n+1}.

Théorème 1.4. *Si $x_1, x_2, ..., x_n$ sont les zéros de p_n et $y_1, y_2, ..., y_{n+1}$ sont les zéros de p_{n+1}, on a que les zéros $(y_k)_{1 \leqslant g \leqslant n+1}$ sont entrelacés par les zéros $(x_k)_{1 \leqslant k \leqslant n}$:*

$$y_1 < x_1 < y_2 < x_2 < y_3 < ... < y_n < x_n < y_{n+1}.$$

Démonstration.

Par induction : le zéro de p_1 est $b_0 \in \mathbb{R}$. Observez que (1.24) donne

$$a_2 p_2(b_0) + a_1 p_0(b_0) = 0,$$

et donc le signe de $p_2(b_0)$ est négatif. Ça veut dire que p_2 change de signe entre $-\infty$ et b_0 et entre b_0 et $+\infty$, donc p_2 a deux zéros réels y_1 et y_2 avec $y_1 < b_0 < y_2$.

Supposons que p_n et p_{n-1} ont des zéros entrelacés et $x_1, x_2, ..., x_n$ sont les zéros de p_n, alors le signe de $p_{n-1}(x_j)$ est $(-1)^{n-k}$. La relation de récurrence (1.24) donne $a_{n+1} p_{n+1}(x_k) + a_n p_{n-1}(x_k) = 0$, donc le signe de $p_{n+1}(-\infty)$ est $(-1)^{n+1}$. Donc p_{n+1} change de signe entre $-\infty$ et x_1, entre x_k et x_{k+1} pour $1 \leqslant k < n$ et entre x_n et $+\infty$, et ça donne l'entrelacement des zéros de p_{n+1} avec les zéros de p_n.

□

On écrit les zéros de p_n comme $\{x_{k,n} \mid 1 \leqslant k \leqslant n\}$, avec

$$a < x_{1,n} < x_{2,n} < ... < x_{n,n} < b.$$

Une autre méthode est un résultat de l'algèbre linéaire. Nous introduisons la **matrice de Jacobi**

$$J_n = \begin{pmatrix} b_0 & a_1 & 0 & 0 & \cdots & 0 \\ a_1 & b_1 & a_2 & 0 & \cdots & 0 \\ 0 & a_2 & b_2 & a_3 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-2} & b_{n-2} & a_{n-1} \\ 0 & \cdots & 0 & 0 & a_{n-1} & b_{n-1} \end{pmatrix} \tag{1.28}$$

avec les coefficients de la récurrence (1.24). Cette relation de récurrence donne

$$\begin{pmatrix} b_0 & a_1 & 0 & 0 & \cdots & 0 \\ a_1 & b_1 & a_2 & 0 & \cdots & 0 \\ 0 & a_2 & b_2 & a_3 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-2} & b_{n-2} & a_{n-1} \\ 0 & \cdots & 0 & 0 & a_{n-1} & b_{n-1} \end{pmatrix} \begin{pmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{pmatrix} = x \begin{pmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ \vdots \\ p_{n-2}(x) \\ p_{n-1}(x) \end{pmatrix} - a_n p_n(x) \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \tag{1.29}$$

Quand $x = x_{k,n}$ est un zéro de p_n on a donc

$$\begin{pmatrix} b_0 & a_1 & 0 & 0 & \cdots & 0 \\ a_1 & b_1 & a_2 & 0 & \cdots & 0 \\ 0 & a_2 & b_2 & a_3 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n-2} & b_{n-2} & a_{n-1} \\ 0 & \cdots & 0 & 0 & a_{n-1} & b_{n-1} \end{pmatrix} \begin{pmatrix} p_0(x_{k,n}) \\ p_1(x_{k,n}) \\ p_2(x_{k,n}) \\ \vdots \\ p_{n-2}(x_{k,n}) \\ p_{n-1}(x_{k,n}) \end{pmatrix} = x_{k,n} \begin{pmatrix} p_0(x_{k,n}) \\ p_1(x_{k,n}) \\ p_2(x_{k,n}) \\ \vdots \\ p_{n-2}(x_{k,n}) \\ p_{n-1}(x_{k,n}) \end{pmatrix}$$

Cela veut dire que $x_{k,n}$ est une valeur propre de J_n avec un vecteur propre

$$v_k = \begin{pmatrix} p_0(x_{k,n}) \\ p_1(x_{k,n}) \\ p_2(x_{k,n}) \\ \vdots \\ p_{n-2}(x_{k,n}) \\ p_{n-1}(x_{k,n}) \end{pmatrix}.$$

Comme J_n est une matrice symétrique, toutes les valeurs propres sont réels et donc tous les zéros de p_n sont réels.

