

Fontainebleau

N-449

NOTE GEOSTATISTIQUE N° 133

MODELES ISOFACTORIELS DISCRETS

ET MODELE DE WALSH

=====

G. MATHERON

Octobre 1975

MODELES ISOFACTORIELS DISCRETS, et MODELE DE WALSH

NOTE GEOSTATISTIQUE N° 133

Dans une note antérieure, j'ai indiqué comment on pouvait obtenir des modèles à facteurs polynomiaux et à lois marginales continues (Gauss, gamma ou beta) à partir de processus de Markov dont le générateur infinitésimal possède un caractère local. Ici, je me propose de faire quelquechose d'analogue dans le cas discret : le caractère local du générateur A est ici remplacé par une condition exprimant que les seules transitions possibles sont $i \rightarrow i+1$ et $i \rightarrow i-1$, soit :

$$(1) \quad (Af)_i = - (a_i + b_i) f_i + a_i f_{i+1} + b_i f_{i-1}$$

Suivant les valeurs de a_i et b_i , on distingue trois types de processus possibles :

~ Si a_i et b_i sont strictement positifs pour $i = 0, \pm 1, \pm 2, \dots$, i varie de $-\infty$ à $+\infty$. Je n'examinerai pas ce cas.

~ Si $b_0 = 0$ et $b_i > 0$ pour $i = 1, 2, \dots$, $a_i > 0$ pour $i = 0, 1, \dots$, i varie de 0 à $+\infty$ (cas infini).

~ Si $b_0 = 0$, $a_N = 0$, $b_i > 0$ pour $i = 1, \dots, N$ et $a_i > 0$ pour $i = 0, \dots, N-1$, i varie de 0 à N (cas fini).

Il faut ensuite que ce processus soit ergodique, donc qu'il existe une probabilité $W = \{W_i\}$ telle que :

$$- (a_i + b_i) W_i + a_{i-1} W_{i-1} + b_{i+1} W_{i+1} = 0$$

En écrivant cette relation sous la forme :

$$b_{i+1} W_{i+1} - a_i W_i = b_i W_i - a_{i-1} W_{i-1}$$

ceci implique

$$(2) \quad b_{i+1} W_{i+1} = a_i W_i$$

Si a_i et b_i sont strictement positifs, (2) reste vrai pourvu que, par exemple, $a_i W_i$ et $b_i W_i$ tendent vers 0 si i tend vers $-\infty$. Dans les deux cas qui nous intéressent (cas infini et cas fini), on a $b_0 = 0$ et (2) est vraie sans hypothèse supplémentaire. Donc, la loi limite, si elle existe, est définie par :

$$(3) \quad W_n = \frac{a_0 a_1 \dots a_{n-1}}{b_1 b_2 \dots b_n} W_0$$

et la condition d'existence est $\sum W_n < \infty$.

La condition (2) exprime aussi que le processus est réversible ($W_i P_{ij}(t) = W_j P_{ji}(t)$).

Condition pour que les facteurs soient polynomiaux.

Le processus étant réversible, il est facile de voir que l'on aura des facteurs polynomiaux si et seulement si :

(i) les polynômes sont dans $L^2(\mathbb{R}, W)$, soit $\sum W_i i^n < \infty$.

(ii) pour chaque n , $A i^n$ est un polynôme en i de degré n

($n = 0, 1, \dots$ dans le cas infini, $n = 0, 1, \dots, N$ dans le cas fini).

Considérons d'abord la condition (ii), soit

$$A(i)^n \equiv a_i[(i+1)^n - i^n] - b_i[i^n - (i-1)^n] = \text{polynome en } i \text{ de degré } n$$

Pour $n = 1$ et $n = 2$, cette condition donne :

$$a_i - b_i = \text{polynome de degré } 1$$

$$a_i + b_i = \text{polynome de degré } 2$$

Donc :

$$(4) \quad \begin{cases} a_i = a_0 + \alpha i + \gamma i^2 \\ b_i = b_0 + \beta i + \gamma i^2 \end{cases}$$

avec d'ailleurs $b_0 = 0$ (puisque nous n'examinons pas le cas où i varie de $-\infty$ à $+\infty$). Il est facile de voir que si a_i et b_i sont de cette forme, la condition (ii) est effectivement remplie pour tout n .

Il reste à examiner (dans le cas infini) si la condition (i) est également satisfaite. La classification ci-dessous montre dans quels cas il en est bien ainsi (dans le cas non retenu où a_i et b_i sont > 0 et où i varie de $-\infty$ à $+\infty$, on trouverait $E(i^n) = \infty$ pour n assez grand, de sorte qu'il n'y a pas de modèles à facteurs polynomiaux : c' est la raison pour laquelle je n'ai pas retenu ce cas).

Modèles Linéaires ($\gamma = 0$)

$$a_i = a_0 + \alpha i$$

$$b_i = \beta i$$

a_0 doit être strictement > 0 (sinon, la chaîne n'est plus irréductible) et de même $\beta > 0$. Il y aura donc 3 cas à envisager : $\alpha > 0$,

$\alpha = 0$ et $\alpha < 0$.

1 - Cas $\alpha > 0$ (lois binomiales négatives).

