NOTE GEOSTATISTIQUE N° 76

LOIS STABLES ET LOI LOGNORMALE

INTRODUCTION

Dans cette note, un peu disparate, je fixe, avant ~~XXX~~ qu'elles ne s'évaporent, quelques reflexions inspirées de la lecture du Volume II de Feller. Ces reflexions tournent autour du probleme, ancien et ennuyeux, de la permanence de la lognormalité. On sait qu'une moyenne arithmétique, ou, plus généralement, une combinaison linéaire à coëfficients constants et positifs, de variables lognormales (indépendantes ou non) ne peut pas être elle-même lognormale. Expérimentalement, cependant, on observe souvent une lognormalité approchée aussi bonne, parfois même meilleure, que celle des variables initiales. De ce paradoxe apparent, on doit tirer deux conclusions :

1- Aucune de ces lois n'est réellement lognormale. Elles sont seulement lognormaloïdes.

2- Cette permanence paradoxale de la forme de la loi est une propriété caractéristique des lois dites stables. On doit penser que ces soit-disant lois lognormales sont en réalité des lois stables, ou, plus généralement, des lois appartenant au domaine d'attraction d'une loi stable (ce qui expliquerait pourquoi les combinaisons linéaires sont souvent plus proches du type idéal que les variables initiales elles-mêmes).

Il y a une expérience bien facile à réaliser : construire, en coordonnées gausso-logarithmiques, la "droite de Henry" associée à la loi stable de paramètre $\alpha = 1/2$: cette loi décrit la probabilité d'absorption dans un Wiener-Lévy, et sa tabulation se déduit sans peine des tables de Gauss. On obtient effectivement une droite de qualité si bonne que, s'il s'agissait d'un ajustement empirique, on n'hésiterait pas une seconde à décreter la lognormalité.

Cependant, cette substitution des lois stables à la loi lognormale traditionnelle fait apparaitre quelques difficultés inattendues. Les lois stables de paramètre

$\alpha < 1$ et de transformée de Laplace $e^{-\lambda^\alpha}$ (les seules qui nous interessent) n'ont pas d'espérance, ni, a fortiori, de variance. C'est dejà gênant, mais il y a plus grave . Puisque toute combinaison linéaire $\sum \lambda_i X_i$ est équivalente en loi à aX (les X_i et X obéissant à la même loi satble) on voit que $\log \sum \lambda_i X_i$ équivaut en loi à $\log a + \log X_i$: ces deux variables auraient même loi, à une translation près, et, en particulier, même variance logarithmique. D'où ce nouveau paradoxe, que l'on peut formuler comme suit dans le langage de la géostatistique : on aurait un gisement dans lequel les teneurs des échantillons ne possèderaient ni espérance ni variance, mais où par contre les logarithmes des teneurs possederaient la même variance, indépendamment de la taille de l'echantillon.

Le paradoxe provient du fait que l'on a implicitement raisonné sur un gisement de taille infinie. La théorie des variables régionalisées nous a déjà familiarisés avec cette idée que les moments d'une distribution empirique dépendent du champ géometrique, et cessent d'exister, en général, si le champ devient infini (pour un variogramme non borné). Mais on travaille toujours dans un champ fini. Ce que l'on observe expérimentalement, ce ne sont jamais les lois stables elles-mêmes, mais toujours ces mêmes lois prises conditionnellement lorsque la teneur X globale du champ est fixée. Si le champ est constitué de la réunion de volumes v_i disjoints de teneurs X_i , on observe la loi des X_i liées par $\frac{1}{N} \sum X_i = X$ où X , moyenne générale, est fixée, même si on ne la connait pas. Les X_i pris conditionnellement, sont bornés, et possèdent donc des moments de tous les ordres, (qui dépendent de la valeur numérique x de X) . Leur calcul n'est pas facile, mais on peut vérifier que la variance conditionnelle de X_i est effectivement une fonction décroissante de la taille du support.

1°/ Transformée de Laplace, et formule d'inversion.

Soit X une variable ne prenant que des valeurs positives, F(dx) sa loi. La transformée de Laplace de la loi F est :

(1) $$\Phi(\lambda) = E[e^{-\lambda X}] = \int_0^\infty e^{-\lambda x} F(dx) \qquad (\lambda \geq 0)$$

Comme $1 - e^{-\lambda x}$ (à x fixé) est la fonction de répartition d'une variable aléatoire Λ d'espérance $1/x$ et de loi exponentielle on peut interpréter $1-e^{-\lambda x}$ comme la loi conditionnelle de Λ à X = x fixé : la loi a priori de Λ lorsque X redevient aléatoire, admet alors la fonction de répartition :

$$1 - \Phi(\lambda) = \int_0^\infty (1 - e^{-\lambda x}) \, F(dx)$$

Cette interprétation probabiliste simple de la transformée de Laplace conduit à une formule d'inversion classique. Supposons qu'à X = x fixé nous prenions la somme $\Lambda_1 + \ldots + \Lambda_n$ de n variables exponentielles indépendantes d'espérance $1/x$. Cette somme obéit à la loi gamma de densité : $\frac{x^n \lambda^{n-1}}{(n-1)!} e^{-\lambda x}$

