N-231

NOTE GEOSTATISTIQUE N° 115

LES FORMULES D'APPROXIMATION DE LA GEOSTATISTIQUE

Par

G. MATHERON

Mai 1971

NOTE GEOSTATISTIQUE N° 115

LES FORMULES D'APPROXIMATION DE LA GEOSTATISTIQUE

Table des Matières

NOTE GEOSTATISTIQUE N° 115

LES FORMULES D'APPROXIMATION DE LA GEOSTATISTIQUE

Je me propose ci-après d'examiner le degré de validité des formules d'approximation utilisées couramment en géostatistique pour le calcul des variances d'estimation. Ce scrupule tardif a été suscité par le résultat d'un calcul "scolaire" effectué sur le variogramme linéaire. Il apparaît, en effet, dans ce cas, que le terme d'effet de bordure qui vient majorer la variance d'estimation de la teneur moyenne de l'intervalle [0,L] lorsque l'on utilise une maille régulière à implantation non préférentielle est du même ordre de grandeur que le terme principal. Ce résultat n'était pas à proprement parler inattendu, car la théorie en prévoyait l'existence et permettait d'en calculer l'ordre de grandeur (sinon la valeur numérique exacte). Mais, dans les applications pratiques, l'usage s'était instauré depuis longtemps d'oublier ce terme et de le négliger purement et simplement. Il m'a donc semblé utile de reprendre l'étude de cette question, et, par la même occasion, de préciser les conditions de validité d'un certain nombre de formules usuelles. Il apparaîtra, en particulier, que la structure de l'effet de bordure est beaucoup plus complexe qu'on ne le soupçonnait.

1 - EXEMPLE INTRODUCTIF:

L'EFFET DE BORDURE POUR UN VARIOGRAMME LINEAIRE.

Soit $Z(x)$ une FAI sans dérive sur la droite, et $\gamma(h) = |h|$ son demi-variogramme. On veut estimer la teneur du segment $(0,L)$ de longueur $L = (n+\varepsilon)a$, $0 \leq \varepsilon < 1$ à l'aide d'échantillons prélevés selon une maille a régulière à implantation non préférentielle. On sait que l'on formalise cette situation en admettant que tout se passe comme si le premier échantillon positif était implanté en un point x_o choisi au hasard sur $(0,a)$ selon une loi de probabilité uniforme. L'estimateur utilisé $Z^*(x_o)$ est donc une variable aléatoire définie sur la σ-algèbre $\mathcal{A} \otimes \mathcal{B}$, où $\mathcal{A}$ est la σ-algèbre sur laquelle est définie notre FAI, et $\mathcal{B}$ la tribu borélienne de $[0,a]$. Explicitement, $Z^*(x_o)$ s'écrit :

$$Z^* = \begin{cases} \dfrac{1}{n+1} \sum\limits_{i=0}^{n} Z(x_o+ia) & \text{si } 0 \leq x_o \leq \varepsilon a \\ \dfrac{1}{n} \sum\limits_{i=0}^{n+1} Z(x_o+ia) & \text{si } \varepsilon a < x_o < a \end{cases}$$

La variance d'estimation est $\sigma_E^2 = D^2 \left(Z^* - \frac{1}{L} \int_0^L Z(x)\, dx\right)$. A x_o fixé, l'erreur $Z^*(x_o) - \frac{1}{L}\int_0^L Z(x)\, dx$ a une espérance conditionnelle nulle, de sorte que l'on a :

$$\sigma_E^2 = \frac{1}{a} \int_0^a \sigma_E^2 (x_o)\, dx_o \tag{1-1}$$

en désignant par $\sigma_E^2(x_o)$ la variance d'estimation conditionnelle en x_o.

Dans le cas $0 \leq x_o \leq \varepsilon a$, on trouve :

$$\sigma_E^2(x_o) = -\frac{L}{3} - \sum_{i,j} |i-j| \frac{a}{(n+1)^2} + \frac{2}{(n+1)} \sum_{i=0}^{n} \frac{(x_o+ia)^2 + (L-x_o-ia)^2}{2\,L}$$

On calcule sans peine les expressions :

$$\sum_{i,j} |i-j| = \frac{n(n+1)(n+2)}{3}$$

$$\frac{1}{n+1} \sum_{i=0}^{n} \frac{(x_o+ia)^2 + (L-x_o-ia)^2}{L} = \frac{1}{L} \left[n(2n+1) \frac{a^2}{3} + n\, a(2x_o-L) + 2x_o^2 - 2x_o L + L \right.$$

On en déduit pour $0 \leq x_o \leq \varepsilon a$:

$$\sigma_E^2(x_o) = \frac{1}{L} \left[n(2n+1) \frac{a^2}{3} + n\, a(2x_o-L) + 2x_o^2 - 2x_o L + L^2 \right] - \frac{L^2}{3} - \frac{n(n+2)}{n+1} \frac{a}{3}$$

En remplaçant n par n-1, on trouve pour $\varepsilon a < x_o < a$:

$$\sigma_E^2(x_o) = \frac{1}{L} \left[n(2n-1) \frac{a^2}{3} + (n-1)\, a(2x_o-L) + 2x_o^2 - 2x_o L + L^2 \right] - \frac{L^2}{3} - \frac{n^2-1}{n} \frac{a}{3}$$

En reportant ces expressions dans (1-1), on obtient, après quelques calculs élémentaires :

$$\sigma_E^2 = \frac{1}{L} \left[\frac{a^2}{3} (n-1)(2n-1) + \varepsilon \frac{a^2}{3} (4n-1) + a^2\varepsilon^2 - aL\varepsilon + (n-\frac{1}{3})a^2 - naL + L^2 \right.$$

$$- \frac{L}{3} - \frac{a}{3} \left[\frac{n^2-1}{n} + \varepsilon \frac{n^2+n+1}{n(n+1)} \right]$$

Enfin, en remplaçant L par $(n+\varepsilon)a$, cette expression se simplifie, et on trouve :

$$\sigma_E^2 = \frac{a}{3(n+\varepsilon)} \left[1 - \varepsilon(1-\varepsilon)\left(1 - \frac{1}{n(n+1)}\right)\right]$$

Lorsque n est grand, la variance d'estimation admet donc la partie principale :

$$\sigma_E^2 \simeq \frac{a}{3n} \left[1 - \varepsilon(1-\varepsilon)\right] \tag{1-2}$$

Dans les applications, on ne connaît jamais ε. En assimilant ε à une V.A. uniformément répartie sur $[0,1]$, on voit que la "valeur moyenne" de cette variance - valeur qu'il est raisonnable d'utiliser en pratique - est :

$$\tilde{\sigma}_E^2 = \frac{5}{18}\,\frac{a}{n}$$

Dans le cas d'une maille à implantation préférentielle (dispositif centré ou fermé), la variance d'estimation est $\frac{1}{6}\,\frac{a}{n}$: la majoration due à l'effet de bordure est donc égale aux deux tiers du terme principal et ne peut donc pas être négligée.

Dans ma thèse ([1], p. 209), j'avais bien prévu l'existence de cet effet de bordure, et donné son ordre de grandeur à l'aide d'un raisonnement approximatif valable d'ailleurs seulement dans le cas où il existe une covariance et non pas seulement un variogramme. Cet ordre de grandeur est $\sigma^2 \frac{\sigma_V^2}{V^2}$, $\frac{\sigma_V^2}{V^2}$ représentant la variance relative de l'estimation du champ géométrique, et σ^2 la variance de la F.A. stationnaire. A dire vrai, la nature même de ce raisonnement approximatif ne permettait pas de préciser si σ^2 représente effectivement la variance a priori $D^2(Z_x^2)$ de la FA, ou sa variance $\sigma^2(O|V)$ dans le champ V à estimer. C'est seulement dans le cas où l'on retient cette dernière interprétation que l'on peut espérer transposer le résultat obtenu au cas où il n'existe qu'un variogramme. L'effet de bordure serait alors $F(V)\ \frac{\sigma_V^2}{V^2}$. Dans le cas particulier traité ci-dessus, on a $\sigma_V^2/V^2 = 1/6\ n^2$ et $F(V) = L/3$, ce qui conduirait à la valeur $\frac{1}{18}\ \frac{a}{n}$ au lieu de la valeur correcte $\frac{2}{18}\ \frac{a}{n}$. L'erreur est du simple au double, mais cependant l'ordre de grandeur est correct, puisque l'on trouve un terme en $1/n$.

2 - LA FORMULE D'EULER-Mac LAURIN

Les formules rigoureuses donnant les variances d'estimation font apparaître, comme on sait, des différences entre valeurs exacte et approchée de certaines intégrales. On peut donc obtenir des expressions approchées en utilisant une version modifiée de la formule classique d'Euler. Je vais donc commencer par établir cette formule classique sous une forme un peu plus générale que la forme habituelle.

Soit f une fonction définie sur l'intervalle $[0,L]$ et y admettant des dérivées continues jusqu'à l'ordre n+1. Aux points 0 et L, l'existence et la continuité de ces dérivées ne sont requises qu'à droite et à gauche respectivement. Sous ces hypothèses, f(x) admet un développement en série de Fourier convergente :

(2-1) $$f(x) = \sum_{p=-\infty}^{+\infty} C_p\ e^{2ip\pi \frac{x}{L}} \qquad (0 < x < L)$$

En $x = 0$ comme en $x = L$, la série de Fourier vaut $\frac{f(o) + f(L)}{2}$. Le coefficient C_o est la valeur moyenne de f sur $(0,L)$, soit :

$$C_o = \frac{1}{L} \int_0^L f(x)\, dx$$

Pour $p \neq 0$, on trouve, en effectuant une intégration par partie :

$$C_p = \frac{1}{L} \int_0^L f(x)\, e^{-2i\pi p \frac{x}{L}} dx = - \frac{1}{L} \frac{L}{2i\pi p} \left(f(L)-f(0)\right) +$$

$$+ \frac{1}{L} \frac{L}{2i\pi p} \int_0^L f'(x)\ e^{-2i\pi p \frac{x}{L}} dx$$

et, en réitérant cette opération :

(2-3) $$C_p = - \frac{1}{L} \sum_{k=0}^{n} \left(\frac{L}{2i\pi p}\right)^{k+1} \left(f^{(k)}(L) - f^{(k)}(o)\right) +$$

$$+ \frac{1}{L} \left(\frac{L}{2i\pi p}\right)^{n+1} \int_0^L f^{(n+1)}(x)\ e^{-2i\pi p \frac{x}{L}} dx$$

Pour abréger les notations, posons :

(2-4) $$\begin{cases} \Delta_k = f^{(k)}(L) - f^{(k)}(o) \\ \\ \rho_p = \int_0^L f^{(n+1)}(x)\ e^{-2i\pi px} dx \end{cases}$$

en notant la majoration :

(2-4') $$|\rho_p| \leq \int_0^L |f^{(n+1)}(x)|\ dx$$

D'après les hypothèses faites sur la fonction f, la dérivée $f^{(n+1)}$ est continue, et admet donc un développement en série de Fourier convergent, dont les coefficients sont les $L\,\rho_p$. En portant (2-3) dans (2-1), on obtient donc un développement convergent, égal à f(x) pour $0 < x < L$, et à $\frac{1}{2}\,[f(o) + f(L)]$ en $x = 0$, ou en $x = L$.

