LA SÉLECTIVITÉ DES DISTRIBUTIONS

G. MATHERON

N–686

Décembre 1981

CENTRE DE GÉOSTATISTIQUE

35 rue Saint-Honoré, 77305 FONTAINEBLEAU (France)

ECOLE DES MINES
DE PARIS

LA SELECTIVITE DES DISTRIBUTIONS

Table

LA SELECTIVITE DES DISTRIBUTIONS

0 - INTRODUCTION

Je propose, dans ce qui suit, un formalisme général pour rendre compte des effets de support et d'information, et des dégradations qu'ils entrainent pour les relations tonnage/teneur. Ce formalisme, qui met en jeu une relation d'ordre très forte (F_X plus sélective que F_Y) repose sur un théorème très important du à CARTIER. Le premier paragraphe rappelle, dans le cas continu, les propriétés des fonctions Q, T, m et V de la teneur de coupure, et introduit une notion nouvelle : l'indicateur S de dispersion (à partir duquel on définit ensuite de manière naturelle un indice de sélectivité). Le paragraphe suivant présente les mêmes notions dans le cas général, en utilisant la dualité entre la fonction convexe V(y) et la fonction concave Q(T).

Ces notions prennent leur intérêt au paragraphe III, où sont définies 2 relations d'ordre (F_X plus sélective que F_Y, et F_X est une dispersée de F_Y) dont le théorème de CARTIER montre l'équivalence. Fondamentales pour la Géostatistique (effets de support et d'information), ces relations d'ordre me semblent en outre présenter un intérêt général en théorie des probabilités. Le dernier paragraphe concerne directement la Géostatistique minière, et, je l'espère, suscitera un certain intérêt chez les praticiens de cette discipline.

I - LES RELATIONS TONNAGE/TENEUR, DANS LE CAS D'UNE INFORMATION PARFAITE.

Dans ce premier paragraphe, nous nous plaçons dans le cas idéal où il n'y a ni effet de support, ni effet d'information. Autrement dit, nous supposons parfaitement connues les teneurs y_i des différents blocs sélectionnables V_i. Il est bien entendu qu'il s'agit là d'un idéal toujours inaccessible en pratique. Les relations tonnage/teneur calculées dans ce cas idéal constituent donc un maximum absolu (et inaccessible) qui peut servir de référence pour évaluer l'ampleur des dégradations ou des pertes dont sont responsables les différents effets de support et d'information qui affectent nécessairement les relations tonnage/teneur réelles.

L'information étant supposée parfaite, il est clair que, si nous sélectionnons, nous choisirons les blocs les meilleurs, c'est-à-dire les plus riches. Les teneurs y_i de ces blocs étant supposées rangées en ordre décroissant : $y_1 \geq y_2 \geq \ldots$, nous retiendrons les k premiers blocs $V_1, \ldots V_k$ et abandonnerons les suivants. La valeur de k, que nous pouvons choisir à notre gré, fixera le tonnage du minerai sélectionné, soit

$$T_k = \sum_{i=1}^{k} V_i$$

et le tonnage métal correspondant sera

$$Q_k = \sum_{i=1}^{k} V_i \, y_i$$

Au paramétrage en k (nombre des blocs), il est souvent commode de substituer un paramétrage en y (<u>teneur de coupure</u>) : la teneur de coupure y étant donnée, on sélectionne les blocs V_i dont la teneur y_i est supérieure ou égale à y, et on désigne par T(y) et Q(y) les tonnages de minerai et de métal correspondant. Il est clair que l'on a

$$T(y) = T_k \text{ et } Q(y) = Q_k \text{ pour } y_{k+1} < y \leq y_k$$

de sorte .../

que ces deux formes de paramétrages seront équivalentes. $T(y)$ est évidemment une fonction décroissante de y. Il est commode pour ce qui suit de poser $T(0) = 1$, c'est-à-dire de prendre comme unité de tonnage le tonnage total, associé à la teneur de coupure $y = 0$. Alors la fonction F défini par

$$F(y) = 1 - T(y) \qquad (y \geqslant 0)$$

est une fonction de répartition. Elle représente la loi de probabilité de la variable aléatoire $Y = y_i$ lorsque l'indice i est choisi au hasard selon la loi

$$p_i = \frac{V_i}{V_1+V_2+\dots}$$

La quantité de métal $Q(y)$, de son côté, s'exprime à l'aide de la loi F selon la relation

$$Q(y) = \int_y^\infty \eta \, F(d\eta)$$

Ceci conduit à la définition générale suivante :

Soit Y une <u>variable aléatoire</u> positive d'espérance $E(Y)$ finie, $F(dy)$ sa loi de probabilité sur $(0, \infty)$, $F(y) = P(Y < y)$ sa fonction de répartition. Pour tout $y \geqslant 0$, nous poserons :

$$\begin{cases} T(y) = 1 - F(y) \\ Q(y) = \int_y^\infty \eta \, F(d\eta) \end{cases}$$

Si nous introduisons l'indicatrice :

$$1_{Y \geqslant y} = \begin{cases} 0 & \text{si } Y < y \\ 1 & \text{si } Y \geqslant y \end{cases}$$

associée à la teneur de coupure y, nous trouvons aussi bien

$$\begin{cases} T(y) = E(1_{Y \geq y}) \\ Q(y) = E(Y\ 1_{Y \geq y}) \end{cases}$$

Il est commode d'introduire aussi la teneur moyenne m(y) du minerai sélectionné à la teneur de coupure y :

$$m(y) = Q(y)/T(y)$$

Il est clair que m(y) est l'espérance conditionnelle de Y lorsque Y est $\geq$ y

$$m(y) = E(Y/Y \geq y)$$

Enfin, nous introduirons la fonction V(y) définie par :

$$V(y) = E((Y-y)_+) = \int_y^{\infty} (z-y)\, F(dz)$$

Cette fonction V(y) admet une interprétation économique, et nous dirons qu'elle représente la valeur conventionnelle du minerai sélectionné à la teneur de coupure y.

En effet, si on admet (en 1ère approximation) que la valeur d'un lot de minerai est proportionnelle à la quantité de métal qu'il contient, et que les frais d'extraction et de traitement sont proportionnels au tonnage, cette valeur sera :

$$(bz - \lambda)\, \delta T$$

(δT : tonnage, z = teneur, b : valeur du point de métal, λ : frais d'extraction et de traitement à la tonne). Cette valeur est positive pour $z \geq \lambda/b$.
On doit donc sélectionner un lot selon le critère de coupure $z \geq y$, avec $y = \lambda/b$, et la valeur de ce lot est alors :

$$b\,(z-y)_+\, \delta T = \begin{cases} 0 & \text{si } z < y \\ b\,(z-y)\,\delta T & \text{si } z \geq y \end{cases}$$

Moyennant le choix b = 1 de l'unité de valeur, on voit que V(y) représente effectivement la valeur économique de minerai pour la teneur de coupure y.

D'après la définition de V et de Q, on a évidemment :

$$V(y) = Q(y) - y\, T(y)$$

Notons aussi la relation évidente :

$$(z - y)_+ = \int_y^\infty 1_{z \geq x}\, dx$$

D'après le théorème de Fubini, nous en déduisons :

$$V(y) = \int (z-y)_+ \, F(dz) = \int_y^\infty dx \int_x^\infty F(dz)$$

c'est-à-dire :

$$V(y) = \int_y^\infty T(x) dx$$

En résumé, à notre variable aléatoire Y positive et d'espérance finie, nous avons associé les quatre fonctions suivantes :

$$(I - 1) \quad \begin{cases} T(y) = E\,(1_{Y \geq y}) \\ Q(y) = E\,(Y\, 1_{Y \geq y}) \\ m(y) = E\,(Y / Y \geq y) \\ V(y) = E\,((Y - y)_+) \end{cases}$$

Ces fonctions s'expriment à l'aide de la fonction de répartition F(y) selon les formules :

$$(I - 2) \quad \begin{cases} T(y) = 1 - F(y) \\ Q(y) = \int_y^\infty x\, F(dx) \\ m(y) = Q(y)/T(y) \\ V(y) = \int_y^\infty (x-y)\, F(dx) \end{cases}$$

et vérifient de plus les relations :

$$
(\text{I - 3}) \qquad \begin{cases} V(y) = Q(y) - y\,T(y) \\ V(y) = \displaystyle\int_y^{\infty} T(x)dx \end{cases}
$$

Propriétés des fonctions T, Q, m, et V

D'après les 2 premières relations (I-2), T(y) et Q(y) sont des fonctions décroissantes de y (mais non continues, en général, à moins que la loi F n'admette une densité f). De même, la teneur moyenne m(y) est une fonction croissante de la teneur de coupure y : en effet, supposons $y_1 \leq y_2$, et donc $T(y_1) = T_1 \geq T(y_2) = T_2$. Comme $m_1 = m(y_1)$ est l'espérance conditionnelle de Y si $Y \geq y_1$, on peut écrire :

$$
m_1 = \frac{T_1 - T_2}{T_1} E\,(Y / y_1 \leq Y < y_2) + \frac{T_2}{T_1} E\,(Y / Y \geq y_2)
$$

Or, on a nécessairement

$$
E(Y / y_1 \leq Y < y_2) \leq y_2
$$

$$
m_2 = E\,(Y / y \geq y_2) \geq y_2
$$

Il vient donc

$$
m_1 \leq \frac{T_1 - T_2}{T_1} y_2 + \frac{T_2}{T_1} m_2 \leq m_2
$$

La fonction V(y), de son côté, est convexe et décroissante. Pour le voir, il suffit de remarquer que $(Y-y)_+$ est déjà elle-même décroissante et convexe en y : comme $V(y) = E(Y-y)_+$, V(y) possède ces mêmes propriétés. Si la loi F admet une densité f, T(y) est continue et, d'après (I-3), V(y) admet la dérivée:

$$\frac{dV}{dy} = - T$$

négative et croissante (comme il convient, puisque V est convexe et décroissante). Dans le cas général, utilisant la continuité à gauche de $T(y) = 1-F(y)$, on trouve que V admet une dérivée à gauche égale à $-T(y) = -P(Y \geq y)$ et une dérivée à droite égale à $-P(Y > y) = -T(y+0)$

La fonction V(y) (cas continu)

Pour plus de simplicité, nous allons maintenant supposer que la loi F admet une densité $f = dF/dx$ strictement positive sur $\mathbb{R}_+$ (le cas général sera examiné ensuite).