Si maintenant X redevient aléatoire, on obtient pour cette somme la loi a priori de densité $\varphi_n(\lambda) = (-1)^n \frac{\lambda^{n-1}}{(n-1)!} \frac{d^n}{d\lambda^n} \Phi(\lambda)$

On peut aussi s'interesser à la moyenne arithmétique $Y_n = \frac{1}{n}(\Lambda_1 + \cdots + \Lambda_n)$ Conditionnellement, à X = x, Y_n converge p.s. vers son espérance $1/x$ lorsque n tend vers l'infini. Par suite (nous en donnerons la démonstration rigoureuse dans un instant) la loi a priori de Y_n converge vers la loi de $1/X$. Autrement dit, désignant par

$$\Phi_n(\lambda) = \int_0^\lambda \varphi_n(u) \, du$$

la loi a priori de $\Lambda_1 + \ldots + \Lambda_n$, celle de Y_n est :

$$P(Y_n < y) = \Phi_n(ny) .$$

et on a, en tout point de continuité de F(x) :

$$(2) \qquad 1 - F(x) = \lim_{n \to \infty} \Phi_n\left(\frac{n}{x}\right)$$

puisque l'évènement "$Y_n < 1/x$" converge vers l'évènement "$X > x$" : telle est la formule d'inversion annoncée. On notera surtout, sur le plan théorique, qu'elle ne nécessite en aucune manière le recours à un itinéraire dans le plan complexe.

Démonstration de la convergence en loi de Y_n vers $1/X$: la transformée de Λ, à X = x fixé, est $\int_0^\infty x\, e^{-\lambda x - \lambda\mu}\, d\lambda = \frac{x}{\mu + x}$

et celle de $\Lambda_1 + \ldots + \Lambda_n$ est $\left(\frac{x}{\mu+x}\right)^n$. Pour X aléatoire, la loi a priori de $\Lambda_1 + \ldots + \Lambda_n$ admet donc la transformée :

$$\int_0^\infty \left(\frac{x}{\mu+x}\right)^n F(dx) = E\left[\left(\frac{X}{\mu+X}\right)^n\right]$$

Pour Y_n cette transformée devient $E\left(\frac{X}{\mu/n + X}\right)^n = E\left[\left(1 - \frac{\mu}{nX+\mu}\right)^n\right]$

Lorsque n tend vers l'infini, on vérifie facilement que le passage à la limite sous le symbole E est légitime, et on obtient la limite $E\left[e^{-\mu/X}\right]$, qui est justement la transformée de Laplace de $1/X$.

2º/ Passage de la loi de X à celle de logX.

Si X admet un moment d'ordre β négatif ou inférieur à 1, on vérifie sans peine que celui-ci se déduit de la transformée de Laplace selon la relation :

$$(3) \qquad E(X^\beta) = -\int_0^\infty \frac{\lambda^{-\beta}}{\Gamma(1-\beta)} \cdot \Phi(d\lambda) \qquad (\beta < 1)$$

Il suffit de remplacer β par iu pour obtenir la transformée de Fourier $E(X^{iu}) = E(e^{iu\log X})$ de $\log X$. Pour éviter les difficultés qui accompagnent un voyage dans le plan complexe, nous allons faire un raisonnement direct.

A X = x fixé, Λ est de la forme Z/x, où Z est la variable exponentielle réduite, de loi e^{-z}. Par suite, $\log \Lambda$ est de la forme $\log Z - \log x$ et admet la fonction caractéristique $e^{-iu\log x}\Gamma(1 + iu)$. De cette loi conditionnelle en X=x, on déduit la fonction caractéristique a priori de Λ qui est : $\Gamma(1+iu)\, E\left[e^{-iu\log X}\right] = -\int_0^\infty e^{iu\log\lambda}\, \Phi(d\lambda)$

D'où la relation cherchée :

$$(4) \qquad E\left[e^{iu\log X}\right] = -\frac{1}{\Gamma(1-iu)} \int_0^\infty e^{-iu\log\lambda}\, \Phi(d\lambda)$$

Cette relation exprime simplement que la variable aléatoire Λ est de la forme $\Lambda = Z/X$, où Z est une variable exponentielle réduite indépendante de X.