1.9 Théorème spectral

Chaque suite de polynômes orthogonaux satisfait une relation de récurrence comme (1.24). Le résultat inverse est le théorème spectral pour des polynômes orthogonaux :

Théorème 1.5. *Supposons que la suite* $(p_n)_{n\in\mathbb{N}}$ *des polynômes est donnée par la récursion*

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_np_n(x) + a_np_{n-1}(x), \quad n \geqslant 0$$

avec $a_{n+1} > 0$ *et* $b_n \in \mathbb{R}$ *pour tous* $n \in \mathbb{N}$, *et avec des valeurs initiales* $p_0 = 1$ *et* $p_{-1} = 0$. *Alors il existe une mesure de probabilité* μ *sur* $\mathbb{R}$ *pour laquelle*

$$\int p_n(x)p_m(x)d\mu(x) = \delta_{n,m}.$$

Démonstration.

Nous savons déjà que les zéros de p_n sont les valeurs propres de J_n. On peut donc diagonaliser J_n en utilisant les valeurs propres et les vecteurs propres comme

$$V^tJ_nV = \begin{pmatrix} x_{1,n} & 0 & \cdots & 0 \\ 0 & x_{2,n} & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & x_{n,n} \end{pmatrix},$$

où

$$V = \begin{pmatrix} \frac{v_1}{\|v_1\|_2} & \frac{v_2}{\|v_2\|_2} & \cdots & \frac{v_n}{\|v_n\|_2} \end{pmatrix}$$

contient comme colonnes les vecteurs propres normalisés. La matrice V est orthogonale, c'est à dire $V^tV = I_n$. Mais alors on a aussi $VV^t = I_n$, et donc $(VV^t)_{i,j} = \delta_{i,j}$, ou

$$\sum_{k=1}^{n} \frac{1}{\|v_k\|_2^2} p_i(x_{k,n})p_j(x_{k,n}) = \delta_{i,j}, \qquad i,j \leqslant n-1, \tag{1.30}$$

Nous introduisons les mesures de probabilité discrètes

$$\mu_n = \sum_{k=1}^{n} \frac{\delta_{x_{k,n}}}{\|v_k\|_2^2},$$

avec des points de masse aux zéros

$$\{x_{k,n} \mid 1 \leqslant k \leqslant n\}.$$

La relation (1.30) devient

$$\int p_i(x)p_j(x)d\mu_n(x) = \delta_{i,j}, \qquad i,j \leqslant n-1.$$

La suite des mesures de probabilité $(\mu_n)_{n\in\mathbb{N}_0}$ contient une sous-suite $(\mu_{n_k})_{k\in\mathbb{N}}$ qui converge faiblement vers une mesure μ (principe de sélection de Helley). On peut montrer que la mesure μ est de nouveau une mesure de probabilité. Donc si on prend la limite lorsque $n \longrightarrow \infty$, on trouve

$$\int p_i(x)p_j(x)d\mu(x) = \delta_{i,j}, \qquad i, j \geqslant 0,$$

ce qui est une relation d'orthogonalité pour la suite de polynômes $(p_n)_{n\in\mathbb{N}}$. $\square$

Chapitre 2

Processus stochastique

Les références principale de ce chapitre est [5, 8]

Dans ce chapitre, nous introduisons les notions qui vont nous aider à comprendre la notion de processus de naissance et de mort.

2.1 Introduction aux problèmes stochastiques

On dit qu'on a affaire à un problème stochastique lorsque le hasard y intervient : c'est la difficulté principale de ces problèmes. Si on a une connaissance statistique du passé récent et si on estime que l'avenir proche ressemblera à ce passé, on peut songer à utiliser son information pour se prémunir contre des conséquences fâcheuses du hasard.

La première idée de la théorie des processus stochastiques revient sans doute à A. Einstein qui, en 1905, soutint sa thèse sur le mouvemententent brownien. Ensuite, A.Markov, vers 1910, en étudiant l'œuvre de Pouchkine intitulé " Eugène Onéguine ", qui décrivait de manière probabiliste alternance des consonnes et des voyelles sous la forme d'une chaîne qui porte son nom. Le Danois Erlang créa, vers 1914, la théorie des files d'attente (pour l'aider à résoudre des problèmes de dimensionnement de standards téléphoniques). C'est à A. Kolmogorov, que l'on doit la formalisation générale du processus stochastique (1933). Des mathématiciens comme D.G. Kendall, P.Lévy, A. Khintchine, W. Feller, J. Doob ont entrepris, dans l'intervalle ou par la suite, de développer encore cette théorie.

2.2 Définition d'un processus stochastique

Un processus stochastique (ou processus aléatoire) est une famille de variable aléatoires $X(t) : \{X(t), t \in T\}$ définies sur un même espace de probabilité, où t parcourt l'ensemble T qui représente le plus souvent un ensemble de temps. Si T est discret, on parle de suite stochastique, les processus stochastiques étant pour le cas de T continu.

Dans le cas des espaces d'états finis ou dénombrables, les variables $X(i)$, i $\in \mathbb{N}$ sont appelées discrètes .