Considérons alors le graphe de la fonction V(y). Il présente un résumé des quatre fonctions (I-3).

En effet, au point d'abscisse y_0, la tangente admet la pente $dV/dy=-T$. L'équation de cette tangente est donc

$$v - V(y_0) = -T(y_0)\ (y-y_0)$$

Cette tangente coupe donc l'axe des v au point d'ordonnée $v = V(y_0) + y_0 T(y_0)$, c'est-à-dire $v = Q(y_0)$ (voir Fig.1). De même, elle coupe l'axe des y au point d'abscisse

$$y = y_0 + V(y_0)/T(y_0) = Q(y_0)/T(y_0) = m(y_0)$$

L'intégrale de la fonction V(y) présente un intérêt particulier. On trouve, en effet :

$$\int_0^\infty V(y)dy = \int_0^\infty dy \int_0^\infty (z-y)_+ F(dz)$$
$$= \int_0^\infty F(dz) \int_0^\infty (z-y)_+ dy = \frac{1}{2} \int_0^\infty z^2\ F(dz)$$

.../

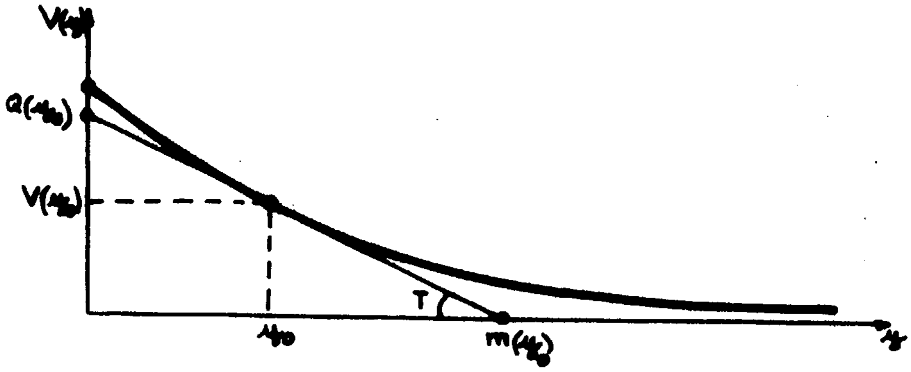


Fig.1 - Graphe de la fonction convexe V(y)

La tangente au point d'abcisse y_0 admet la pente - $T(y_0)$, coupe l'axe des V au point d'ordonnée $Q(y_0)$ et l'axe des y au point d'abscisse $m(y_0)$.

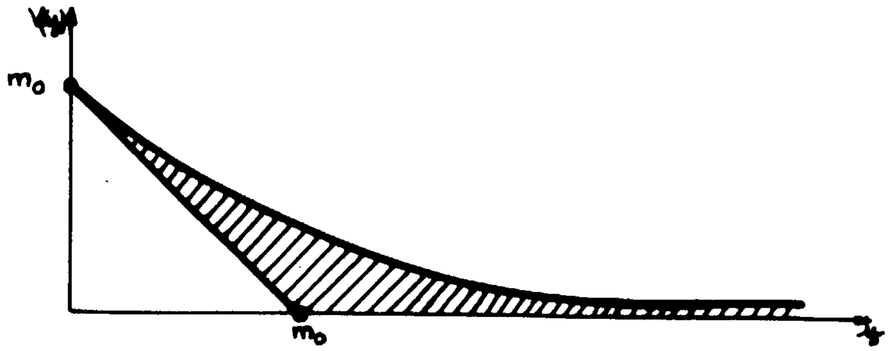


Fig 1 - Bis

La variance de Y est égale à deux fois l'aire hachurée (comprise entre le graphe de V(y), la tangente à l'origine, dont la pente est - 1, et l'axe des y) c'est-à-dire, en posant $m_o = E(Y)$ et $\sigma^2 = Var(Y)$:

$$\int_0^{\infty} V(y)dy = \frac{1}{2}(m_o^2 + \sigma^2)$$

En $y = 0$, on a $V(o) = m_o$ et $dV/dy = -1$, et l'équation de cette tangente est $v = m_o - y$. Comme on a

$$\int_0^{m_o} (m_o - y)dy = \frac{1}{2} m_o^2$$

il en résulte que l'aire de la surface hachurée de la Fig.1-B.z est égale à la <u>moitié de la variance</u> :

$$\text{(I-4)} \qquad \int_0^{\infty} \left[V(y) - (m_o - y)_+\right] dy = \frac{1}{2}\sigma^2$$

<u>Le paramétrage en T</u> (cas continu)

La fonction $T(y)$ est continue (lorsque la loi F admet une densité f) et décroit de 1 à 0 lorsque y varie de 0 à l'infini. (Nous supposons pour l'instant $f > 0$ sur $\mathbb{R}+$). On peut donc définir la fonction inverse $y(T)$ (T variant de 0 à 1), et exprimer Q, m et V en fonction du paramètre T. Examinons les propriétés de ces fonctions $y(T)$, $Q(T)$, $m(T)$ et $V(T)$.

Tout d'abord, $y(T)$ (c'est-à-dire le (I-T) quantile est continue et décroissante de T, en tant que fonction inverse de $T(y)$. Elle varie de $+\infty$ à 0 lorsque T varie de 0 à 1. Pour étudier les autres fonctions, le plus simple est de différentier les relations (I-2). Il vient :

$$dT = -f(y)dy \qquad ; \qquad dQ = -y\ f(y)dy \qquad ; dV = -Tdy$$

Ainsi :

$$\text{(I-5)} \qquad \begin{cases} dQ = y\ dT \\ dV = -Tdy \end{cases}$$

La première relation donne :

$$\frac{dQ}{dT} = y(T) \qquad ; \qquad \frac{d^2Q}{dT^2} = \frac{dy}{dT}$$

La dérivée première de Q(T) est donc ≥ 0, et la dérivée seconde ≤ 0. Ainsi, Q(T) est une f nction croissante et concave de T. De même, la relation

$$\frac{dV}{dT} = - T \frac{dy}{dT}$$

montre que V(T) est croissante.

En ce qui concerne m(T), partant de Q = mT, nous trouvons dQ = mdT + Tdm. Comme dQ = ydT, il vient ainsi :

(I-6)
$$\frac{dm}{dT} = \frac{y-m}{T} = - \frac{V(T)}{T^2}$$

Cette dérivée étant négative, on voit que m(T) est, comme il se doit, une fonction décroissante de T. Comme dV = -Tdy, nous trouvons ensuite :

(I-7)
$$\frac{d}{dT} \left(T^2 \frac{dm}{dT} \right) = T \frac{dy}{dT}$$

Si, au lieu de T, nous introduisons son inverse U = 1/T , les relations précédentes s'écrivent aussi bien:

$$\left\{ \begin{array}{l} \dfrac{dm}{dU} = V \\ \\ \dfrac{d^2m}{dU^2} = - \dfrac{1}{U} \dfrac{dy}{dU} \end{array} \right.$$

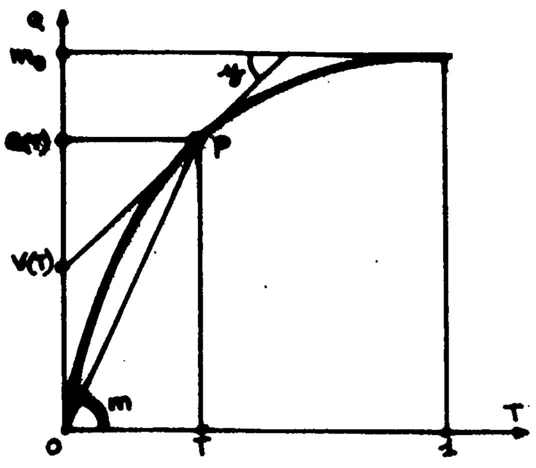


Fig.2 - Graphe de Q(T)

La tangente au point courant P a pour pente y(T) et coupe l'axe des Q au point d'ordonnée V(T). La droite OP a pour pente m(T)

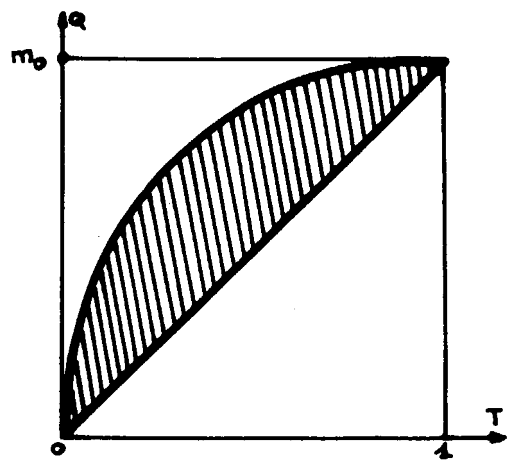


Fig.2 - Bis

L'aire hachurée est égale à la moitié de l'indicateur de dispersion

$$S = \int_0^\infty F(y) \ (1-F(y) \ ) \, dy$$

On voit ainsi que m est une fonction croissante et concave de $U = I/T$.

N.B : La relation (I-7) permet de retrouver facilement la loi $F(y)$ connaissant la fonction $m(T)$. Supposons par exemple que $m(T)$ soit de la forme (dite "loi de Lasky")

$$m(T) = A - b \log T$$

On trouve $dm/dT = - b/T$, et, en portant dans (I-7) :

$$T \frac{dy}{dT} = - b$$

On en tire

$$y(T) = C - b \log T, \text{ et donc}$$

$$T(y) = B\, e^{- y/b}$$

avec d'ailleurs $B = I$, puisque $T(0) = 1$: Y admet une loi exponentielle d'espérance b.