Exemples. a/ Loi gamma. Si X admet la densité $\frac{b^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-bx}$ on a $\Phi(\lambda) = \left(\frac{b}{b+\lambda}\right)^\alpha$ mais, plutot que d'utiliser (4), il vaut mieux faire un calcul direct, qui donne :

$$E[e^{iu\log X}] = e^{-iu\log b} \frac{\Gamma(\alpha+iu)}{\Gamma(\alpha)} \tag{5}$$

b/ Lois stables Les lois stables concentrées sur la demi-droite positive sont caractérisées par leurs transformées de Laplace qui sont de la forme : $e^{-a\lambda^\alpha}$ – a est un simple paramètre d'échelle, tandis qu' α (toujours compris entre 0 et 1) caractérise le type de la loi. Il suffit d'appliquer la relation (4) pour obtenir :

$$E[e^{iu\log X}] = e^{\frac{iu}{\alpha}\log a} \frac{\Gamma(1-\frac{iu}{\alpha})}{\Gamma(1-iu)} \tag{6}$$

On déduit de cette formule que – contrairement à la variable X elle-même – log X admet des moments de tous les ordres. Les developpements classiques :

$$\left\{\begin{aligned} -\log\Gamma(z) &= \log z + \sum_{n=1}^{\infty}\left[\log\left(1+\frac{z}{n}\right)-\frac{z}{n}\right] + Cz \\ -\frac{\Gamma'(z)}{\Gamma(z)} &= \frac{1}{z} + \sum_{n=1}^{\infty}\left(\frac{1}{n+z}-\frac{1}{n}\right) + C \\ \left(\frac{\Gamma'(z)}{\Gamma(z)}\right)' &= \sum_{n=0}^{\infty}\frac{1}{(n+z)^2} \end{aligned}\right. \tag{7}$$

permettent le calcul de l'espérance et de la variance de log X :

$$\left\{\begin{aligned} E[\log X] &= \frac{1}{\alpha}\log a + C\left(\frac{1}{\alpha}-1\right) \\ D^2[\log X] &= \frac{\pi^2}{6}\left(\frac{1}{\alpha^2}-1\right) \end{aligned}\right. \tag{8}$$

<u>Remarque</u>. Du dévelpppement de la fonction eulérienne, on déduit une propriété inattendue de la loi gamma et des lois stables : non seulement ces lois sont indéfiniment divisibles, mais si X obéit à l'une de ces lois, log X est également une variable indéfiniment divisible. Nous posons donc le problême suivant :

<u>Problême</u>. Déterminer la classe la plus générale des lois de probabilité telles que toute variable X soumise à l'une d'elles soit indéfiniment divisible ainsi que son logarithme log X .

3º/ Caractérisation de la loi lognormale.

La loi lognormale n'est ni stable ni indéfiniment divisible. Elle est pour cette raison inapte à décrire d'une manière théoriquement satisfaisante un phénomène régionalisé. Son efficacité (relative) sur le plan pratique provient du fait qu'elle est sans doute (relativement) voisine d'une loi stable. Elle possède pourtant une propriété interessante qui suffit à la caractériser complétement :

<u>Critère de lognormalité</u>. La loi lognormale est la seule loi qui conserve son type lorsque l'on pondère les fréquences par la variable aléatoire elle-même.

Autrement dit, si X est une variable et $F(dx)$ sa loi (admettant une espérance $m = \int_0^\infty x\, F(dx)$) on peut trouver une constante $a \neq 0$ telle que la variable aX admette la loi $x\, F(dx)/m$ si et seulement si F est une loi lognormale.

On sait que la loi lognormale vérifie bien cette propriété. Inversement, si une loi $F(dx)$ possède cette propriété, elle admet une espérance, et, par une récurrence immédiate, des moments de tous les ordres. Le moment d'ordre n : $m_n = \int_0^\infty x^n F(dx)$ s'explicite ainsi :

$$m_n = m \int_0^\infty x^{n-1} \frac{x}{m} F(dx) = m\, a^{n-1} m_{n-1}$$

On en tire aussitot : $m_n = m^n a^{\frac{n(n-1)}{2}}$

Ceci implique $a > 1$, et F est alors la loi lognormale de moyenne m et de variance log a .

Conséquence. De ce critère résulte qu'aucune loi indéfiniment divisible ne peut conserver son type par autopondération. Cependant, il existe une propriété un peu plus générale qui leur est applicable.On sait que toute variable positive et indéfiniment divisible admet une transformée de Laplace de la forme :

$$\Phi(\lambda) = e^{-a\int_0^\infty \frac{1-e^{-\lambda y}}{y} G(dy)}$$

Limitons nous au cas où la mesure positive G est sommable : $\int_0^\infty G(dy) = 1$
Il est interessant d'interpreter cette loi en termes de processus à accroissements indépendants et stationnaires : X(a) est une fonction aléatoire croissante du "temps" a , variant par sauts positifs (à ce ci près qu'un intervalle fini peut contenir une infinité de sauts très petits). La granulométrie en longueur de ces sauts est donnée par G(dy) . La granulométrie en nombre n'existe que si l'intégrale $\int_0^y \frac{1}{u} G(du)$ est convergente.

