COMPARAISON DE QUELQUES DISTRIBUTIONS DU POINT DE VUE DE LA SELECTIVITE

Georges MATHERON*

TABLE DES MATIERES

RESUME 2

A - INTRODUCTION 3

B - LA SYMETRIE EN MODE T 3

C - LA LOI DE SICHEL 6

D - QUELQUES LOIS EXTREMALES 8

ABSTRACT 2

ANNEXE 1 : LES NORMES 14

ANNEXE 2 : LES CONVERGENCES 17

ANNEXE 3 : LES INEGALITES 19

BIBLIOGRAPHIE 21

ILLUSTRATIONS

Figure 1 - Courbe Q = Q(T) 11

Figure 1 bis - Courbe Q = Q(T) 11

Figure 2 - Evolution de l'indice de normalité S $\sqrt{\pi/\sigma}$ 12

Figure 3 - Evolution de l'indice de sélectivité S/m 12

Figure 4 - Evolution de l'indice de maximalité S/Smax 13

----

*Centre de Géostatistique et de Morphologie Mathématique, Ecole des Mines de Paris,
35 rue Saint Honoré, FONTAINEBLEAU CEDEX - FRANCE

"ETUDES GEOSTATISTIQUES", Séminaire CFSG - 17-18 Juin 1985 - Fontainebleau, France, in *Sci. de la Terre, Sér. Inf., n° 24*

RESUME

Dans le cas de distributions à coefficient de variation σ/m très élevé, l'indice de sélectivité S/m constitue peut-être un meilleur paramètre. Dans tous les cas, l'évolution de S/m en fonction de σ/m présente un grand intérêt. Sous l'effet d'un changement de support, S/m décroît, mais moins vite que σ/m, et l'indice de normalité augmente. Cet effet est mis en évidence pour la loi lognormale, la loi gamma, la loi de Poisson, la loi exponentielle poissonisée et deux autres lois à caractère extrémal.

ABSTRACT

Study of the selectivity index properties in the case of different statistical distributions

In the case of distributions with high coefficient of variation σ/m, the selectivity index S/m is probably a better parameter. In any case, the evolution of S/m with regard to σ/m is of great interest. When affected by a change of support, S/m decreases less rapidly than σ/m, and the distribution tends towards normality. This is being demonstrated for the lognormal, gamma, Poisson, Poisson exponential distributions, and also for two other extremal distributions.

KEY WORDS : Coefficient of variation, selectivity index, Sichel distribution, normality index

A – INTRODUCTION

On sait que, dans le cas des distributions à coefficient de variation σ/m très élevé, la variance n'est plus un paramètre tout-à-fait satisfaisant, en raison de sa trop grande sensibilité à l'influence des valeurs élevées : il en résulte, en pratique, de sérieuses difficultés pour l'inférence statistique, et aussi bien, une certaine inadéquation de ce paramètre pour servir de guide dans le choix des modèles. D'où l'idée de lui substituer un indicateur plus robuste, comme l'indice de sélectivité par exemple. En géostatistique, cependant, on ne peut faire l'économie des moments d'ordre deux, pour la raison décisive que la variance est le seul paramètre dont on sache prévoir exactement l'évolution sous l'effet d'un changement de support. D'où l'intérêt que présentent les relations pouvant exister entre l'indice de sélectivité S/m et le coefficient de variation σ/m. On sait que, dans les changements de support réels, comme dans les modèles usuels, S/m décroît avec σ/m, mais moins vite que lui, en raison de la convergence vers la normalité : l'indice de normalité, que nous définirons, à un facteur près, comme le rapport S/σ , augmente donc avec le support.

Dans ce qui suit, on examine des relations pour quelques modèles simples : loi lognormale, loi gamma, loi de Sichel (dans sa version continue), loi de Poisson et loi exponentielle poissonisée. Nous y joignons deux autres lois satisfaisant un critère d'extremalité : la loi S-Max, qui maximise S à m et σ^2 fixés, et une autre loi qui vérifie le même critère dans le cas des lois présentant la symétrie en mode T définie ci-dessous.

B – LA SYMETRIE EN MODE T

Dans ce qui suit, nous utilisons les notations usuelles en géostatistique, dont la définition est rappelée en Annexe 1. Nous disons qu'une loi de probabilité F est symétrique en mode T, ou, plus brièvement, qu'elle est T-symétrique, si la fonction $Q(T)$ qui lui est associée admet un graphe symétrique par rapport à la droite d'équation $Q + T = 1$. Voir, par exemple, Figure 1.

Comme on a toujours $Q(o) = 0$, la T-symétrie entraine $Q(1) = 1$, c'est-à-dire $m = 1$: toute loi T-symétrique admet la moyenne $m = 1$. Cette condition $m = 1$ n'est pas réellement restrictive, puisque changer Z en Z/m revient à changer $Q(T)$ en $Q(T)/m$. De même, les inégalités $0 \leqq T \leqq 1$

entrainant $0 \leqq Q \leqq 1$ s'il y a T-symétrie, la fonction concave Q(T) doit être non décroissante, et cela entraine que la loi F est concentrée sur $\mathbb{R}_+$: toute variable T-symétrique est positive. Dans ce qui suit, et sauf mention explicite du contraire, nous ne considérons plus que des variables $\geqq 0$

On obtient un critère plus commode de T-symétrie en utilisant le paramètrage en z, et les deux fonctions usuelles :

$$(1) \qquad T(z) = \int_z^\infty F(dx) \quad ; \quad Q(z) = \int_z^\infty x\, F(dx)$$

(Noter que la même lettre Q désigne deux fonctions, Q(T) et Q(z), bien différentes mais représentant la même grandeur physique (la "quantité de métal"). Nous suivons en cela l'usage des physiciens. Pour un mathématicien, la fonction Q(z) serait notée $Q \circ T : z \to Q(T(z))$.)

De fait, s'il y a T-symétrie, deux tangentes, ou, plus généralement, deux droites d'appui à la courbe Q(T) symétriques par rapport à la diagonale $Q + T = 1$ ont des pentes z et $1/z$ inverses l'une de l'autre. Il en résulte aussitôt qu'en variable z la condition nécessaire et suffisante de T-symétrie s'écrit :

$$(C1) \qquad T(1/z) = 1 - Q(z)$$

Lorsqu'il existe une densité f(z), le critère (C1) prend la forme :

$$(C'1) \qquad f(1/z) = z^3 f(z)$$

On peut remarquer que $T(1/z) = P(Z > 1/z)$ est la fonction de répartition de la variable inverse $1/Z$, tandis que $1 - Q(z)$ est l'intégrale de $z\, F(dz)$. D'une manière générale, si Z est une variable positive, F sa loi et m son espérance (finie), nous désignerons par Z' une variable admettant la loi

$$F'(dz) = (z/m)\, F(dz)$$

et nous dirons que Z' est l'autopondérée de Z. Dans ces conditions, nous pouvons énoncer comme suit le critère (C1) :

Une variable Z est T-symétrique si et seulement si son inverse 1/Z est équivalent en loi à son autopondérée Z', soit symboliquement :

$$(C2) \qquad 1/Z \overset{\mathcal{L}}{\equiv} Z'$$

En effet, on vient de voir que (C1) entraine (C2). La réciproque est évidente dès que l'on remarque que (C2) entraine $m = E(1/Z')$, et donc $m = 1$ puisque $E(1/Z') = 1/m$.