Pour $0 < x < L$, donc, on peut écrire :

$$(2\text{-}5)\quad \left\{ \begin{aligned} f(x) &= \frac{1}{L}\int_0^L f(x)\,dx - \frac{1}{L}\sum_{k=0}^{n}\sum_{p\neq 0}\left(\frac{L}{2i\pi p}\right)^{k+1}\Delta_k\, e^{2i\pi p\frac{x}{L}} \\ &+ \frac{1}{L}\sum_{p\neq 0}\left(\frac{L}{2i\pi p}\right)^{n+1}\rho_p\, e^{2i\pi p\frac{x}{L}} \end{aligned} \right.$$

Soit maintenant N un entier > 0, et posons :

$$a = \frac{L}{N}$$

Prenons un point x_o sur $(0,a)$, soit : $0 < x_o < a$, et proposons-nous de calculer l'expression :

$$S_N = S_N\,(x_o;a,f) = a\sum_{k=0}^{N-1} f(x_o+ka) - \int_0^L f(x)\,dx$$

Pour celà, nous pouvons utiliser (2-5), en notant pour $p \neq 0$:

$$\sum_{k=0}^{N-1} e^{2i\pi p\frac{x_o+ka}{L}} = \begin{cases} 0 \text{ si } p \text{ non multiple de } N \\ N\, e^{2i\pi p\frac{x_o}{L}} \text{ si } p \text{ multiple de } N \end{cases}$$

D'après (2-5), on trouve donc :

$$(2\text{-}6) \qquad S_N = - \sum_{k=0}^{n} \Delta_k \sum_{p\neq 0} \left(\frac{a}{2i\pi p}\right)^{k+1} e^{2i\pi p \frac{x_o}{a}} + R_N$$

avec

$$R_N = \sum_{p\neq 0} \left(\frac{a}{2i\pi p}\right)^{n+1} \rho_p \; e^{2i\pi p \frac{x_o}{a}}$$

Compte tenu de (2-4'), ce reste R_N admet la majoration :

$$R_N \leq 2\, a^{n+1} \sum_{p=1}^{\infty} \frac{1}{(2\pi p)^{n+1}} \int_0^L |f^{(n+1)}(x)|\,dx$$

On voit que S_N va s'exprimer de manière naturelle au moyen des nombres et des polynomes de Bernouilli. Pour λ réel > 1, on pose :

$$(2\text{-}7) \quad \begin{cases} B_\lambda = 2\,\Gamma(1+\lambda) \displaystyle\sum_{p=1}^{\infty} \frac{1}{(2\pi p)^\lambda} \\ \\ B_\lambda(x) = 2\,\Gamma(1+\lambda) \displaystyle\sum_{p=1}^{\infty} \frac{\cos\,(2\pi p x - \lambda \frac{\pi}{2})}{(2\pi p)^\lambda} \end{cases}$$

B_λ est le nombre de Bernouilli d'indice λ ($B_2 = \frac{1}{6}$, $B_4 = \frac{1}{30}$, $B_6 = \frac{1}{42}$, $B_8 = \frac{1}{30}$, etc...), et $B_\lambda(x)$ la fonction de Bernouilli correspondante. En $x = 0$, on a évidemment :

$$B_\lambda(0) = B_\lambda \cos \lambda \frac{\pi}{2}$$

Lorsque λ est entier, $B_\lambda(x)$ coïncide sur $(0,1)$ avec un polynome de degré λ, appelé encore polynome de Bernouilli. Pour $\lambda = 1$, le nombre de Bernouilli n'existe pas, mais la fonction $B_1(x)$ existe :

$$B_1(x) = 2 \sum_{p=1}^{\infty} \frac{\sin 2\pi px}{2\pi p} = \frac{1}{2} - x$$

La seconde égalité n'est valable que pour $0 < x < 1$. Par contre, les relations suivantes sont vraies pour $0 \leq x \leq 1$.

$$B_2(x) = -\frac{1}{6} + x(1-x)$$

$$B_3(x) = -\frac{x}{2} + \frac{3x^2}{2} - x^3$$

$$B_4(x) = \frac{1}{30} - x^2 + 2x^3 - x^4$$

Compte tenu de (2-7), on voit que (2-6) peut s'écrire :

$$\text{(2-8)} \quad \begin{cases} S_N(x_o;\ a;\ f) = a \sum_{k=0}^{N-1} f(x_o+ka) - \int_o^L f(x)\,dx \\ \\ = -\sum_{k=1}^{n+1} \frac{a^k}{k!} B_k\left(\frac{x_o}{a}\right) [f^{(k-1)}(L) - f^{(k-1)}(o)] + R_N \end{cases}$$

avec la majoration

$$\text{(2-8')} \qquad |R_N| \leq a^{n+1} \frac{B_{n+1}}{(n+1)!} \int_o^L |f^{(n+1)}(x)|\,dx$$

En remarquant que, pour $0 < x_o < a$, $B_1\left(\frac{x_o}{a}\right)$ vaut $\frac{1}{2} - \frac{x_o}{a}$, on obtient une variante de (2-8) en faisant passer dans le premier membre le terme d'ordre $k = 1$, soit :

$$a\left(\frac{1}{2}-\frac{x_o}{a}\right)[f(L)-f(o)]+a\sum_{k=0}^{N-1} f(x_o+ka)-\int_0^L f(x)\,dx =$$

$$= -\sum_{k=2}^{n+1}\frac{a^k}{k!}\,B_k\left(\frac{x_o}{a}\right)[f^{(k-1)}(L)-f^{(k-1)}(o)]+R_N$$

Sous cette forme, la relation reste valable pour $x = 0$ ou $x = L$ et donne :

$$(2\text{-}9)\quad \begin{cases} a\left[\frac{1}{2}f(o)+f(a)+\ldots+f((N-1)a)+\frac{1}{2}f(Na)\right]-\int_0^{Na} f(x)\,dx = \\ -\sum_{k=1}^{n'}(-1)^k\,\frac{a^{2k}\,B_{2k}}{(2k)!}\left[f^{(2k-1)}(L)-f^{(2k-1)}(o)\right]+R_N \end{cases}$$

(n' désigne le plus grand entier tel que $2n' \leq n+1$. En général, on supposera n pair, d'où $n = 2\,n'$).

Ces formules d'Euler (2-8) et (2-9) permettent le calcul des variances d'estimation de la géostatistique si le variogramme ou le covariogramme utilisé admet des dérivées continues jusqu'à l'ordre $n+1$ en tout point x tel que $0 < x < L$, et si de plus en $x = 0$ et en $x = L$ ces mêmes dérivées existent encore à droite en $x = 0$ et à gauche en $x = L$. En particulier, donc, ces formules sont directement utilisables lorsque la partie irrégulière ne comporte que des termes <u>de degrés impairs</u>.

<u>Exemple</u> - Par exemple, pour $f(x) = e^{-x}$, la formule (2-9) donne :

$$\frac{a}{1-e^{-a}}-\frac{a}{2}-1=\sum_{k=1}^{n'}(-1)^{k+1}\frac{B_{2k}}{(2k)!}\,a^{2k}+R$$

$$|R| \leq a^{2n'+1}\,\frac{B(2n'+1)}{(2n'+1)}$$

Pour a < 2π, d'après (2-7), ce reste tend vers 0 si n' → ∞, d'où la relation :

$$\frac{a}{1 - e^{-a}} - \frac{a}{2} - 1 = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{a^{2k}}{(2k)!} B_{2k}$$

qui sert parfois de définition aux nombres de Bernouilli d'indice pair. On en déduit aussi la variance d'estimation transitive associée au covariogramme $g(h) = e^{-b|h|}$:

$$a \sum_{k=-\infty}^{+\infty} e^{-|k|ba} - \int_{-\infty}^{+\infty} e^{-b|h|} dh = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{a^{2k} b^{2k-1}}{(2k)!} B_{2k}$$

$$(ab < 2\pi)$$

Le coefficient du terme en h^{2k-1} dans le développement de e^{-bh} étant $b^{2k-1}/(2k-1)!$, on retrouve sur le développement précédent la règle de correspondance qui associe à un terme en h^{2k-1} dans le développement de g(h) autour de h = 0 le terme $(-1)^{k+1}\ 2\ a^{2k}\ B_{2k}/2k$.

Toutefois, ce raisonnement ne justifie pas la règle de correspondance elle-même, et d'autre part ne donne pas le terme relatif à $|h|^{\alpha}$ pour α non entier.

3 - LA REGLE DE CORRESPONDANCE

Nous allons maintenant considérer une fonction g continue sur la demi-droite positive et vérifiant de plus les deux conditions suivantes :

a/ en tout $x > 0$, g admet des dérivées continues jusqu'à l'ordre $n + 1$ (mais <u>les dérivées n'existent pas nécessairement en $x = 0$</u>)

b/ Pour $c > 0$, on a :

$$\int_c^{\infty} |g^{(n+1)}(x)| \, dx < \infty$$

Exemple : pour $0 \leq \alpha < n$, la fonction $g(x) = x^{\alpha}$ vérifie ces deux conditions, même si α n'est pas entier.

Nous nous proposons d'établir que les limites suivantes existent : pour $0 < x_o < a$,

$$S(x_o;a) =$$

(3-1)

$$= \lim_{N \to \infty} \left[a \sum_{k=0}^{N-1} g(x_o+ka) - \int_0^{Na} g(x) \, dx + \sum_{k=1}^{n+1} \frac{a^k}{k!} \, g^{(k-1)}(Na) \, B_k\left(\frac{x_o}{a}\right) \right]$$

et pour $x_o = 0$, avec $n = 2n'$ entier pair :

$$(3\text{-}1')\left\{\begin{aligned} S(a) &= \lim_{N \to \infty} \Big[a \, [\tfrac{1}{2} \, g(o)+g(a)+\dots+g((N-1)a)+ \tfrac{1}{2} \, g(Na) \Big| - \int_0^{Na} g(x)dx \\ &\quad - \frac{a^2}{2!} \, B_2 \, g'(Na) + \dots + (-1)^{n'} \frac{a^{2n'}}{(2n')!} \, B_{2n'} \, g^{(2n'-1)}(Na) \Big] \end{aligned}\right.$$

Nous chercherons ensuite à montrer que le principe de correspondance s'applique pour $S(x_o;a)$ et $S(a)$.

Pour établir (3-1), par exemple, posons :

$$S_N = a \sum_{k=0}^{N-1} g(x_o+ka) - \int_o^{Na} g(x)\, dx + \sum_{k=1}^{n+1} \frac{a^k}{k!}\, g^{(k-1)}(Na)\, B_k\left(\frac{x_o}{a}\right)$$

D'après le critère de Cauchy, il faut montrer que $S_N - S_{N'}$ tend vers 0 lorsque N et N' tendent vers l'infini. Soit alors m un entier > 0, et considérons l'expression $S_{N+m} - S_m$. Appliquons la formule (2-8) à la fonction $f(x) = g(x+ma)$, qui vérifie les conditions voulues, d'après a/. Comme $f^{(k)}(Na) - f^{(k)}(o) = g^{(k)}\big((N+m)a\big) - g^{(k)}(ma)$, (2-8) nous donne :

$$S_{N+m} - S_n = R_n'$$

et, d'après (2-8'), on a la majoration :

$$|S_{N+m} - S_m| \leq \frac{a^{n+1}}{(n+1)!}\, B_{n+1} \int_{ma}^{(N+m)a} |g^{(n+1)}(x)|\, dx$$

D'après la condition b/, cette expression tend vers 0 uniformément en N pour $m \to \infty$. Ainsi est établie l'existence de la limite (3-1). Celle de (3-1') s'établit de la même manière.