La relation (3) donne :

$$W_n = \left(\frac{\alpha}{\beta}\right)^n \cdot \frac{a_0}{\alpha} \left(1 + \frac{a_0}{\alpha}\right) \dots \left(\frac{a_0}{\alpha} + n-1\right) \frac{W_0}{n!}$$

Posons :

$$p = \frac{\alpha}{\beta} \quad ; \quad v = \frac{a_0}{\alpha}$$

Alors $\sum W_n s^n = (1 - p.s.)^{-v} W_0$ ($s < 1/p$)

Donc, la loi limite W existe si et seulement si $p < 1$, soit

$$\alpha < \beta$$

et admet la fonction génératrice

$$G(s) = \frac{q^v}{(1 - p.s.)^v}$$

avec :

$$p = \frac{\alpha}{\beta} \quad , \quad q = 1-p \quad , \quad v = \frac{a_0}{\alpha}$$

Comme G admet des dérivées de tous les ordres en $s = 1$, les polynomes sont dans $L^2(R, W)$. Sous la seule condition $\alpha < \beta$, il existe donc bien un modèle à facteurs polynomiaux admettant les lois marginales (binomiales négatives)

$$W_n = q^v \frac{\Gamma(v+n)}{\Gamma(v)} \frac{p^n}{n!}$$

Il reste à déterminer explicitement les facteurs polynomiaux.

Cherchons d'abord les valeurs propres λ_n de l'opérateur (1) :

pour le monome i^n on trouve

$$\begin{aligned} A(i^n) &= (a_0 + \alpha i) ((i+1)^n - i^n) - \beta i(i^n - (i-1)^n) \\ &= n(\alpha - \beta) i^n + \text{polynome de degré } < n \end{aligned}$$

Par conséquent :

$$\lambda_n = -n(\beta - \alpha) = -n \beta q$$

Donc, si $\chi_n(i)$ est le facteur normalisé d'ordre n , la loi $F_{ij}(t)$ à 2 variables i et j sera

$$(5) \quad \begin{cases} F_{ij}(t) = W_i W_j \sum_{n=0}^{\infty} \rho^n \chi_n(i) \chi_n(j) \\ \rho = e^{-\beta q t} \end{cases}$$

D'autre part, on a :

$$F_{ij}(t) = W_i P_{ij}(t)$$

où $P_{ij}(t)$ est la matrice de transition, solution de la seconde équation de Kolmogorov.

$$(K_2) \quad P'_{ij}(t) = - (a_j + b_j) P_{ij} + a_{j-1} P_{i,j+1} + b_{j+1} P_{i,j-1}$$

Posons $G_i(s, t) = \sum_j P_{ij}(t) s^j$. (K_2) entraîne

$$\frac{\partial G}{\partial t} + (1-s)(\alpha s - \beta) \frac{\partial G}{\partial s} = -a_0(1-s)$$

avec $G_i(s, 0) = s^i$. Ceci s'intègre :

$$(6) \quad G_i(s, t) = \left(\frac{1 - p s - \rho(1-s)}{1 - p s - p \rho(1-s)} \right)^i \left(\frac{q}{1 - p s - p \rho(1-s)} \right)^v$$

avec comme ci-dessus $\rho = e^{-\beta q t}$ et $p = \alpha/\beta$. De (6) nous pouvons déduire la fonction génératrice à 2 variables s et σ (t est sous-entendu)

$$G(s, \sigma) = \sum_{i,j} F_{ij} s^j \sigma^i = \sum W_i \sigma^i G_i(s, t)$$

$$= \left(\frac{q}{1 - p s - p \rho(1-s)} \right)^v \left(\frac{q}{1 - p \sigma \frac{1 - ps - \rho(1-s)}{1 - ps - p \rho(1-s)}} \right)^v$$

soit :

$$(7) \quad G(s, \sigma) = \left(\frac{q^2}{(1-p s)(1-p \sigma) - \rho p(1-s)(1-\sigma)} \right)^v$$

Pour $\sigma = 1$, on trouve évidemment $G(s, 1) = G(s) = (q/1-p s)^v$,
et pour $\rho = 1$, $(q/1 - p s \sigma)^v$. Cette fonction génératrice se développe en série entière en ρ :

$$G(s, \sigma) = \left(\frac{q^2}{(1-p s)(1-p \sigma)} \right)^v \sum_{n=0}^{\infty} \frac{\Gamma(v+n)}{\Gamma(v)} \frac{p^n \rho^n}{n!} \frac{(1-s)^n (1-\sigma)^n}{(1-p s)^n (1-p \sigma)^n}$$

Mais d'après (5) on a aussi :

$$G(s, \sigma) = \sum \rho^n \sum_i W_i \chi_n(i) s^i \sum_j W_j \chi_n(j) \sigma^j$$

Comme $\rho = \exp(-\beta q t)$ varie de 0 à $+\infty$, nous pouvons identifier terme à terme ces deux développements, et cela nous donne la fonction génératrice des $\chi_n(i) W_i$:

$$\sum_{i=0}^{\infty} \chi_n(i) W_i s^i = q^v \sqrt{\frac{\Gamma(n+v)}{\Gamma(v)}} \frac{p^n}{n!} \frac{(1-s)^n}{(1-p s)^{n+v}}$$

On en tire l'expression du polynome χ_n qui est - sauf erreur de calcul de ma part :

$$(8) \quad \chi_n(i) = \left(\frac{\Gamma(v)}{\Gamma(n+v)} \frac{p^n}{n!} \right)^{1/2} \sum_{k=0}^n (-1)^k C_n^k \frac{\Gamma(v+i+n-k)}{\Gamma(v+i)} \frac{i(i-1)\dots(i-k+1)}{p^k}$$

2 - Cas $\alpha = 0$: Lois de Poisson.