Graphe de Q(T)

La fonction concave $Q(T)$ croît de 0 à $m_0 = E(Y)$ lorsque T varie de 0 à 1 et admet une tangente horizontale en $T= 1$ (puisque $dQ/dT = y(T)$ s'annule en $T = I$)

Son graphe met en évidence les quatre fonctions Q, m, y et V. En effet, $m(T) = Q/T$ est la pente de la droite joignant l'origine au point courant (Fig.2), $y(T)$ est la pente de la tangente et $V(T)$ l'ordonnée du point où cette tangente coupe l'axe des Q.

L'indicateur S de dispersion

L'intégrale $\int_0^1 Q(T)\, dT$ présente un certain intérêt. Compte tenu de (I-7), nous trouvons en intégrant par partie :

$$\int_0^1 Q\, d\, T = \int_0^1 m\, T\, dT = \frac{1}{2} m_o - \frac{1}{2} \int_0^{m_o} T^2 dm$$

$$= \frac{1}{2}\; m_o + \frac{1}{2} \int_0^1 V(T)\; dT$$

Posant :

$$S = \int_0^1 V(T)\; dT$$

et remarquant que

$$\frac{1}{2}\; m_o = \int_0^1 m_o\; T\; dT$$

on voit ainsi que l'aire de la surface comprise entre le graphe de la fonction Q(T) et la droite $q = m_o T$ (aire hachurée de la fig. 2-Bis) est :

$$\int_0^1 (Q - m_o T)\; dT = \frac{1}{2}\; S$$

Le paramètre S apparait comme un indicateur de dispersion. On peut l'exprimer à l'aide de la fonction de répartition F, en écrivant :

$$S = \int_0^1 V(T)\; dT = \int_0^\infty V(y)\; F(dy) = \int_0^\infty F(dy) \int_y^\infty (I-F(z)\;)\; dz$$

$$= \int_0^\infty (I - F(z) \quad dz \int_0^z F(dy)$$

d'où

$$S = \int_0^\infty F(y)\;\; (\; I- F(y)\,)\; dy$$

Sous cette forme, on voit que S = 0 si et seulement si la loi F(dy) est concentrée sur un seul point (c'est une mesure de Dirac)

Cet indicateur de dispersion S vérifie une <u>inégalité isopérimétrique</u> simple. De fait, partant de

$$\frac{1}{2} S = \int_0^1 (Q - m_0 T)\, dT$$

et effectuant une intégration par partie, nous trouvons

$$\frac{1}{2} S = T(Q - m_0 T) \Big|_0^1 - \int_0^1 (y - m_0)\, T\, dT$$
$$= \int_0^1 (m_0 - y)\, T\, dT$$

soit encore, puisque $\int_0^1 (m_0 - y)\, dT = 0$, :

$$\frac{1}{2} S = \int_0^1 (m_0 - y)\, (T - \frac{1}{2})\, dT$$

Appliquons l'inégalité de Schwarz à cette relation. Cela donne :

$$\frac{1}{2} S \leq \left(\int_0^1 (m_0 - y)^2\, dT \right)^{1/2} \left(\int_0^1 (T - \frac{1}{2})^2 dT \right)^{1/2}$$

avec égalité si et seulement si $m_0 - y(T)$ est proportionnel à $(T-1/2)$.
Comme on a :

$$\int_0^1 (T - \frac{1}{2})^2\, dT = \frac{1}{12} \quad ; \quad \int_0^1 (m_0 - y)^2\, dT = \sigma^2$$

Il vient ainsi

(I-8) $$S \leq \frac{\sigma}{\sqrt{3}}$$

avec égalité si et seulement si $(m_0 - y(T)) = a\, (T - 1/2)$

Cette égalité correspond à une loi F <u>uniforme</u> sur le segment $(m_0 - \frac{a}{2}, m_0 + \frac{a}{2})$.

$Q(T)$ est alors une parabole. Ainsi, pour une variance σ^2 donnée, les lois F qui maximisent S sont les lois uniformes sur un segment quelconque de longueur $\sigma\sqrt{12}$. Ou, si on préfère, à S donné, les lois qui minimisent la variance sont les lois uniformes sur des segments de longueur $6S$.

Inégalités de type Tchebychev pour les quantiles

Voici maintenant des inégalités qui montrent que le paramètre S a bien la signification d'un indicateur de dispersion. Dans le cas où il existe une densité strictement positive la médiane μ et les quantiles sont définis sans ambiguité par la fonction $y(T)$:

$$\mu = y(1/2)$$

D'autre part, pour $0 \leq \varepsilon \leq 1$, on a

$$P(Y \geq y(\varepsilon)) = 1-\varepsilon$$

de sorte que $y(\varepsilon)$ est le $1-\varepsilon$ quantile. Mais ce qui suit est valable pour une loi $F(dx)$ quelconque.

Supposons par exemple (quitte à changer ε en $1-\varepsilon$)

$$0 < \varepsilon < 1/2$$

de sorte que $y(\varepsilon) > \mu > y(1-\varepsilon)$. Partant de la relation

$$S = \int_0^\infty F(y)\,(1 - F(y)\,dy$$

établie ci-dessus, et remarquant que l'on a :

$$F(y)\,(1-F(y)) \geq \varepsilon(1-\varepsilon)$$

pour $y(\varepsilon) \geq y \geq y(1-\varepsilon)$, on trouve :

$$S \geq \int_{y(1-\varepsilon)}^{y(\varepsilon)} F(y)\,(1-F(y)\,dy \geq \varepsilon(1-\varepsilon)\left[y(\varepsilon) - y(1-\varepsilon)\right]$$

Ainsi, pour tout p $(0 < p < 1)$, les quantiles y_p et y_{1-p} vérifient l'inégalité :

(I-9) $$|y_p - y_{1-p}| \leq \frac{S}{p(1-p)}$$

Telles sont les inégalités cherchées. On note qu'elles subsistent dans le cas d'une loi quelconque, continue ou non, non nécessairement concentrée sur $\mathbb{R}+$ (dans ce dernier cas, on définit évidemment S par une intégrale de $-\infty$ à $+\infty$). $|y_p - y_{1-p}|$ désigne alors le plus grand possible entre les p-quantiles et les $(1-p)$ - quantiles admissibles.

Condition d'existence pour S

L'indicateur S existe (a une valeur finie) si et seulement si Y admet une espérance m_0.

En effet, pour une variable $Y \geq 0$, l'intégrale

$$S = \int_0^\infty F(y)\ (1-F(y))\ dy$$

converge en même temps que l'intégrale

$$E(Y) = m_0 = \int_0^\infty (1 - F(y))dy$$

Comme $F(1-F)$ est majoré par $1-F$, on a d'ailleurs aussi

(I-10) $$S \leq m_0$$

Si Y est quelconque (non nécessairement positive), chacune des deux intégrales:

$$\int_{-\infty}^0 F(y)dy \qquad \text{et} \qquad \int_0^\infty (1-F(y))dy$$

doit converger. Comme la somme de ces deux intégrales vaut $E(|Y|)$, on voit que S existe si et seulement si $E(|Y|) < \infty$ (ce qui est, par définition, la condition d'existence de $E(Y)$) et de plus :

$$S \leq E(|Y|)$$

Exemples :

1 - Pour une loi de Gauss d'écart type σ (et de moyenne quelconque) on trouve

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

On note que cette valeur diffère relativement peu du maximum théorique $\sigma / \sqrt{3}$: du point de vue de l'inégalité isopérimétrique (I-8), la loi de Gauss apparaît comme assez proche de la loi uniforme. En ce qui concerne les quantiles, l'inégalité (I-9) s'écrit (avec $z \geq 0$)

$$2z \leq \frac{1}{\sqrt{\pi}\; G(z)\; \left[1-G(z)\right]}$$

G(z) désignant la fonction de répartition de la loi de Gauss réduite. Cette inégalité est lâche lorsque z est grand. En fait, on a l'inégalité plus stricte (pour z grand) :

$$z \leq \frac{g(z)}{G(z)\; \left[1-G(z)\right]}$$

qui est en même temps une équivalence. Par contre, le terme $\left[2z\sqrt{\pi}\; G(z)\; (1-G(z))\right]^{-1}$ est de l'ordre de 2 pour z de l'ordre de 0. 8 ou 1.