Passons maintenant à la règle de correspondance. Il est clair, en premier lieu, que les fonctions g vérifiant les conditions a/ et b/ ci-dessus constituent un espace vectoriel, que nous désignerons par $\mathcal{G}_n$. D'autre part, S(a) et $S(x_o;a)$ à a et x_o fixés, définissent manifestement des fonctionnelles linéaires sur $\mathcal{G}_n$. En troisième lieu, si une fonction $f \in \mathcal{G}_n$ admet en x = 0 des dérivées à droite d'ordre 1, 2,...n continues à droite, les formules (2-8) et (2-9) donnent aussitôt :

$$S(x_o;a;f) = \sum_{k=1}^{n+1} \frac{a^k f^{(k-1)}(o)}{k!} B_k\left(\frac{x_o}{a}\right) + R$$

$$S(a;f) = \sum_{k=1}^{n'} (-1)^k \frac{a^{2k}}{(2k)!} B_{2k} f^{(2k-1)}(o) + R \qquad (n = 2n')$$

avec les majorations :

$$|R| \leq a^{n+1} \frac{B_{n+1}}{(n+1)!} \int_o^{\infty} |f^{(n+1)}(x)| \, dx$$

Ainsi, si la fonction $f \in \mathcal{G}_n$ a toutes ses dérivées jusqu'à l'ordre $n+1$ nulles en $x = 0$, $S(x_o,a,f)$ est un infiniment petit d'ordre $n+1$ en a. Pour $S(a,f)$, il suffit même que seules les dérivées d'ordre pair s'annulent en $x = 0$ pour que $S(a,f)$ soit de l'ordre de $a^{2n'+1}$.

D'après le caractère linéaire de ces fonctionnelles, on voit que la relation :

$$g = \sum_{i=1}^{p} \lambda_i g_i + f$$

avec g_i et $f \in \mathcal{G}_n$, et f admettant des dérivées continues à droite et nulles en $x = 0$, entraîne :

$$S(x_o;a;g) = \sum_{i=1}^{p} \lambda_i S(x_o;a;g_i) + O(a^{n+1})$$

De même, en supposant seulement les dérivées paires de f nulles en $x = o$:

$$S(a;g) = \sum_{i=1}^{p} \lambda_i S(a;g_i) + O(a^{2n+1})$$

Tel est l'énoncé général du principe de correspondance. Dans les applications, on utilise surtout des fonctions g_i de la forme x^α $(0 \leq \alpha < n)$ ou $x^\alpha \log x$ $(0 < \alpha < n)$. Nous poserons :

$$(3\text{-}2)\qquad \begin{cases} S_\alpha\,(x_o;a) = S(x_o;a;x^\alpha) \\[2ex] S_\alpha\,(a) = S(a;x^\alpha) \end{cases}$$

Il reste à expliciter le calcul de ces expressions.

<u>Calcul des S_α</u>.- Pour $0 \leq \alpha < n$, il est visible sur les formules de définition que l'on a :

$$\begin{cases} S_\alpha(x_o;a) = S_\alpha\left(\dfrac{x_o}{a}\right)\; a^{1+\alpha} \\[2ex] S_\alpha\,(a) = S_\alpha\; a^{1+\alpha} \end{cases}$$

avec des coefficients $S_\alpha\left(\dfrac{x_o}{a}\right)$ et S_α indépendants de a (le premier dépendant seulement su rapport $\varepsilon = x_o/a$, $0 < \varepsilon < 1$). Calculons donc ces coefficients, en procédant par identification à partir du résultat explicite obtenu pour une fonction particulière. Le plus simple semble être de prendre la fonction g définie par :

$$g(x) = x^\alpha\, e^{-bx} \qquad\qquad (0 \leq x < \infty)$$

La transformée de Fourier de g(x) (en supposant g(x) = 0 pour x < 0) est :

$$G(u) = \int_0^\infty e^{-2i\pi ux}\, x^\alpha\, e^{-bx} dx = \frac{\Gamma(1+\alpha)}{(b + 2i\pi u)^{1+\alpha}}$$

La formule sommatoire de Poisson donne ensuite :

$$a \sum_{k=-\infty}^{+\infty} g(ka) - \int_{-\infty}^{+\infty} g(h)\,dh = \sum_{p=1}^{\infty} \left[\frac{\Gamma(1+\alpha)}{(b + 2i\pi \frac{p}{a})^{1+\alpha}} + \frac{\Gamma(1+\alpha)}{(b - 2i\pi \frac{p}{a})^{1+\alpha}} \right]$$

$$= 2 \sum_{p=1}^{\infty} \left[\frac{\Gamma(1+\alpha) \cos \frac{1+\alpha}{2}\pi}{(2\pi p)^{1+\alpha}} a^{1+\alpha} - \frac{b}{1!} \frac{\Gamma(2+\alpha) \cos \frac{2+\alpha}{2}\pi}{(2\pi p)^{2+\alpha}} a^{2+\alpha} + \dots \right]$$

A cause de la décroissance rapide à l'infini de $x^{\alpha}\, e^{-bx}$ et de toutes ses dérivées, le premier membre est $S(a;g)$. Par identification, on trouve donc :

$$(3\text{-}3) \qquad S_{\alpha} = 2\,\Gamma(1+\alpha) \cos \frac{1+\alpha}{2}\pi \sum_{p=1}^{\infty} \frac{1}{(2\pi p)^{1+\alpha}} = \frac{1}{1+\alpha} B_{1+\alpha}(0)$$

Ce coefficient est donc lié au nombre de Bernouilli d'ordre $1 + \alpha$, soit :

$$S_{\alpha} = \frac{1}{1+\alpha} \cos\left(\frac{1+\alpha}{2}\pi\right) B_{1+\alpha}$$

Lorsque $g(h)$ est un covariogramme transitif, donc une fonction symétrique définie sur la droite entière, la formule (3-3) permet le calcul de la variance d'estimation :

$$\sigma^2(a) = a \sum_{k=-\infty}^{+\infty} g(ka) - \int_{-\infty}^{+\infty} g(h)\,dh$$

Il suffit, en effet, de poser $g_{+}(x) = g(x)$ si $x \geq 0$, $g_{+}(x) = 0$ si $x < 0$ pour obtenir

$$\sigma^2(a) = 2\,S(a,g_{+})$$

pourvu que g_+ soit dans $\mathcal{C}_n$ et que ses dérivées décroissent suffisamment vite à l'infini. En particulier, le terme $T_\alpha \, a^{1+\alpha}$ associé dans $\sigma^2(a)$ à un terme $|h|^\alpha$ figurant dans $g(h)$ est défini par le coefficient:

$$T_\alpha = 2 \, S_\alpha = \frac{2}{1+\alpha} \, B_{1+\alpha} \cos \frac{1+\alpha}{2} \pi$$

(résultat bien connu)

Passons maintenant au calcul de la fonction $S_\alpha(\varepsilon)$ $(0 < \varepsilon < 1)$, avec $\varepsilon = \frac{x_o}{a}$. Nous allons ici encore procéder par identification. Posons, à nouveau :

$$g(x) = \begin{cases} x^\alpha \, e^{-bx} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

Evaluons $S_\alpha(x_o;g)$. La décroissance rapide de g et de ses dérivées montre que l'on a

$$S_\alpha(x_o;g) = a \sum_{k=0}^{\infty} g(x_o+ka) - \int_0^\infty g(x) \, dx$$

Considérons alors la fonction f définie par :

$$f(x) = g(x_o+x) = \begin{cases} (x_o+x)^\alpha \, e^{-bx} & (x \geq -x_o) \\ 0 & (x < -x_o) \end{cases}$$

Pour $0 < x_o < a$, on a clairement encore :

$$S_\alpha \, (x_o;g) = a \sum_{k=-\infty}^{+\infty} f(ka) - \int_{-\infty}^{+\infty} f(x) \, dx$$

La transformée de Fourier de f est :

$$F(u) = e^{2i\pi u x_0}\ G(u) = \frac{\Gamma(1+\alpha)\ e^{2i\pi u x_0}}{(b + 2i\pi u)^{1+\alpha}}$$

Il reste alors à appliquer la formule sommatoire de Poisson :

$$S_\alpha(x_0;g) = \sum_{p=1}^{\infty} \left[F\left(\frac{p}{a}\right) + F\left(-\frac{p}{a}\right)\right]$$

pour obtenir par un calcul identique à celui qui a été effectué ci-dessus :

$$(3\text{-}4) \qquad S_\alpha(\varepsilon) = 2 \sum_{p=1}^{\infty} \frac{\Gamma(1+\alpha)}{(2\pi p)^{1+\alpha}} \cos\left[2\pi p\varepsilon - (1+\alpha)\frac{\pi}{2}\right] = \frac{1}{1+\alpha}\ B_{1+\alpha}(\varepsilon)$$

Cette relation (3-4) permet de retrouver sans peine l'expression du Zitterbewegung qui figure dans ma thèse (Formule (F-6), p. 301).

Pour les covariogrammes transitifs usuels, dont les singularités sont localisées en h = 0 et (en cas de Zitterbewegung) en h = b, b désignant la portée, les résultats de ce paragraphe justifient entièrement les procédés usuels de calcul des variances d'estimation transitive. Examinons maintenant le cas des variances d'estimation de la théorie intrinsèque.

4 - LES VARIANCES D'ESTIMATION DE LA THEORIE INTRINSEQUE

Soit γ un variogramme, χ et F les fonctions auxiliaires habituelles. Nous supposons $\gamma \in \mathcal{G}_n$, quitte à introduire ultérieurement des hypothèses complémentaires, au fur et à mesure qu'elles seront requises. Rappelons d'abord la formule générale donnant la variance d'estimation σ_E^2 de la teneur moyenne de l'intervalle (0,L) lorsque l'on utilise comme estimateur la teneur moyenne de n échantillons implantés en des points x_i appartenant à (0,L) :

$$(4\text{-}1) \quad \sigma_E^2 = -\frac{1}{N^2}\sum_{i,j}\gamma(x_i-x_j) - F(L) + \frac{2}{N}\sum_{i=1}^{N}\frac{x_i\ \chi(x_i)+(L-x_i)\ \chi(L-x_i)}{L}$$

Nous examinerons le cas des mailles régulières à implantation préférentielle (centrée ou fermée) puis à implantation flottante, avec comme objectif, dans ce dernier cas, de mettre en évidence l'influence de l'effet de bordure.

Posons d'abord un lemme. Pour $g \in \mathcal{G}_n$ et $N > 0$, on a :

$$(4\text{-}2) \quad \begin{cases} a\,[\frac{1}{2}\,g(o) + g(a) + \ldots + g((N-1)a) + \frac{1}{2}\,g(Na)] - \displaystyle\int_0^{Na} g(x)\,dx = \\ \\ S(a;g) + \frac{a^2}{2!}\,B_2\,g'(Na)\ldots - (-1)^{n'}\,\frac{a^{2n'}}{(2n')!}\,B_{2n'}\,g^{(2n'-1)}(Na) + R \end{cases}$$

avec $n = 2n' - 1$ et la majoration :

$$(4\text{-}2') \qquad |R| \leq \frac{a^{2n'+1}}{(2n'+1)!}\,B_{2n'+1}\int_{Na}^{\infty}|g^{(2n'+1)}(x)|\,dx$$

Celà résulte aussitôt, en effet, du raisonnement même qui nous a permis d'établir l'existence de la limite S(a) dans la relation (3-1'). De la même manière, pour $0 < x_0 < a$, on trouve :

$$(4\text{-}3)\quad \left\{ \begin{aligned} & a \sum_{k=0}^{N-1} g(x_0+ka) - \int_0^{Na} g(x)\,dx = \\ & = S(x_0;a;g) - \sum_{k=1}^{n+1} \frac{a^k}{k!}\, g^{(k-1)}(Na)\, B_k\left(\frac{x_0}{a}\right) + R' \\ & |R'| \leq \frac{a^{n+1}}{(n+1)!}\, B_{n+1} \int_{Na}^{\infty} |g^{(n+1)}(x)|\,dx \end{aligned} \right.$$

Ces formules ont l'intérêt de séparer nettement l'influence des irrégularités à l'origine (terme S(a) ou $S(x_0;a)$, auquel s'applique la règle de correspondance établie ci-dessus) et la partie de l'erreur qui se trouve "localisée" à l'extrémité droite de l'intervalle (0,Na) (valeurs en Na des dérivées de la fonction g). Ces formules généralisent donc les résultats classiques (2-8) et (2-9) au cas d'une fonction présentant des irrégularités à l'origine.

