Corrélations discrètes de certaines suites pseudo-aléatoires sur un alphabet de taille k

Pierre-Adrien Tahay

Université de Lorraine, IECL, Nancy

08 mars 2019

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

$$\begin{array}{cccccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \cdots \\ & \uparrow & & & \uparrow & & & & & & & \\ & i & & & j & & & & & & & \end{array}$$

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

$$\begin{array}{ccccccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \cdots \\ & & \uparrow & & & \uparrow & & & & & & \\ & & i+1 & & & j+1 & & & & & & \end{array}$$

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

$$\begin{array}{ccccccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \cdots \\ & & & \uparrow & & & \uparrow & & & & & \\ & & & i+2 & & & j+2 & & & & & \end{array}$$

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

$$\begin{array}{cccccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \cdots \\ & & & & \uparrow & & & \uparrow & & & & \\ & & & & i+3 & & & j+3 & & & & \end{array}$$

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

$$\begin{array}{ccccccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \cdots \\ & & & & & \uparrow & & & \uparrow & & & \\ & & & & & i+4 & & & j+4 & & & \end{array}$$

Coefficient de corrélation

Définition

Soit $k \geq 2$ un entier et $x = x_0 x_1 \cdots$ un mot infini sur l'alphabet $\{0, 1, \ldots, k-1\}$.
Pour un vecteur (i, j) tel que $0 \leq i < j$ on définit le coefficient de corrélation discret $\delta(i, j)$ d'ordre 2 par :

$$\delta(i,j) = \begin{cases} 0, & \text{si } x_i = x_j, \\ 1, & \text{sinon.} \end{cases}$$

$$\begin{array}{ccccccccccccc} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \cdots \\ & & & & & & \uparrow & & & \uparrow & & \\ & & & & & & i+5 & & & j+5 & & \end{array}$$

Remarque

Pour une suite aléatoire où chaque lettre est tirée indépendamment avec probabilité $\frac{1}{k}$ on a $\mathbb{P}(x_i = x_j) = \frac{1}{k}$.

Ainsi

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} \delta(n+i, n+j) = 1 - \frac{1}{k} \text{ avec probabilité } 1$$

.

Suite de Rudin-Shapiro

Définition

La suite de Rudin-Shapiro $(r_n)_{n\geqslant 0} = 0, 0, 0, 1, 0, 0, 1, 0, \ldots$ est définie pour tout $n \in \mathbb{N}$ par :

$r_n =$ (nombre de blocs 11 dans l'écriture binaire de n) mod 2

Suite de Rudin-Shapiro

Définition

La suite de Rudin-Shapiro $(r_n)_{n\geqslant 0} = 0, 0, 0, 1, 0, 0, 1, 0, \ldots$ est définie pour tout $n \in \mathbb{N}$ par :
$r_n =$ (nombre de blocs 11 dans l'écriture binaire de n) mod 2

Pour $n = 55 = (110111)_2$ on a $r_{55} = 3 \text{ mod } 2 = 1$

Suite de Rudin-Shapiro

Définition

La suite de Rudin-Shapiro $(r_n)_{n\geqslant 0} = 0, 0, 0, 1, 0, 0, 1, 0, \ldots$ est définie pour tout $n \in \mathbb{N}$ par :
$r_n =$ (nombre de blocs 11 dans l'écriture binaire de n) mod 2

Pour $n = 55 = (110111)_2$ on a $r_{55} = 3 \bmod 2 = 1$

Remarque

On en déduit directement les relations

$$r_{2n} = r_n \text{ et } r_{2n+1} = \begin{cases} r_n + 1, & \text{si } n \equiv 1 \bmod 2, \\ r_n, & \text{si } n \equiv 0 \bmod 2. \end{cases}$$

Remarque

Avec la remarque précédente on peut définir la suite de Rudin-Shapiro de la manière suivante :

Remarque

Avec la remarque précédente on peut définir la suite de Rudin-Shapiro de la manière suivante :

$$r_0 = 0 \text{ et } r_{2n+j} = (r_n + g(j,n)) \text{ mod } 2$$

Remarque

Avec la remarque précédente on peut définir la suite de Rudin-Shapiro de la manière suivante :

$$r_0 = 0 \text{ et } r_{2n+j} = (r_n + g(j,n)) \text{ mod } 2$$

avec $g(j,n) = \begin{cases} 1, & \text{si } j = 1,\ n \equiv 1 \text{ mod } 2, \\ 0, & \text{sinon.} \end{cases}$

Remarque

Avec la remarque précédente on peut définir la suite de Rudin-Shapiro de la manière suivante :

$$r_0 = 0 \text{ et } r_{2n+j} = (r_n + g(j,n)) \bmod 2$$

avec $g(j,n) = \begin{cases} 1, & \text{si } j = 1,\ n \equiv 1 \bmod 2, \\ 0, & \text{sinon.} \end{cases}$