Si l'on introduit la transformée de Laplace de G : $\gamma(\lambda) = \int_0^\infty e^{-\lambda y} G(dy)$
on a aussi bien :

$$\Phi(\lambda) = e^{-a\int_0^\lambda \gamma(\mu)\, d\mu}$$

d'où l'on tire l'équation différentielle :

$$\Phi' = -a\gamma\Phi$$

Si X est une variable dont la loi F admet la transformée Φ , on a $m = E(X) = a$ et l'équation différentielle ci-dessus exprime que x/m F(dx) est la loi de probabilité d'une variable $Z = X + Y$, somme de deux variables indépendantes : X, obéissant à la loi indéfiniment divisible F , et Y dont la loi est la granulométrie en longueur G(dy) des sauts du processus associé.

Si Y est elle-même indéfiniment divisible, le processus pourra se repeter, et $x^n/m_n F(dx)$ sera là loi d'une variable du type $X + Y_1 + Y_2 + \dots + Y_n$.

On observe un exemple interessant dans le cas des lois gamma. Si X est la variable exponentielle réduite, on a $\Phi(\lambda) = 1/(1+\lambda)$, d'où $\gamma(\lambda) = 1/(1+\lambda) = \Phi(\lambda)$: $X, Y_1, \dots, Y_n$ obéissent à la même loi exponentielle, de sorte qu'il est équivalent de prendre la somme de $n+1$ variables exponentielles indépendantes, ou de pondérer par x^n la loi de l'une d'elles.

On notera toutefois que les lois stables n'admettent pas de décomposition de ce type. Pour $\Phi(\lambda) = e^{-a\lambda^\alpha}$, on obtient en effet $\gamma(\lambda) = \alpha\lambda^{\alpha-1}$ et, comme $0 < \alpha < 1$, on trouve pour la mesure G l'expression $G(dy) = \frac{\alpha y^{-\alpha}}{\Gamma(1-\alpha)} dy$ L'intégrale $\int_0^b G(dy)$ ne converge pas lorsque b tend vers l'infini. Les sauts du processus associé n'admettent de granulométrie ni en nombre ni en longueur. On vérifiera, par contre, que les sauts d'amplitude supérieure à une valeur b donnée admettent une granulométrie en nombre (mais non en longueur), et qu'inversement les sauts d'amplitude inférieure à b admettent une granulométrie en longueur, mais non en nombre : un très grand nombre de sauts très petits coexistent avec un nombre limité de sauts extremement grands.

4º/ Caractérisation des lois gamma.

Nous venons de caractériser la loi gamma par la propriété que l'autopondération équivaut, pour les variables qui lui sont soumises, à l'addition d'une variable exponentielle. Nous allons maintenant indiquer un deuxième critère, également caractéristique, qui nous conduira à une interpretation géostatistique instructive.

Critère de la loi gamma. Soient X et Y deux variables positives indépendantes, et $Z = X + Y$. Pour que la variable X/Z soit indépendante de Z ,il faut et il suffit que X et Y obéissent à des lois gamma admettant le même paramètre d'échelle (de densités $\frac{b^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1 - 1} e^{-bx}$ et $\frac{b^{\alpha_2}}{\Gamma(\alpha_2)} x^{\alpha_2 - 1} e^{-bx}$ avec le même paramètre b).

La loi gamma vérifie bien cette propriété. Posant $T = X/(X+Y)$, on sait qu'à $Z = X + Y$ fixé T obéit à la loi béta de paramètres α_1 et α_2, donc est indépendante de Z.

C'est la réciproque qui présente le plus d'interêt. On exprime l'indépendance de $X/(X+Y)$ et de $(X+Y)$ en écrivant :

$$E\left[\left(\frac{X}{X+Y}\right)^n e^{-\lambda(X+Y)}\right] = E\left[\left(\frac{X}{X+Y}\right)^n\right] E\left[e^{-\lambda(X+Y)}\right]$$

En dérivant n fois en λ, on voit que cette condition implique que les transformées de Laplace Φ_1 et Φ_2 de X et de Y vérifient la relation :

$$\Phi_2(\lambda)\frac{d^n}{d\lambda^n}\Phi_1(\lambda) = m_n \frac{d^n}{d\lambda^n}\left[\Phi_1(\lambda)\Phi_2(\lambda)\right]$$

(avec $m_n = E\left[\left(\frac{X}{X+Y}\right)^n\right]$ Pour $n = 1$, on trouve :

$$\Phi_2 \Phi_1' = m_1(\Phi_2\Phi_1' + \Phi_1\Phi_2')$$

d'où l'on déduit que Φ_1 et Φ_2 sont de la forme :

$$\begin{cases} \Phi_1(\lambda) = e^{m_1 \psi(\lambda)} \\ \Phi_2(\lambda) = e^{(1-m_1)\psi(\lambda)} \end{cases}$$

Pour $n = 2$, et compte tenu des expressions de Φ_1 et de Φ_2 ainsi obtenues, il vient :

$$m_2\left[\psi'' + (\psi')^2\right] = m_1\psi'' + (m_1\psi')^2$$

ce qui donne $\psi(\lambda) = -\alpha \log\left(1 + \frac{\lambda}{\sigma}\right)$

avec $\alpha = \dfrac{m_1 - m_2}{m_2 - (m_1)^2}$ et b arbitraire.