4-1 Maille régulière, dispositif centré.

Si a est la maille, on a ici $L = Na$ et $x_i = ia - \frac{a}{2}$. La formule (4-1) s'écrit :

$$\sigma_E^2 = \frac{4}{NL} \sum_{i=1}^{N} (i-\tfrac{1}{2})a\,\chi(ia-\tfrac{a}{2}) - \frac{2}{N^2} \sum_{i=1}^{N-1} (N-i)\,\gamma(ia) - F(L)$$

Il est commode de composer cette expression en 3 termes que nous évaluerons séparément :

$$
(4\text{-}4)\quad
\begin{cases}
\sigma_E^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 \\[2mm]
\sigma_1^2 = -\frac{2}{N}\left[\frac{1}{2}\gamma(o) + \gamma(a) + \ldots + \gamma\big((N-1)a\big) + \frac{1}{2}\gamma(Na)\right] + \\[2mm]
\qquad + \frac{2}{Na}\displaystyle\int_o^{Na} \gamma(h)\,dh \\[2mm]
\sigma_2^2 = \frac{2}{N^2}\left[1\,\gamma(a) + \ldots + (N-1)\,\gamma\big((N-a)a\big) + \frac{1}{2}N\,\gamma(Na)\right] - \\[2mm]
\qquad - \frac{2}{N^2a^2}\displaystyle\int_o^{Na} h\,\gamma(h)\,dh \\[2mm]
\sigma_3^2 = \frac{4}{N^2}\displaystyle\sum_{i=1}^{N} \left(i-\frac{1}{2}\right)\chi\left(ia-\frac{a}{2}\right) - \frac{4}{L^2}\int_o^{Na} x\,\chi(x)\,dx
\end{cases}
$$

D'après la formule (4-2), le premier terme vaut :

$$
(4\text{-}5)\quad
\begin{cases}
\sigma_1^2 = -\frac{2}{Na}\left[S(a,\gamma) + \frac{a^2}{2!}B_2\,\gamma^1(Na) - \ldots - (-1)^{n'}\frac{a^{2n'}}{(2n')!}B_{2n'}\gamma^{(2n'-1)}(Na)\right] + R_1 \\[2mm]
|R_1| \le \frac{2}{N}\; a^{2n'}\; B_{2n'+1}\displaystyle\int_{Na}^{\infty}|\gamma^{(2n')}(x)|\,dx
\end{cases}
$$

Le deuxième terme s'obtient en appliquant le même algorithme à la fonction $h\,\gamma(h)$, pourvu que celle-ci soit également dans $\mathcal{C}_{2n'-1}$, ce que nous supposerons. Il vient ainsi, en écrivant $\gamma_1(h)$ pour $h\,\gamma(h)$:

$$
(4\text{-}5')\quad
\begin{cases}
\sigma_2^2 = \frac{2}{N^2a^2}\left[S(a,\gamma_1) + \frac{a^2}{2!}B_2\,\gamma_1'(Na) - \ldots - (-1)^{n'}\frac{a^{2n'}}{(2n')!}B_{2n'}\gamma_1^{(2n'-1)}(Na)\right] + R_2 \\[2mm]
|R_2| \le \frac{2}{N^2}\; a^{2n'-1}\; B_{2n'+1}\displaystyle\int_{Na}^{\infty}|\gamma_1^{(2n')}(x)|\,dx
\end{cases}
$$

Si la fonction $x\,\chi(x) = \chi_1(x)$ est aussi dans $\mathcal{C}_n$, enfin, on peut évaluer le troisième terme à l'aide de la formule (4-3) écrite pour $x_o = \frac{a}{2}$. Il vient alors :

$$
(4\text{-}5'') \quad \left\{
\begin{aligned}
&\sigma_3^2 = \frac{4}{N^2a^2}\left[S(\tfrac{a}{2};a;\chi_1) - \sum_{k=1}^{n+1} a^k\, \chi_1^{(k-1)}(Na)\, B_k(\tfrac{1}{2})\right] + R_3 \\
&|R_3| \leq \frac{4}{N^2a^2}\,\frac{a^{n+1}}{(n+1)!}\, B_{n+1}\int_{Na}^{\infty} |\chi_1^{(n+1)}(x)|\,dx
\end{aligned}
\right.
$$

En nous reportant à (2-7), nous trouvons :

$$
B_k(\tfrac{1}{2}) = \frac{2\,\Gamma(1+k)}{(2\pi)^k}\sum_{p=1}^{\infty}\frac{\cos(\pi p - k\frac{\pi}{2})}{p^k}
$$

Ce terme est donc nul pour k impair

$$
B_{2k+1}(\tfrac{1}{2}) = 0
$$

et pour un entier pair 2k, on trouve :

$$
B_{2k}(\tfrac{1}{2}) = \frac{2\,\Gamma(1+2k)}{(2\pi)^{2k}}(-1)^k\sum_{p=1}^{\infty}\frac{(-1)^p}{p^{2k}}
$$

d'où l'on déduit

$$
B_{2k}(\tfrac{1}{2}) = (-1)^{k+1}\left(1 - \frac{1}{2^{2k-1}}\right) B_{2k}
$$

Par suite, (4-5") devient :

$$
(4\text{-}6) \quad \sigma_3^2 = \frac{4}{N^2a^2}\left[S(\tfrac{a}{2};a;\chi_1) - \sum_{k=1}^{n'}\frac{a^{2k}}{(2k)!}\,\chi_1^{(2k-1)}(Na)\left(1-\frac{1}{2^{2k-1}}\right)(-1)^{k+1}B_{2k}\right] + R_3
$$

Il reste à recoller les morceaux. Ce regroupement fait apparaître deux termes et un reste. Le premier terme, lié uniquement au comportement de γ autour de $h = 0$, s'exprime à l'aide des fonctions S. Explicitement, ce premier terme s'écrit :

$$\text{(4-7)} \qquad \sigma^2_{E_o} = - \frac{2}{Na} \; S(a,\gamma) + \frac{2}{N^2a^2} \; S(a,\gamma_1) + \frac{4}{N^2a^2} \; S(\frac{a}{2}\, ,a,\chi_1)$$

D'après ce que l'on a vu au paragraphe précédent, le principe de correspondance est valable pour le calcul de $\sigma^2_{E_o}$. A un terme h^α figurant dans le développement de $\gamma(h)$ répondent les termes $h^{1+\alpha}$ et $h^{1+\alpha}/(\alpha+1)$ dans les développements de γ_1 et χ_1 respectivement, et par suite le terme :

$$T_\alpha \frac{a^\alpha}{N} + T'_\alpha \; \frac{a^\alpha}{N^2}$$

dans l'expression de $\sigma^2_{E_o}$, avec :

$$\begin{cases} T_\alpha = - 2\, S_\alpha \\ \\ T'_\alpha = 2\, S_{1+\alpha} + \dfrac{4}{\alpha+1} \; S_{1+\alpha}\,(\frac{1}{2}) \end{cases}$$

Ces coefficients, à leur tour, s'expriment à l'aide des nombres de Bernouilli d'indice $1+\alpha$ et $2+\alpha$. On a déjà vu, en effet, que S_λ vaut :

$$S_\lambda = \frac{1}{1+\lambda} \; B_{1+\lambda} \cos \left(\frac{1+\lambda}{2} \pi\right)$$

Pour calculer $S_{1+\alpha}\,(\frac{1}{2})$, soit

$$S_{1+\alpha}\,(\frac{1}{2}) = \frac{1}{2+\alpha} \;\; B_{\alpha+2}\,(\frac{1}{2})$$

Calculons d'abord la valeur en $x = \frac{1}{2}$ de la fonction de Bernouilli d'indice λ, soit, d'après (2-7) :

$$B_\lambda\left(\tfrac{1}{2}\right) = 2\,\Gamma(1+\lambda)\cos\frac{\lambda\pi}{2}\sum_{p=1}^{\infty}\frac{(-1)^p}{(2\pi p)^\lambda}$$

Mais les relations

$$\sum_{p=1}^{\infty}\frac{(-1)^p}{p^\lambda} = \frac{1}{2^\lambda}\sum_{p=1}^{\infty}\frac{1}{p^\lambda} - \sum_{p=1}^{\infty}\frac{1}{(2p-1)^\lambda}$$

$$\Sigma\,\frac{1}{p^\lambda} = \frac{1}{2^\lambda}\,\Sigma\,\frac{1}{p^\lambda} + \Sigma\,\frac{1}{(2p-1)^\lambda}$$

donnent :

$$\Sigma\,\frac{1}{(2p-1)^\lambda} = (1-2^{-\lambda})\,\Sigma\,\frac{1}{p^\lambda}$$

puis :

$$\Sigma\,\frac{(-1)^p}{p^\lambda} = (2^{1-\lambda}-1)\,\Sigma\,\frac{1}{p^\lambda}$$

On en déduit aussitôt :

$$B_\lambda\left(\tfrac{1}{2}\right) = (2^{1-\lambda}-1)\,B_\lambda(o) = (2^{1-\lambda}-1)\,B_\lambda\cos\frac{\lambda\pi}{2}$$

D'où l'expression de $S_{1+\alpha}\left(\frac{1}{2}\right)$:

$$S_{1+\alpha}\left(\tfrac{1}{2}\right) = \frac{B_{\alpha+2}}{\alpha+2}\left(1-\frac{1}{2^{1+\alpha}}\right)\cos\frac{\alpha\pi}{2}$$

D'où finalement les coefficients T_α et T'_α du principe de correspondance :

$$(4-8)\quad \begin{cases} T_\alpha = - \dfrac{2\ B_{1+\alpha}}{1+\alpha} \cos\left(\dfrac{1+\alpha}{2}\pi\right) \\[2ex] T'_\alpha = \dfrac{2\ B_{2+\alpha}}{(\alpha+1)(\alpha+2)} (1 - \alpha - 2^{-\alpha}) \cos \dfrac{\alpha\pi}{2} \end{cases}$$

Après ce premier terme $\sigma^2_{E_0}$ donné par (4-7) et (4-8), terme qui prend en charge l'effet du comportement de γ en $h = 0$, apparaît un second terme $\sigma^2_{E_1}$, lié au comportement de γ en $h = Na$. Pour obtenir ce terme, évaluons le coefficient de B_{2k} dans la somme $\sigma_1^2 + \sigma_2^2 + \sigma_3^2$.

En nous reportant à (4-5), (4-5') et (4-6), et en notant :

$$\chi_1^{(k)}(h) = \gamma^{(k-1)}(h)$$

$$\gamma_1^{(k)}(h) = k\ \gamma^{(k-1)}(h) + h\ \gamma^{(k)}(h)$$

nous trouvons que le coefficient de B_{2k} est :

$$\frac{2}{Na} (-1)^k \frac{a^{2k}}{(2k)!} \gamma^{(2k-1)} - \frac{2}{N^2a^2} (-1)^k \left[(2k-1)\ \gamma^{(2k-2)} + Na\ \gamma^{(2k-1)}\right] \frac{a^{2k}}{(2k)!}$$

$$+ \frac{4}{N^2a^2} \left(1 - \frac{1}{2^{2k-1}}\right) \frac{a^{2k}}{(2k)!} (-1)^k \gamma^{(2k-2)} =$$

$$= (-1)^k \frac{2}{N^2a^2} \frac{a^{2k}}{(2k)!} (3 - 2k - 2^{2-2k})\ \gamma^{(2k-2)}(Na)$$

Ce coefficient est nul pour $k = 1$, de sorte que $\sigma^2_{E_1}$ ne contient que des termes en γ'', $\gamma^{(4)}$ etc... Explicitement :

$$(4\text{-}9)\quad \begin{cases} \sigma^2_{E_1} = \dfrac{2}{N^2 a^2} \displaystyle\sum_{k=2}^{n'} C_k \dfrac{a^{2k}}{(2k)!} \gamma^{(2k-2)}(Na) \\ C_k = (-1)^k (3 - 2k - 2^{2-2k}) B_{2k} \end{cases}$$

Finalement, la variance d'estimation est :

$$\sigma^2_E = \sigma^2_{E_0} + \sigma^2_{E_1} + R$$

avec un terme R admettant une majoration en a^{n+1}/N dès que $\gamma(h)$, $h\,\gamma(h)$ et $\chi(h)$ sont dans $\mathcal{G}_n$ (on a posé $n = 2n'$).