Remarque

On peut représenter $g(j,n)$ comme le coefficient de la j-ième colonne et de la n-ième ligne de la matrice $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Suites généralisées de Rudin-Shapiro

On se donne une fonction $g : \{0, 1, \ldots, k-1\} \times \mathbb{Z} \to \mathbb{Z}$, $(j, n) \longmapsto g(j, n)$ k-périodique en n et telle que pour tous entiers u, i avec $0 \leq u < u + i \leq k - 1$ on a

$$\{(g(u+i, n) - g(u, n)) \bmod k : 0 \leq n < k\} = \{0, 1, \ldots, k-1\}.$$

Suites généralisées de Rudin-Shapiro

On se donne une fonction $g : \{0, 1, \ldots, k-1\} \times \mathbb{Z} \to \mathbb{Z}$, $(j, n) \longmapsto g(j, n)$ k-périodique en n et telle que pour tous entiers u, i avec $0 \leq u < u + i \leq k - 1$ on a

$$\{(g(u+i, n) - g(u, n)) \bmod k : 0 \leq n < k\} = \{0, 1, \ldots, k-1\}.$$

Définition [Grant, Shallit, Stoll (2009)]

Une suite $(\hat{\mathrm{a}}(n))_{n\geq 0}$ sur l'alphabet $\{0, 1, \ldots, k-1\}$ est une suite généralisée de Rudin-Shapiro s'il existe une suite d'entiers $(a(n))_{n\geq 0}$ telle que $\hat{\mathrm{a}}(n) \equiv a(n) \bmod k$ et

$$a(nk + j) = a(n) + g(j, n) \quad 0 \leq j \leq k-1, \quad n \geq 1.$$

Exemple

Pour $k = 2$ et

$$g(j,n) = \begin{cases} 1, & \text{si } j = 1,\ n \equiv 1 \bmod 2, \\ 0, & \text{sinon,} \end{cases}$$

on retrouve la suite de Rudin-Shapiro sur l'alphabet $\{0,1\}$,
$(\hat{\mathrm{a}}(n))_{n\geqslant 0} = 0,0,0,1,0,0,1,0,\ldots$

Exemple

Pour $k = 2$ et

$$g(j,n) = \begin{cases} 1, & \text{si } j = 1,\ n \equiv 1 \bmod 2, \\ 0, & \text{sinon,} \end{cases}$$

on retrouve la suite de Rudin-Shapiro sur l'alphabet $\{0,1\}$, $(\hat{\mathrm{a}}(n))_{n\geqslant 0} = 0,0,0,1,0,0,1,0,\ldots$

$$\begin{pmatrix} g(0,0) & g(1,0) \\ g(0,1) & g(1,1) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Exemple

Pour $k = 3$ et

$$g(j,n) = \begin{cases} 1, & \text{si } j \equiv n \bmod 3, \\ 0, & \text{sinon,} \end{cases}$$

avec comme condition initiale : $\hat{\mathrm{a}}(0) = 0$.

Exemple

Pour $k = 3$ et

$$g(j,n) = \begin{cases} 1, & \text{si } j \equiv n \text{ mod } 3, \\ 0, & \text{sinon,} \end{cases}$$

avec comme condition initiale : $\hat{\mathrm{a}}(0) = 0$.

La suite $\hat{\mathrm{a}}(n)_{n\geq 0} = 0,0,0,0,1,0,0,0,1,1,0,0,1,2,1,0\ldots$ compte le nombre cumulé (mod 3) des sous-blocs (00),(11) et (22).

Exemple

Pour $k = 3$ et

$$g(j,n) = \begin{cases} 1, & \text{si } j \equiv n \bmod 3, \\ 0, & \text{sinon,} \end{cases}$$

avec comme condition initiale : $\hat{\mathrm{a}}(0) = 0$.

La suite $\hat{\mathrm{a}}(n)_{n\geq 0} = 0,0,0,0,1,0,0,0,1,1,0,0,1,2,1,0\ldots$ compte le nombre cumulé (mod 3) des sous-blocs (00),(11) et (22).

$$\begin{pmatrix} g(0,0) & g(1,0) & g(2,0) \\ g(0,1) & g(1,1) & g(2,1) \\ g(0,2) & g(1,2) & g(2,2) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Cas k premier

Théorème [Grant, Shallit, Stoll (2009)]

Soit $(\hat{\mathrm{a}}(n))_{n\geq 0}$ une suite généralisée de Rudin-Shapiro sur $\{0,1,\ldots,k-1\}$ avec k premier, et soit $r \geq 1$.
Alors, quand $N \to \infty$, on a

$$\sum_{n<N} \delta(n, n+r) = N\left(1-\frac{1}{k}\right) + O_k\left(r\log\left(\frac{N}{r}\right)+r\right),$$

où la constante implicite dépend uniquement de k.