Autrement dit, X et Y obéissent aux lois gamma de paramètres $\alpha_1 = m_1\alpha$ et $\alpha_2 = (1-m_1)\alpha$ respectivement, et admettant le même paramètre d'échelle b .

Interprétation géostatistique. Imaginons un gisement à effet de pépite pur, c'est à dire à loi indéfiniment divisible et accroissements indépendants. Désignons par v_i des échantillons ou des panneaux disjoints, et par $Q(v_i)$ la quantité de métal contenue dans v_i . Les deux propriétés suivantes sont équivalentes :

a/ Q(v) obéit à la loi gamma (de paramètre α proportionnel au volume v de l'échantillon)

b/ Si $v \cap v' = \emptyset$, les variables $Q(v)/Q(v \cup v')$ et $Q(v \cup v')$ sont indépendantes.

Remarquons que la relation

$$Q_1 = \frac{\varphi_1}{\varphi_1 + \varphi_2} \; \frac{\varphi_1 + \varphi_2}{\varphi_1 + \varphi_2 + \varphi_3} \; \cdots \; (\varphi_1 + \cdots + \varphi_n)$$

montre que Q_1 se présente sous la forme du produit d'un nombre (aussi grand que l'on veut) de variables indépendantes de variances relatives aussi petites que l'on veut. Il n'en résulte en aucune façon que la variable Q_1 obéisse à une loi lognormale.

Nous avons raisonné comme si la quantité de métal $Q = Q_1 + \dots + Q_n$ du gisement entier était aléatoire. En fait, dès que l'on a déterminé le champ V dans lequel on va prélever des échantillons, Q se trouve fixé (même si on ne le connait pas). Le rapport Q_1/Q , c'est à dire, à un facteur près, la teneur de v , obéit à une loi béta de paramètres α_1 et β_1 proportionnels à v /V et à (1-v /V) , et celà quelle que soit la valeur (fixée) de Q . Si la loi a priori de Q est la loi gamma de paramètre $\alpha = aV$, on a $\alpha_1 = av$ et $\beta_1 = a(V - v)$.

La factorisation de Q_1 peut ici se mettre sous la forme :

$$\frac{Q_1}{Q} = \frac{Q_1}{Q_1+Q_2}\;\frac{Q_1+Q_2}{Q_1+Q_2+Q_3}\;\cdots\;\frac{Q_1+\cdots+Q_{n-1}}{Q}$$

d'un produit de variables béta <u>indépendantes</u>, de variances aussi modestes que l'on veut : mais Q_1/Q obéit à la loi béta, et non à la loi lognormale.

Analytiquement, il est instructif d'in[illegible]nction caractéristique $\Phi(u;\alpha,\beta)$ du logarithme d'une variable béta :

$$\Phi(u;\alpha,\beta) = E[X^{iu}] = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, x^{\alpha-1+iu}\,(1-x)^{\beta-1}\,dx$$

On trouve :

$$\Phi(u;\alpha,\beta) = \frac{\Gamma(\alpha+iu)}{\Gamma(\alpha)}\;\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+iu)} \tag{9}$$

En particulier, si $Y = X_1X_2\ldots X_n$, où les X_i sont des béta indépendantes, de paramètres :

X_1 : α_1, β_1
X_2 : $\alpha_2 = \alpha_1+\beta_1$, β_2
X_{n-1} : $\alpha_{n-1} = \alpha_{n-2}+\beta_{n-2}$, β_{n-1}

il est immédiat que Y est elle-même une variable béta de paramètres α_1 et $(\beta_1+\beta_2+\ldots+\beta_{n-1})$: tel était bien le sens de la factorisation de Q_1/Q.

<u>Anomalies pour la variance du rapport.</u> Un variable béta possède, comme on sait, la variance : $\dfrac{\alpha}{\alpha+\beta}\left[\dfrac{1+\alpha}{1+\alpha+\beta} - \dfrac{\alpha}{\alpha+\beta}\right]$

Pour $Q(v)/Q(V)$ on a : $\alpha = av$ et $\alpha+\beta = aV$, d'où la variance :

$$D^2\left[\frac{Q(v)}{Q(V)}\right] = \frac{v}{V}\;\frac{V-v}{V(1+aV)}$$

Si $m = E[Q(v)/Q(V)] = v/V$ désigne l'espérance de ce rapport, la variance est donc de la forme $m(1 - m)/(1 + aV)$, et la variance relative est : $\frac{1}{1+aV}\left(\frac{V}{v} - 1\right)$

Si aV est grand, on retrouve la formule de type usuel $1/a\,(1/v - 1/V)$. Par contre, si aV est petit , c'est à dire si la loi a priori de Q est très dissymétrique, on a une formule en $V(1/v - 1/V)$ qui croit linéairement avec la taille V du champ. Cette anomalie, qui se manifeste ici dans un exemple pourtant très simple, (loi gamma et effet de pépité à l'état pur) doit inciter à une grande prudence dans l'usage de raisonnements heuristiques à la de Wijs. C'est seulement à la différence $Q(v)/v - Q(V)/V$ que s'applique la formule géostatistique de la variance. On n'a pas le droit de l'utiliser pour le rapport $Q(v)/Q(V)$. Ces anomalies apparaitront sous un aspect plus net encore si nous examinons directement les variances logarithmiques.