Si $\alpha = \gamma = b_0 = 0$ dans (4), il reste :

$$a_i = a_0 \quad ; \quad b_i = \beta i \quad (\beta > 0)$$

D'après (3), la loi limite existe toujours, et c'est la loi de Poisson de paramètre

$$\theta = \frac{a_0}{\beta}$$

Le processus correspondant décrit une file d'attente à une infinité de guichets (arrivées poissonniennes a_0 , temps de service exponentiel β). Le générateur infinitésimal est :

$$(Af)_i = a_0(f_{i+1} - f_i) - \beta i(f_i - f_{i-1})$$

de sorte que la valeur propre λ_n associée au facteur polynomial χ_n de degré n est

$$\lambda_n = - n \beta$$

Par conséquent, la loi à 2 variables $F_{ij}(t) = W_i P_{ij}(t)$ admet la représentation :

$$(9) \quad \begin{cases} F_{ij} = W_i W_j \sum_{n=0}^{\infty} \rho^n \chi_n(i) \chi_n(j) \\ \rho = e^{-\beta t} \end{cases}$$

D'autre part, on trouve sans peine

$$\sum_j P_{ij}(t) s^j = (1-\rho + \rho s)^j e^{-\theta(1-\rho)(1-s)}$$

d'où l'on déduit la fonction génératrice à 2 variables :

$$G(s, \sigma) = \sum_{ij} F_{ij} s^j \sigma^i = \sum_{ij} W_i P_{ij} \sigma^i s^j$$

a savoir :

$$G(s, \sigma) = e^{-\theta(1-s) - \theta(1-\sigma) + p \theta(1-s)(1-\sigma)}$$

Comparons avec l'expression déduite directement de (9) :

$$G(s, \sigma) = \sum_n \rho^n \sum_i W_i \chi_n(i) \sigma^i \sum_j W_j \chi_n(j) s^j$$

En identifiant les termes en ρ^n , il vient

$$\begin{aligned} \sum_i W_i \chi_n(i) s^i &= \sqrt{\frac{\theta^n}{n!}} (1-s)^n e^{-\theta(1-s)} \\ &= (-1)^n \sqrt{\frac{\theta^n}{n!}} \frac{d^n}{d\theta^n} e^{-\theta(1-s)} \end{aligned}$$

Par conséquent

$$W_i \chi_n(i) = (-1)^n \sqrt{\frac{\theta^n}{n!}} \frac{d^n}{d\theta^n} \left(\frac{\theta^i}{i!} e^{-\theta} \right)$$

soit :

$$(10) \quad \chi_n(i) = (-1)^n \sqrt{\frac{\theta^n}{n!}} \times \frac{i! e^\theta}{\theta^i} \frac{d^n}{d\theta^n} \left(\frac{\theta^i}{i!} e^{-\theta} \right)$$

3 - Cas $\alpha < 0$: Lois Binomiales.

Si α est négatif, $a_i = a_0 + \alpha i$ deviendrait négatif, dans le cas infini, pour i assez grand. Par conséquent, on doit se placer dans le cas fini : a_i doit s'annuler pour une valeur $i = N$, et le processus reste confiné à l'intervalle $(0, N)$. On a donc ici :

$$a_i = a(N-i) \quad ; \quad b_i = b i$$

D'après (3), la loi limite est binomiale (p, N) avec

$$p = \frac{a}{a+b}$$

L'expression du générateur

$$(Af)_i = a(N-i)(f_{i+1} - f_i) - b_i(f_i - f_{i-1})$$

montre que la valeur propre associée à χ_n est

$$\lambda_n = -n(a+b)$$

Par conséquent, la loi à 2 variables est

$$(11) \quad F_{ij}(t) = W_i W_j \sum_{n=0}^N \rho^n \chi_n(i) \chi_n(j)$$

avec

$$\begin{cases} \rho = e^{-(a+b)t} \\ W_i = C_N^i p^i (1-p)^{N-i} \end{cases}$$

D'autre part, on trouve directement

$$\sum_{j=0}^N P_{ij}(t) s^j = [q(1-p) + (p+\rho q)s]^i [q + \rho p + p(1-p)s]^{N-i}$$

et

$$\begin{aligned} G(s, \sigma) &= \sum W_i P_{ij}(t) s^j \sigma^i = \\ &= [q + \rho p + p(1-p)s]^N \left[q + p \frac{q(1-p) + (p+\rho q)s}{q + \rho p + p(1-p)s} \sigma \right]^N \\ &= [(q+ps)(q+p\sigma) + \rho p q(1-s)(1-\sigma)]^N \end{aligned}$$

On obtient donc le développement suivant en ρ :

$$G(s, \sigma) = (q + ps)^N (q+p\sigma)^N \sum_{n=0}^N \rho^n p^n q^n C_N^n \left(\frac{1-s}{q+ps} \right)^n \left(\frac{1-\sigma}{q+p\sigma} \right)^n$$

qui doit être identifié à celui que l'on déduit directement de (11).
Cela donne :

$$\sum_{i=0}^N w_i \chi_n(i) s^i = \sqrt{C_N^n p^n q^n} (1-s)^n (q+ps)^{N-n}$$

$$\stackrel{\substack{\rightarrow \\ (N-n)! \\ N!}}{=} \frac{n!}{N!} \sqrt{C_N^n p^n q^n} \frac{d^n}{d q^n} (q(1-s) + s)^N$$