Cas k premier

Théorème [Grant, Shallit, Stoll (2009)]

Soit $(\hat{\mathrm{a}}(n))_{n\geq 0}$ une suite généralisée de Rudin-Shapiro sur $\{0,1,\ldots,k-1\}$ avec k premier, et soit $r \geq 1$.
Alors, quand $N \to \infty$, on a

$$\sum_{n<N} \delta(n, n+r) = N\left(1-\frac{1}{k}\right) + O_k\left(r\log\left(\frac{N}{r}\right)+r\right),$$

où la constante implicite dépend uniquement de k.

Remarque [Grant, Shallit, Stoll (2009)]

Dans le même papier les auteurs donnent également un résultat pour k sans facteurs carrés.

Généralisations ?

Question : Peut-on généraliser à k arbitraire ?

Généralisations ?

Question : Peut-on généraliser à k arbitraire ?

Le premier cas non pris en compte par les résultats précédents est $k = 4$.

Généralisations ?

Question : Peut-on généraliser à k arbitraire ?

Le premier cas non pris en compte par les résultats précédents est $k = 4$.

Problème : On ne peut pas définir de matrice de différence sur $\{0, 1, 2, 3\}$ comme précédemment.

Matrices de différences

Définition

Une **matrice de différence** est une matrice à coefficients dans un groupe additif fini G et telle que pour deux colonnes C_i et C_j avec $i \neq j$, la différence $C_i - C_j$ est constituée de tous les éléments de G qui apparaissent chacun avec le même nombre d'occurrence.

On note $D(k, G)$ l'ensemble des matrices de différence de taille k à coefficients dans le groupe G.

Matrices de différences

Exemple

$$D(2, \mathbb{Z}_2) : \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Matrices de différences

Exemple

$$D(2,\mathbb{Z}_2) : \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Exemple

$$D(3,\mathbb{Z}_3) : \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Matrices de différence

Proposition

Soit k un nombre premier. Alors la matrice carrée $A = (ij \mod k)_{1 \leq i,j \leq k}$ est une matrice appartenant à $D(k, \mathbb{Z}_k)$.

Proposition

Soit k un nombre premier. Alors la matrice carrée $A = (ij \bmod k)_{1 \leq i,j \leq k}$ est une matrice appartenant à $D(k, \mathbb{Z}_k)$.

Exemple

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 1 & 3 \\ 0 & 3 & 1 & 4 & 2 \\ 0 & 4 & 3 & 2 & 1 \end{pmatrix}$$

Remarque

L'ensemble $D(4, \mathbb{Z}_4)$ est vide.

Remarque

L'ensemble $D(4, \mathbb{Z}_4)$ est vide.

Exemple

$$D(4, \mathbb{Z}_2 \times \mathbb{Z}_2) : \begin{pmatrix} (0,0) & (0,0) & (0,0) & (0,0) \\ (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,1) & (0,1) & (1,0) \end{pmatrix}$$

Remarque

L'ensemble $D(4, \mathbb{Z}_4)$ est vide.

Exemple

$$D(4, \mathbb{Z}_2 \times \mathbb{Z}_2) : \begin{pmatrix} (0,0) & (0,0) & (0,0) & (0,0) \\ (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,1) & (0,1) & (1,0) \end{pmatrix}$$

Théorème

Pour tout nombre premier p et tout entier $k \geq 1$, l'ensemble $D(p^k, \mathbb{Z}_p^k)$ est non vide.

Définition

Soit p un nombre premier, $k \geq 1$ et M une matrice de différence de $D(p^k, \mathbb{Z}_p^k)$.

Définition

Soit p un nombre premier, $k \geq 1$ et M une matrice de différence de $D(p^k, \mathbb{Z}_p^k)$.

On définit la suite $(a(n))_{n\geq 0} = (a_1(n), \ldots, a_k(n))_{n\geq 0}$ par $a(0) = (0, 0, \ldots, 0)$ et

$$a(p^k n + j) = a(n) + g(j, n), \quad 0 \leq j \leq p^k - 1, \quad n \geq 1.$$

Cas k puissance d'un nombre premier

Théorème

Soit p un nombre premier et m un entier. On se place sur l'alphabet $\{0, 1, \ldots, p^m - 1\}$.

Cas k puissance d'un nombre premier

Théorème

Soit p un nombre premier et m un entier. On se place sur l'alphabet $\{0, 1, \ldots, p^m - 1\}$.

Soit $(a(n))_{n\geq 0}$ la suite de la définition précédente et soit $r \geq 1$.