Anomalies pour les variances logarithmiques.

Les developpements (7) et la formule (8) permettent de determiner la variance logarithmique d'une variable béta :

$$D^2(\log X) = \sum_{n=0}^{\infty} \frac{1}{(n+\alpha)^2} - \sum_{n=0}^{\infty} \frac{1}{(n+\alpha+\beta)^2}$$

Dans le cas de la variable $Q(v)/Q(V)$, on a $\alpha = av$, et $(\alpha + \beta) = aV$. Si av est déjà grand, cette variance diffère peu de $1/a\,(1/v - 1/V)$, c'est à dire de la variance relative calculée ci-dessus. Par contre, si av est petit, la somme $\sum 1/(n + av)^2$ équivaut à son premier terme, qui est $1/(av)^2$: pour ces très petits échantillons, la variance logarithmique est en raison inverse du carré du volume. En pareil cas, l'application brutale des formules de la géostatistique, (établies pour les variances arithmétiques) aux logarithmes des teneurs et à leurs variances peut conduire à des erreurs grossières.

<u>Critère complémentaire pour la loi gamma</u>. La loi gamma possède également une autre propriété caractéristique : elle est la seule à conserver son type lorsque l'on pondère les fréquences par une exponentielle. Autrement dit, $e^{-\alpha x}F(dx)/\Phi(\alpha)$ est la loi de aX (pour une constante a fonction de α) si et seulement si la loi F(dx) de X est une loi gamma.

Il est immédiat que la loi gamma vérifie bien ce critère. Inversement, supposons que pour tout $\alpha > 0$ on puisse trouver une constante $a(\alpha)$ telle que $e^{-\alpha\lambda}F(dx)/\phi(\alpha)$ soit la loi de aX , ou, ce qui revient au même, telle que :

$$\Phi(a\lambda) = \frac{\Phi(\alpha+\lambda)}{\Phi(\alpha)} \tag{10}$$

Posons $\psi = \log \phi$ et dérivons en α :

$$\lambda\ a'(\alpha)\ \psi'(a\lambda) = \psi'(\alpha+\lambda) - \psi'(\alpha)$$

En $\alpha = 0$, on a $a = 1$ et :

$$\lambda\ a'(0)\ \psi'(\lambda) = \psi'(\lambda) - \psi'(0) \tag{11}$$

(l'existence de $\psi'(0) = -E(X) = -m$ résulte de la relation fonctionnellle (10) elle-même,: il suffit de la dériver en λ et de faire $\lambda = 0$. A droite, on obtient $\psi'(\alpha)$ qui existe toujours pour $\alpha > 0$, et, à gauche, a $\Phi'(0)$)

Ainsi la fonction $a(\alpha)$ est donnée par

$$a = -\frac{1}{m}\ \frac{\Phi'(\alpha)}{\Phi(\alpha)} = -\frac{1}{m}\ \psi'(\alpha)$$

En particulier $a'(0)$ est négatif. Prenant $a'(0) = -1/b$, l'équation (11) s'intègre sans difficulté sous la forme :

$$\psi(\lambda) = -mb\ \log\left(1 + \frac{\lambda}{b}\right)$$

ce qui achève de démontrer le critère.

Remarque. Etant donnée une fonction $h(x)$, on pourrait chercher à determiner la classe des lois F stable pour les pondérations par $h(\alpha x)$, c'est à dire telles que pour tout $\alpha > 0$ il existe une constante $a(\alpha)$ telle que $h(\alpha x)\ F(dx)/E[h(\alpha X)]$ soit la loi de $a(\alpha)\ X$. Moyennant des hypothèses raisonnables sur la régularité de $h(x)$ en $x = 0$, on peut montrer que le probleme n'est soluble que dans deux cas : $h(x)$ est un monome x^{λ} ,et on tombe sur la loi lognormale, ou $h(x)$ est une exponentielle, et on retrouve la loi gamma.

5º/ Les lois stables prises conditionnellement.