Le théorème suivant fournit un troisième critère.

Une variable Z est T-symétrique si et seulement si son espérance conditionnelle à $Z + 1/Z = S$ fixé est égale à 1 quel que soit S :

(C3) $$E(Z|Z + 1/Z) = 1$$

En effet, si (C2) est vrai, Z équivaut à $1/Z'$ et $\varphi(Z+1/Z)$ à $\varphi(Z'+1/Z')$, pour toute fonction φ mesurable et bornée. On a donc :

$$E(Z\ \varphi(Z + 1/Z)) = E((1/Z')\ \varphi(Z'+1/Z')) =$$

$$= \int_0^\infty \frac{1}{x}\ \varphi(x + \frac{1}{x})\ x\ F(dx) = E(\ \varphi(Z+1/Z))$$

Mais cela équivaut à (C3).

Inversement, supposons (C3) vérifié. A $Z + 1/Z = s$ fixé (avec d'ailleurs nécessairement $|s| \geqq 2$), Z ne peut prendre que deux valeurs, à savoir les racines de l'équation $x^2 - sx + 1 = 0$. Ces racines ayant toutes deux le même signe que s, il s'ensuit $s > 0$ et donc $s \geqq 2$ pour que (C3) soit possible. Désignons par

$$x_o = x_o(s) = \frac{1}{2}(s - \sqrt{s^2-4})$$

la plus petite de ces deux racines, avec donc $0 < x_o \leqq 1$. D'après (C3), la loi de Z à $Z + 1/Z = s$ fixé est nécessairement:

(2) $$F(dz|s) = \frac{1}{1+x_o}\,\delta_{x_o} + \frac{x_o}{1+x_o}\,\delta_{\frac{1}{x_o}}$$

L'équivalence (C2) de l'inverse et de l'autopondérée est donc assurée pour cette loi conditionnelle $F(dz|s)$ et subsiste manifestement par déconditionnalisation, ce qui achève la démonstration.

La loi lognormale de moyenne $m = 1$ fournit un exemple simple de loi symétrique en mode T. En effet, pour la variable

$$Z = \exp(sY - s^2/2)$$

où Y est une gaussienne réduite et s un paramètre > 0, le paramètrage en z est donné par :

$$T(z) = 1 - G(y) \quad ; \quad Q(z) = 1 - G(y-s) \qquad (y = (1/s)\log z + s/2)$$

où G est la fonction de répartition de la gaussienne réduite Y. Or, changer z en $1/z$ revient ici à changer y en $s-y$, de sorte que le critère (C1) est vérifié : pour toute variable lognormale, la courbe $Q(T)/m$ vérifie la symétrie en mode T.

Rappelons que, pour toute loi lognormale, le coefficient de variation σ/m et l'indice de sélectivité S/m ont pour valeurs :

$$\sigma/m = (\exp(\sigma^2) - 1)^{1/2} \quad ; \quad S/m = 2\, G(s/\sqrt{2}) - 1$$

C - LA LOI DE SICHEL

La loi de Sichel (version continue) constitue un second exemple de symétrie en mode T. La manière la plus simple de l'introduire consiste à partir du mouvement brownien (voir P. Lévy, [1]) Si X_t est le processus gaussien d'espérance nulle et de variance $\sigma^2(t) = t$, on sait que l'instant aléatoire où la trajectoire de ce processus atteint pour la première fois la droite $x = \alpha - \beta\, t$ ($\alpha > 0$ et $\beta \geq 0$ donnés) admet la loi de densité

(3) $$f(z) = \frac{\alpha}{\sqrt{2\pi}}\, \frac{1}{z^{3/2}} \exp\left(\alpha\beta - \frac{\alpha^2}{2z} - \frac{\beta^2}{2} z\right)$$

Moyenne m et variance σ^2 n'ont de valeurs finies que si $\beta > 0$ strictement. On trouve :

$$m = \alpha/\beta \quad ; \quad \sigma^2 = \alpha/\beta^3$$

Le cas $\alpha = \beta$ correspond à $m = 1$ (et, quitte à changer Z en Z/m, on peut toujours se ramener à ce cas). Mais, pour $\alpha = \beta$, il est clair que la densité f définie en (3) vérifie le critère (C'1). Ainsi, toute variable de Sichel de moyenne $m = 1$ vérifie la symétrie en mode T.

Cette symétrie est aussi apparente sur les formules ci-après, qui donnent la quantité de minerai $T(z)$ et la quantité de métal $Q(z)$, définies selon les relations (1) ci-dessus :

(4) $$\begin{cases} T(z) = 1 - G\left(\dfrac{\beta z - \alpha}{\sqrt{z}}\right) - e^{2\alpha\beta}\left[1 - G\left(\dfrac{\beta z + \alpha}{\sqrt{z}}\right)\right] \\[2ex] Q(z)/m = 1 - G\left(\dfrac{\beta z - \alpha}{\sqrt{z}}\right) + e^{2\alpha\beta}\left[1 - G\left(\dfrac{\beta z + \alpha}{\sqrt{z}}\right)\right] \end{cases}$$

Dans ces relations, G est la fonction de répartition de la loi de Gauss réduite. Ces formules remarquables permettent un calcul très rapide des quantités de minerai et de métal. On peut les justifier, sans difficulté, par dérivation directe. Elles admettent aussi, en termes de mouvement brownien, une interprétation probabiliste simple : en écrivant que l'évènement "$X(z) < \alpha - \beta z$" peut se réaliser de deux façons selon que le processus a, ou non, atteint la droite $x = \alpha - \beta t$ antérieurement à $t = z$, on obtient, en effet, une relation de la forme :

$$G\left(\frac{\alpha - \beta z}{\sqrt{z}}\right) = P(T \geq z) + P(T < z \text{ et } X(z) < \alpha - \beta z)$$

qui s'identifie à la première relation (4).