Alors, quand $N \to \infty$, on a

$$\sum_{n<N} \delta(n, n+r) = N\left(1 - \frac{1}{p^m}\right) + O_{p,m}\left(r\log\left(\frac{N}{r}\right) + r\right),$$

où la constante implicite ne dépend que de p et de m.

Cas général

Définition

Soit $d \geq 2$, $p_1, \ldots, p_d$ des nombres premiers et $k_1, \ldots, k_d$ des entiers. On pose $k = p_1^{k_1} \cdots p_d^{k_d}$ et on se place sur l'alphabet $\{0, 1, \ldots, k-1\}$.

Cas général

Définition

Soit $d \geq 2$, $p_1, \ldots, p_d$ des nombres premiers et $k_1, \ldots, k_d$ des entiers. On pose $k = p_1^{k_1} \cdots p_d^{k_d}$ et on se place sur l'alphabet $\{0, 1, \ldots, k-1\}$.

Pour chaque $1 \leq i \leq d$ on considère M_i une matrice de différence de $D(p_i^{k_i}, \mathbb{Z}_{p_i}^{k_i})$ à laquelle on associe une fonction $g_i(j, n)$ et une suite $(a_i(n))_{(n \geq 0)}$ que l'on définit comme précédemment.

Cas général

Définition

Soit $d \geq 2$, $p_1, \ldots, p_d$ des nombres premiers et $k_1, \ldots, k_d$ des entiers. On pose $k = p_1^{k_1} \cdots p_d^{k_d}$ et on se place sur l'alphabet $\{0, 1, \ldots, k-1\}$.

Pour chaque $1 \leq i \leq d$ on considère M_i une matrice de différence de $D(p_i^{k_i}, \mathbb{Z}_{p_i}^{k_i})$ à laquelle on associe une fonction $g_i(j, n)$ et une suite $(a_i(n))_{(n \geq 0)}$ que l'on définit comme précédemment.

On définit la suite

$$(\hat{\mathrm{a}}(n))_{n \geq 0} \text{ par } \hat{\mathrm{a}}(n) = (a_1(n) \bmod p_1, \ldots, a_d(n) \bmod p_d)$$

Cas général

Théorème

Soit $r \geq 1$ et $0 < \gamma < 1$. Alors, quand $N \to \infty$, on a

$$\sum_{n<N} \delta(n, n+r) = N\left(1 - \frac{1}{k}\right) + O\left(rN^{1-\frac{\gamma}{d}} \log\left(\frac{N^{\frac{\gamma}{d}}}{r}\right) + rN^{1-\frac{\gamma}{d}}\right)$$

Bibliographie

- J.P. Allouche and P. Liardet. Generalized Rudin-Shapiro sequences. *Acta Arithmetica*, 1991.
- Gennian Ge, On $(g, 4; 1)$-difference matrices. *Discrete Mathematics* 301(2-3) 164-174, October 2005.
- E. Grant, J. Shallit and T. Stoll. Bounds for the discrete correlation of inifinite sequences on k symbols and generalized Rudin-Shapiro sequences. *Acta Arithmetica*, 140(4) : 345-368, 2009.
- A.S. Hedayat, N.J.A Sloane, and J. Stufken. *Orthogonal Arrays*. Springer, New-York, 1999.
- Pekka H. J. Lampio Classification of difference matrices and complex Hadamard matrices. "Thèse de doctorat, 2015".

Questions ouvertes

- Optimisation des termes d'erreur ?

Questions ouvertes

- Optimisation des termes d'erreur ?
- Les résultats portent sur les corrélations d'ordre 2. Que se passe-t-il pour les corrélations d'ordre supérieur ?

Questions ouvertes

- Optimisation des termes d'erreur ?
- Les résultats portent sur les corrélations d'ordre 2. Que se passe-t-il pour les corrélations d'ordre supérieur ?
- Les preuves utilisent toutes des matrices de différence carrées. Peut-on utiliser d'autres matrices de différence ?

Questions ouvertes

- Optimisation des termes d'erreur ?
- Les résultats portent sur les corrélations d'ordre 2. Que se passe-t-il pour les corrélations d'ordre supérieur ?
- Les preuves utilisent toutes des matrices de différence carrées. Peut-on utiliser d'autres matrices de différence ?
- Étude d'autres suites pseudo-aléatoires ?

Questions ouvertes

- Optimisation des termes d'erreur ?
- Les résultats portent sur les corrélations d'ordre 2. Que se passe-t-il pour les corrélations d'ordre supérieur ?
- Les preuves utilisent toutes des matrices de différence carrées. Peut-on utiliser d'autres matrices de différence ?
- Étude d'autres suites pseudo-aléatoires ?

Merci pour votre attention !