Imaginons un gisement dans lequel $Q(v)$ obéit à la loi stable $e^{-av\lambda^{\alpha}}$ et où $Q(v)$ et $Q(v')$ sont indépendants dès que v et v' sont disjoints. La loi a priori de $Q(v)$ n'a ni espérance ni variance, celle de $\log Q(v)$ admet une variance indépendante de la taille v de l'échantillon. Mais, v étant contenu dans un champ V, on étudie en réalité $Q(v)$ conditionnellement à $Q(V)$ fixé. Comme $Q(v)$ est inférieur à $Q(V)$, il admet des moments conditionnels de tous les ordres. Pour l'espérance, on trouve sans peine :

$$E\left[Q(v) \mid Q(V) = Q \right] = Q \frac{v}{V}$$

C'est d'ailleurs là une propriété générale pour toutes les lois indéfiniment divisibles. Le calcul de la variance est plus difficile. Pour obtenir au moins une indication d'ordre général, nous calculerons seulement la variance a priori $D^2[Q(v)/Q(V)]$ du rapport de ces deux variables : contrairement à ce qui se passe dans le cas de la loi gamma, ce rapport n'est pas indépendant de $Q(V)$, et la variance a priori de ce rapport ne permet pas de reconstituer la variance conditionnelle de $Q(v)$, qui est une fonction complexe de $Q(V)$.

Prenons $X = Q(v)$, $Y = Q(V - v)$, $X + Y = Q(V)$. X et Y sont indépendantes et admettent les lois $e^{-a_1\lambda^{\alpha}}$ et $e^{-a_2\lambda^{\alpha}}$ ($a_1 = av$, $a_2 = a(V-v)$) ; $X + Y$ admet la loi $e^{-(a_1+a_2)\lambda^{\alpha}}$, et nous cherchons les moments du rapport $X/(X+Y)$: la formule (3), avec β négatif, montre que nous les déduirons de $E\left[e^{-\lambda \frac{X+Y}{X}} \right] = e^{-\lambda}\Phi(\lambda)$, $\Phi(\lambda)$ étant la transformée de Y/X .

a/ Calcul de $\Phi(\lambda)$. A $X = x$ fixé, on a $E(e^{-\lambda \frac{Y}{x}}) = e^{-a_2 \lambda^{\alpha} x^{-\alpha}}$
On en tire, sous réserve de la convergence de cette série (que nous constaterons dans un instant) :

$$\Phi(\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (a_2)^n \lambda^{n\alpha} E[X^{-n\alpha}]$$

Pour calculer $E(X^{-n\alpha})$, nous utiliserons (3) qui donne :

$$E[X^{-n\alpha}] = -\frac{1}{\Gamma(1+n\alpha)} \int_0^{\infty} \mu^{n\alpha} d(e^{-a_1 \mu^{\alpha}}) = \frac{n!}{(a_1)^n \, \Gamma(1+n\alpha)}.$$

et nous trouvons :

(12) $$\Phi(\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(1+n\alpha)} \left(\frac{a_2}{a_1}\right)^n \lambda^{n\alpha}$$

c'est à dire une fonction apparentée à celle de Mittag-Leffler.

b/ Calcul des moments. Comme $e^{-\lambda}\Phi(\lambda)$ est la transformée de $(X+Y)/X$, la même formule (3) nous donne aves $\beta = -k$:

$$E\left[\left(\frac{X}{X+Y}\right)^k\right] = -\frac{1}{k!} \int_0^{\infty} \lambda^k \, d[e^{-\lambda}\Phi(\lambda)]$$

Laissons de coté la discussion analytique relative à la légitimité du passage à la limite sous le signe somme (il faut sans doute supposer $a_2 < a_1$, ce qui est toujours loisible, quitte à échanger ces deux paramètres, c'est à dire à étudier $Y/(X+Y)$ au lieu de $X/(X+Y)$). Compte tenu de (12), et en intégrant terme à terme, on trouve

(13) $$E\left[\left(\frac{X}{X+Y}\right)^k\right] = \frac{1}{(k-1)!} \sum_{n=0}^{\infty} (-1)^n \frac{\Gamma(k+n\alpha)}{\Gamma(1+n\alpha)} \left(\frac{a_2}{a_1}\right)^n$$

$$E\left[\left(\frac{X}{X+Y}\right)^{\lambda}\right] = 1 - x \frac{\sin \pi\lambda}{\pi} \int_0^1 \frac{1}{1+x t^{\alpha}} t^{\lambda+\alpha-1} (1-t)^{-\lambda} dt \qquad \left(x = \frac{a_2}{a_1}\right)$$

$$E\left[\log \frac{X}{X+Y}\right] = -\frac{1}{\alpha} \log\left(\frac{a_1+a_2}{a_1}\right)$$

On en déduit la moyenne et la variance :

$$\begin{cases} m = E\left(\dfrac{X}{X+Y}\right) = \dfrac{a_1}{a_1+a_2} \\ \sigma^2 = D^2\left[\dfrac{X}{X+Y}\right] = (1-\alpha)\,\dfrac{a_1 a_2}{(a_1+a_2)^2} \end{cases}$$