a - TRANSFORMEE DE LAPLACE DE LA LOI DE SICHEL

Si, dans la relation

$$\int_0^{\infty} \exp\left(-\frac{\alpha^2}{2z}-\frac{\beta^2}{2}z\right)dz = \frac{\sqrt{2\pi}}{\alpha}\exp(-\alpha\beta)$$

qui exprime que (3) est une loi de probabilité, nous remplaçons β par $(\beta^2+2\lambda)^{1/2}$ et α par $(\alpha^2+2\mu)^{1/2}$, nous obtenons sans peine :

$$(5) \qquad E\left[e^{-\lambda Z-\mu/Z}\right] = \frac{\alpha}{\sqrt{\alpha^2+2\mu}}\, e^{\alpha\beta-\sqrt{(\beta^2+2\lambda)(\alpha^2+2\mu)}}$$

En faisant $\mu = 0$ dans (5), nous obtenons la transformée de Laplace de la loi de Sichel (3), à savoir :

$$(6) \qquad \Phi(\lambda) = e^{\alpha(\beta-\sqrt{\beta^2+2\lambda})}$$

Sous cette forme, il est clair que la loi de Sichel est indéfiniment divisible, comme cela résultait d'ailleurs déjà de l'interprétation en termes de mouvement brownien.

Plaçons-nous dans le cas $\alpha = \beta$, c'est-à-dire $m = 1$. Pour $\mu = \lambda$, la relation (5) donne :

$$E\left[e^{-\lambda(Z+1/Z-2)}\right] = \frac{\alpha}{\sqrt{\alpha^2+2\lambda}}$$

Ainsi, la variable $S = Z + 1/Z - 2$ obéit à une loi gamma de paramètre 1/2. Compte tenu de (2), on en déduit un procédé simple de simulation : Partant d'une gaussienne réduite X, de sorte que $X^2/2$ est une variable gamma réduite, on prendra :

$$S = X^2/\alpha^2 \quad ; \quad z_o(S) = \frac{1}{2}(S+2-\sqrt{S^2+4S})$$

puis $Z = z_o(S)$ avec la probabilité $1/(1+z_o)$ ou $Z = 1/z_o(S)$ avec la probabilité complémentaire : la variable Z ainsi construite obéit à la loi de Sichel (3), avec $\alpha = \beta$.

Du point de vue numérique, on constate qu'à moyenne et variance égales, la loi de Sichel et la loi lognormale diffèrent relativement peu tant que le coefficient de variation σ/m est $\geqq 1$ (Voir Tableau 1 et Figure 1). Pour $\sigma/m \geqq 2$, par contre, les différences deviennent importantes. Compte tenu des grands avantages que lui confère son indéfinie divisibilité, on peut se demander si la loi de Sichel ne devrait remplacer plus souvent la loi lognormale dans les applications.

Contrairement au cas lognormal, il n'existe pas, semble-t-il, pour la loi de Sichel, de formule explicite simple donnant l'indice de sélectivité S/m, et les valeurs numériques présentées sur la Figure 2 et les Tableaux 2 et 3 ont été calculées par simple intégration numérique.

D – QUELQUES LOIS EXTREMALES

On sait (Voir Annexe 3) que toute variable aléatoire, positive ou non, vérifie une inégalité de Schwarz de la forme :

$$S \leqq \sigma/\sqrt{3} \tag{7}$$

où S est la sélectivité et σ l'écart-type, avec égalité si et seulement si la distribution est uniforme sur un intervalle (a,b). En fait, le calcul donne $a = m - \sigma\sqrt{3}$, de sorte que la valeur extrême $S = \sigma/\sqrt{3}$ n'est possible, pour une variable positive, que si la variance relative σ^2/m^2 est inférieure à 1/3. Pour $\sigma^2/m^2 > 1/3$, l'indice de sélectivité S/m d'une variable positive vérifie une inégalité plus sévère que (7); à savoir :

$$S/m \leqq 1 - \frac{8}{9}\frac{m^2}{m^2+\sigma^2} \tag{8}$$

avec égalité si et seulement si la distribution comporte une probabilité > 0, ou atome, placée en z = 0, plus une densité uniforme sur un intervalle (o,b). Sur les figures et les tableaux ci-joints, la grandeur Smax correspond à cette loi extrêmale (loi uniforme avec ou sans atome en z = 0 selon que le coefficient de variation est ou non supérieur à $1/\sqrt{3}$).

Cependant, cette loi extrêmale n'est nullement symétrique en mode T. On peut en fait montrer sans grande difficulté le théorème suivant : parmi les lois T-symétriques de variance σ^2 donnée (et de moyenne m = 1 naturellement), il en existe une qui maximise la sélectivité S. Elle admet une densité de la forme :

$$f(z) = C(1 + 1/z^3) \qquad (\varepsilon \leqq z \leqq 1/\varepsilon)$$

sur un intervalle de la forme $(\varepsilon, 1/\varepsilon)$, pour un ε compris entre 0 et 1. La constante de normalisation $C = C(\varepsilon)$ est donnée par

$$C(\varepsilon) = \frac{2\,\varepsilon^2}{(1-\varepsilon)(1+\varepsilon)^3}$$

La moyenne est m = 1. Variance σ^2 et sélectivité S sont données par :

$$(9)\quad \begin{cases} \sigma^2 = C\left(\dfrac{1}{3\varepsilon^3} + \dfrac{\varepsilon^3}{3} - 2\log\varepsilon\right) - 1 \\ S = C\left(\dfrac{1}{2\varepsilon^3} + \dfrac{\varepsilon^2}{2} - \varepsilon - \dfrac{1}{\varepsilon} + 1 + \sigma^2\right) \end{cases}$$

On notera que ni l'une ni l'autre des deux lois extrêmales précédentes n'est indifiniment divisible. On verra, en Annexe 3, que, dans le cas d'une variable indéfiniment divisible, positive ou non, l'inégalité (7) est remplacée par :

$$(10)\qquad S \leqq \sigma/\sqrt{\pi}$$

avec égalité si et seulement si la distribution est gaussienne.

Cette inégalité conduit à définir un indice de normalité I_n, en posant

$$I_n = \frac{S}{\sigma}\sqrt{\pi}$$

Dans un modèle de changement de support, la croissance de cet indice reflète la convergence vers la normalité. On peut voir sur les tableaux et figures ci-après que, dans les cas étudiés ici, cet indice de normalité augmente lorsque la variance relative diminue. On notera que l'on n'a $I_n \leqq 1$ que pour les lois indéfiniment divisibles, mais seulement $I_n \leqq \sqrt{\pi/3} = 1{,}023\ \ldots$ en général, avec égalité pour la loi uniforme.