Par suite :

$$w_i \chi_n(i) = \frac{n!}{N!} \sqrt{C_N^n p^n q^n} \frac{d^n}{d q^n} C_N^i (1-q)^i q^{N-i}$$

et finalement (sauf erreur de calcul!) :

$$(12) \quad \chi_n(i) = \sqrt{\frac{n! p^n q^n}{N! (N-n)!}} \frac{1}{q^{N-i} (1-q)^i} \frac{d^n}{d q^n} q^{N-i} (1-q)^i$$

$\sqrt{\frac{n! p^n q^n}{N! (N-n)!}}$

Il reste à examiner le cas où γ n'est pas nul.

4 - Cas $\gamma \neq 0$: Lois Hypergéométriques.

Ici $a_i = a_0 + \alpha i + \gamma i^2$ et $b_i = \beta i + \gamma i^2$ avec γ non nul. Dans le cas infini ($\gamma > 0$, $a_0 > 0$ et $a_i, b_i > 0$ pour $i = 1, 2, \dots$) on peut, moyennant certaines conditions sur les coefficients, trouver des lois limites W , mais les polynômes ne peuvent pas être dans $L^2(R, W)$ au-delà d'un degré n fini, de sorte que nous n'obtenons pas de modèles à facteurs polynomiaux.

Par contre, dans le cas fini, on obtient une solution. Soit $(0, N)$ l'intervalle de variation de i , ce qui correspond (moyennant un changement de notation) à :

$$a_i = (N-i) [a + \gamma(N-i)] \quad ; \quad b_i = i(b + \gamma i)$$

avec les conditions

$$a + \gamma > 0 \quad , \quad a + N\gamma > 0 \quad ; \quad b + \gamma > 0 \quad , \quad b + N\gamma > 0$$

D'après (3), la loi limite est de la forme :

$$W_n = \frac{N(N-1)\dots(N-n+1)}{n!} \frac{\left(\frac{a}{\gamma} + N\right)\left(\frac{a}{\gamma} + N-1\right)\dots\left(\frac{a}{\gamma} + N-n+1\right)}{\left(1 + \frac{b}{\gamma}\right)\left(1 + \frac{b}{\gamma}\right)\dots\left(n + \frac{b}{\gamma}\right)} W_0$$

Soit, en posant

$$\frac{a}{\gamma} + N = \alpha, \quad \frac{b}{\gamma} = \beta$$

$$W_n = \frac{(-N)(-N+1)\dots(-N+n-1)(-\alpha)(-\alpha+1)\dots(-\alpha+n-1)}{\beta(\beta+1)\dots(\beta+n-1) n!} W_0$$

La fonction génératrice $G = \sum W_n s^n$ est donc l'hypergéométrique :

$$G(s) = W_0 F(-N, -\alpha, \beta ; s)$$

D'après les formules classiques, en $s = 1$, on a :

$$F(-N, -\alpha, \beta ; 1) = \frac{\Gamma(\beta) \Gamma(\alpha+\beta+N)}{\Gamma(\beta+N) \Gamma(\alpha+N)}$$

d'où la valeur de W_0 :

$$W_0 = \frac{\Gamma(\alpha+N) \Gamma(\beta+N)}{\Gamma(\beta) \Gamma(\alpha+\beta+N)}$$

et

$$(13) \quad W_n = C_N^n \frac{\Gamma(\beta+N)}{\Gamma(\beta+n)} \frac{\Gamma(\alpha+N)}{\Gamma(\alpha+\beta+N)} \alpha(\alpha-1)\dots(\alpha-n+1)$$

Il resterait à déterminer les facteurs polynomiaux. Les calculs sont plus compliqués que dans les trois cas précédents, et il ne semble pas urgent de les expliciter avant de rencontrer un exemple d'application qui nous y incite.

Au lieu de cela, je vais terminer cette note en décrivant un

modèle bien différent, où les facteurs sont les fonctions de Walsh, et d'où l'on peut déduire toute une famille de modèles discrets.

5 - Les Loïs de Walsh à 2 variables.

Nous allons nous intéresser à des lois à deux variables $F(dx, dy)$ dont les lois marginales sont uniformes sur l'intervalle $(0, 1)$. Il est facile de donner des exemples de modèles isofactoriels admettant ces lois marginales.

~ nous connaissons déjà un modèle à facteurs polynomiaux (les polynômes de Jacobi).

~ le mouvement brownien sur le cercle, dont le générateur est du type $Af(\theta) = \frac{1}{2} f''(\theta)$ conduit à un modèle dont les facteurs sont les $\cos n \theta$ et les $\sin n \theta$: les lois conditionnelles correspondantes sont du type "loi de Gauss enroulée". Il existe d'ailleurs bien d'autres lois sur $(0, 1)$ admettant les mêmes facteurs trigonométriques.

~ plus généralement, si on a un modèle admettant les facteurs $\chi_n(z)$ et les lois marginales absolument continues $G(dz)$, le changement de variable $X = G(Z)$ donne un modèle admettant des lois marginales uniformes sur $(0, 1)$ et les facteurs $\chi_n[G^{-1}(x)]$.