2 - Pour une loi lognormale, il est possible également de calculer S. Avec :

$$Y = m\, e^{\sigma Z - \sigma^2/2}$$

où Z est une gaussienne réduite, on trouve d'abord en effectuant une intégration par parties :

$$S = m\, e^{-\sigma^2/2} \int_{-\infty}^{+\infty} G(z)\; (1-G(z))\; \frac{d}{dz}\, e^{\sigma z}\, dz$$

$$= m\, e^{-\sigma^2/2} \int (2\, G(z)-1)\;\; g(z)\, e^{\sigma z}\;\; dz$$

$$= \qquad m \int (2\, G(z)-1)\;\; g(z - \sigma)\, dz$$

Remarquant ensuite que S s'annule en $\sigma = 0$, on trouve

$$S = m \int_{-\infty}^{+y} \left[2\, G(z + \sigma) - 1 \right] g(z)\, dz$$

$$= 2m \int_0^{\sigma} f(\tau)\, d\tau$$

avec

$$f(\tau) = \int_{-\infty}^{+\infty} g(z + \sigma)\, g(z)\, dz = \sqrt{2}\; g\left(\frac{\sigma}{\sqrt{2}}\right)$$

Par suite :

$$S = m \left[2\, G\left(\frac{\sigma}{\sqrt{2}}\right) - 1 \right]$$

3 - Plus généralement, supposons que Y soit donnée par une anamorphose gaussienne (transformation croissante) :

$$Y = \varphi(z) = \sum_{n=0}^{\infty} \frac{C_n}{n!}\, H_n(z)$$

Il faut alors calculer :

$$S = \int_{-\infty}^{+\infty} G(z) \left[1 - G(z) \right] \frac{d}{dz}\, \varphi(z)\, dz$$

$$= \int_{-\infty}^{+\infty} \left[2\,G(z) - 1 \right] \varphi(z)\; g(z)\, dz$$

On peut le faire numériquement, ou bien utiliser les développements en polynome d'Hermite. Dans ce dernier cas, partant de :

$$2\,G(z) - 1 = -\frac{1}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k\, 2^{-2k}}{(2k+1)\; k!}\, H_{2k+1}(z)$$

on trouve :

$$S = -\frac{1}{\sqrt{\pi}} \sum_{k=0} \frac{(-1)^k\, C_{2k+1}}{2^{2k}\,(2k+1)\; k!}$$

Formulaire pour l'indicateur de dispersion S

$$S = E\,(\,V(Y)\,) = \int_0^1 V(T)\,dT$$

$$S = \int_{\infty}^{+\infty} F(y)\,\left[\,1 - F(y)\right] dy$$

$$\frac{1}{2}\,S = \int_0^1 (Q(T) - m_0 T)\,dT$$

$$S \leq \frac{\sigma}{\sqrt{3}} \qquad\qquad S \leq E(|Y|)$$

$$|\,y_p - y_{1-p}\,| \leq \frac{S}{p\,(1-p)}$$

$$S = -\,\frac{1}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k\, C_{2k+1}}{2^{2k}\,(2k+1)\,k!}$$

Estimation de S

Supposons que nous disposions de n réalisations $y_1, y_2 \dots y_n$. Désignons par $x_1 \leq x_2 \dots$ les mêmes valeurs numériques rangées par ordre croissant.
Si ces valeurs sont toutes distinctes ($x_1 < x_2 < \dots < x_n$), on prendra l'estimateur "naïf" $F^*(x)$ de la fonction de répartition

$$F^*(x) = \frac{1}{n} \sum_i 1_{x_i < x}$$

et on formera l'estimateur S^* de S :

$$S^* = \int F^*(x)\,\left[1 - F^*(x)\right] dx = \frac{1}{n^2} \sum_{i<j} (x_j - x_i)$$

On trouve sans difficulté :

(I-11) $$S^* = \frac{1}{n^2} \sum_{j=1}^{n} (2j - n - 1) x_j$$

Mais il peut arriver aussi que les valeurs des x_i ne soient pas toutes distinctes (soit que la loi F présente des atomes, soit que l'on ait procédé à des regroupements en classes). Changeant les notations, nous désignerons par $x_1 < x_2 < \cdots < x_k$ ($k \leq n$) les valeurs <u>distinctes</u> prises par les y_i, et par n_j le nombre d'occurences de la valeur x_j. Alors :

$$F^* (x) = \frac{1}{n} \sum_i n_i \, 1_{x_i < x}$$

et

$$S^* = \int F^*(1 - F^*) \, dx = \frac{1}{n^2} \sum_{i<j} n_i \, n_j(x_j - x_i)$$

Pour chaque j, désignons par N_j^- le nombre des valeurs strictement inférieures à x_j, et par N_j^+ le nombre des valeurs inférieures ou égales à x_j :

$$N_j^- = \sum_i n_i \, 1_{i < j} \quad , \quad N_j^+ = \sum n_i \, 1_{i \leq j} = N_j^- + n_j$$

Alors l'estimateur S^* peut s'écrire :

(I-11bis) $$S^* = \frac{1}{n^2} \sum_{j=1}^{k} n_j \, x_j \, (N_j^+ + N_j^- - n)$$

Ayant défini (conventionnellement) $F^*(x)$ à l'aide des indicatrices $1_{x_i < x}$ relatives aux inégalités <u>strictes</u>, on a d'ailleurs par définition :

$$N_j^- = n \, F^*(x_j)$$

Autrement dit encore avec $\nu_j = n_j/n$:

$$S^* = \sum_{j=1}^{k} \nu_j \left[2 \, F^*(x_j) + \nu_j - 1 \right] x_j$$

Nous allons maintenant montrer que - sous les hypothèses habituelles de la statistique classique (ie : les y_i dont des réalisations de n VA indépendantes et identiquement distribuées) - on a :

(I-12) $$E(S^*) = \frac{n-1}{n} S$$

Ceci suggère de remplacer S^* par $(n/(n-1))\ S^*$, exactement comme pour une variance. Cette correction, d'ailleurs peu importante si n n'est pas très petit, n'a cependant plus beaucoup de sens dans le cas le plus fréquent où les Y_i ne peuvent pas être considérées comme indépendantes.

Pour établir la relation (I-12), remarquons que $F^*(x)$ peut être considérée comme une réalisation de la VA

$$\frac{1}{N} \sum_{i=1}^{N} 1_{y_i < x}$$

On a donc, sous les hypothèses de la statistique classique :

$$E(F^*(x)) = F(x)$$

$$E\left[(F^*(x))^2\right] = \frac{1}{n} F(x)\left[(1-F(x)\right] + (F(x))^2$$

Par suite

$$E\left[F^*(x)\ (1-F^*(x)\right] = \frac{n-1}{n} F(x)\left[1-F(x)\right]$$

et (I-12) en résulte -

N.B : Dans le cas continu la formule

$$S = \int (2\,F(x) - 1)\ x\ F(dx)$$

est correcte (on l'obtient par une intégration par parties) mais, dans le cas général d'une loi non continue, elle doit être remplacée par :

$$S = \int (F_-(x) + F_+(x) - 1) \times F(dx)$$

($F_-(x) = P(Y < x)$, $F_+(x) = P(Y \leq x)$ désignent les valeurs à gauche et à droite de la fonction de répartition). On le voit par un calcul identique à celui qui a conduit à la formule (II-11-Bis) - En effet, si $\delta F = F_+(x) - F_-(x)$ est le saut de F en un point x, celui de $F^2(x)$ est

$$(F_- + \delta F)^2 - F_-^2 = \delta F(2F_- + \delta F) = \delta F (F_+ + F_-)$$

II - CAS GENERAL (LOIS NON NECESSAIREMENT CONTINUES).

Nous allons maintenant aborder le cas général où la loi F n'est plus nécessairement continue. Bien que la plupart des résultats subsistent si la VA Y n'est pas supposée positive, nous continuerons cependant, pour des raisons de simplicité, à supposer $Y \geq 0$ (les seuls changements notables qui apparaissent lorsque l'on supprime cette restriction concernant la monotonie des fonctions Q(x) et Q(T), qui ne subsiste pas, ces fonctions admettant respectivement un unique minimum et un unique maximum, et aussi le remplacement de m_0 par $E(|Y|)$ dans l'inégalité $S \leq E(|Y|)$.

Nous poserons :

$$F_-(y) = P(Y < y) \qquad F_+(y) = P(Y \leq y)$$

$$T_-(y) = P(Y \geq y) \qquad T_+(y) = P(Y > y)$$

$$Q_-(y) = E(Y\, 1_{Y \geq y}) \qquad Q_+(y) = E(Y\, 1_{Y > y})$$

$$V(y) = E\left[(Y-y)\, 1_{Y \geq y}\right] = E\left[(Y-y)\, 1_{Y > y}\right] = E\left[(Y-y)_+\right]$$

L'indice moins désigne ainsi la valeur à gauche d'une fonction non continue mais monotone, et l'indice + la valeur à droite (une autre notation usuelle, mais plus lourde est F(y-0) pour $F_-(y)$ et F(y+0) pour F_+). Nos fonctions T et Q étant non croissantes, on a toujours

$$T_- \geq T_+ \quad , \quad Q_- \geq Q_+$$

Plus généralement, pour : $y_1 < y_2$

$$T_-(y_1) \geqslant T_+(y_1) \geqslant T_-(y_2) \geqslant T_+(y_2)$$

et (Y étant $\geqslant 0$) des relations analogues pour Q. Lorsque nous écrirons T ou Q sans préciser l'indice, cela signifiera que la relation correspondante est valable soit en mettant (partout) des indices +, soit au contraire des indices moins.

La fonction $V(y)$ est continue, puisque $(Y - y)\, 1_{Y \geqslant y}$ ne diffère pas de $(Y - y)\, 1_{Y > y}$, et il n'y a pas lieu de distinguer V_+ et V_-. Nous avons déjà vu que cette fonction est convexe et non croissante en y. Comme toute fonction convexe d'une variable, elle admet une dérivée à droite, notée V'_+ ou $d^+ V/dy$, et une dérivée à gauche V'_- ou $d^- V/dy$. On a d'ailleurs vu que l'on a :

$$\frac{d^+ V(y)}{dy} = - T^+(y) \qquad ; \qquad \frac{d^- V(y)}{dy} = - T_-(y)$$

La dérivée n'existant pas, on ne peut plus parler de tangente, mais, la fonction étant convexe, la notion de tangente est remplacée par celle de <u>droite d'appui</u> : une droite d'appui est une droite qui rencontre (en un ou plusieurs points) le graphe de la fonction convexe sans qu'aucun point de ce graphe n'appartienne au demi-plan ouvert situé au-dessous de cette droite. En tout point $(y, V(y))$, les droites d'appui passant par ce point sont celles dont les pentes sont comprises entre $V'_+(y)$ et $V'_-(y)$, c'est-à-dire ici $-T'_+(y)$ et $-T'_-(y)$

La figure 3 résume la situation.