Dans l'inégalité (10), l'égalité n'est jamais réalisée pour une variable à la fois positive et indéfiniment divisible, de sorte que, dans ce cas, il doit exister une majoration plus sévère, et même beaucoup plus sévère pour les lois à haute dispersion, puisque l'inégalité (8), plus stricte que (10) en ce cas, doit elle-même être renforcée. Mais je ne connais pas la forme exacte de cette majoration.

Faute de la connaître, on a porté sur les tableaux et figures ci-après les valeurs de l'indice de sélectivité S/m pour trois familles de lois : tout d'abord la loi gamma, pour laquelle, en fonction du paramètre $\alpha = m^2/\sigma^2$, on a :

$$S/m = \frac{\Gamma(\alpha+1/2)}{\sqrt{\pi}\ \Gamma(\alpha+1)}$$

puis deux autres lois : la loi de Poisson, et une autre loi que l'on peut appeler loi exponentielle poissonisée définie par sa transformée de Laplace :

$$E(e^{-\lambda Z}) = e^{-a\frac{\lambda}{c+\lambda}}$$

c'est-à-dire que Z est la somme de N variables exponentielles, où N est elle-même une variable de Poisson.

Il est remarquable qu'à variances relatives égales, ces deux lois (Poisson et exponentielle poissonisée) ont le même indice de sélectivité, donné par :

$$S/m = e^{-4/V} \sum_{n \geq 1} \left(\frac{4}{V}\right)^n \frac{\Gamma(n+1/2)}{n!\ (n-1)!\ \sqrt{\pi}}$$

où V est la variance relative, soit $V = 1/\theta$ pour une variable de Poisson et $V = 2/a$ pour l'exponentielle poissonisée. (Ces formules se déduisent sans peine de la relation (1-8) de l'Annexe 1).

Sur les tableaux et figures ci-joints, on peut voir qu'à variance relative donnée, c'est toujours la loi lognormale qui présente le plus faible indice de sélectivité S/m. Viennent ensuite, dans l'ordre des indices de sélectivité croissants, la loi de Sichel continue, la loi en $1 + 1/z^3$ (maximum de S dans le cas T-symétrique), ensuite la loi gamma, puis la loi de Poisson ou l'exponentielle poissonisée, et enfin la loi uniforme avec atome éventuel, qui réalise le maximum de S pour les variables positives. Il n'y a qu'une seule interversion, aux très petites valeurs de σ^2/m^2, la loi gamma et la loi de Poisson précédant alors la loi en $1 + 1/z^3$.

TABLEAU 1

Quelques valeurs de Q(T)/m, pour différentes lois de même variance relative $\sigma^2/m^2 = 1$

Lois	T = .05	T = .25	T = .5	T = .75
Lognormale	.201	.563	.798	.934
Sichel	.211	.577	.807	.936
en $1 + 1/z^3$	.181	.641	.832	.929
exponentielle	.200	.597	.847	.966
S Max.	.144	.609	.937	1.000

TABLEAU 2

Evolution de l'indice de sélectivité S/m en fonction de la variance relative $V = \sigma^2/m^2$

Lois	V = 0.1	V = 1	V = 10	V = 100
Lognormale	.173	.444	.726	.871
Sichel	.173	.457	.800	.948
en $1 + 1/z^3$	.179	.488	.862	.983
gamma	.176	.500	.883	.986
Poisson	.177	.524	.909	.990
S Max.	.183	.556	.919	.991

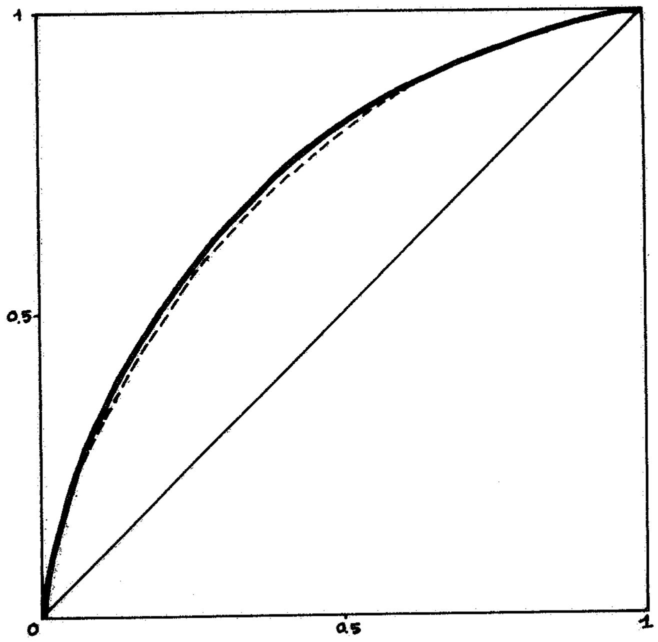

Figure 1

Courbe Q = Q(T), pour la loi de Sichel (en traits pleins) et la loi lognormale dans le cas $\sigma^2/m^2 = 1$.

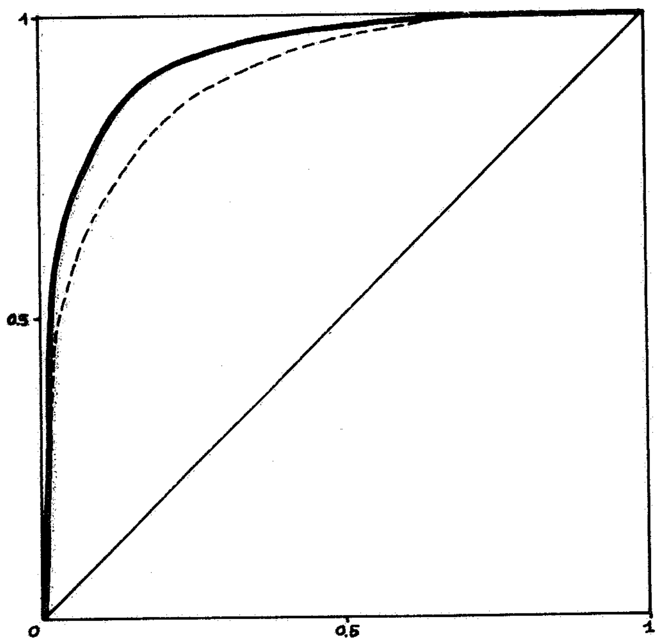

Figure 1 bis

Courbe Q = Q(T); pour la loi de Sichel (en traits pleins) et la loi lognormale (en pointillés), dans le cas $\sigma^2/m^2 = 20$.