<u>Discussion</u>.- On sait que la croissance d'un variogramme à l'infini est nécessairement moins rapide que celle de h^2, soit :

$$\lim_{h \to \infty} \frac{\gamma(h)}{h^2} = 0$$

Les variogrammes que l'on utilise dans les applications possèdent tous une propriété plus forte, à savoir :

$$\lim_{h \to \infty} h^{k-2} \gamma^{(k)}(h) = 0$$

Autrement dit, la dérivée seconde de γ tend vers 0 à l'infini, et la dérivée d'ordre $k > 2$ converge vers 0 plus vite que h^{2-k}. D'après (4-9), donc, et sous réserve que le champ $L = Na$ <u>soit suffisamment grand</u> pour que l'effet de cette décroissance des dérivées soit déjà sensible en $x = L$, le terme $\sigma^2_{E_1}$ sera de l'ordre de :

$$\sigma^2_{E_1} \simeq - \frac{5}{48} \; B_4 \; \frac{a^2 \, \gamma''(Na)}{N^2} = - \frac{1}{288} \; \frac{a^2 \, \gamma''(Na)}{N^2}$$

donc négligeable vis-à-vis du terme principal $\sigma^2_{E_0}$.

Ainsi, dans tous les cas usuels, pour que la variance d'estimation associée à une maille a petite ne dépende que du comportement du $\gamma(h)$ au voisinage de l'origine et puisse être évaluée par application du principe de correspondance, la seule condition requise est que le champ $L = Na$ soit suffisamment grand pour qu'en $h = L$ le variogramme $\gamma(h)$ soit déjà entré dans sa zone de croissance ralentie (moins rapide que h^2). En particulier, le principe de correspondance est toujours applicable au variogramme $\gamma(h) = |h|^\alpha$ $(0 < \alpha < 2)$. Pour $\gamma(h) = h^2$, qui n'est d'ailleurs plus un variogramme acceptable, on vérifierait que σ^2_E est nulle.

Mais le principe de correspondance peut être mis en défaut si le champ Na n'est pas suffisamment grand. Par exemple, considérons le variogramme <u>sphérique</u> d'équation:

$$\gamma(h) = \frac{3}{2} \; \frac{h}{b} - \frac{1}{2} \left(\frac{h}{b}\right)^3 \qquad (h \leq b)$$

Pour $Na \leq b$, on trouvera :

$$\sigma^2_{E_0} = \frac{1}{N} \left(\frac{1}{4} \frac{a}{b} + \frac{1}{120} \; \frac{a^3}{b^3} \right)$$

par application de la règle de correspondance, mais $\sigma^2_{E_1}$ vaut :

$$\sigma^2_{E_1} = \left(\frac{1}{2} \; \frac{1}{188} \; 6 \; \frac{a^3}{b^3}\right) \frac{N}{N^2} = \frac{1}{96N} \; \frac{a^3}{b^3}$$

Ce terme correctif est donc du même ordre de grandeur que le second terme de $\sigma^2_{E_o}$, et on trouve :

$$\sigma^2_{E_o} = \frac{1}{N}\left(\frac{1}{4} \; \frac{a}{b} + \frac{3}{160} \; \frac{a^3}{b^3}\right)$$

4-2 Maille régulière, dispositif fermé.

Ici, l'on utilise l'estimateur $Z^* = \left(\frac{1}{2} Z_o + Z_1 + \ldots + Z_{N-1} + \frac{1}{2} Z_N\right) \frac{1}{N}$ pour évaluer $\frac{1}{L} \int_0^L Z(x)\, dx$, avec $Z_i = Z(ia)$ et $L = Na$.- De :

$$(Z^*)^2 = \left(\sum_{i=0}^{N} Z_i - \frac{Z_o+Z_N}{2}\right)^2 = (\Sigma\, Z_i)^2 - (Z_o+Z_N)\, \Sigma\, Z_i + \frac{(Z_o+Z_N)^2}{4}$$

on déduit l'expression suivante de la variance d'estimation :

$$\sigma^2_E = -\frac{1}{N^2}\left[\sum_{i,j} \gamma_{ij} - 2\sum_i \gamma_{oi} + \frac{1}{2}\gamma_{oN}\right] - \frac{2}{L^2}\int_0^L (L-x)\, \gamma(x)\, dx$$

$$+ \frac{4}{NL}\left[\frac{1}{2}\chi_1(o) + \chi_1(a) + \ldots + \chi_1\big((N-1)a\big) + \frac{1}{2}\chi_1(Na)\right]$$

(avec $\chi_1(x) = x\, \chi(x)$). Explicitons le premier terme :

$$\sum_{i,j} \gamma_{ij} - 2\sum_i \gamma_{oi} + \frac{1}{2}\gamma_{oN} = 2\sum_{i=1}^{N} (N+1-i)\, \gamma(ia)$$

$$- 2\sum_{i=1}^{N} \gamma(ia) + \frac{1}{2}\gamma(Na) = 2\sum_{i=1}^{N} (N-i)\, \gamma(ia) + \frac{1}{2}\gamma(Na)$$

Nous pouvons donc mettre σ_E^2 sous la forme de la somme de 4 termes :

$$\sigma_E^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 - \frac{1}{2N^2}\ \gamma(Na) \tag{4-10}$$

avec :

$$\left\{\begin{aligned}
\sigma_1^2 &= -\frac{2}{N}\left(\frac{1}{2}\gamma(o) + \gamma(a) + \ldots + \gamma\big((N-1)a\big) + \frac{1}{2}\gamma(Na)\right) + \frac{2}{Na}\int_o^{Na}\gamma(h)dh \\
\sigma_2^2 &= \frac{2}{N^2}\left(\gamma(a) + 2\,\gamma(2a) + \ldots + (N-1)\,\gamma\big((N-1)a\big) + \frac{1}{2}N\,\gamma(Na)\right) - \\
&\qquad - \frac{2}{N^2a^2}\int_o^{Na} h\ \gamma(h)dh \\
\sigma_3^2 &= \frac{4}{NL}\left(\frac{1}{2}\chi_1(o) + \chi_1(a) + \ldots + \chi_1\big((N-1)a\big) + \frac{1}{2}\chi_1(Na)\right) - \\
&\qquad - \frac{4}{L^2}\int_o^{Na}\chi_1(x)dx
\end{aligned}\right. \tag{4-11}$$

Les deux premiers termes, σ_1^2 et σ_2^2, sont identiques à ceux qui figuraient déjà en (4-4), et qui ont été évalués en (4-5) et (4-5'). D'après la formule (4-2), le troisième terme peut s'écrire :

$$\sigma_3^2 = \frac{4}{N^2a^2}\left[S(a,\chi_1) + \frac{a^2}{2!}\ B_2\ \chi_1'(Na) - \ldots - (-1)^{n'}\ \frac{a^{2n'}}{(2n')!}\ B_{2n'}\ \chi_1^{(2n'-1)}(Na)\right] + R_3$$

On effectue ensuite le regroupement des termes comme au paragraphe précédent, soit :

$$\sigma_E^2 = \sigma_{E_o}^2 + \sigma_{E_1}^2 + R$$

Le terme $\sigma^2_{E_o}$, lié au comportement de γ en $h = 0$, vaut ici :

$$\sigma^2_{E_o} = - \frac{2}{Na} S(a,\gamma) + \frac{2}{N^2a^2} S(a,\gamma_1) + \frac{4}{N^2a^2} S(a,\chi_1)$$

On peut l'évaluer par application de la règle de correspondance : au terme h^α figurant dans $\gamma(h)$ répond dans σ^2_E le terme :

$$T_\alpha \frac{a^\alpha}{N} + T''_\alpha \frac{a^\alpha}{N^2}$$

avec ici :

$$(4\text{-}12) \quad \begin{cases} T_\alpha = - 2 S_\alpha = - \dfrac{2 B_{1+\alpha}}{1+\alpha} \cos \dfrac{1+\alpha}{2} \pi \\[2ex] T''_\alpha = 2 \left(1+ \dfrac{2}{1+\alpha}\right) S_{1+\alpha} = - \dfrac{2 (\alpha+3)}{(\alpha+1)(\alpha+2)} B_{2+\alpha} \cos \dfrac{\alpha\pi}{2} \end{cases}$$

Le terme $\sigma^2_{E_1}$ regroupe les termes de σ^2_1, σ^2_2 et σ^2_3 où figurent les valeurs en Na de γ et de ces dérivées. Le terme $- \frac{1}{2N^2} \gamma(Na)$, dont la présence en (4-10) pouvait paraître insolite, se révèle ici nécessaire pour assurer l'élimination du terme en $\gamma(Na)$. Seuls subsistent donc les termes en $\gamma''(Na)$, $\gamma^{(4)}(Na)$ etc... comme dans le cas du dispositif centré. Explicitement, on trouve :

$$(4\text{-}12') \qquad \sigma^2_{E_1} = \frac{6}{N^2a^2} \sum_{k=2}^{n'} (-1)^{k+1} B_{2k} \frac{a^{2k}}{(2k)!} \gamma^{(2k-2)}(Na)$$

Les conclusions pratiques sont les mêmes que dans le cas du dispositif centré. Si l'on se contente du premier terme $T_\alpha \frac{a^\alpha}{N}$ de la règle de correspondance, il y a même équivalence entre les deux dispositifs fermé et centré.

5 - L'EFFET DE BORDURE

Passons maintenant au cas le plus difficile, celui de l'effet de bordure. Comme dans l'exemple introductif de la première partie, mais avec un variogramme $\gamma(h)$ quelconque, on veut estimer la teneur moyenne de l'intervalle $(0,L)$, $L = (N+\varepsilon)a$ à l'aide de la moyenne arithmétique:

$$Z^* = \begin{cases} \dfrac{1}{n+1} \displaystyle\sum_{i=0}^{n} Z(x_o+ia) & (0 \leq x_o < \varepsilon a) \\ \dfrac{1}{n} \displaystyle\sum_{i=0}^{n-1} Z(x_o+ia) & (\varepsilon a < x_o < a) \end{cases}$$

associée à une maille régulière a dont l'origine est prise en un point x_o quelconque de l'intervalle $(0,a)$. Le problème va comporter 3 parties :

~ calculer la variance d'estimation $\sigma_E^2(x_o)$ à x_o fixé.

~ prendre l'espérance de $\sigma_E^2(x_o)$ lorsque x_o est choisi au hasard sur $(0,a)$.

~ calculer la valeur moyenne de cette espérance lorsque ε est assimilé à une V.A. uniformément répartie sur $(0,1)$.

5-1 Calcul de $\sigma_E^2(x_o)$.