2.3 Chaînes de Markov à espace d'état temps discret

C'est ainsi qu'on nomme une suite stochastique (le temps est donc discret) à espace d'états discret et vérifiant la propriété sans mémoire. Nous supposerons, en outre, le processus homogène. On a donc $T = t_0, t_1, t_2, \cdots, t_n$; le plus souvent on confondra T avec $\mathbb{N}$.

Par définition, une chaîne de Markov possède la propriété « sans mémoire », si

$P[X_n = j|X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1} = P[X_n = j|X_{n-1} = i_{n-1}] = P_{ij}^{(n)}$. Soit $P_{ij}^{(n)}$ la probabilité qu'une chaîne de Markov passe de l'état i à l'état j en n transitions ou étapes :

$P_{ij}^{(n)} = P(X_n = j|X_0 = i) = P(X_{n+1} = j|X_k = i), \quad (n \geq 1,\, k \geq 1)$, avec $P_{ij}^{(1)} = P_{ij}$. On note $P^{(n)}$ la matrice des probabilités de transition à n étapes : $P^{(n)} = (P_{ij}^{(n)})$. Alors pour tout $n \geq 1$, $P^{(n)} = P^n$. En effet,

$$\begin{aligned}
p_{ij}^{(n)} &= \mathbb{P}(X_n = j|X_0 = i) = \sum_{k\in S} P(X_n = j, X_{n-1} = k|X_0 = i)(n \geq 2 \\
&= \sum_{k\in S} P(X_n = j|X_{n-1} = k, X_0 = i).P(X_{n-1} = k|X_0 = i) \\
&= \sum_{k\in S} P(X_n = j|X_{n-1} = k).P(X_{n-1} = k|X_0 = i) \\
&= \sum_{k\in S} p_{ik}^{n-1} p_{kj}^{1},
\end{aligned}$$

ce qui signifie que $P^{(n)} = P^{(n-1)}P$ et par itération on obtient

$P^{(n)} = P^n$.

Plus généralement on a donc $P^m P^n = P^{m+n}$ et on obtient ainsi les équations de Chapman-Kolmogorov :

$$\textstyle\sum_{k\in S} p_{ik}^{m} p_{kj}^{n} = P^{m+n}.$$

Introduisons maintenant les probabilités d'état $\pi_k(n) = P(X_n = k), \quad k \in S$. La distribution de X_n peut être écrite sous forme vectorielle avec $\pi(n) = (\pi_1(n), \pi_2(n), \cdots)$ dont la somme des termes vaut 1.

D'après le théorème des probabilités totales on a alors

$$\pi_k(n) = \textstyle\sum_{i\in s} \pi_i(0) P_{ik}^{(n)}.$$

Matriciellement cette relation s'écrit $\pi(n) = \pi(0)P^{(n)}$, $\forall n \geq 0$. On a de même $\pi(n+1) = \pi(n)P$.

Ces résultats permettent d'affirmer qu'une chaîne de Markov est complètement définie si l'on connaît sa matrice des probabilités de transition ainsi que la distribution de X_0.

2.4 Processus de Markov à temps discret

On parle de processus de Markov, les processus stochastiques à temps continu et à espace d'états discret .

C'est un processus de Markov à temps continu et à espace d'états discret s'il vérifie les trois propriétés suivantes :

• le temps t est continu. Le plus souvent le temps varie de 0 à l'infini , alors : $t \in T = \mathbb{R}^+$

•$X(t)$ appartient à un espace d'état qui est discret,

• Le processus stochastique $X(t)$ vérifie la propriété sans mémoire (ou propriété de Markov),

• Nous supposerons en outre dans la suite que les processus de Markov sont homogènes (de même que nous l'avons fait pour les chaînes de Markov).

2.5 Processus de Markov particulier

2.5.1 Processus de naissance

Utilisés dès 1940 par Lundberg pour des statistiques d'accidents, puis vers 1943 par Arley à propos du rayonnement cosmique, ils ont été employés vers 1950 par Kendall, Bartlett et Feller en biologie. Ils sont utilisés en recherche opérationnelle, par exemple pour représenter des arrivées aléatoires des clients dans des files d'attente ou encore des occurrences de pannes sous-équipements.

Définition 2.1. *Un processus de Markov homogène $X(t)$ est un processus de naissance si les probabilités de transition entre t et $t + \Delta t$, ont pour expression :*

$$\mathbb{P}(t+\Delta t) = \begin{cases} \lambda_i \Delta t + \circ(\Delta t) \\ 1 - \lambda_i \Delta t + \circ(\Delta t), \end{cases} \tag{2.1}$$

où $\lambda_{i,i+1} = \lambda_i$ est le taux de naissance à partir de l'état E_i.
Les probabilités de transition de i vers $i+2, i+3, \ldots$, sont négligeables.
Celles de i vers $i-1$, $i-2, \ldots$ sont nulles.