Pour introduire le modèle de Walsh, considérons l'écriture digitale d'un $X \in (0, 1)$, soit

$$(14) \quad X = \sum_{n=1}^{\infty} \frac{\varepsilon_n}{2^n}$$

où chaque ε_n est ± 1 . Les ε_n sont déterminés sans ambiguïté presque partout sur $(0, 1)$, puisque seuls les rationnels sont susceptibles

de plusieurs écritures.

Si les ε_n sont des V.A. indépendantes avec $P(\varepsilon_n = 1) = p$, X devient une V.A. Pour $p \neq 1/2$ la loi de X appartient au type aberrant bien connu (fonction de répartition continue mais non absolument continue). Par contre, pour $p = 1/2$, la loi de X est uniforme sur $(0,1)$, et c'est ce cas qui nous retiendra.

Pour le démontrer de manière simple, on peut introduire le processus Markovien à temps discret défini par :

$$X_{n+1} = \frac{X_n + \varepsilon_n}{2}$$

($\varepsilon_n = 0$ ou 1 avec la probabilité $1/2$). Ce processus est manifestement ergodique, et sa loi limite ω coïncide avec la loi du X défini en (14). Cette loi est donc caractérisée par :

$$X \equiv \frac{X + \varepsilon}{2}$$

ou, en terme de transformée de Laplace

$$\Phi(\lambda) = \frac{1}{2} \Phi\left(\frac{\lambda}{2}\right) (1 + e^{-\frac{\lambda}{2}})$$

Cette équation admet la solution $\Phi(\lambda) = \frac{1 - e^{-\lambda}}{\lambda}$, et, d'après l'unicité de la probabilité ω invariante pour le processus, X est distribué uniformément sur $(0,1)$.

Définissons maintenant le processus de Walsh $X(t)$ en posant

$$(15) \quad X(t) = \sum_{n=1}^{\infty} \frac{\varepsilon_n(t)}{2^n}$$

les $\varepsilon_n(t)$ étant des Markov indépendants à 2 états (0 et 1) admettant

la même matrice de transition

$$\begin{bmatrix} \frac{1+\rho}{2} & \frac{1-\rho}{2} \\ \frac{1-\rho}{2} & \frac{1+\rho}{2} \end{bmatrix} \quad \text{avec } \rho = e^{-t}$$

Du fait que les $\varepsilon_n(o)$ sont déterminés par la donnée de $X(o)$, le processus de Walsh $X(t)$ est lui-même Markovien. La probabilité de transition de $X(t)$ ne peut pas s'écrire de manière simple (elle est de type continu mais non absolument continu) mais nous allons voir qu'elle admet une représentation factorielle avec comme facteurs les fonctions de Walsh.

Notons que $X(t)$ est un processus réversible, puisque les $\varepsilon_n(t)$ possèdent cette propriété. Autrement dit, si $X = X_0$ admet la loi uniforme sur $(0,1)$, la loi de (X,Y) où $X = X_0$ et $Y = X(t)$ est symétrique en (X,Y) . Par conséquent, les facteurs de cette loi sont caractérisés par la relation

$$E[f(X_t) \mid X_0] = \lambda f(X_0)$$

Montrons que ce sont les fonctions de Walsh.

Les Fonctions de Walsh.

Tout $x \in (0,1)$ admet une écriture décimale de type (14) et, inversement (pour presque tout $x \in (0,1)$) le terme $\varepsilon_n = \varepsilon_n(x)$ est univoquement déterminé par la donnée de x . Posons

$$F_0(x) = 1 \quad ; \quad F_n(x) = 2 \varepsilon_n(x) - 1 \quad (x \in (0,1))$$

Ces fonctions ne prennent que les valeurs ± 1 et sont

orthonormées dans $H = L^2((0,1), dx)$. Plus généralement, on introduira les fonctions

$$F_{i_1, \dots, i_k} = F_{i_1} F_{i_2} \dots F_{i_k} \quad (0 \leq i_1 < \dots < i_k)$$

Ce sont les fonctions de Walsh, et nous allons montrer qu'elles forment une base orthonormée de H .

a) Comme chaque F_{i_k} ne prend que les valeurs ± 1 , $F_{i_1, \dots, i_k}$ ne prend elle-même que ces deux valeurs, et donc a une norme unité dans H .

b) X étant uniformément distribué sur $(0,1)$, les $\varepsilon_n(X)$ sont indépendants et $P(\varepsilon_n = 1) = P(\varepsilon_n = 0) = 1/2$. Donc $E[F_n F_m] = \delta_{nm}$ et, plus généralement

$$E[F_{i_1, \dots, i_k} F_{j_1, \dots, j_p}] = 0$$

dès l'une des deux suites ordonnées $\{i_1, \dots, i_k\}$ et $\{j_1, \dots, j_p\}$ contient un élément qui ne figure pas dans l'autre, c'est-à-dire que les deux fonctions de Walsh correspondantes sont distinctes : ces fonctions sont donc orthonormées.