Supposons d'abord $0 \leq x_o < \varepsilon a$, et calculons :

$$\sigma_E^2(x_o) = D^2 \left[\frac{1}{N+1} \sum_{i=0}^{N} Z(x_o+ia) - \frac{1}{L}\int_0^L Z(x)dx \right]$$

Il vient en premier lieu :

$$(5\text{-}1)\quad \begin{cases} \sigma_E^2(x_o) = -\dfrac{2}{(N+1)^2}\displaystyle\sum_{i=1}^{N}(N+1-i)\,\gamma(ia) - \dfrac{2}{L^2}\int_0^L \gamma(x)\,(L-x)dx \\ \qquad + \dfrac{2}{N+1}\displaystyle\sum_{i=0}^{N}\dfrac{(x_o+ia)\,\chi(x_o+ia) + (L-x_o-ia)\,\chi(L-x_o-ia)}{L} \end{cases}$$

Nous utiliserons les notations habituelles :

$$\begin{cases} \gamma_1(x) = x\,\gamma(x) \quad ; \quad \chi_1(x) = \displaystyle\int_0^x \gamma(y)dy \quad ; \quad \chi(x) = \dfrac{1}{x}\,\chi_1(x) \\ F(x) = \dfrac{2}{x^2}\displaystyle\int_0^x y\,\chi(y)dy = \dfrac{2}{x^2}\int_0^x (x-y)\,\gamma(y) \\ G(x) = \dfrac{x^2}{2}\,F(x) = \displaystyle\int_0^x y\,\chi(y)dy \end{cases}$$

Pour faciliter l'étude de σ_E^2, nous allons décomposer (5-1) en quatre termes :

$$\sigma_E^2(x_o) = H_1(x_o) + H_2(x_o) + H_3(x_o) + H_4(x_o)$$

avec pour $0 \leq x_o \leq \varepsilon a$

$$(5\text{-}2)\quad \begin{cases} H_1(x_o) = -\dfrac{2}{(N+1)^2}\displaystyle\sum_{i=1}^{N}(N+1-i)\,\gamma(ia) + \dfrac{2}{(N+1)^2}\int_0^{(N+1)a}(Na+a-x)\gamma(x)dx \\ H_2(x_o) = \dfrac{2}{(N+1)L}\displaystyle\sum_{i=o}^{N}\chi_1(x_o+ia) - \dfrac{2}{L^2}\int_0^L \chi_1(x)dx \\ H_3(x_o) = H_2(\varepsilon a-x_o) \\ H_4(x_o) = F(L) - F((N+1)a) \end{cases}$$

Pour $\varepsilon a < x_o < a$, on remplace partout N+1 par N, et la troisième relation par

$$H_3(x_o) = H_2(a+\varepsilon a-x_o)$$

Le premier terme, $H_1(x_o)$, se révèle identique à la somme $\sigma_1^2 + \sigma_2^2$ des deux termes ainsi désignés dans (4-4), à condition de changer N en N+1. Il résulte donc de (4-5) et (4-5') que l'on a (pour $x_o \leq \varepsilon a$) :

$$(5\text{-}3) \qquad \left\{ \begin{aligned} H_1(x_o) = &- \frac{2}{(N+1)a} S(a,\gamma) + \frac{2}{(N+1)^2 a^2} S(a,\gamma_1) - \\ &- \frac{2}{(N+1)^2 a^2} \sum_{k=1}^{n'} (-1)^k \frac{(2k-1)}{(2k)!} B_{2k}\, a^{2k}\, \gamma^{(2k-2)}((N+1)a) \end{aligned} \right.$$

à un reste près que nous n'écrirons plus. Il serait possible d'évaluer de la même manière H_2 et H_3, mais nous verrons que celà n'est pas nécessaire.

5-2 Passage à l'espérance en x_o.

En supposant x_o uniformément distribué sur (0,a), on a maintenant :

$$\sigma_E^2 = \frac{1}{a}\int_0^a \sigma_E^2(x_o)\, dx_o = \sum_{i=1}^{4} E(H_i)$$

Evaluons d'abord $E(H_2)$, soit :

$$E(H_2) = \frac{1}{a}\int_0^{\varepsilon a} \frac{2}{(N+1)L} \sum_{i=0}^{N} \chi_1(x_0+ia)\,dx_0 + \frac{1}{a}\int_{\varepsilon a}^{a} \frac{2}{NL} \sum_{i=0}^{N-1} \chi_1(x_0+ia)dx_0 - F(L)$$

$$= \frac{2}{(N+1)aL} \sum_{i=0}^{N} [G(ia+\varepsilon a) - G(ia)] + \frac{2}{NaL} \sum_{i=0}^{N-1} [G(a+ia)-G(ia+\varepsilon a)] - $$
$$- F(L)$$

Il vient en regroupant les termes :

$$E(H_2) = \frac{2}{N(N+1)aL} \left(\sum_{i=0}^{N} G(ia) - \sum_{i=0}^{N-1} G(ia+\varepsilon a) \right) + \frac{2}{(N+1)aL} G(L) - F(L)$$

$$= \frac{2}{N(N+1)aL} \left\{ \left[\frac{1}{2} G(o)+G(a)+\ldots+G\big((N-1)a\big) + \frac{1}{2} G(Na) \right] - \sum_{i=0}^{N-1} G(ia+\varepsilon a) \right\}$$

$$+ \frac{1}{N(N+1)aL} G(Na) + \frac{2}{(N+1)aL} G(L) - F(L)$$

Le terme entre crochets s'évalue au moyen des formules (4-2) et (4-3). En tenant compte aussi de $G(L) = \frac{L^2}{2} F(L)$, on obtient ainsi :

$$(5\text{-}4\text{-}2) \quad \left\{ \begin{aligned} E(H_2) = \frac{2}{N(N+1)a^2L} \Bigg[& S(a,G) - S(\varepsilon a,a,G) + \\ & + \sum_{k=1}^{n'} (-1)^{k+1} \frac{a^{2k}}{(2k)!} B_{2k} G^{2k+1}(Na) \\ & + \sum_{k=1}^{n+1} \frac{a^k}{k!} G^{(k-1)}(Na)\, B_k(\varepsilon) \Bigg] + \frac{1}{2} \frac{N}{(N+1)(N+\varepsilon)} F(Na) - \frac{1-\varepsilon}{N+1} F(L) \end{aligned} \right.$$

On vérifie sans peine que H_3 donne le même résultat que H_2 :

$$(5\text{-}4\text{-}3) \qquad E(H_3) = E(H_2)$$

En ce qui concerne H_4, il vient immédiatement :

$$E(H_4) = F(L) - \varepsilon F\big((N+1)a\big) - (1-\varepsilon)\ F(Na)$$

En utilisant un développement de Taylor, on trouve :

$$(5\text{-}4\text{-}4) \qquad E(H_4) = -\ \varepsilon(1-\varepsilon)\ \frac{a^2}{2!}\ F'' - \varepsilon(1-\varepsilon^2)\ \frac{a^3}{3!}\ F'''(Na) - \varepsilon(1-\varepsilon^3)\ \frac{a^4}{4!}\ F^{(4)}(Na) - \ldots$$

Venons-en, enfin, au terme en H_1. D'après (5-3), on trouve :

$$(5\text{-}4\text{-}1) \quad \begin{cases} E(H_1) = \left(-\dfrac{2}{Na} + \dfrac{2\ \varepsilon}{N(N+1)a}\right) S(a,\gamma) + \left(\dfrac{2\ \varepsilon}{(N+1)^2 a^2} + \dfrac{2(1-\varepsilon)}{N^2 a^2}\right) S(a,\gamma_1) \\ \\ -2\ \displaystyle\sum_{k=1}^{n'} (-1)^k\ \frac{2k-1}{(2k)!}\ B_{2k}\ a^{2k-2}\ \left(\frac{\varepsilon}{(N+1)^2}\ \gamma^{(2k-2)}((N+1)a) + \frac{1-\varepsilon}{N^2}\ \gamma^{(2k-2)}(Na)\right) \end{cases}$$

En regroupant ces quatre termes, on obtient :

$$\sigma_E^2 = \sigma_{E_o}^2 + \sigma_{E_1}^2$$

Le terme $\sigma_{E_o}^2$ regroupe les fonctions S et prend en charge l'influence du comportement du $\gamma(h)$ au voisinage de l'origine. Il vaut :

$$(5\text{-}5) \quad \begin{cases} \sigma_{E_o}^2 = -\ \dfrac{2}{Na}\ \left(1-\dfrac{\varepsilon}{N+1}\right)\ S(a,\gamma) + \dfrac{2}{a^2}\ \left(\dfrac{\varepsilon}{(N+1)^2} + \dfrac{1-\varepsilon}{N^2}\right) S(a,\gamma_1) \\ \\ \qquad + \dfrac{4}{N(N+1)(N+\varepsilon)a^3}\ [S(a,G) - S(\varepsilon a,a,G)] \end{cases}$$

Si nous nous limitons aux termes en $1/N$ et $1/N^2$, il reste :

(5-5') $$\sigma^2_{E_0} \simeq - \frac{2}{Na} \left(1 - \frac{\varepsilon}{N}\right) S(a,\gamma) + \frac{2}{a^2N^2} S(a,\gamma_1)$$

L'effet de bordure ne se répercute qu'à l'ordre 2 en 1/N. La partie principale de $\sigma^2_{E_0}$ n'est donc pas affectée par l'effet de bordure. Le terme $\sigma^2_{E_0}$ se calcule par application de la règle de correspondance, un terme h^α dans $\gamma(h)$ donnant le terme :

(5-5") $$- 2\, S_\alpha \frac{a^\alpha}{N} + 2\,(\varepsilon S_\alpha + S_{1+\alpha}) \frac{a^\alpha}{N^2}$$

Le terme $\sigma^2_{E_1}$ regroupe tous les termes qui, dans les H_i, dépendent des valeurs prises au voisinage de L par les différentes fonctions auxiliaires. C'est lui, donc, qui prend en charge l'effet de bordure. Nous nous contenterons de calculer sa valeur moyenne en ε.

5-3 Valeur moyenne de l'effet de bordure.

Intégrons donc en ε les différents termes de bordure figurant dans les relations (5-4), en désignant par H'_i la contribution de $E(H_i)$. Pour H'_1, (5-4-1) donne immédiatement :

(5-6-1) $$H'_1 = - \sum_{k=1}^{n'} (-1)^k \frac{2k-1}{(2k)!} B_{2k}\, a^{2k-2} \left(\frac{\gamma^{(2k-2)}((N+1)a)}{(N+1)^2} + \frac{\gamma^{(2k-2)}(Na)}{N^2} \right)$$

La partie principale de ce terme est :

(5-6-1') $$H'_1 \simeq \frac{1}{6N^2}\, \gamma(Na)$$

Passons à H'_2 et H'_3. En remarquant que l'intégrale de $B_k(\varepsilon)$ est nulle, on déduit de (5-4-2) la partie principale de $H_2^{'}$:

$$E(H_2) = \frac{2}{a^3\ N^2(N+1)}\ (1-\frac{\varepsilon}{N})\ \left[\frac{1}{12}\ Na^3\ \chi(Na) + \frac{N^2a^2}{2}\ F(Na)\ (\frac{1}{2}-\varepsilon) + \ldots\right]$$

$$+ \frac{1}{2}\ \frac{1}{N+1}\ \ (1-\frac{\varepsilon}{N}+\ldots)\ F(Na) - \frac{(1-\varepsilon)}{N+1}\ \left[F(Na) + \frac{2\varepsilon}{n}\ \left(\chi(Na)-F(Na)\right)\right]$$

En intégrant en ε, on trouve :

$$H_2^{'} = \frac{1}{N^2}\left[\frac{\chi(Na)}{6} + \frac{1}{12}\ F(Na) - \frac{1}{4}\ F(Na) - \frac{\chi(Na)}{3} + \frac{F(Na)}{3}\right]$$

soit :

(5-6-2) $$H_2^{'} = \frac{1}{6N^2}\ \left(F(Na) - \chi(Na)\right)$$

De même

(5-6-3) $$H_3^{'} = \frac{1}{6N^2}\ \left(F(Na) - \chi(Na)\right)$$

Pour $H_4^{'}$, enfin, (5-4-4) donne :

$$H_4^{'} = -\frac{1}{12}\ a^2\ F''(Na) - \left(\frac{1}{2}-\frac{1}{4}\right)\ \frac{a^3}{3!}\ F'''(Na) - \ldots$$

Les relations :

$$F^{'}(x) = \frac{2}{x}\ (\chi - F) \quad ; \quad F^{''}(x) = \frac{6}{x^2}\ F - \frac{8}{x^2}\ \chi + \frac{2}{x^2}\ \gamma$$

permettent de trouver la partie principale de $H_4^{'}$, qui est :

3- Limites Inductives de fermés aléatoires

Considérons dans E une suite croissante d'ouverts relativement compacts B_n recouvrant E en vérifiant $\overline{B}_n \subset B_{n+1}$. Donnons nous, pour chaque n, un fermé aléatoire $A_n \subset \overline{B}_n$, défini par la fonctionnelle T_n telle que $T_n(K) = P(A_n \in V_K)$, $K \in \mathcal{K}(\overline{B}_n)$. On peut aussi bien, d'ailleurs, prolonger T_n sur $\mathcal{K}(E)$ entier en posant $T_n(K) = T_n(K \cap \overline{B}_n)$ si $K \not\subset \overline{B}_n$. Si les fonctionnelles T_n vérifient la condition :

$$(3\text{-}1) \quad T_{n+m}(K) = T_n(K) \qquad (n, m > 0,\ K \in \mathcal{K},\ K \subset \overline{B}_n)$$

l'ensemble $A_{n+m} \cap \overline{B}_n$ est équivalent à A_n (c.a.d. admet la même probabilité que lui). Il est alors possible de définir un fermé aléatoire A de $\mathcal{F}(E)$ tel que, pour tout n, $A \cap \overline{B}_n$ soit équivalent à A_n. On dira que ce fermé aléatoire A est la limite inductive des fermés aléatoires A_n dont les fonctionnelles T_n vérifient la condition (3-1).