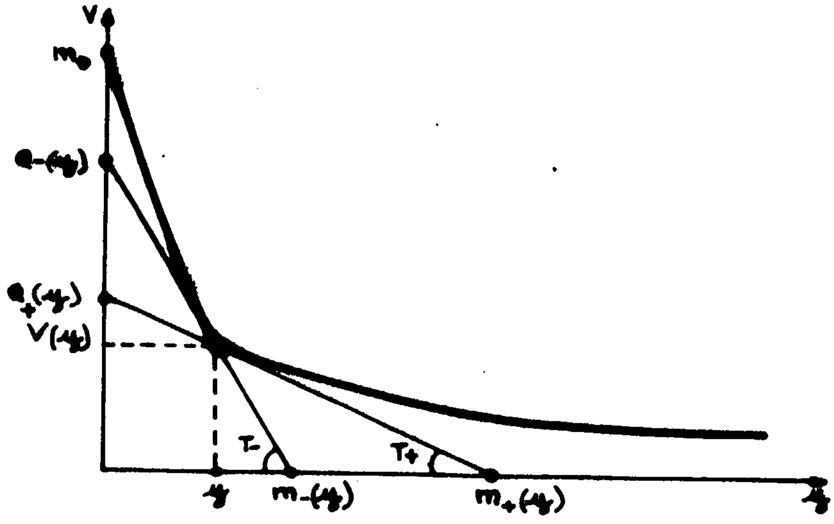


Fig.3 - Graphe de V(y)

montrant les deux droites d'appui extremales au point y.

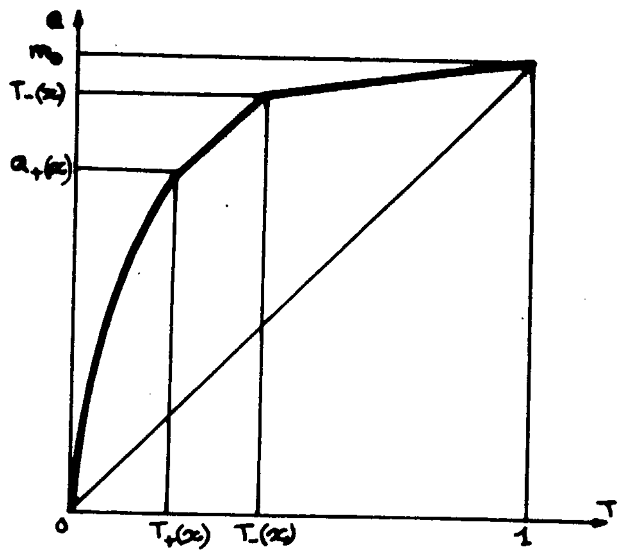


Fig.4 - Graphe de Q(T)

Plaçons nous en un point donné y_0, et prenons t tel que :

$$T_+(y_0) \leq t \leq T_-(y_0)$$

Alors la droite d'équation

$$v = V(y_0) - t(y - y_0)$$

est une droite d'appui, et on a donc pour tout y :

$$V(y_0) + t\, y_0 - ty \leq V(y)$$

pourvu que t soit compris entre $T_+(y_0)$ et $T_-(y_0)$. Pour $t = T_+(y_0)$, on a

$V(y_0) + T_+(y_0)\, y_0 = Q_+(y_0)$, et de même $V(y_0) + y_0 T_-(y_0) = Q_-(y_0)$.

Ainsi, quels que soient y et y_0 :

(II - 1) $$Q(y_0) - y\, T(y_0) \leq V(y)$$

Cette inégalité devenant une égalité pour $y_0 = y$, nous pouvons ainsi écrire

(II- 2) $$V(y) = \operatorname*{Sup}_{x} \left[Q(x) - y\, T(x)\right]$$

Partant de

$$Q(x) \leq V(y) + y\,(T(x)$$

Nous trouvons de même

(II - 3) $$Q(x) = \operatorname*{Inf}_{y} \left[V(y) + y\, T(x)\right]$$

Passons maintenant à la définition de la fonction $Q(T)$. Si T coïncide avec $T_+(x)$ ou $T_-(x)$ pour une valeur de x, on prendra évidemment $Q(T) = Q_+(x)$ ou $Q_-(x)$ respectivement. Pour définir $Q(T)$ lorsque $T_+(x) < T < T_-(x)$

pour un x (nécessairement unique), nous allons utiliser les propriétés de dualité des fonctions convexes. On sait que si l'on pose

$$\chi(t) = \operatorname*{Inf}_{y} (V(y) + yt)$$

la fonction $\chi(t)$ est concave (comme Inf de fonctions linéaires) et que l'on a, réciproquement

$$V(y) = \operatorname*{Sup}_{t} (\chi(t) - yt)$$

si et seulement si V(y) est convexe (si la fonction initiale V n'était pas convexe, on obtiendrait sa plus grande minorante convexe)

D'après (II-3), on aura $\chi(t) = Q(x)$ s'il existe un x tel que $t = T(x)$ (les indices + ou - étant sous-entendus). Cette fonction réalise donc l'interpolation cherchée, et il est naturel de prendre par définition

$$Q(T) = \chi(T) \qquad (0 \leq T \leq 1)$$

Pour un x donné, posons pour abréger $T_+ = T_+(x)$ etc...
Comme on l'a vu, pour T tel que $T_+ \leq T \leq T_-$, on a

$$V(x) + Tx \leq V(y) + Ty$$

pour tout autre y. Par suite

$$(\mathrm{I} - \quad) \qquad Q(T) = \operatorname*{Inf}_{y} (V(y) + Ty) = V(x) + xT$$

Autrement dit, la fonction Q(T) <u>interpole linéairement</u> entre les points "réels" (T_+, Q_+) et (T_-, Q_-) (on peut aussi obtenir ce résultat en remarquant, d'après (II - 2) que la famille des points (T(x), Q(x) contient tous les points extremaux du graphe de V)

Maintenant, étant concave, la fonction Q(T) admet une dérivée à droite $y_+(T)$ et une dérivée à gauche $y_-(T)$, soit

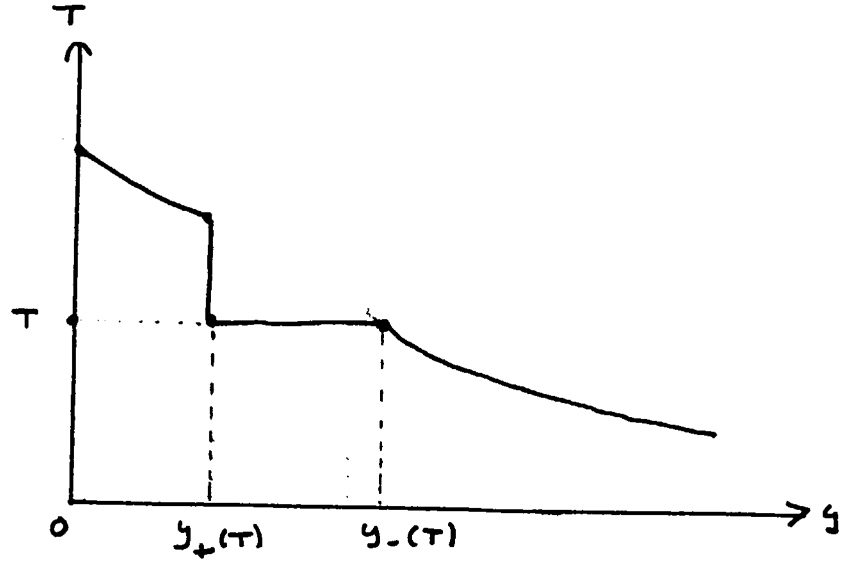
$$y_+(T) = \frac{d^+}{dT} Q(T) \quad ; \qquad y_-(T) = \frac{d^-}{dt} Q(T)$$

avec $y_+(T) \leq y_-(T)$, et, plus généralement, pour $T_1 < T_2$

$$y_-(T_1) \geq y_+(T_1) \geq y_-(T_2) \geq y_+(T_2)$$

puisque Q est concave (De plus y_+ et y_- sont ≥ 0, puisque Q(T) est croissante, Y étant supposée ≥ 0).

Il est facile de voir que $y_+(T)$ est le plus petit des y tels que l'on ait $T_+(y) \leq T$, et de même $y_-(T)$ est le plus grand des y tels que $T_-(y) \geq T$. Ils représentent, respectivement, le plus petit et le plus grand (I - T) - quantiles.



On a d'ailleurs $y_+(T) = y_-(T)$ sauf pour un ensemble au plus dénombrable de valeurs de T (à savoir les valeurs prises par la fonction T(x) sur les intervalles ouverts où elle reste constante). En particulier, si T est une VA uniformément distribuée sur (0,1), $y_+(T)$ et $y_-(T)$ sont presque sûrement égales et équivalentes à Y (méthode de Monte Carlo)

Le paramètre S, défini en posant :

$$S = E\left[V(Y)\right] = \int_{-} F(y)\,(1 - F(y))\,dy$$

vérifie par suite la relation :

$$S = \int_0^1 V\,(y(T)\,)\,dT$$

Mais (pour presque tout T), on a

$$V(y(T)\,) = Q(T) - T\,y(T)$$

Ainsi :

$$S = \int_0^1 Q(T)\,dT - \int_0^1 y(T)\,T\,dT$$

Mais : $\int y(T)\,T\,d\,T = \int T\,d\,Q = QT\Big|_0^1 - \int Q\,d\,T$

et donc :

$$S = 2\int_0^1 Q(T)dT - \bar{m}_o$$

c'est-à-dire :

$$\int_0^1 (Q(T) - m_o\,T)\,d\,T = \frac{1}{2}\,S$$

On démontre de la même façon que toutes les propriétés du paramètre S établies dans le cas continu restent valables dans le cas général.

III - COMPARAISON DE DEUX DISTRIBUTIONS

Nous allons maintenant considérer deux VA X et Y positives (Cette restriction n'a d'ailleurs rien d'essentiel), et désigner par F_X, F_Y, T_X, T_Y etc... les diverses fonctions attachées à ces variables.