c) Soit $\mathcal{B}_n$ la σ -algèbre engendrée par les intervalles $I_k(n) = \{x : k 2^{-n} \leq x < (k+1) 2^{-n}\}$, $k = 0, 1, \dots, (2^n - 1)$. L'espace $H_n = L^2((0,1), \mathcal{B}_n, dx)$, constitué de fonctions constantes sur les $I_k(n)$, est de dimension 2^n , et les 2^n fonctions de Walsh

$$F_0 \text{ et } F_{i_1, \dots, i_k} \quad (k = 0, 1, \dots, n, \quad 0 < i_1 < \dots < i_k \leq n)$$

étant orthonormées dans H_n en constituent une base.

d) $H = L^2((0,1), \mathcal{B}, dx)$, où $\mathcal{B}$ est la σ -algèbre de Borel, est

la fermeture de UH_n . (cela résulte sans difficulté de ce que toute VA est limite monotone de VA étagée). En particulier, tout $f \in H$ est limite dans H de la suite $\{f_n\}$, où f_n est la projection de f dans H_n . D'ailleurs $f_n = E(f|\mathcal{B}_n)$ est l'espérance de f relativement à $\mathcal{B}_n$:

$$\forall f \in H, E[f|\mathcal{B}_n] \rightarrow f \quad \text{pour } n \rightarrow \infty$$

Or

$$E[f|\mathcal{B}_n] = E(f) + \sum_{0 < i_1 < \dots < i_k \leq n} \langle f, F_{i_1, \dots, i_k} \rangle F_{i_1, \dots, i_k}$$

puisque les fonctions de Walsh d'ordre $\leq n$ forment une base de H_n . Comme ces fonctions sont orthonormées, il en résulte :

$$f = E(f) + \sum_{0 < i_1 < \dots < i_k} \langle f, F_{i_1, \dots, i_k} \rangle F_{i_1, \dots, i_k}$$

et les fonctions de Walsh constituent bien une base orthonormée de H lui-même.

Les facteurs de $X(t)$.

Il reste à montrer que les fonctions de Walsh sont les facteurs de la loi $F(dx, dy)$, où $X = X_0$ est uniformément distribué sur $(0, 1)$ et $Y = X(t)$. Pour cela, calculons l'espérance conditionnelle :

$$E[\varepsilon_n(t) | \varepsilon_n(0)] = \frac{1+\rho}{2} \varepsilon_n(0) + \frac{1-\rho}{2} (1 - \varepsilon_n(0)) = \rho \varepsilon_n(0) + \frac{1-\rho}{2}$$

donc

$$\begin{aligned} E[F_n(X_t) | X_0] &= E[2 \varepsilon_n(t) - 1 | \varepsilon_n(0) = \varepsilon_n(X_0)] \\ &= \rho [2 \varepsilon_n(X_0) - 1] = \rho F_n(X_0) \end{aligned}$$

et, plus généralement :

$$\begin{aligned} E[F_{i_1, \dots, i_k}(X_t) | X_0] &= E[(2 \varepsilon_{i_1}(t)-1) \dots (2 \varepsilon_{i_k}(t)-1) | X_0] \\ &= \rho^k (2 \varepsilon_{i_1}(X_0)-1) \dots (2 \varepsilon_{i_k}(X_0)-1) \end{aligned}$$

puisque les processus $\varepsilon_n(t)$ sont indépendants, c'est-à-dire :

$$(16) \quad E[F_{i_1, \dots, i_k}(X_t) | X_0] = \rho^k F_{i_1, \dots, i_k}(X_0)$$

Les fonctions de Walsh sont donc bien les facteurs, les valeurs propres correspondantes étant les $\rho^k = e^{-kt}$. D'où la représentation factorielle de la "densité" $f(x, y)$ de la loi de Walsh à 2 variables:

$$(16') \quad \left\{ \begin{aligned} f(x, y) &= 1 + \sum_{k=1}^{\infty} \rho^k \sum_{0 < i_1 < \dots < i_k} F_{i_1, \dots, i_k}(x) F_{i_1, \dots, i_k}(y) \\ (0 \leq x < 1 \ ; \ 0 \leq y < 1) \end{aligned} \right.$$

(N.B.: il s'agit d'une écriture symbolique, la série n'étant pas convergente. La relation caractéristique est (16) : elle suffit à déterminer la loi $F(dx, dy)$, puisque les fonctions de Walsh forment une base hilbertienne de H . Ici, les espérances conditionnelles ne sont pas des opérateurs compacts, puisque les espaces propres associés à chaque valeur propre ρ^k sont de dimension infinie. C'est cette circonstance qui est responsable de la non-convergence de la représentation (16') et de l'absence de densité).

Cette loi à deux variables possède la propriété remarquable suivante : si on arrête le développement à $k = n$, soit explicitement :

$$(17) \quad f_n(x, y) = 1 + \sum_{k=1}^n \rho^k \sum_{0 < i_1 < \dots < i_k \leq n} F_{i_1, \dots, i_k}(x) F_{i_1, \dots, i_k}(y)$$

on obtient la version discrétisée de la loi précédente, et, en particulier $f_n(x, y) \geq 0$.

En effet

$$X_n(t) = E(X(t) | \mathcal{B}_n) = \sum_{p=1}^n \frac{\varepsilon_p(t)}{2^p}$$

est encore un Markov, et en prenant $X_n = \sum_{p=0}^n \frac{\varepsilon_p(0)}{2^p} = E(X_0 | \mathcal{B}_n)$ et $Y_n = X_n(t)$, (X_n, Y_n) admet la loi de densité f_n constante sur les carreaux de côté 2^{-n} : c'est la loi de Walsh discrétisée.