2.5.2 Processus de Poisson

Par définition, un processus de Poisson est un processus de naissance ouvert pour lequel le taux de naissance ne dépend pas de la taille de la population :

$\lambda = \lambda_i, \forall i = 0, 1, \cdots$.

Supposons qu'à $t = 0$, la taille de la population soit nulle : $X_0 = 0$ et exprimons les probabilités des états à $t + \Delta t$ en fonction de celles-ci à t, et du taux de naissance λ.

$$\begin{cases} \pi_0(t+\Delta t) = \pi_0(t)(1-\lambda\Delta t) + \circ(\Delta t) \\ \pi_n(t+\Delta t) = \pi_{n-1}(t)(\lambda\Delta t) + \pi_n(t)(1-\lambda\Delta t) + \circ(\Delta t). \end{cases} \tag{2.2}$$

La première relation de l'équation (2.2) montre que $\pi_0'(t) = -\lambda\pi_0(t)$, d'où $\pi_0(t) = ke^{-\lambda t}$. Puisque $\pi_0(0) = 1$, il vient que $k = 1$ et $\pi_0(t) = e^{-\lambda t}$.

La deuxième relation de (2.2) montre que :

• Pour n=1,

$\pi_1'(t) = \lambda\pi_0(t) - \lambda\pi_1(t)$, soit $\pi_1'(t) + \lambda\pi_1(t) = \lambda e^{-\lambda t}$ qui a pour solution

$$\pi_1(t) = \lambda t e^{-\lambda t}.$$

•Pour n=2, $\pi_2'(t) = \lambda\pi_1(t) - \lambda\pi_2(t)$, soit $\pi_2'(t) + \lambda\pi_2(t) = \lambda e^{-\lambda t}$

qui a pour solution

$$\pi_2(t) = \frac{(\lambda t)^2}{2!}e^{-\lambda t}.$$

• Pour n=3, $\pi_3'(t) = \lambda\pi_2(t) - \lambda\pi_3(t)$, soit $\pi_3'(t) + \lambda\pi_3(t) = \lambda e^{-\lambda t}$,

qui a pour solution

$$\pi_3(t) = \frac{(\lambda t)^3}{3!}e^{-\lambda t}.$$

On fait alors hypothèse de récurrence :

$$\pi_n(t) = \tfrac{(\lambda t)^n}{n!}e^{-\lambda t}$$

et l'on montre que $\pi_{n+1}(t)$ a la forme analogue :

$$\pi_{n+1}(t) = \tfrac{(\lambda t)^{n+1}}{n+1!}e^{-\lambda t}.$$

Avec les probabilités des états d'un processus de Poisson, on peut calculer l'espérance du nombre de naissances entre 0 et t puisque le processus est homogène :

$$
\begin{aligned}
E(X_t) &= \sum_{n=0}^{\infty} X(t)\pi_n(t) \\
&= \sum_{n=0}^{\infty} n\frac{(\lambda t)^n}{n!}e^{-\lambda t} \\
&= e^{-\lambda t}\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{(n-1)!} \\
&= e^{-\lambda t}\sum_{n=1}^{\infty} \lambda t\frac{(\lambda t)^{n-1}}{(n-1)!} \\
&= e^{-\lambda t}\lambda t\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\lambda t e^{\lambda t} \\
&= \lambda t.
\end{aligned}
\tag{2.3}
$$

Nous dirons que le « nombre moyen » de naissances sur un intervalle de temps amplitude t est λt.

$P_n(t) =$ probabilité que l'état du système à l'époque t soit égal à n, $P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$. On interprète alors comme la probabilité que, à un instant quelconque dans le long terme, exactement n clients soient présents dans le système, $P_n(t) = \lim_{t\longrightarrow\infty} P_n(t)$.

On appellera processus de Poisson unidimensionnel homogène (en partant de $\mathbb{N} = 0$) le processus qui compte le nombre des naissances pendant [0, t]

2.5.3 La loi exponentielle

Soit un processus de Poisson de taux μ . Considérons un intervalle de temps T séparant deux événements (naissances) consécutifs.

En remarquant que T sera supérieur à t si et seulement si aucune naissance n'a lieu entre 0 et t, il vient :

$$P[T > t] = \pi_0(t) = P[0 \quad naissance \quad sur \quad l'amplitude \quad de \quad temps \quad t] = e^{-\mu t},$$

la variable aléatoire T est donc régie par une loi exponentielle.