Il semble assez naturel de dire que <u>la loi</u> F_X <u>est meilleure que la loi</u> F_Y

$$\text{si } Q_X(T) \geq Q_Y(T) \quad \text{pour tout } T \in (0,1)$$

ou bien encore si $V_X(z) \geq V_Y(z)$ pour tout $z \geq 0$.

Ces deux définitions sont d'ailleurs équivalentes :

$$Q_X \geq Q_Y \iff V_X \geq V_Y$$

Pour le voir, il suffit d'utiliser les formules de dualité entre les fonctions V et Q :

$$\begin{cases} V(z) = \sup\limits_T \,(Q(T) - zT) \\ \\ Q(T) = \inf\limits_z \,(V(z) + zT) \end{cases}$$

Supposons, par exemple, $Q_Y \leq Q_X$. Alors, pour tout T et tout z :

$$Q_Y(T) - zT \leq Q_X(T) - zT \leq \sup\limits_T \,(Q_X(T) - zT) = V_X(z)$$

Par suite aussi

$$V_Y(z) = \sup\limits_T \,(Q_Y(T) - zT) \leq V_X(z)$$

Ainsi $Q_Y \leq Q_X$ entraine $V_Y \leq V_X$, et la réciproque se démontre de la même manière.

La relation d'ordre : F_X plus sélective que F_Y

En Géostatistique minière, c'est surtout entre distributions de même moyenne que la comparaison présente de l'intérêt.

Nous conviendrons de dire que la loi F_X est plus sélective que la loi F_Y si :

(III - 1) $\qquad E(X) = E(Y) \quad$ et $V_X \geqslant V_Y \quad$ (ou $Q_X \geqslant Q_Y$)

Voici une condition nécessaire et suffisante :

<u>Pour que la loi F_X soit plus sélective que la loi F_Y, il faut et il suffit que l'on ait</u> $E\left[\varphi(X)\right] \geqslant E\left[\varphi(Y)\right]$, c'est-à-dire :

(III - 2) $$\int \varphi(x)\, F_X(dx) \geqslant \int \varphi(y)\, F_Y(dy)$$

<u>pour toute fonction convexe</u> φ.

En effet, toute fonction convexe sur $\mathbb{R}_+$ se met sous la forme

$$\varphi(x) = a + bx + \int_0^x (x - t)\, \mu(dt)$$

pour une mesure positive μ sur $\mathbb{R}_+$. Par conséquent, pour toute VA X, on trouve

$$E\left[\varphi(X)\right] = a + b\,E(X) + \int_0^\infty F_X(dx) \int_0^x (x - t)\, \mu(dt)$$
$$= a + b\,E(X) + \int_0^\infty V_X(t)\, \mu(dt)$$

Par suite les conditions (III - 1) entrainent (III - 2).

Inversement, si (III - 2) est vrai, en prenant $\varphi(x) = bx$ il vient $b\,E(X) \geqslant b\,E(Y)$. Comme b peut être choisi positif ou négatif, cela entraine $E(X) = E(Y)$. prenant ensuite $\varphi(x) = (x - x_0)_+$, la relation (III - 2) donne $V_X(x_0) \geqslant V_Y(x_0)$, et donc (III - 1) est vérifiée.

L'effet de support

Voici un exemple simple et très important : soit $Z(x)$, $x \in \mathbb{R}^n$ une FA (stationnaire <u>ou non</u>), $\Pi = \bigcup \varpi_i$ un panneau constitué de la réunion de blocs ϖ_i disjoints. Sans restreindre la généralité, nous pouvons prendre le volume de Π comme unité de volume, et désigner par p_i le volume du bloc ϖ_i, de sorte que $\sum p_i = 1$.

Désignons par Y la valeur moyenne de $Z(x)$ sur le volume Π, et de même par X_i celle de $Z(x)$ sur le bloc ϖ_i, de sorte que ces VA vérifient identiquement

$$Y = \sum p_i \; X_i \tag{III - 3}$$

Comme les p_i constituent une probabilité sur l'espace des indices, désignons par I la VA telle que $P(I = i) = p_i$, et par $X = X_I$ la VA composée obtenue par substitution de I à i dans X_i (randomisation de i). On trouve pour l'espérance conditionnelle de X en Y

$$E\,(X\,/Y) = \sum p_i \; E(X_i \;/\; Y)$$

et, par suite de la relation (III - 3), il vient :

$$E(X \;/\; Y) = Y \tag{III - 4}$$

Il en résulte que <u>la loi de X (loi des blocs) est plus sélective que la loi de Y (loi des panneaux)</u>. Ce phénomène, fondamental en Géostatistique, est connu sous le nom <u>d'effet de support</u>.

En effet, pour toute fonction convexe φ et toute loi de probabilité F, on a

$$\varphi\left[\int zF(dz)\right] \leq \int \varphi\,(z)\;F(dz)$$

En particulier, appliquons ceci à la loi conditionnelle de X. Il vient, d'après (III - 4) :

$$\varphi\,(Y) = \varphi\left[E(X/Y)\right] \leq E\left[\varphi\,(X)/Y\right]$$

Prenant ensuite l'espérance en Y, on trouve :

$$E\,[\varphi(Y)] \leq E[\varphi(X)]$$

Mais, d'après le critère (III - 2) cela signifie justement que F_X est plus sélective que F_Y .

La relation d'ordre : F_X est une dispersée de F_Y

Plus généralement, nous dirons qu'une loi F_X se déduit par dispersion d'une loi F_Y (ou est une dispersée de F_Y) si l'on peut trouver une loi à deux variables $F(dx, dy)$ admettant les lois marginales F_X et F_Y et telle que l'on ait :

$$E(X/Y) = Y$$

Il revient au même de dire que pour (F_Y presque) tout y on peut trouver une probabilité de transition $P_y(dx)$ vérifiant la condition barycentrique :

$$\int x\, P_y(dx) = y$$

et telle que l'on ait

$$\int F_Y(dy)\, P_y(A) = F_X\,(A)$$

pour tout borelien A. Sous cette forme markovienne, il suffit d'un jeu d'écritures pour vérifier que la relation : F_X est une dispersée de F_Y est une relation d'ordre .

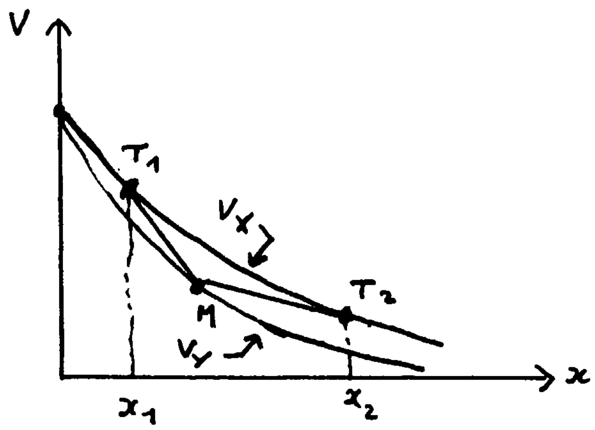
Dans l'exemple précédent (effet de support), la loi des blocs était une dispersée de la loi des panneaux. D'ailleurs, le raisonnement que nous avons fait ensuite pour montrer que la loi des blocs est plus sélective que celle des panneaux n'a utilisé rien d'autre que la propriété (III - 4). Par conséquent, nous pouvons dire : Si F_X est une dispersée de F_Y, alors F_X est plus sélective que F_Y.

En fait, la réciproque est vraie (et même dans des espaces vectoriels beaucoup plus généraux que $\mathbb{R}$). Ce théorème très important est dû à Cartier (on trouvera la démonsteation dans le livre d'Alfsen) sur les compacts convexes). Ainsi :

Une loi F_X est plus sélective qu'une loi F_Y si et seulement si F_X se déduit de F_Y par dispersion. Ou encore :

On a $E\,\varphi(X) \geq E\,\varphi(Y)$ pour toute fonction convexe φ (c-a-d : $E\,X = E\,Y$ et $V_X \geq V_Y$) si et seulement si on peut trouver une loi à deux variables $F(dx, dy)$ admettant les marginales F_X et F_Y et telle que $E(X/Y) = Y$.

REMARQUE : Sans donner une démonstration rigoureuse de la réciproque du théorème de Cartier, j'indique seulement ici une marche possible. Supposons F_X plus sélective que F_Y, c'est-à-dire $m_X = m_Y = m_0$, et $V_X \geq V_Y$, et



considérons le graphe de ces deux fonctions. Par un point M du graphe de V_Y passent 2 droites d'appui MT_1 et MT_2 au graphe de V_X. En remplaçant l'arc $T_1\,T_2$ de V_X par les segments MT_1 et MT_2, nous obtenons le graphe d'une nouvelle fonction convexe V_1, associée à une loi F_1. On voit facilement que F_1 se déduit de F_X en concentrant en son barycentre la masse $F_X(x_2) - F_X(x_1)$ comprise entre les points x_1 et x_2. Autrement dit F_X se déduit de F_1 par dispersion. D'autre part, F_1 est plus sélective que F_Y (puisque $V_1 \geq V_Y$ et que la moyenne reste invariante).

Dans le cas où la loi F_Y est concentrée sur un nombre fini de points $y_1, y_2 \ldots y_n$, il suffit d'itérer n fois l'opération ci-dessus (en prenant pour les points d'abscisse $y_1, y_2 \ldots$) pour amener le graphe supérieur en coïncidence avec V_Y : dans le cas fini, donc, cela suffit pour montrer que F_X est une dispersée de F_Y. Dans le cas général, il faut effectuer un passage à la limite (ce qui est possible) pour obtenir la conclusion. Mais, s'agissant d'un théorème connu, je n'insisterai pas sur ce point.

L'indice de sélectivité S/m_o

Parmi toutes les lois F admettant la même moyenne m_o, il y en a une qui est moins sélective que toutes les autres, à savoir la loi δm_o où toute la masse est concentrée au barycentre m_o (toute loi F de moyenne m_o est bien une dispersée de δm_o). Pour cette loi minimale, on a :

$$V_o(x) = (x - m_o)_+ \quad ; \quad Q_o(T) = m_o T$$

et les paramètres S et σ^2 sont nuls.