On a là un exemple tout à fait exceptionnel de loi à 2 variables où le fait de ne conserver que les 2^n premiers termes de la représentation factorielle équivaut à discrétiser la loi, ou, plus précisément, à remplacer la densité f par la densité f_n constante sur chaque carreau de côté 2^{-n} , c'est-à-dire X et Y par $E(X | \mathcal{B}_n)$ et $E(Y | \mathcal{B}_n)$.

Espérance et Variance Conditionnelles de X si Y

Pour calculer l'espérance conditionnelle $E(X|Y)$ pour la loi de Walsh (16) partons de l'identité

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{2^n} = \sum_{n=1}^{\infty} \frac{1 + F_n(x)}{2^{n+1}} = \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{F_n(x)}{2^n}$$

soit

$$(18) \quad x = \frac{1}{2} \sum_{n=0}^{\infty} \frac{F_n(x)}{2^n}$$

Les F_n étant des facteurs associés aux valeurs propres $\lambda = 0$ pour $n = 0$ et $\lambda = \rho$ pour $n = 1, 2, \dots$, il en résulte aussitôt :

$$E(X|Y) = \frac{1}{2} F_0(Y) + \frac{\rho}{2} \sum_{n=1}^{\infty} \frac{F_n(Y)}{2^n} = \frac{1}{2} + \frac{\rho}{2} \sum_{n=1}^{\infty} \frac{F_n(Y)}{2^n}$$

Mais - d'après le développement (18) lui-même :

$$\frac{1}{2} \sum_{n=1}^{\infty} \frac{F_n(Y)}{2^n} = Y - \frac{1}{2}$$

et par suite :

$$(19) \quad E(X|Y) = \frac{1}{2} + \rho \left(Y - \frac{1}{2} \right)$$

Autrement dit, et en dépit de la nature analytique plutôt compliquée de la loi de Walsh, la régression est linéaire. Ou encore, si l'on préfère, le polynôme de degré 1 : $x - \frac{1}{2}$ est un facteur associé à la valeur propre ρ . Cela résulte de ce que $x - \frac{1}{2}$ se développe en fonction des seuls facteurs F_n , $n \geq 1$, tous associés à la même valeur propre ρ .

Considérons maintenant le développement :

$$\left(x - \frac{1}{2}\right)^2 = \frac{1}{4} \left(\sum_{n=1}^{\infty} \frac{F_n(x)}{2^n} \right)^2 = \frac{1}{4} \sum_{n=1}^{\infty} \frac{(F_n(x))^2}{2^{2n}} + \frac{1}{2} \sum_{0 < i < j} \frac{F_i(x) F_j(x)}{2^{i+j}}$$

Comme $F_n(x)^2 = 1$, ceci se réduit à :

$$\left(x - \frac{1}{2}\right)^2 = \frac{1}{12} + \frac{1}{2} \sum_{0 < i < j} \frac{F_i(x) F_j(x)}{2^{i+j}}$$

Ici les $F_i(x) F_j(x)$ sont des facteurs associés à la même valeur propre ρ^2 , et par suite :

$$\begin{aligned} E\left[\left(X - \frac{1}{2}\right)^2 | Y\right] &= \frac{1}{12} + \frac{\rho^2}{2} \sum_{0 < i < j} \frac{F_i(Y) F_j(Y)}{2^{i+j}} \\ &= \frac{1-\rho^2}{12} + \rho^2 \left(Y - \frac{1}{2}\right)^2 \end{aligned}$$

D'après (19), on a aussi $(E(X - \frac{1}{2}|Y))^2 = \rho^2(Y - \frac{1}{2})^2$. Par suite la variance conditionnelle est

$$D^2(X|Y) = \frac{1 - \rho^2}{12}$$

Exactement comme dans le cas Gaussien, la variance conditionnelle de X à Y fixé est indépendante de Y et vaut $\sigma^2(1-\rho^2)$. Ici s'arrête l'analogie, et l'on n'obtient pas de résultat simple pour les polynomes de degré supérieur à 2.

REMARQUE - Le processus $X(t)$ à temps continu n'autorise que des valeurs $\rho \in (0,1)$, puisque $\rho = e^{-t}$. Mais en partant d'un processus $X(k)$ à temps k discret, soit

$$X(k) = \sum \varepsilon_n(k) 2^{-n}$$

où les $\varepsilon_n(k)$ sont des Markov discrets à 2 états, on voit qu'en fait ρ peut varier de - 1 à + 1.