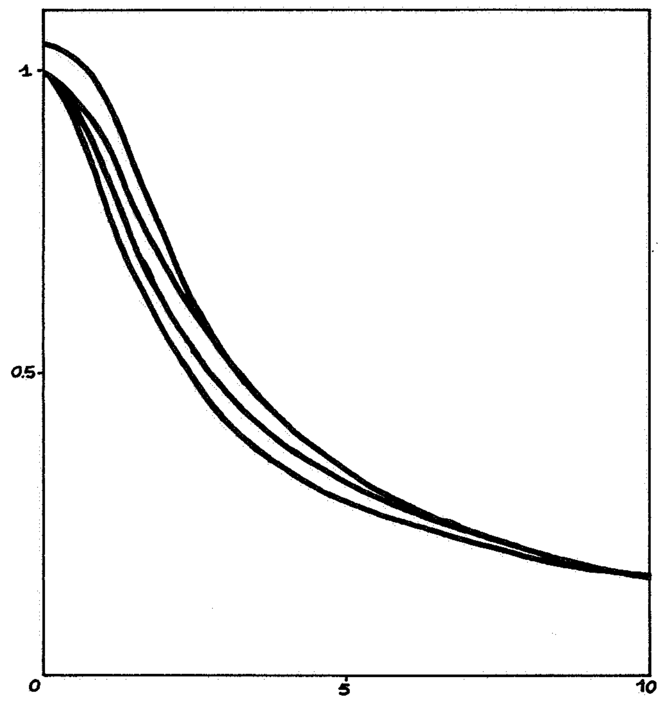

Figure 2

Evolution de l'indice de normalité $S\sqrt{\pi}/\sigma$ en fonction du coefficient de variation σ/m. De haut en bas : loi S-max, loi gamma, loi de Sichel et loi lognormale.

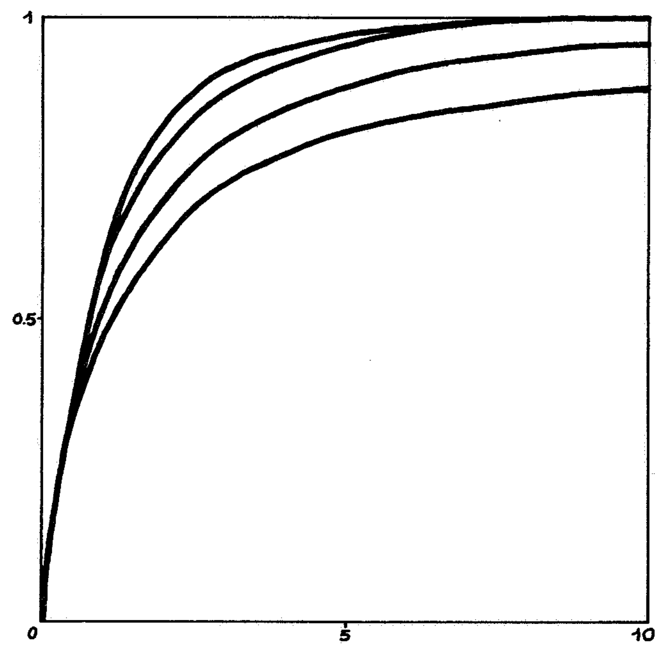

Figure 3

Evolution de l'indice de sélectivité S/m en fonction du coefficient de variation σ/m. De haut en bas : loi S-max, loi gamma, loi de Sichel, loi lognormale.

TABLEAU 3				
Evolution de l'indice de normalité I_n en fonction de la variance relative $V = \sigma^2/m^2$				
Lois	V = 0.1	V = 1	V = 10	V = 100
Lognormale	.969	.787	.407	.154
Sichel	.971	.810	.448	.168
en $1 + 1/z^3$	1.001	.865	.483	.174
gamma	.988	.886	.495	.175
Poisson	.994	.928	.510	.175
S Max.	1.023	.985	.515	.176

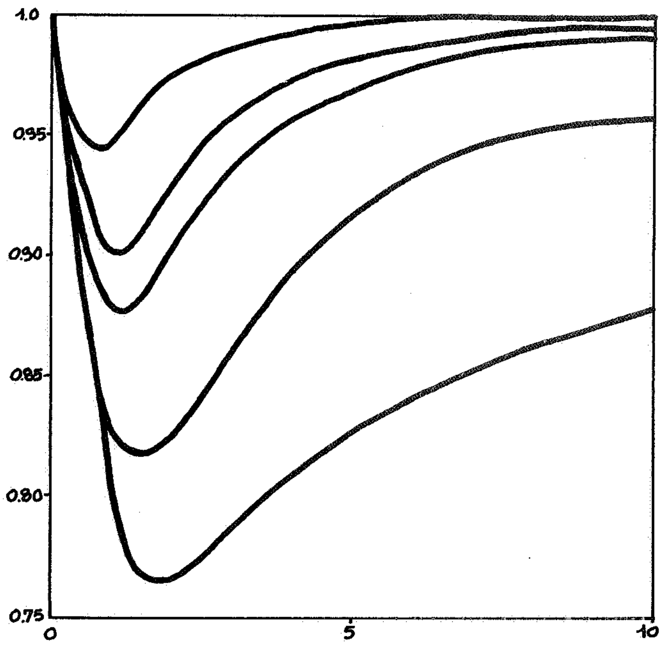

Figure 4

Evolution de l'indice de maximalité S/S-max en fonction du coefficient de variation σ/m. De haut en bas : Loi de Poisson (ou loi exponentielle poissonisée), loi gamma, loi en $1 + 1/z^3$, loi de Sichel et loi lognormale.

ANNEXE 1 : LES NORMES

a - DEFINITION ET RAPPELS

Pour des raisons de simplicité, on se limite ici au cas des variables Z positives. Pour plus de détails, on se reportera à G. Matheron, [2] . A toute probabilité F(dx) sur $\mathbb{R}^+$, on associe la fonction non croissante $T_F(z)$, (ou simplement T(z), s'il n'y a pas ambiguité), définie presque partout sur $\mathbb{R}^+$, et admettant une version T^- continue à gauche, et une version T^+ continue à droite :

$$T^-(z) = P(Z \geqq z) \quad ; \quad T^+(z) = P(Z > z)$$

De même, $z_F(t)$, ou simplement z(t), désignera la fonction inverse de T(z), définie presque partout sur (0,1). C'est une fonction non croissante, admettant deux versions z^- et z^+ continues à gauche et à droite respectivement. Si U est une variable uniformément distribuée sur (0,1), $z_F(U)$ admet la loi F.

La moyenne m_F, finie ou non, est donnée par

(1-1) $$m_F = \int_0^\infty T_F(z)\ dz = \int_0^1 z_F(t)\ dt$$

et la paramètre S_F de sélectivité est :

(1-2) $$S_F = \int_0^\infty T_F(1-T_F)\ dz = \int_0^1 z_F(t)(1-2T)\ dt$$

Il a une valeur finie si et seulement si la moyenne m_F est $< \infty$. On utilise dans ce cas l'indice de sélectivité défini par le rapport S/m, avec l'inégalité évidente $S/m \leqq 1$.