La densité de probabilité de T notée f(t), vaut :

$$\begin{aligned} f(t)dt &= P[t < T \leq T + dt] \\ &= d(P[T \leq t]) = d(1 - e^{-\mu t}dt \\ &= \mu e^{-\mu t}. \end{aligned} \tag{2.4}$$

La durée moyenne entre deux naissance consécutives est

$$\begin{aligned} E(T) &= \int_0^{\infty} tf(t)dt \\ &= t\mu e^{-\mu t}dt \\ &= \frac{1}{\mu}. \end{aligned} \tag{2.5}$$

Cherchons la probabilité d'occurrence d'un événement entre t et $t + \varDelta t$, sachant qu'il y a eu un événement à $t = 0$, puis aucun événement entre 0 et t :

$$\begin{aligned} P[t < T \leq T + \varDelta t] \,|\, T > t &= \frac{P[t < T \leq T + \varDelta t, T > t]}{P[T > t]} \\ &= \frac{\mu e^{-\mu t}\varDelta t}{e^{-\mu t}} \\ \\ &= \mu \varDelta t. \end{aligned} \tag{2.6}$$

2.5.4 Processus de mort

Soit le processus stochastique $X(t)$, où $X(t)$ s'interprète comme la taille d'une population à l'instant t, initialement non nulle : $X(0) = M$.

La probabilité de disparition d'un individu (on dit aussi mort, ou départ) entre t et $t+\varDelta t$), lorsque la popution comporte i individu est $\mu_i \varDelta t + \circ(\varDelta t)$. Il ne se produit pas de naissance.

Un processus de Markov homogène $X(t)$ est un processus de mort si les probabilités de transition entre t et $t + \varDelta t$, ont pour expression :

$$\mathbb{P}(t + \varDelta t) = \begin{cases} 1 - \mu_i \varDelta t + \circ(\varDelta t) \\ \mu_i \varDelta t + \circ(\varDelta t), \end{cases} \tag{2.7}$$

où $\mu_{i-1} = \mu_i$ est le taux de mort à partir de l'état E_i.
Les probabilités de transition de i vers $i - 2$, $i - 3, \ldots$ sont négligeables.
Celles de i vers $i + 1$, $i + 2, \ldots$ sont nulles.

2.6 Processus de naissance et de mort

Nous passons maintenant à ce processus, très employé pour modéliser notamment des phénomènes d'attente ou encore des système sujets à des pannes réparables.

Le processus de naissance et de mort est obtenu en superposant un processus de naissance et un processus de mort.

Définition 2.2. *On appelle processus de naissance et mort une chaîne de Markov homogène à temps continu $X(t)_{t\geqslant 0}$, à valeurs dans $\mathbb{N}$ telle que :*

$$\mathbb{P}_n(t+\Delta t) = \begin{cases} \lambda_i \Delta t + \circ(\Delta t), & une \quad naissance \quad entre \quad t+\Delta t, \\ 1 - r_i \Delta t + \circ(\Delta t), & ni \quad naissance, \quad ni \quad mort \quad entre \quad t+\Delta t, \\ \mu_i \Delta t + \circ(\Delta t), & une \quad mort \quad entre \quad t+\Delta t. \end{cases} \tag{2.8}$$

Voici le générateur infinitésimal associée, noté A :

$$\begin{pmatrix} -r_0 & \lambda_0 & 0 & 0 & 0 & \ldots 0 \\ \mu_1 & -r_1 & \lambda_1 & 0 & 0 & \ldots 0 \\ 0 & \mu_2 & -r_2 & \lambda_2 & 0 & \ldots 0 \\ \vdots & \ldots & \ddots & \ddots & \ddots & \ddots \\ \end{pmatrix}. \tag{2.9}$$

Si tous les λ_i sont nuls, on parle de processus de mort.
Si tous les μ_i sont nuls, on parle de processus de naissance.

Chapitre 3

Polynômes orthogonaux et processus de naissance et de mort

Les références principales de ce chapitre sont [5, 8, 15, 16]

Considérons le processus aléatoire $\{X(t), t \in [0, \infty)\}$, où $X(t)$ est le nombre d'éléments dans une population. Pour chaque $t > 0$ la variable $X(t)$ est une variable aléatoire avec des valeurs en $\mathbb{N}$. On veut savoir la distribution $P_n(t) = \mathbb{P}(X(t) = n)$ quand on connaît la population initiale $X(0)$. Le dynamique de cette processus est que les probabilités conditionnelles satisfont

- Pour $t > s$ la quantité $P_{m,n}(t-s) = \mathbb{P}(X(t) = n|\ X(s) = m)$ dépend seulement de la différence $t - s$: le processus est stationnaire ;

- La probabilité conditionnelle $\mathbb{P}(X(t) = n|\ X(u), u \in [0, s])$ dépend seulement de l'histoire la plus récente et donc est égale à $\mathbb{P}\Big(X(t) = n|\ X(s)\Big)$: propriété de Markov.