Pour établir l'existence de A, considérons un compact $K \in \mathcal{K}$. Comme les ouverts B_n recouvrent E, K est contenu dans $\overline{B}_n$ dès que n est supérieur à un n_0, et on a $T_n(K) = T_{n_0}(K)$ pour $n \geq n_0$ d'après la condition (3-1). On peut donc définir une fonctionnelle T sur $\mathcal{K}$ entier en posant $T(K) = T_{n_0}(K)$. On a évidemment $T(\emptyset) = 0$ et $T \leq 1$ sur $\mathcal{K}$. T est scs, car si $K_n \downarrow K$ dans $\mathcal{K}$, les K_n sont contenus dans un $\overline{B}_{n_0}$ et T_{n_0} est scs. De même, les fonctions S_n associées à T sont ≥ 0, puisque toute famille finie de compacts est contenue dans un $\overline{B}_{n_0}$ et que T_{n_0} est une capacité. D'après le théorème 1,

Effet de bordure

$$\varphi^* = a \sum k(x_0+ia)\, Y(x+ia) \qquad V^* = a \sum k(x_0+ia)$$

Variance de $\dfrac{\varphi^*}{V^*} - \dfrac{\varphi}{V}$?

$$\frac{\varphi^*}{V^*} - \frac{\varphi}{V} = \frac{V\varphi^* - \varphi V^*}{V V^*} = \frac{V\varphi^* - \varphi V^*}{V^2}\left(1 - \frac{\varepsilon_V}{V} + \frac{\varepsilon_V^2}{V^2} \dots\right)$$

Se limiter au 1er ordre : $\dfrac{V\varphi^* - \varphi V^*}{V^2}$

$$Z(x) = (Y(x) - \bar{Y})\, k(x) \qquad \left(\bar{Y} = \frac{1}{V}\int_V Y(x)\right)$$

$$a \sum Z(x_0+ia) - \int k(x) Z(x)\,dx = \varphi^* - \bar{Y} V^* - \cancel{\varphi} + \cancel{\bar{Y} V} = \varphi^* - \frac{\varphi}{V} V^*$$

$$= \frac{V\varphi^* - \varphi V^*}{V}$$

$$m^* - \bar{Y} = \frac{V\varphi^* - \varphi V^*}{V^2} = \frac{1}{V}\left[a \sum Z(x_0+ia) - \int k(x) Z(x)\,dx\right]$$

$$D^2(m^* - \bar{Y}) = \frac{1}{V^2}\, \mathcal{E}_a\, g$$

$$g = \frac{1}{V^2} E \int k(x) k(x+h) (Y(x) - \bar{Y})(Y(x+h) - \bar{Y})\,dx$$

$$\{ -\gamma(h) + \bar{\gamma}(x,V) + \bar{\gamma}(x+h, V) - \bar{\gamma}(V,V)$$

$$\int k(x) k(x+h)\, \bar{\gamma}(x, V)\,dx = K(h)\, \bar{\gamma}(V_h, V)$$

$$g(h) = \frac{1}{V^2}\left[-\gamma(h) K(h) + K(h)\, \bar{\gamma}(V_h, V) + K(h)\, \bar{\gamma}(V_{-h}, V) - K(h) F(V)\right]$$

$$K(h)\bar{\gamma}(V_h, V) = K(0)\, \bar{\gamma}(\bar{V}, \bar{V}) - (K(0) - K(h))\, \bar{\gamma}(P, V)$$

$$\frac{1}{V^2}\, \mathcal{E}_a \left[-\gamma K + 2(\bar{\gamma}(P,V) - F(V))\, K(h)\right]$$

$$\varphi^* - \varphi = m^*(V^* - V) + V(m^* - \bar{Y})$$

$\frac{1}{V}\ (Y(x) - \bar{Y})\, I(x) \rightarrow D^2(m^* - \bar{Y})$

$\varphi(\omega) \quad \varphi^*(\omega, x_0) \quad V^*(x_0)$

$E(V^*) = V$

$E(\varphi^* | \omega) = \varphi(\omega)$

$g(h) = K(h)[m^2 + C(h)]$

$g_{12}(h) = m K(h)$

$D^2(\varepsilon_q) = \mathcal{E}_a\, g$

$E(\varepsilon_q \varepsilon_v) = m\, E(\varepsilon_v^2)$

$E(\varepsilon_v^2) = \mathcal{E}_a K$

$\varepsilon_q = m\, \varepsilon_v + \varepsilon \qquad \langle \varepsilon, \varepsilon_v \rangle = 0$

$E(\varepsilon^2) = \mathcal{E}_a g - m^2 \mathcal{E}_a K = \mathcal{E}_a K(h) C(h)$

<u>Remarque</u> – L'intérêt de ce corollaire vient de ce que (E étant LCD) on peut toujours trouver une famille $\mathcal{B}'$ dénombrable vérifiant les conditions de l'énoncé. Dans $\mathbb{R}^N$, par exemple, on pourra prendre pour $\mathcal{B}$ la famille des réunions finies des boules ouvertes de centres et de rayons rationnels. Il est alors possible de définir la notion de <u>fermé aléatoire conditionnel</u>. En effet, soit A le fermé aléatoire défini par une probabilité P sur $\sigma(\mathcal{G})$, u une application mesurable de $(\mathcal{F}, \sigma(\mathcal{G}))$ dans un espace $(\Omega, \mathcal{Q})$, et F la probabilité sur $\mathcal{Q}$ déduite de P par u. Si $\mathcal{B}'$ est dénombrable, ~~[illegible]~~ on peut pour F-presque tout u définir sur $\mathcal{B}'$ une fonction T_u en posant :

$$T_u(B) = E(1_{V_B} \mid u) \qquad (B \in \mathcal{B}')$$

On vérifie sans peine que T_u satisfait aux conditions ~~[illegible]~~ du corollaire. Il existe donc pour presque tout u un fermé aléatoire $A(u)$ défini par une probabilité P_u sur $\sigma(\mathcal{G})$ mesurable en u et vérifiant :

$$(2\text{-}2) \qquad P(A \in V \text{ et } u \in U) = \int_U P_u(V)\, F(du) \qquad (V \in \sigma(\mathcal{G}),\ U \in \mathcal{Q})$$

Voici une autre application du Théorème 1 – Soit $\mathcal{I}$ l'ensemble des parties finies de E. Une indicatrice aléatoire $Y(x), x \in E$ (c'est-à-dire une fonction aléatoire p.s. égale à 0 ou 1 en chaque $x \in E$) est définie par sa loi spatiale, c'est-à-dire par une fonctionnelle φ sur $\mathcal{I}$ telle que $\varphi(I) = P(Y(x) = 0,\ x \in I)$ pour $I \in \mathcal{I}$. Inversement, une fonctionnelle φ sur $\mathcal{I}$ est une loi spatiale si et seulement si ~~[illegible]~~ les fonctions S_n construites à partir de $T = 1 - \varphi$ sont ≥ 0 sur $\mathcal{I}$. On peut alors montrer que cette loi spatiale se prolonge par une probabilité sur $\sigma(\mathcal{G})$, autrement dit qu'il existe un fermé aléatoire admettant cette loi spatiale, si et seulement si T est SCS pour la restriction à $\mathcal{I}$ de la topologie myope (la continuité monotone séquentielle $\downarrow$ ne suffit évidemment plus). Mais ce prolongement n'est pas unique, comme le montre l'exemple de la fonctionnelle T identiquement nulle sur $\mathcal{I}$ qui peut être associée à un processus de Poisson ponctuel aussi bien qu'au fermé p.s. vide.

(5-6-4) $$H'_4 = \frac{1}{N^2} \left(- \frac{1}{2} F(Na) + \frac{2}{3} \chi(Na) - \frac{1}{6} \gamma(Na) \right)$$

En regroupant les quatre relations (5-6), on trouve donc finalement la partie principale de l'effet de bordure, qui est :

(5-7) $$\sigma^2_{E_1} = \frac{1}{6N^2} \left(2 \chi(Na) - F(Na) \right)$$

5-4 Conclusion sur l'effet de bordure

Le résultat exact (5-7) confirme, dans une certaine mesure, la formule $\sigma^2 \sigma_V^2/V^2$ à laquelle conduit le raisonnement approximatif rappelé dans l'introduction de cette étude. On a ici, en effet, $\sigma_V^2/V^2 = 1/6N^2$, et l'on obtient le résultat exact à condition de remplacer σ^2 par la variance d'extension dans $L = Na$ de l'échantillon le plus mal placé. Lorsqu'il y a une portée finie b et que L est plus grand que b, $2\chi-F$ diffère peu de la variance a priori σ^2, et (5-7) se réduit à $\sigma^2/6N^2 = \sigma^2 \sigma_V^2/V^2$. Dans les applications, donc, cette formule simplifiée sera souvent utilisable.