La seconde relation (1-2) montre que S est le double de la covariance de z(U) et 1-U, pour U uniforme sur (0,1). D'où l'inégalité se Schwarz :

(1-3) $$S \leqq \sigma/m$$

où σ est l'écart-type, avec égalité si et seulement si la loi F est uniforme (sur un intervalle quelconque). On voit que S/2 représente, en somme, la covariance de la variable et de son rang. De fait, si l'on définit le rang d'une valeur z comme la quantité $(F_+(z) + F_-(z))/2$, on trouve, en effet, exactement :

$$S = \int_0^\infty \left[F_+(z) + F_-(z) \right] \ z\ F(dz) - m$$

On utilise aussi les fonctions $Q_F(T)$ et $B_F(z)$, définies par :

(1-4) $$B(z) = E\left[(Z-z)_+\right] = \int_z^{\infty} T(u)\,du \quad ; \quad Q(T) = \int_0^{T} z(t)\,dt$$

et vérifiant les relations de dualité

$$B(z) = \operatorname*{Sup}_T \{Q(T)-zT\} \quad ; \quad Q(T) = \operatorname*{Inf}_z \{B(z)+zT\}$$

La sélectivité S est liée à ces fonctions par les relations :

$$S = \int_0^{\infty} B(z)\,F(dz) = 2\int_0^{1} \left[Q(T) - mT\right] dT$$

b - LA NORME EN T ET LA NORME EN z

Dans ce contexte, il est naturel d'introduire les normes hilbertiennes

$$\|T\|^2 = \int_0^{\infty} T^2(z)\,dz \quad ; \quad \|z\|^2 = \int_0^{1} z^2(t)\,dt$$

et de munir l'espace des lois F de la norme en T, soit $\|T_F\|^2$, ou, aussi bien, de la norme en z, soit $\|z_F\|^2$. Si $m_F < \infty$, on a :

(1-5) $$\|T_F\|^2 = m_F - S_F \quad ; \quad \|z_F\|^2 = m_F^2 + \sigma_F^2$$

D'après les relations

(1-6 $$\|T\|^2 = 2\int_0^{1} z(u)\,u\,du \quad ; \quad \|z\|^2 = 2\int_0^{\infty} T(z)\,z\,dz$$

il apparaît de plus que $\|T_F\|^2$ est une fonctionnelle linéaire de z_F , $\|z_F\|^2$ une fonctionnelle linéaire de T_F . Mais, de ces deux fonctionnelles linéaires, seule la première est continue. On verra en Annexe 2 les conséquences de ce point.

A ces normes sont associés des produits scalaires. Si F et G sont deux lois de probabilités, on écrira :

$$\langle T_F\, T_G \rangle = \int_0^{\infty} T_F(z)\,T_G(z)\,dz \; ; \; \langle z_F\, z_G \rangle = \int_0^{1} z_F(t)\,z_G(t)\,dt$$

Le produit scalaire en T admet l'interprétation suivante :

$$\langle T_F\, T_G \rangle = E(X_F \wedge X_G)$$

où X_F et X_G sont deux variables indépendantes admettant respectivement les lois F et G. Lorsque les moyennes existent, ceci s'écrit aussi :

$$< T_F \; T_G > = \frac{1}{2} \; [m_F + m_G - E\,(|X_F - X_G|\,)]$$

En particulier, si F = G, compte tenu de $\| T \|^2 = m - S$, nous obtenons une expression intéressante du paramètre S de sélectivité

$$S = \frac{1}{2}\, E(\,|\,X - X'|\,) \tag{1-7}$$

où X et X' sont deux variables indépendantes admettant la même loi F. Si $\Phi(u) = E(\exp(iuX))$ est la fonction caractéristique associée à la loi F, celle de la variable X - X' est $|\Phi(u)|^2$. Compte tenu de (1-7) et de la relation :

$$|x-x'| = \frac{2}{\pi} \int_0^\infty [1 - \cos u(x-x')] \, \frac{du}{u^2}$$

on obtient donc, par application du théorème de Fubini, la formule remarquable :

$$S = \frac{1}{\pi} \int_0^\infty [1 - |\Phi(u)|^2] \, \frac{du}{u^2} \tag{1-8}$$

Concernant le produit scalaire en z, le théorème ci-dessous exprime que $< z_F \; z_G > - m_F \, m_G$ est la plus grande valeur possible de la covariance de deux variables X et Y admettant les marginales F et G.

THEOREME - Pour toute loi bivariable (X,Y), admettant les marginales F et G, on a $< X \; Y > \leqq < z_F \; z_G >$, et l'égalité est atteinte si $X = z_F(U)$ et $Y = z_G(u)$, pour U uniforme sur (0,1).

En effet, la relation simple $Z = \int_0^\infty 1_{Z \geqq z} \, dz$ entraine pour tout couple X,Y :

$$< X \; Y > = \int_0^\infty \int_0^\infty P(X \geqq x \text{ et } Y \geqq y) \; dx \; dy$$

Mais on a toujours $P(X \geqq x \text{ et } Y \geqq y) \leqq T_F(x) \wedge T_G(y)$, et de plus cette inégalité devient une égalité dans le cas $X = z_F(U)$, $Y = z_G(U)$, car $z_F(U) \geqq x$ équivaut à $U \leqq T_F(x)$. En intégrant, il vient donc :

$$< X \; Y > \leqq \int \int (T_F(x) \wedge T_G(Y)) \; dx \; dy = < z_F \; z_G >$$

ANNEXE 2 : LES CONVERGENCES

Grâce aux deux normes ci-dessus, l'espace des lois F se trouve muni de quatre convergences (les convergences forte et faible associées à chacune de ces normes), auxquelles s'ajoutent la convergence étroite (ou en loi) et la convergence vague des probabilités sur $\mathbb{R}^+$.

On rappelle que la convergence vague $F_n \to G$ a lieu si $T_{F_n}(z) \to T_G(z)$ en tout point z de continuité de T_G. Mais la mesure G peut être défective (i.e. $T_G(\infty) > 0$). La convergence en loi correspond au cas où l'on a de plus $T_G(\infty) = 0$, c'est-à-dire où la limite G est une probabilité.

LEMME 1 - Une suite F_n de probabilités sur $\mathbb{R}^+$ converge au sens vague vers une mesure G (éventuellement défective) si et seulement si la suite z_{F_n} converge vers z_G presque partout sur (0,1). Cette convergence a lieu en loi si et seulement si de plus z_G est finie presque partout sur (0,1).