Comme a_1 et $a_1 + a_2$ sont proportionnels aux volumes v et V, on a aussi :

$$\text{(L4)} \qquad \begin{cases} m = \dfrac{v}{V} \\ \sigma^2 = (1-\alpha)\; m(1-m) = (1-\alpha)\,\dfrac{v}{V}\left(1-\dfrac{v}{V}\right) \end{cases}$$

Ce sont là des formules d'aspect rassurant. N'oublions pas, cependant, qu'il s'agit de la variance a priori du rapport $Q(v)/Q(V)$, et non de la variance conditionnelle à $Q(V)$ fixé: Cette dernière nous réserverait sans doute bien des surprises. C'est en passant aux logarithmes que nous allons rencontrer ces anomalies.

c/ Variance logarithmique. La fonction caractéristique de $\log(X/X+Y)$ se déduit de (13) en changeant k en iu :

$$\psi(u) = \frac{1}{\Gamma(iu)} \sum_{n=0}^{\infty} (-1)^n \left(\frac{a_2}{a_1}\right)^n \frac{\Gamma(iu+n\alpha)}{\Gamma(1+n\alpha)}$$

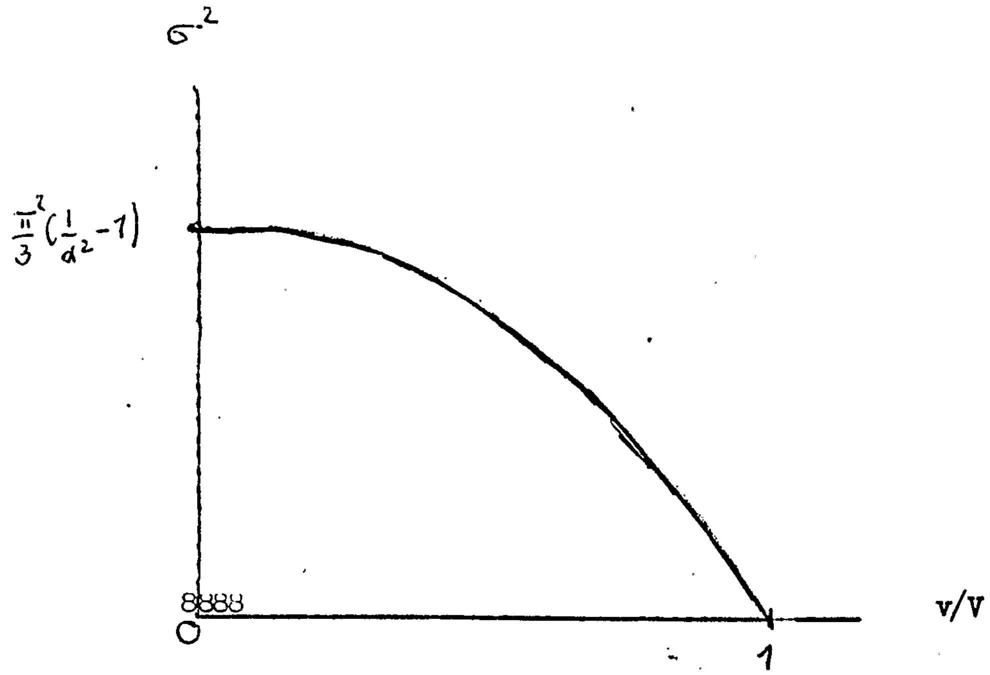
(avec $a_2 < a_1$) . On en déduit sans difficulté le moment d'ordre 1 :

$$E\left[\log\frac{X}{X+Y}\right] = -\frac{a_2}{a_1+a_2} = \frac{v}{V} - 1$$

Mais nous ne savons pas si cette formule reste valable pour $a_1 < a_2$. En dérivant une deuxième fois en u , on pourrait obtenir la variance sous la forme d'une série double : mais cette expression est compliquée, et de plus elle n'est valable que pour $a_2 < a_1$, alors que nous nous interessons plutot à l'autre cas (échantillon petit vis à vis du champ V).

Nous nous contenterons donc d'évaluer la variance de $\log(X/Y)$, qui ne doit pas différer beaucoup de celle de $\log(X/X+Y)$ lorsque v est petit. Mais, logX et log Y étant indépendants, on peut utiliser la formule (8) : cette variance est donc $\frac{\pi^2}{3}\left(\frac{1}{\alpha^2} - 1\right)$, et ne dépend pas de v ni de V .

Ainsi, les petits échantillons auront une variance pratiquement constante. Ce n'est que lorsque v deviendra suffisamment grand vis à vis de V que la liaison Q(V) = constante donnée fera sentir son effet, et que l'on verra la variance logarithmique commencer à décroitre, pour s'annuler enfin en v = V . Cette loi, que nous représentons de manière simplement indicative sur la figure ci-après, diffère profondément des lois usuelles en 1/v - 1/V .



G. Matheron - Novembre 1967