Les probabilités de transition sont

$$P_{m,n}(t) = \mathbb{P}(X(t) = n|\ X(0) = m),$$

et la matrice $\mathbb{P}(t) = \Big(P_{i,j}(t)\Big)_{i,j\in\mathbb{N}}$ est la matrice de transition. On a toujours par la loi d'espérance totale

$P_n(t) = \sum_{m=0}^{\infty} \mathbb{P}(X(t) = n|\ X(0) = m)\mathbb{P}(X(0) = m) = \sum_{m=0}^{\infty} P_{m,n}(t)P_m(0) = [P(0)\mathbb{P}(t)]_n$
,

où $P(0) = (P_0(0), P_1(0), P_2(0), \cdots)$ est le vecteur contenant les probabilités de la population initiale.

Pour les processus de naissance et de mort, les probabilités de transitions en plus satisfont

$$\mathbb{P}(X(t+\Delta t)=j|\ X(t)=i)=\begin{cases}\lambda_i\Delta t+\circ(\Delta t), & j=i+1,\\ \mu_i\Delta t+\circ(\Delta t), & j=i-1,\\ 1-r_i\Delta t+\circ(\Delta t), & j=i,\\ \circ(\Delta t), & |j-i|>1,\end{cases} \tag{3.1}$$

avec $\lambda_i, \mu_i > 0$ pour $i \geq 0$, $\mu_0 \geq 0$, et $r_i = \lambda_i + \mu_i$ et

$\lim_{\Delta t \longrightarrow 0} \frac{\circ(\Delta t)}{\Delta t} = 0.$

Les λ_i sont les coefficients de naissance et les μ sont les coefficients de mort.

Nous savons que

$\mathbb{P}'(t) = \lim_{\Delta t \longrightarrow 0} \frac{\mathbb{P}(t+\Delta t)-\mathbb{P}(t)}{\Delta t}.$

Comme $\mathbb{P}(s+t) = P(s)P(t)$ (dynamique de naissance et de mort),

on a :

$$\begin{aligned}\mathbb{P}'(t) &= \lim_{\Delta t\longrightarrow 0}\frac{\mathbb{P}(t)\mathbb{P}(\Delta t)-\mathbb{P}(t)}{\Delta t}\\ &= \mathbb{P}(t)\lim_{\Delta t\longrightarrow 0}\frac{\mathbb{P}(\Delta t)-I}{\Delta t} \quad \Big(I=\mathbb{P}(0)\Big)\\ &= \mathbb{P}(t)A,\end{aligned}$$

où $\qquad A = \lim_{\Delta t \longrightarrow 0} \frac{\mathbb{P}(\Delta t)-I}{\Delta t},$

est la génératrice infinitésimale.

$$\Big[\frac{\mathbb{P}(\Delta t)-I}{\Delta t}\Big]_{i,j}=\begin{cases}\lambda_i+\frac{\circ(\Delta t)}{\Delta t}, & j=i+1,\\ \mu_i+\frac{\circ(\Delta t)}{\Delta t}, & j=i-1,\\ -r_i, & j=i,\\ \frac{\circ(\Delta t)}{\Delta t}, & |j-i|>1,\end{cases}$$

$$A = \mathbb{P}'(0) = \begin{cases} \lambda_i, & j = i+1, \\ \mu_i, & j = i-1, \\ -r_i, & j = i, \\ 0, & |j-i| > 1, \end{cases} \tag{3.2}$$

ce qui donne une matrice tridiagonale

$$A = \begin{pmatrix} -r_0 & \lambda_0 & 0 & 0 & 0 & \ldots 0 \\ \mu_1 & -r_1 & \lambda_1 & 0 & 0 & \ldots 0 \\ 0 & \mu_2 & -r_2 & \lambda_2 & 0 & \ldots 0 \\ 0 & 0 & \mu_3 & -r_3 & \lambda_3 & \ldots 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{3.3}$$

L'équation différentielle d'ordre 1 (pour des matrices infinies)

$$\mathbb{P}'(t) = \mathbb{P}(t)A,$$

avec la condition initiale $\mathbb{P}(0) = I$ a comme solution

$$P(t) = e^{At}, \tag{3.4}$$

où l'exponentielle de la matrice At est définie par

$e^{At} = \sum_{k=0}^{\infty} A^k \frac{t^k}{k!}$.

La formule (3.4) donne

$P_{n,m}(t) = \left(e^{At}\right)_{n,m}$,

mais ce n'est pas tout à fait simple parce qu'on doit calculer l'exponentielle de la matrice At.

Karlin et McGregor (1958) ont proposé d'introduire une suite de polynôme $(Q_n)_{n\in\mathbb{N}}$ qui satisfont la relation de récurrence

$$-xQ_n(x) = \lambda_n Q_{n+1}(x) - r_n Q_n(x) + \mu_n Q_{n-1}(x), \quad n \geqslant 0 \tag{3.5}$$

avec $Q_{-1} = 0$ et $Q_0 = 1$.