A toute fonction positive semi-continue supérieurement sur $\mathbb{R}_+$, on associe son sous-graphe (ensemble des couples (x,t) tels que $t \leqq f(x)$) qui est fermé dans $\mathbb{R}_+ \times \mathbb{R}_+$. Il est facile de voir que la convergence au sens presque partout d'une suite de fonctions f_n s.c.s. et décroissantes équivaut à la convergence de leurs sous-graphes dans l'espace des fermés de $\mathbb{R}_+ \times \mathbb{R}_+$. le lemme en découle, en remarquant que l'on peut identifier les sous-graphes des fonctions s.c.s. T_F^- et z_F^- .

THEOREME 1 - La convergence faible $T_{F_n} \to T_G$ pour la norme T entraine la convergence en loi $G_n \to G$.

En effet, supposons $T_{F_n} \to T_G$ faiblement. D'après le théorème de Helly, on peut extraire une suite partielle F_{n_k} convergeant vaguement vers une limite G', ce qui implique $T_{F_{n_k}} \to T_{G'}$ presque partout sur $\mathbb{R}^+$. Comme on a aussi $T_{F_n} \to T_G$ faiblement, un théorème classique garantit l'identité de ces deux limites, soit G' = G. Comme $T_{G'} = T_G$ a une norme finie, on a $T_{G'}(\infty) = 0$ et la convergence vers G' a lieu en loi (lemme 1). Enfin, l'unicité de la limite G' = G entraine que la suite F_n elle-même converge en loi vers G.

LEMME 2 - La convergence faible $z_{F_n} \to z_G$ pour la norme z entraine la convergence des moyennes $m_{F_n} \to m_G$ et celles des normes en T, $\|T_{F_n}\| \to \|T_G\|$.

En effet, pour toute loi F, on a $m_F = < z_F, 1 >$, où 1 est l'indicatrice de (0,1), au sens de la norme en z, et de même, d'après (1,6), $\| T_F\|^2 = 2 < z_F\ I >$ où I est la fonction identité sur (0,1). Donc la convergence faible $z_{F_n} \to z_G$ entraine l'existence des moyennes et des normes en T ainsi que leur convergence.

Voici maintenant le résultat le plus important :

THEOREME 2 - La convergence faible $z_{F_n} \to z_G$ en norme z entraine la convergence forte $T_{F_n} \to T_G$ en norme T.

Supposons, en effet, qu'une suite z_{F_n} converge faiblement vers z_G dans $L^2((0,1),dx)$. D'après le lemme 2, la suite $\| T_{F_n} \|$ est bornée. D'après la compacité séquentielle de $L^2(\mathbb{R}_+,dz)$, on peut donc extraire une suite $T_{F_{n_k}}$ convergeant faiblement vers une limite $T_{G'}$. D'après le théorème 1, on a alors $F_{n_k} \to G'$ en loi, et donc, d'après le lemme 1, $z_{F_{n_k}} \to z_{G'}$ presque partout sur (0,1). Or, on a aussi $z_{F_{n_k}} \to z_G$ au sens faible, et on en déduit, d'après un résultat classique, $G' = G$. L'unicité de cette limite montre alors que la suite T_{F_n} elle-même converge vers T_G au sens faible. Mais, d'après le lemme 2, on a de plus la convergence des normes $\| T_{F_n} \| \to \| T_G \|$, et cela entraine, classiquement, que la convergence $T_{F_n} \to T_G$ a lieu, en fait, aussi au sens fort.

En résumé, nos 6 convergences sont parfaitement hiérarchisées : la convergence forte en norme z entraine évidemment la convergence faible en norme z, laquelle entraine la convergence forte en norme T (théorème 2), qui à son tour entraine la convergence faible en norme T. Mais cette dernière entraine la convergence en loi (théorème 1), laquelle enfin entraine la convergence vague. Mais les réciproques sont fausses, et cette hiérarchie est stricte.

Voici une application qui sera utile dans l'Annexe 3 :

Désignons par $\mathscr{L}(m,\sigma^2)$ l'ensemble des lois F sur $\mathbb{R}^+$ telles que

$$m_F = m \quad ; \quad \sigma_F^2 \leq \sigma^2$$

Alors $\mathscr{L}(m,\sigma^2)$ est fermé et même compact pour la norme T.

En effet, soit F_n une suite dans $\mathscr{L}$. On a $\| z_{F_n} \|^2 \leq m^2 + \sigma^2$ par hypothèse, de sorte que l'on peut extraire une suite partielle F_{n_k} telle que $z_{F_{n_k}}$ converge vers une limite z_F au sens faible pour la norme en z. D'après le théorème 2, $T_{F_{n_k}}$ converge donc fortement vers T_F pour la norme T. D'après le lemme 2, la suite $m_{F_n} = m$ converge vers m_F, et par suite $m_F = m$. De même, la convergence faible en norme z donne $\| z_F \|^2 \leq \underline{\lim} \| z_{F_n} \|^2 \leq m^2 + \sigma^2$, d'où $\sigma_F^2 \leq \sigma^2$. On a donc $F \in \mathscr{L}(m,\sigma^2)$, ce qui prouve la compacité de $\mathscr{L}$.

Noter aussi que $\mathscr{L}(m,\sigma^2)$, en tant qu'ensemble de fonctions T_F, est convexe. On vérifie, en effet, sans difficulté que pour F_o, $F_1 \in \mathscr{L}(m,\sigma^2)$ et $0 < \lambda < 1$, la loi F_λ définie en posant

$$T_{F_\lambda} = (1-\lambda)\, T_{F_o} + \lambda T_{F_1}$$

est encore dans $\mathscr{L}(m,\sigma^2)$.

ANNEXE 3 : LES INEGALITES

Nous venons de voir (Annexe 2) que l'ensemble $\mathcal{L}(m,\sigma^2)$ des lois F sur $\mathbb{R}_+$ de moyenne m donnée et de variance $\sigma_F^2 \leqq \sigma^2$ est convexe et fermé pour la norme T. D'après le théorème des projections, il existe donc un élément unique T_{F_o} réalisant le minimum de la norme $\|T_F\|$ dans $\mathcal{L}(m,\sigma^2)$. D'après la relation $\|T\|^2 = m - S$, cet élément unique est aussi celui qui réalise le maximum de la sélectivité S_F pour une moyenne $m_F = m$ et une variance $\sigma_F^2 \leqq \sigma^2$.