EXEMPLE - Discrétisé à l'ordre $n = 2$, le processus a quatre valeurs correspondant aux classes $(0, 1/4)$, $(1/4, 1/2)$, $(1/2, 3/4)$, $(3/4, 1)$ sur lesquelles les 4 fonctions de Walsh correspondantes sont :

$$\begin{aligned} F_0 &= (1, 1, 1, 1) & F_1 &= (-1, -1, 1, 1) \\ F_2 &= (-1, 1, -1, 1) & F_{12} &= (1, -1, -1, 1) \end{aligned}$$

Voici la densité de la loi discrète f_2 correspondante :

$(1-\rho)^2$	$1-\rho^2$	$1-\rho^2$	$(1+\rho)^2$
$1-\rho^2$	$(1-\rho)^2$	$(1+\rho)^2$	$1-\rho^2$
$1-\rho^2$	$(1+\rho)^2$	$(1-\rho)^2$	$1-\rho^2$
$(1+\rho)^2$	$1-\rho^2$	$1-\rho^2$	$(1-\rho)^2$

La loi de formation apparaît assez clairement : il s'agit de superposer aux diverses échelles le même motif

$1-p$	$1+p$
$1+p$	$1-p$

Prenons par exemple $n = 3$, c'est-à-dire 8 classes pour X et 64 carreaux pour (X,Y) , et $p = 1/2$. La densité, multipliée par 8, apparaît sur le tableau suivant :

1	3	3	9	3	9	9	27
3	1	9	3	9	3	27	9
3	9	1	3	9	27	3	9
9	3	3	1	27	9	9	3
3	9	9	27	1	3	3	9
9	3	27	9	3	1	9	3
9	27	3	9	3	9	1	3
27	9	9	3	9	3	3	1

(si p est quelconque, remplacer 1 par $(1-p)^3$, 3 par $(1-p)^2(1+p)$, 9 par $(1-p)(1+p)^2$ et 27 par $(1+p)^3$).

REMARQUE - On notera l'analogie que présente ce mode de construction (d'une loi de probabilité) avec le modèle à homothétie interne de De Wijs. Ceci suggère - paradoxalement - une interprétation de cette répartition probabiliste en terme de régionalisation dans le plan (x,y) des 2 variables. Compte tenu de la lognormalité (plus exactement : logbinomialité) sous-jacente, remplaçons la densité f par son logarithme ou, ce qui revient au même, remplaçons

$(1-\rho)^{n-k}(1+\rho)^k$ par k , i.e. ici : $1 \rightarrow 0$, $3 \rightarrow 1$, $9 \rightarrow 2$, $27 \rightarrow 3$. On obtient la tableau :

0	1	1	2	1	2	2	3
1	0	2	1	2	1	3	2
1	2	0	1	2	3	1	2
2	1	1	0	3	2	2	1
1	2	2	3	0	1	1	2
2	1	3	2	1	0	2	1
2	3	1	2	1	2	0	1
3	2	2	1	2	1	1	0

auquel on associe le variogramme "expérimental" suivant dans la direction des lignes (ou aussi bien des colonnes)

$$h = \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix}$$

$$2\gamma(h) = \begin{matrix} \frac{11}{7} & \frac{8}{6} & \frac{11}{5} & \frac{4}{4} & \frac{7}{3} & \frac{4}{2} & \frac{3}{1} \end{matrix}$$

C'est une approximation (grossière et chaotique) d'un schéma De Wijsien.

On subodore ici un exemple curieux de l'ambiguïté des interpolations possibles d'un ensemble de données. Si le premier tableau est considéré comme une loi de probabilité, l'interprétation en termes de facteurs (les fonctions de Walsh) s'impose : c'est le point de vue de l'analyse des données. Si, après passage aux logarithmes, le même tableau est considéré comme une régionalisation, on est tenté de l'interpréter en termes de F.A.I. De Wijsienne.

Du fait de la structure très particulière de ce tableau (rien

que des 0 dans la diagonale principale, et des 3 dans l'autre diagonale par exemple) l'interprétation en terme de régionalisation peut paraître un peu artificielle. Mais cette impression va s'atténuer si l'on considère la généralisation suivante :

Généralisation.

Considérons le processus à temps discret :

$$X(k) = \sum_{n=1}^{\infty} \frac{\varepsilon_n(k)}{2^n}$$

avec, à nouveau, comme $\varepsilon_n(k)$ des Markov indépendants à 2 états, mais admettant maintenant des matrices de transition différentes pour chaque n :

$$\begin{bmatrix} \frac{1 + \rho_n}{2} & \frac{1 - \rho_n}{2} \\ \frac{1 - \rho_n}{2} & \frac{1 + \rho_n}{2} \end{bmatrix}$$

On obtient une loi à 2 variables, dont les facteurs sont les fonctions de Walsh, mais les valeurs propres sont différentes :

$$E[F_{i_1, \dots, i_k}(X) | Y] = \rho_{i_1} \rho_{i_2} \dots \rho_{i_k} F_{i_1, \dots, i_k}(Y)$$

La loi à 2 variables discrétisée à l'ordre n admet la densité

$$f_n(x, y) = 1 + \sum_{0 < i_1 < \dots < i_k \leq n} \rho_{i_1} \dots \rho_{i_k} \left[F_{i_1}(x) F_{i_1}(y) \dots F_{i_k}(x) F_{i_k}(y) \right]$$

soit

$$(20) \quad f_n(x, y) = (1 + \rho_1 F_1(x) F_1(y)) (1 + \rho_2 F_2(x) F_2(y)) \dots (1 + \rho_n F_n(x) F_n(y))$$

Si les ρ_i sont donnés, ceci est une loi de probabilité. Mais si les ρ_i sont des V.A. (par exemple indépendantes, et admettant la même loi sur l'intervalle $-1,1$), $f_n(x,y)$ devient une F.A., et

$$\log f_n(x,y) = \sum_{i=1}^n \log [1 + \rho_i F_i(x) F_i(y)]$$

est vraisemblablement assez proche d'une F.A.I. De Wijsienne.