Ceci est dû à une méthode directe :

$AQ = -xQ$,

où $\quad Q = \begin{pmatrix} Q_0(x) \\ Q_1(x) \\ \vdots \\ Q_n(x) \end{pmatrix}.$

Ainsi,

$$A^2 Q = x^2 Q,$$
$$\vdots$$
$$A^k Q = (-x)^k Q, \quad k \geq 0$$

$e^{At}Q = \sum_{k=0}^{\infty} t^k \frac{A^k}{k!} Q = \sum_{k=0}^{\infty} \frac{(-xt)^k}{k!} Q.$

Ça veut dire que

$$\mathbb{P}(t)Q(x)^T Q(x) = e^{-xt} Q(x)^T Q(x). \tag{3.6}$$

Le polynôme Q_n n'est pas monique, et si on appelle le coefficient supérieur k_n, on a $Q_n(x) = k_n x^n + \cdots$, et la récurrence (3.5) donne $-k_n = \lambda_n k_{n+1}$, donc

$\frac{k_n}{k_{n+1}} = -\lambda_n, \quad k_0 = 1.$

Les polynômes moniques sont $P_n(x) = \frac{Q_n(x)}{k_n}$ et ils satisfont la relation

$$P_{n+1}(x) = (x - r_n)P_n(x) - \mu_n \lambda_{n-1} P_{n-1}(x), \quad n \geqslant 0 \tag{3.7}$$

Preuve :

$-xQ_n(x) = \lambda_n Q_{n+1}(x) - r_n Q_n(x) + \mu_n Q_{n-1}(x)$

$\Leftrightarrow -\lambda_n Q_{n+1}(x) = xQ_n(x) - r_n Q_n(x) + \mu_n Q_{n-1}(x)$

$\Leftrightarrow -\lambda_n Q_{n+1}(x) = (x - r_n)Q_n(x) + \mu_n Q_{n-1}(x)$

$\frac{k_n}{k_{n+1}} Q_{n+1}(x) = (x - r_n)Q_n(x) + \mu_n Q_{n-1}(x)$

$\Leftrightarrow \frac{Q_{n+1}(x)}{k_{n+1}} = (x - r_n)\frac{Q_n(x)}{k_n} + \mu_n \frac{Q_{n-1}(x)}{k_n}$

$\Leftrightarrow \frac{Q_{n+1}(x)}{k_{n+1}} = (x - r_n)\frac{Q_n(x)}{k_n} - \mu_n \lambda_{n-1} \frac{Q_{n-1}(x)}{k_{n-1}}$

$P_{n+1} = (x - r_n)P_n(x) - \mu_n \lambda_{n-1} P_{n-1}(x)$

Le théorème spectral des polynômes orthogonaux donne qu'il y a une mesure positive μ sur $\mathbb{R}$ par rapport à laquelle les polynômes $(P_n)_{n\in\mathbb{N}}$ sont des polynômes orthogonaux moniques, et ainsi les polynômes $(Q_n)_{n\in\mathbb{N}}$ sont aussi des polynômes orthogonaux. La condition $r_i = \lambda_i + \mu_i$ implique que toutes les valeurs propres des matrices principales de A sont négatives, donc tous les zéros de Q_n(et P_n) sont positifs, et ça veut dire que la mesure μ est supportée sur $[0, \infty)$. Alors on a

$\int_0^{\infty} Q_n(x)Q_m(x)d\mu(x) = 0, \quad m \neq n.$

En utilisant la mesure μ, on peut prendre maintenant l'intégrale en (3.6) et on trouve

$\mathbb{P}(t) \int_0^{\infty} Q(x)^T Q(x) d\mu(x) = \int_0^{\infty} e^{-xt} Q(x)^T Q(x) d\mu(x).$

L'orthogonalité donne

$$\int_0^\infty Q(x)^T Q(x) d\mu(x) = \begin{pmatrix} \pi_0 & 0 & 0 & 0 & \cdots \\ 0 & \pi_1 & 0 & 0 & \cdots \\ 0 & 0 & \pi_2 & 0 & \cdots \\ 0 & 0 & 0 & \ddots & \cdots \end{pmatrix}, \tag{3.8}$$

$\pi_n = \int_0^\infty Q_n^2(x) d\mu(x)$ et alors l'élément sur la ligne n et la colonne m donne

$$P_{n,m}(t) = \frac{1}{\pi_m} \int_0^\infty e^{-xt} Q_n(x) Q_m(x) d\mu(x). \tag{3.9}$$

On appelle (3.9) la représentation de Karlin-McGregor pour les probabilités de transition d'un processus de naissance et de mort. Nous avons donné une dérivation assez formelle de cette représentation : les matrice que nous avons introduite ne sont pas toujours des matrices bornées et on doit être prudent avec les multiplications et les intégrations, qui ne sont toujours valables pour les matrices non-bornées. En effet, on a besoin du théorème spectral pour des opérateurs sur l'espace de Hilbert $l^2(\mathbb{N})$. En plus, le théorème spectral des polynômes orthogonaux stipule qu'il existe une mesure positive sur $\mathbb{R}$ par rapport à laquelle les polynômes $(Q_n)_{n\in\mathbb{N}}$ sont orthogonaux, mais il est possible (pour des matrices A non-bornées) que cette mesure n'est pas unique .