D'après le théorème des projections, cet élément unique est caractérisé par la relation

(3-1) $$< T_F\, T_{F_o} > \;\geqq \|T_{F_o}\|^2 \qquad \forall\, F \in \mathcal{L}(m,\sigma^2)$$

En Annexe 1, la relation de Schwarz (1-3) nous indique que le maximum de S, parmi toutes les lois sur $(-\infty, +\infty)$ admettant une moyenne et une variance données, est atteint pour une loi uniforme sur un segment (a, a+L). D'après les relations

$$m = a + L/2 \quad ; \quad \sigma^2 = L^2/12$$

on trouve $a = m - \sigma\sqrt{3}$. En particulier, on aura $a < 0$ pour $\sigma^2/m^2 > 1/3$, et dans ce cas la majoration (1-3) sera remplacée par une majoration plus forte $S \leqq S_{F_o}$.

Pour $\sigma^2/m^2 \leqq 1/3$, au contraire, on a $a \geqq 0$, et la loi F_o, solution de (3-1), est donc la loi uniforme ci-dessus. Dans le cas $\sigma^2/m^2 > 1/3$, montrons que F_o comporte un atome α en $z = 0$ et une partie uniforme sur un segment (0,L). Ce qui s'écrit :

$$T_{F_o} = (1-\alpha)\ \left[1 - z/L\right]_+$$

En effet, choisissons α et L de manière à obtenir la moyenne m et la variance σ^2. De :

$$m = (1-\alpha)\, L/2 \quad ; \quad m^2 + \sigma^2 = (1-\alpha)\, L^2/3$$

on déduit qu'il faut prendre

$$L = \frac{3}{2}\ \frac{m^2+\sigma^2}{m} \quad ; \quad 1 - \alpha = \frac{4}{3}\ \frac{m^2}{m^2+\sigma^2}$$

(et la condition $\sigma^2 > m^2/3$ garantit $\alpha > 0$). Avec cette définition, on a bien $F_o \in \mathcal{L}(m,\sigma^2)$. Soit alors F quelconque dans $\mathcal{L}(m,\sigma^2)$. D'après la définition de F_o, il vient :

(3-2) $$< T_F\, T_{F_o} > \;= (1-\alpha) \int_0^L T_F(z)\left(1 - \frac{z}{L}\right) dz \;\geqq\; (1-\alpha) \int_0^\infty T_F(z)\left(1 - \frac{z}{L}\right) dz$$

et par suite :

(3-3) $$< T_F \, T_{F_o} > \;\geqq (1-\alpha) \left[m - \frac{1}{2} \frac{m^2+\sigma^2}{L} \right]$$

puisque l'on a $\int T_F(z)\,dz = m_F = m$, et aussi

$$\int z \, T_F(z)\,dz = \frac{1}{2}(m_F^2 + \sigma_F^2) \leqq \frac{1}{2}(m^2+\sigma^2)$$

Mais, pour $F = F_o$, l'inégalité en (3-2) devient une égalité, puisque $T_{F_o}(z)$ est nul pour $z > L$. Il vient donc :

$$\| T_{F_o} \|^2 = (1-\alpha) \left[m - \frac{1}{2} \frac{m^2+\sigma^2}{L} \right]$$

En comparant avec (3-3), on conclut que (3-1) est vérifiée, et T_{F_o} est bien l'élément maximal cherché. D'après les valeurs calculées ci-dessus pour L et $(1-\alpha)$, il vient explicitement

$$\| T_{F_o} \|^2 = \frac{8}{9} \frac{m^3}{m^2+\sigma^2}$$

Compte tenu de la relation $\| T \|^2 = m - S$, on en déduit :

<u>THEOREME 1</u> - Pour toute loi F sur $\mathbb{R}_+$, on a les inégalités :

$$S \leqq \frac{\sigma}{\sqrt{3}} \quad ; \quad \frac{S}{m} \leqq 1 - \frac{8}{9} \frac{m^2}{m^2+\sigma^2}$$

La première inégalité devient une égalité si et seulement si F est une loi uniforme (et dans ce cas $\sigma^2/m^2 \leqq 1/3$), et la seconde si et seulement si F est une loi uniforme sur un intervalle (0,L) muni d'un atome à l'origine (et dans ce cas $\sigma^2/m^2 > 1/3$).

On note que ces lois extrêmales ne sont pas <u>indéfiniment divisibles</u> (I.D.), de sorte que, si l'on se limite aux lois I.D., la majoration précédente doit pouvoir être considérablement améliorée. Compte tenu du fait que toute limite de lois I.D. est elle-même indéfiniment divisible, il n'est pas difficile de démontrer, à l'aide des résultats de l'Annexe 2, qu'il existe effectivement une loi indéfiniment divisible sur $\mathbb{R}_+$ qui maximise la sélectivité S pour m et σ^2 donnés. Mais nous ne connaissons pas sa forme explicite. Nous nous contenterons ici d'un résultat plus faible, concernant les variables I.D. positives ou non.

<u>THEOREME 2</u> - Pour toute loi indéfiniment divisible sur $\mathbb{R}$, de variance σ^2, la sélectivité S vérifie l'inégalité :

$$S \leqq \sigma/\sqrt{\pi}$$

avec égalité si et seulement si la loi est gaussienne.

La démonstration utilise la relation (1-8) de l'Annexe 1. Si Z est I.D., sa fonction caractéristique est du type

$$\Phi(u) = \exp\,(ium + \psi(u))$$

avec
$$\psi(u) = \int_{-\infty}^{+\infty} \frac{e^{iux} - 1 - iux}{x^2}\, H(dx)$$

où H est une mesure $\geqq 0$ telle que $\sigma^2 = \int H(dx)$. Il vient donc

$$|\Phi(u)|^2 = \exp\left(-2\int_{-\infty}^{+\infty} \frac{1 - \cos ux}{x^2}\, H(dx)\right)$$

Mais, pour tout $x \neq 0$, on a l'inégalité stricte :

$$(1 - \cos ux)/x^2 < u^2/2$$

et donc $|\Phi(u)|^2 \geqq \exp\,(-\,u^2\,\sigma^2)$ avec égalité si et seulement si la mesure H est concentrée en $x = 0$, ce qui correspond au cas où la variable Z est gaussienne. Compte tenu de (1-8), on en déduit bien $S \leqq \sigma/\sqrt{\pi}$ avec égalité si et seulement si Z est gaussienne.

BIBLIOGRAPHIE

[1] LEVY, P. (1965) - Processus Stochastiques et Mouvement Brownien. Gauthier-Villars, Paris.

[2] MATHERON, G. (1983) - The selectivity of the distributions and the second principle of Geostatistics. (Proceedings, 2nd NATO A.S.I. "Geostatistics for Natural Resources Characterization", G. Verly et Al. Editors, D. Reidel Publishing Co., Dordrecht, Netherlands, Part 1, pp. 421-433).