Par contre, il n'existe pas de loi plus sélective que toutes les autres. De fait, s'il y en avait une, sa fonction Q serait

$$Q(T) = \begin{cases} 0 & \text{si } T = 0 \\ m_o & \text{si } T > 0 \end{cases}$$

ce qui contredit la continuité de Q(T). Mais il existe des lois réelles aussi proches que l'on veut de ce cas limite (et pour lesquelles par suite le paramètre S est aussi voisin que l'on veut de sa limite supérieure m_o). Par exemple, la VA X_n défini par :

$$P(X_n = 0) = 1 - \frac{1}{n} \quad ; \quad P(X_n = n\, m_o) = \frac{1}{n}$$

admet la moyenne m_o, et Q_n est donné par :

$$Q_n(T) = \begin{cases} m_o & \text{pour } T \geqslant 1/n \\ \dfrac{m_o T}{n} & \text{pour } T < 1/n \end{cases}$$

Son paramètre S vaut $S_n = \dfrac{n-1}{n} m_o$

Ainsi, le rapport S/m_o varie sur l'intervalle (0,1 ((fermé en 0 et ouvert en 1), et il est assez naturel de lui donner le nom <u>d'indice de sélectivité</u> de la loi correspondante. Il est clair que S_1/m_o est plus grand que S_2/m_o dès que la loi F_1 est plus sélective que la loi F_2 (la réciproque n'étant d'ailleurs pas vraie). Comme Q_1 est $\geqslant Q_2$, on a bien, en effet :

$$S_1 = \int_0^1 (Q_1(T) - m_o T)\, dT \geqslant \int_0^1 (Q_2(T) - m_o T)\, dT = S_2$$

<u>REMARQUE</u> : Parmi toutes les lois F <u>concentrées sur l'intervalle</u> (0,L) et admettant la même moyenne m_o ($0 < m_o < L$), il en existe une qui est plus sélective que toutes autres : c'est la loi

$$F_{Max} = \left(1 - \frac{m_o}{L}\right) \delta_o + \frac{m_o}{L} \delta_L$$

concentrée sur les deux points 0 et L et admettant la barycentre m_o. Son existence résulte immédiatement d'un théorème général de Choquet sur les compacts convexes, théorème qui se réduit dans notre cas particulier à quelque chose de très simple : Pour tout $y \in (0,L)$, la probabilité de transition

$$P_y = \left(1 - \frac{y}{L}\right) \delta_o + \frac{y}{L} \delta_L$$

vérifie la condition barycentrique. Si F est une loi concentrée sur (OL) et admettant la moyenne m_o, on trouve

$$\int F(dy)\, P_y = \left(1 - \frac{m_o}{L}\right) \delta_o + \frac{m_o}{L} \delta_L = F_{Max}$$

ce qui montre que F_{Max} est une dispersée de F, donc est plus sélective que F. La variance maximale correspondante à F_{Max} est

$$\sigma^2_{Max} = m_o(L - m_o)$$

et le paramètre S correspondant est

$$S_{Max} = m_o\left(1 - \frac{m_o}{L}\right)$$

Ainsi, pour <u>toute loi concentrée sur (0,L)</u>, <u>l'indice de sélectivité admet la majoration</u> :

(III - 5) $$\frac{S}{m_o} \leq 1 - \frac{m_o}{L}$$

L'effet d'information

Voici maintenant un autre problème, inspiré lui aussi de la pratique minière. Soit Y une VA représentant la teneur d'un bloc que l'on peut, à volonté, sélectionner ou non, et soient X_i, i = 1, 2 ... n les VA représentant les teneurs des échantillons dont on dispose pour prendre cette décision. Supposons que la teneur de coupure soit y. Il s'agit de trouver un critère, c'est-à-dire de choisir un borélien A de $\mathbb{R}^n$ et de sélectionner le bloc si $(X_1, \dots X_n) \in A$, de le rejeter dans le cas contraire. Parmi tous les boréliens A possibles, le meilleur, pour la teneur de coupure y, est le borélien A_y qui réalise le Sup de

$$E\left[(Y-y) 1_A \right) = \int_A E\left[(Y-y) / x_1 \ldots x_n \right] F(dx_1 \ldots dx_n)$$

où $F(dx_1 \ldots dx_n)$ est la loi multivariable des X_i. Posons, pour abréger

$$h(x_1,\ldots x_n) = E (Y/X_1 = x_1,\ldots X_n = x_n)$$

On a donc

$$E \left[(Y-y) 1_A \right] = \int_A (h(x) - y) F(dx_1 \ldots dx_n)$$

On prendra donc comme borelien A_y l'un ou l'autre des deux suivants :

$$A^+_y = \left\{ h(x) \geq y \right\} \quad \text{et} \quad A^-_y = \left\{ h(x) > y \right\}$$

(ou, aussi bien, tout autre borelien compris entre les deux précédents).
On voit que tout se ramène à travailler sur la variable aléatoire

$$H = h(X_1,\ldots X_n) = E(Y/X_1,\ldots X_n)$$

En effet, c'est sur cette variable H que porte la <u>sélection optimale</u>, et la fonction

$$V_H(y) = E ((H-y)_+)$$

représente bien la valeur maximale récupérable pour une teneur de coupure y. De même, pour les tonnages et les quantités de métal :

$$T_H^-(y) = E(1_{H > y}) \quad ; \quad T_H^+(y) = E(1_{H \geq y})$$

$$Q_H^-(y) = E(H\, 1_{H > y}) = E(Y\, 1_{H > y})$$

$$Q_H^+(y) = E(H\, 1_{H \geq y}) = E(Y\, 1_{H} \geq y)$$

(d'après la définition même de l'espérance conditionnelle H). De même encore, en fonction de $T \in (0,1)$, la fonction $Q_H(T)$ représentera le métal récupérable pour un tonnage T donné.

Comme H est une espérance conditionnelle, on a bien $E(H) = E(Y) = m_o$. Pour montrer qua la variable H est moins sélective que Y, il reste à vérifier $V_Y \geq V_H$. Mais cela est évident, puisque $V_Y\ (y_o)$ réalise le Sup de

$$\int_B (y - y_o) \quad F(dy,\ dx_1, \ldots\ dx_n)$$

sur l'ensemble des boreliens B de $\mathbb{R}^{n+1}$, d'où résulte

$$V_Y(y_o) \geq \int_{\mathbb{R} \times A} (y - y_o) \quad F(dy,\ dx_1 \ldots dx_n)$$

$$= \int_A (h\ (x_1, \ldots x_n)^{-y_o} \quad F(dx_1 \ldots dx_n)$$

pour tout borelien de $\mathbb{R}^n$, et par suite avec $A = A_y$:

$$V_Y\ (y_o) \geq V_H\ (y_o)$$

On peut aussi, plus simplement, remarquer que la définition même de $H = E(Y/X_1, \ldots X_n)$ entraine identiquement $E(Y/H) = H$: Ainsi la loi

F_H est une dispersée de F_Y, donc est plus sélective que F_Y.

Ainsi, la variable H est toujours moins sélective que Y. Ce phénomène porte le nom d'effet d'information : la sélection indirecte (i.e la sélection fondée sur les teneurs des échantillons X_i et non sur la teneur Y elle-même) a toujours pour effet de dégrader les fonctions V(y) et Q(T).

IV - UN EXEMPLE TYPIQUE

Nous désignons maintenant par Y la teneur d'un panneau et par X celle d'un échantillon quasi ponctuel (supposé implanté au hasard dans le panneau selon une loi uniforme), de sorte que F_X est plus sélective que F_Y. Si maintenant H = E(Y/X) désigne l'espérance conditionnelle de Y en X (celle de X en Y est égale à Y), nous avons vu que F_Y est plus sélective que F_H, soit symboliquement :

$$F_X \succcurlyeq F_Y \succcurlyeq F_H$$

La première inégalité représente l'effet de support, et la seconde l'effet d'information.

Analysons la conduite naïve d'un exploitant (ignorant la Géostatistique). Soit y_0 la teneur de coupure (en-dessous de laquelle les frais de traitement l'emportent sur la valeur contenue). Cet exploitant naïf va sélectionner les panneaux pour lesquels l'échantillon correspondant donne $X \geqslant y_0$ (ou $X > y_0$). Il s'attend à récupérer le tonnage $T_X(y_0)$, le métal $Q_X(y_0)$ et la valeur $V_X(x_0)$. Pour le tonnage, cet espoir est raisonnable, mais non pour les deux autres fonctions. Nous poserons donc (avec ill pour illusoire)

$$\begin{cases} T_{ill}(y_0) = T_X(y_0) \\ Q_{ill}(y_0) = Q_X(y_0) \\ V_{ill}(y_0) = V_X(y_0) \end{cases}$$

Les quantités effectivement récupérées (à partir de ce mauvais critère)seront en fait

$$T_{eff}(X \geqslant y_0) = T_X(y_0)$$

$$Q_{eff}(X \geqslant y_0) = \int_{x \geqslant y_0} h(x)\ F_X(dx)$$

$$V_{eff}(X \geqslant y_0) = \int_{x \geqslant y_0} (\ h(x) - y_0)\ F_X(dx)$$

Notons tout de suite que l'on a nécessairement :

$$V_{eff}\ (X \geqslant y_0) \leqslant V_H(y_0)$$

puisque

$$V_H(y_0) = \operatorname*{Sup}_A \int_A (h(x) - y_0)\ F_X(dx)$$

A décrivant la famille des boreliens de l'axe des x : avec A = $\left\{ x \geqslant y_0 \right\}$, on obtient bien l'inégalité ci-dessus.