La représentation de Karlin-McGregor donne une séparation des variables n, m et t, qu'on peut utiliser pour l'analyse asymptotique quand $t \longrightarrow \infty$.

Si on suppose que la mesure spectral μ est composée par un composante absolument continue et une composante discrète

$d\mu(x) = w(x)dx + \sum_{k=0}^\infty w_k \delta_{x_k}, \quad x_0 = 0,$

où δ_{x_k} est la mesure de Dirac avec une masse 1 au point x_k, alors on a de (3.9)

$\lim_{t\longrightarrow\infty} P_{n,m}(t) = \frac{w_0}{\pi_m} Q_n(0) Q_m(0), \quad w_0 = \mu(\{0\}).$

Mais de (3.5) on voit que $Q_n(0) = 1$ pour tout n$\in \mathbb{N}$, et ainsi

$\lim_{t\longrightarrow\infty} P_{n,m}(t) = \frac{w_0}{\pi_m}, \quad w_0 = \mu(\{0\}).$

Ça veut dire que le processus $\{X(t)\}|\ t > 0$ a une distribution équilibré (ou stationnaire) quand $t \longrightarrow \infty$,

$p_m = \frac{w_0}{\pi_m} = \lim_{t\longrightarrow\infty} \mathbb{P}(X(t) = m|\ X(0) = n)$, qui ne dépend pas de la condition initiale $X(0) = n$.

Conclusion

Dans notre travail, nous avons rappelé au premier chapitre les notions de base qui nous conduisent aux généralités sur les polynômes orthogonaux. Les principaux concepts développés sont l'équation différentielle de type hypergéométrique, l'orthogonalité de Gram-schmidt, la relation de récurrence et les zéros des polynômes orthogonaux.

Au second chapitre, nous avons rappelé la notion des processus stochastiques, les chaînes et les processus de Markov à temps discret et la notion des processus de naissance et de mort.

Le dernier chapitre concerne des processus de naissance et de mort et les polynômes orthogonaux où nous avons rappelé la notion de processus stationnaire et la propriété de Markov.

Sûrement, notre travail va inciter les chercheurs à faire des recherches sur les applications du processus de naissance et de mort aux polynômes orthogonaux multiples.

Bibliographie

[1] Arnold F.Nikiforov, Vasilii B. Urarov, Special Function of Mathematical Physics, Birkhäuser, Basel-Boston, 1988.

[2] G. Bangerezako, Discrete Darboux Transformations for Discrete polynomials of hypergeometric type, arXiv puprint, 1998.

[3] G. Bangerezako, The factorization method for the Askey-Wilson polynomials, Journal of Computational and Applied Mathematics, 107(1999), 219-232.

[4] Cristna GREAVU, Luc VINET Les polynômes orthogonaux matricielles et la méthode de factorisation, Montréal, 2014.

[5] R. Faure, B. Lemaire, C. Picouleau, Précis de Recherche Opérationnelle, 7^e édition, Dunod Paris, 2014.

[6] Gabor Szegö, Orthogonal Polynomial, Amer.Math.Soc. Colloquium Publications23, 1939, Fourth Edition 1975.

[7] Hans Sagan, Boundary and Eigenvalue Problems in Mathematical Physics, John Wiley, Sons, New York-London-Sydney, 1961 ; Dover Publications, 1989.

[8] E. C inlar, Introduction to stochastic process, Prentice Hall, 1975.

[9] Koekoek, P. Lesky, R.F.Swarttouw, Hypergeometric Orthogonal Polynomials and Their q-analogues, Spring-Verlag, Berlin, 2010 .

[10] Mourad E.H.Ismail, Classical and quantum orthogonal polynomials in One Variable, Encyclopedia of Mathematics and its Aplications98, Cambrige University Press, 2005.

[11] M.L.Mehta, Random Matrices, revised and enlarged second edition, academic Press, San Diego, 1991.

[12] Nico M. Temme, Special Functions : An Introduction to the Classical Functions of Mathematical Physics, John Wiley, Sons, New York, 1996.

[13] F. Ndayiragije, W.Van Assche Asumptotics for the ration and the zeros of multiple Charlier polynomials, J.Appox. Theory 164(2012), 823-840.

[14] F. Ndayiragije, W.Van Assche, Multiple Meixner polynomials and non- Hermitian oscillator Hamiltonians, J. Phys. A : Math. Theor.46(2013) 505201(17 pp).

[15] NIST Digital Library of Mathematical Functions, hhtps :/dlmf.nist.gov/

[16] Provenzi, Edoardo : Introduction à l'analyse fonctionnelle et ses applications, 2021.