6 - ANNEXE 1 - VARIOGRAMME h^{α}

Dans cette première annexe, nous allons traiter explicitement le cas du variogramme $\gamma(h) = |h|^{\alpha}$. En appliquant les formules (4-2) et (4-3) à cette fonction, on trouve (avec a = 1).

$$(6-1)\quad \left\{ \begin{aligned} & \sum_{i=0}^{n} i^{\alpha} = \frac{n^{\alpha+1}}{\alpha+1} + \frac{1}{2} n^{\alpha} + S_{\alpha} + \frac{\alpha}{12} n^{\alpha-1} - \frac{\alpha(\alpha-1)(\alpha-2)}{720} n^{\alpha-3} + \dots \\ & \sum_{i=0}^{n-1} (x_o+i)^{\alpha} = \frac{n^{\alpha+1}}{\alpha+1} + S_{\alpha}(x_o) - n^{\alpha} B_1(x_o) - \frac{\alpha}{2} n^{\alpha-1} B_2(x_o) - \\ & \qquad - \frac{\alpha(\alpha-1)}{6} n^{\alpha-2} B_3(x_o) - \frac{\alpha(\alpha-1)(\alpha-2)}{24} n^{\alpha-3} B_4(x_o) - \dots \end{aligned} \right.$$

Pour $0 \leq x_o \leq \varepsilon$, $L = (n+\varepsilon)$, calculons la variance (5-1), qui s'écrit :

$$\sigma_E^2(x_o) = - \frac{2}{(n+1)^2} \sum_{i=1}^{n} (n+1-i)\, i^{\alpha} - \frac{2}{(\alpha+1)(\alpha+2)} (n+1)^{\alpha}$$

$$+ \frac{2}{(n+1)(n+\varepsilon)} \sum_{i=0}^{n} \frac{(x_o+i)^{\alpha+1}}{\alpha+1} + \frac{2}{(n+1)(n+\varepsilon)(\alpha+1)} \sum_{i=0}^{n} (n+\varepsilon-x_o-i)^{\alpha+1}$$

En utilisant (6-1), on trouve pour $0 \leq x_o < \varepsilon$

$$H_1(x_o) = - \frac{2}{(N+1)^2} \sum_{i=1}^{n} (n+1-i)\, i^{\alpha} = -2 \left[\frac{(n+1)^{\alpha}}{\alpha+1} + \frac{1}{2} (n+1)^{\alpha-1} + \frac{\alpha}{12} (n+1)^{\alpha-2} \right]$$

$$+ 2 \left[\frac{(n+1)^{\alpha}}{\alpha+2} + \frac{1}{2} (n+1)^{\alpha-1} + \frac{\alpha+1}{12} (n+1)^{\alpha-2} \right] - \frac{2S_{\alpha}}{n+1} + \frac{2\, S_{1+\alpha}}{(n+1)^2}$$

soit :

$$(6-1)\quad H_1(x_o) = -\frac{2}{(\alpha+1)(\alpha+2)}(n+1)^{\alpha} - \frac{1}{6}(n+1)^{\alpha-2} - \dots - \frac{2S_\alpha}{n+1} + \frac{2\,S_{1+\alpha}}{(n+1)^2}$$

Pour $\varepsilon < x_o < 1$, on remplacera n+1 par n. D'où l'espérance de H_1 :

$$H_1 = \int_0^1 H_1(x_o)dx_o = -2\,S_\alpha\left(\frac{1}{n} - \frac{\varepsilon}{n(n+1)}\right) + \frac{2\,S_{1+\alpha}}{(n+1)^2} -$$

$$-\frac{2}{(\alpha+1)(\alpha+2)}\left[\varepsilon(n+1)^{\alpha} + (1-\varepsilon)\,n^{\alpha}\right] + \frac{1}{6}\,n^{\alpha-2}$$

Considérons maintenant le second terme, qui s'écrit pour $0 < \varepsilon \leq x_o$:

$$H_2(x_o) = \frac{2}{(n+1)(n+\varepsilon)}\sum\frac{(x_o+i)^{\alpha+1}}{\alpha+1} =$$

$$= \frac{2}{(\alpha+1)(n+\varepsilon)}\left[\frac{(n+1)^{\alpha+1}}{\alpha+2} + \frac{S_{\alpha+1}(x_o)}{n+1} - (n+1)^{\alpha}B_1(x_o) - \right.$$

$$\left. - \frac{\alpha+1}{2}(n+1)^{\alpha-1}B_2(x_o)\right]$$

Pour $x_o > \varepsilon$, on remplace n+1 par n (dans le crochet seulement), soit

$$H_2'(x_o) = \frac{2}{(\alpha+1)(n+\varepsilon)}\left[\frac{S_{\alpha+1}(x_o)}{n} + \frac{n^{\alpha+1}}{\alpha+2} - n^{\alpha}B_1(x_o) - \frac{\alpha+1}{2}n^{\alpha-1}B_2(x_o)\dots\right]$$

Prenons l'espérance en x_o, soit :

$$H_2 = \int_0^{\varepsilon} H_1(x_o)dx_o + \int_{\varepsilon}^1 H_1'(x_o)dx_o$$

A l'approximation en $1/n^2$, l'intégrale de $S_{\alpha+1}(x_o)$ donne 0. On note :

$$\int_0^{\varepsilon} B_k(x_o)dx_o = - \int_{\varepsilon}^{1} B_k(x_o)dx_o$$

Par suite, à l'approximation $n^{\alpha-2}$, le terme en $B_2(x_o)$ est négligeable. Par contre, le terme en $B_1(x_o)$ ne peut pas être négligé. Avec :

$$\int_0^{\varepsilon} B_1(x_o)dx_o = - \int_{\varepsilon}^{1} B_1(x_o)dx_o = \frac{\varepsilon(1-\varepsilon)}{2}$$

on obtient :

$$H_2 = \frac{2}{(\alpha+1)(n+\varepsilon)} \left[\frac{n^{\alpha+1}}{\alpha+2} + \varepsilon \left(\frac{(n+1)^{\alpha+1} - n^{\alpha+1}}{\alpha+2}\right) - \frac{\varepsilon(1-\varepsilon)}{2} \left((n+1)^{\alpha} - n^{\alpha}\right) \right]$$

$$= \frac{2}{(\alpha+1)(\alpha+2)} \; \frac{1}{n+\varepsilon} \left[n^{\alpha+1} + (\alpha+1)\, n^{\alpha}\varepsilon + \left(\frac{\alpha(\alpha+2)}{2}\, \varepsilon^2 - \frac{\alpha}{2}\,\varepsilon\right) n^{\alpha-1} \right].$$

Enfin, avec $1/n+\varepsilon = (1 - \varepsilon/n + \varepsilon^2/n^2)\; 1/n$, il vient :

$$(6\text{-}2) \quad H_2 = \frac{2}{(\alpha+1)(\alpha+2)} \left[n^{\alpha} + \alpha\, \varepsilon\, n^{\alpha-1} + \left(\frac{\alpha^2}{2}\, \varepsilon^2 - \frac{\alpha}{2}\,\varepsilon\right) n^{\alpha-2} \right]$$

Le troisième terme donne :

$$H_3 = H_2$$

et le quatrième est :

$$H_4 = - \frac{2}{(\alpha+1)(\alpha+2)} (n+\varepsilon)^{\alpha} = - \frac{2}{(\alpha+1)(\alpha+2)} \left(n^{\alpha} + \alpha\, \varepsilon\, n^{\alpha-1} + \frac{\alpha(\alpha-1)}{2}\, \varepsilon^2\, n^{\alpha-2} \right)$$

En regroupant les termes, on trouve :

$$\sigma_E^2 = -\,2\left(\frac{1}{n} - \frac{\varepsilon}{n^2}\right) S_\alpha + \frac{2}{n^2}\, S_{1+\alpha} + \left(\frac{1}{6} - \frac{\alpha}{\alpha+2}\quad(\varepsilon-\varepsilon^2)\right) \tag{6-3}$$

Prenons enfin la valeur moyenne en ε. Il vient :

$$\sigma_E^2 = -\,2\left(\frac{1}{n} - \frac{1}{2n^2}\right) S_\alpha + \frac{2}{n^2}\, S_{1+\alpha} + \frac{1}{6n^2}\;\frac{2}{\alpha+2} \tag{6-4}$$

On retrouve bien le terme $\sigma_{E_1}^2$ prévu par la relation (5-7).

$\frac{2}{\alpha+1}\, h^\alpha - \frac{2}{(\alpha+1)(\alpha+2)}$

$2/\alpha+2$

7 - ANNEXE 2 - EXPRESSIONS DES NOMBRES DE BERNOUILLI

D'après la définition (2-7), le nombre de Bernouilli B_λ est :

$$B_\lambda = 2\,\Gamma(1+\lambda) \sum_{p=1}^{\infty} \frac{1}{(2\pi p)^\lambda} \qquad (\lambda > 1) \tag{7-1}$$

Or, on peut écrire :

$$\int_0^\infty x^{\lambda-1}\; e^{-2\pi p x}\, dx = \frac{\Gamma(\lambda)}{(2\pi p)^\lambda}$$

Il en résulte une nouvelle expression de B_λ :

$$B_\lambda = 2\,\lambda \int_0^\infty \frac{x^{\lambda-1}}{e^{2\pi x}-1}\, dx \tag{7-2}$$

Partons maintenant de :

$$\int_0^\infty y^{\beta-1}\, e^{-xy}\, dy = \frac{\Gamma(\beta)}{x^\beta} \qquad (0 < \beta < 1)$$

et intégrons en x. Il vient :

$$\int_0^\infty (1 - e^{-ux})\, u^{\beta-2}\, du = \frac{\Gamma(\beta)}{1-\beta}\, x^{1-\beta}$$

Soit, avec $\alpha = 1 - \beta$

(7-3) $$x^\alpha = \frac{\sin \alpha\pi}{\pi}\, \Gamma(1-\alpha) \int_0^\infty (1 - e^{-ux})\, \frac{du}{u^{1+\alpha}} \qquad (0 < \alpha < 1)$$

En intégrant n fois en x, on obtient plus généralement :

(7'3) $$x^{\alpha+n} =$$

$$= (-1)^n \frac{\sin \alpha\pi}{\pi}\, \Gamma(n+1+\alpha) \int_0^\infty \left(1 - \frac{ux}{1!} + \frac{(ux)^2}{2!} - \dots + (-1)^n \frac{(ux)^n}{n!} - e^{-ux}\right) \frac{du}{u^{1+N+\alpha}}$$

Calculons à l'aide de (7-3) le coefficient $S_\alpha(a)$ de la relation (3-2). On trouve :

$$S_\alpha(a) = \lim_{N \to \infty} \frac{\sin \alpha\pi}{\pi}\, \Gamma(1-\alpha) \int_0^\infty \Bigg[1 - e^{-Nua} \left[- \frac{a}{1 - e^{ua}} + \frac{a}{2} + \frac{1}{u} \right] -$$

$$- u\, B_2 \frac{a^2}{2!}\, e^{-Nua} + \dots \Bigg] \frac{du}{u^{1+\alpha}}$$

Le passage à la limite n'offre pas de difficulté, et on trouve :

$$S_\alpha(a) = -\Gamma(1+\alpha)\,\frac{\sin\alpha\pi}{\pi}\int_0^\infty\left(\frac{a}{1-e^{-ua}}-\frac{a}{2}-\frac{1}{u}\right)\frac{du}{u^{1+\alpha}}$$

Soit, avec $a = 1$, et $0 < \alpha < 1$

(7-4) $$S_\alpha = -\Gamma(1+\alpha)\,\frac{\sin\alpha\pi}{\pi}\int_0^\infty\left(\frac{x}{1-e^{-x}}-\frac{x}{2}-1\right)\frac{dx}{x^{1+\alpha}}$$

En comparant avec (3-3), on trouve :

$$\sum_{p=1}^\infty\frac{1}{(2\pi p)^{1+\alpha}} = \frac{1}{\pi}\cos\frac{\alpha\pi}{2}\int_0^\infty\left[\frac{x}{1-e^{-x}}-\frac{x}{2}-1\right]\frac{dx}{x^{1+\alpha}}$$

donc :

(7-5) $$B_{1+\alpha} = \frac{2\,\Gamma(1+\alpha)}{\pi}\cos\alpha\frac{\pi}{2}\int_0^\infty\left(\frac{x}{1-e^{-x}}-\frac{x}{2}-1\right)\frac{dx}{x^{1+\alpha}} \qquad (0 < \alpha < 1)$$

De même, en partant de (7-3'), on trouvera :

(7-6) $$S_{\alpha+n} =$$

$$= (-1)^{n+1}\,\frac{\sin\alpha\pi}{\pi}\,\Gamma(n+1+\alpha)\int_0^\infty\left(\frac{x}{1-e^{-x}}-1-\frac{x}{2}+\sum_{k=1}^{n'}(-1)^k\frac{B_{2k}}{(2k)!}x^{2k}\right)\frac{dx}{x^{1+n+\alpha}}$$

n' désignant le plus petit entier tel que $2n' \geq n-1$. Si l'on compare ces relations à (7-2), on voit que les S_χ jouent en quelque sorte le rôle des nombres de Bernoulli d'indices négatifs.