En l'absence d'effet d'information (c'est-à-dire si la teneur Y était connue et si l'on sélectionnait par Y $\geqslant y_0$) on obtiendrait les meilleurs relations tonnage/teneur possibles, compte-tenu de l'effet de support: c'est là un idéal inaccessible en pratique, que nous désignerons par l'indice id :

$$\begin{cases} T_{id}(y_0) = T_Y(y_0) \\ Q_{id}(y_0) = Q_Y(y_0) \\ V_{id}(y_0) = V_Y(y_0) \end{cases}$$

Compte tenu du fait que l'on connait seulement X, et non Y (effet d'information), la meilleure politique consiste, on l'a vu, à couper selon $H \geqslant y_o$, avec $H = h(X) = E(Y/X)$.Cette politique optimale conduit à :

$$T_{opt}(y_o) = T_H(y_o) = \int_{h(x) \geqslant y_o} F_X(dx)$$

$$Q_{opt}(y_o) = Q_H(y_o) = \int_{h(x) \geqslant y_o} h(x)\, F_X(dx)$$

$$V_{opt}(y_o) = V_H(y_o) = \int_{h(x) \geqslant y_o} (h(x)-y_o)\, F_X(dx)$$

C'est ce que l'on peut obtenir de mieux dans la situation réelle. Les inégalités suivantes sont <u>inéluctables</u> :

$$V_{eff}(y_o) \leq V_{opt}(y_o) \leq V_{id}(y_o) \leq V_{ill}(y_o)$$

La première inégalité exprime simplement que $h(x) \geqslant y_o$ est le meilleur critère possible lorsque l'on ne connait que la teneur de l'échantillon X. La perte correspondante $V_{opt} - V_{eff}$ résulte du choix du mauvais critère $X \geqslant y_o$: elle est <u>inexcusable</u>, puisqu'elle peut être évitée sans frais. La seconde inégalité représente <u>l'effet d'information</u> (F_Y est plus sélective que F_H), et la troisième <u>l'effet de support</u> (F_X est plus sélective que F_Y).

Si maintenant nous éliminons le paramètre y_o, de manière à former les fonctions Q(T) correspondantes, nous trouvons ceci : $Q_{eff}(T)$ sera égal à $Q_{opt}(T)$, pourvu seulement que la fonction h(x) soit croissante en x : si h(x) est croissante, on a bien en effet

$$T_{opt}(y_o) = T_{eff}(y_1) \quad ; \quad Q_{opt}(y_o) = Q_{eff}(y_1)$$

pour $h(y_1) = y_o$: ce sont bien les mêmes panneaux qui sont sélectionnés, mais pour des coupures différentes.

Pour le reste, on aura évidemment

$$Q_{opt}(T) \leq Q_{id}(T) \leq Q_{ill}(T)$$

pour les mêmes raisons (sélectivité croissante des variables H, Y et X)

Cas lognormal

Pour illustrer ceci, supposons que X et Y soient des variables log normales de même moyenne m_o, et soient σ^2_X et σ^2_Y les variances de log X et log Y. La condition E(X/Y) = Y montre que le coefficient de corrélation doit être :

$$\rho = \frac{\sigma_Y}{\sigma_X}$$

(ce qui implique évidemment $\sigma^2_Y \leq \sigma^2_X$)

On trouve alors

$$h(x) = E(Y/x) = m_o \left(\frac{x}{m_o}\right)^{\rho^2} e^{\frac{1}{2}\sigma^2_Y(1-\rho^2)}$$

Ainsi H est elle-même lognormale, avec l'espérance m_o, et son logarithme admet la variance

$$\sigma^2_H = \rho^4 \sigma^2_X = \rho^2 \sigma^2_Y$$

(d'où les inégalités $\sigma^2_H \leq \sigma^2_Y \leq \sigma^2_X$, qui traduisent l'effet d'information et l'effet de support)

Pour une teneur de coupure y_o donnée, on posera

$$z = \frac{1}{\sigma} \log \frac{y_o}{m_o} + \frac{1}{2} \sigma$$

(z aura le même indice H, Y ou X que σ) et les fonctions opt, id et ill. sont de la même forme

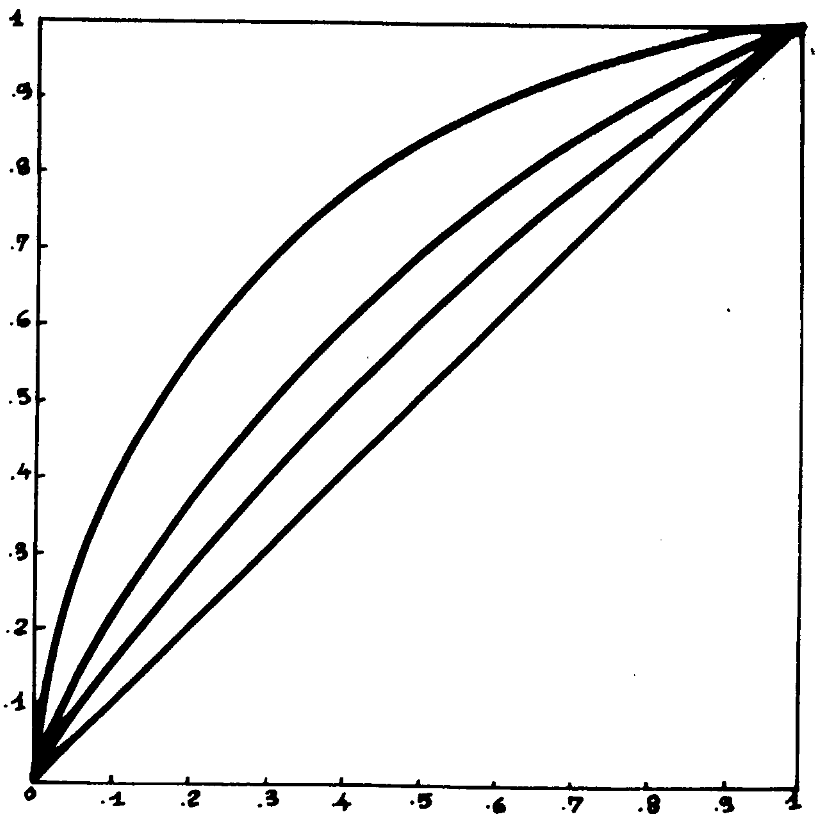


Fig.5 - Q_{ill}, Q_{id} et Q_{opt}

Cas lognormal, $m_o = 1$, $\sigma_X = 1$, $\rho = \sigma_Y = 0.5$

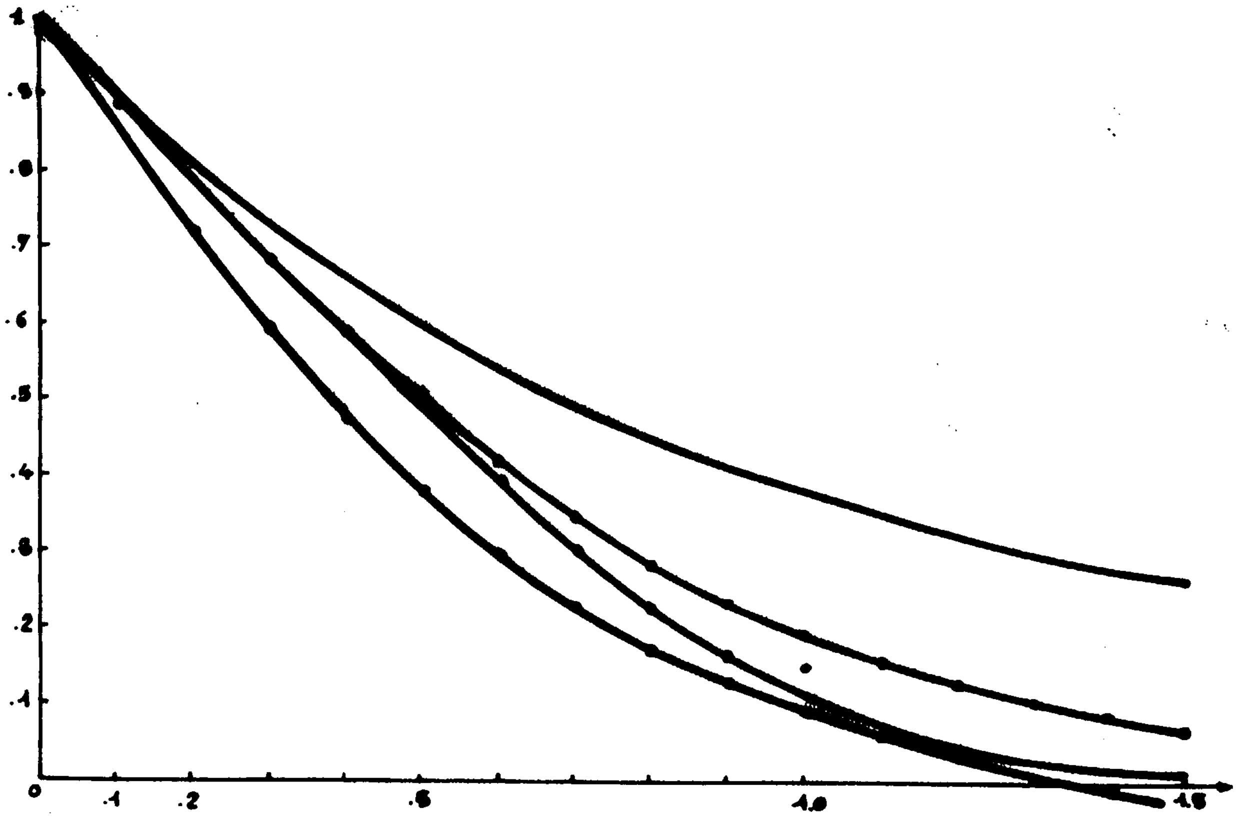


Fig.6 V_{ill}, V_{id}, V_{opt} et V_{eff}

Cas lognormal, $m_o = 1$, $\sigma_X = 1$ $\rho = \sigma_Y = 0.5$

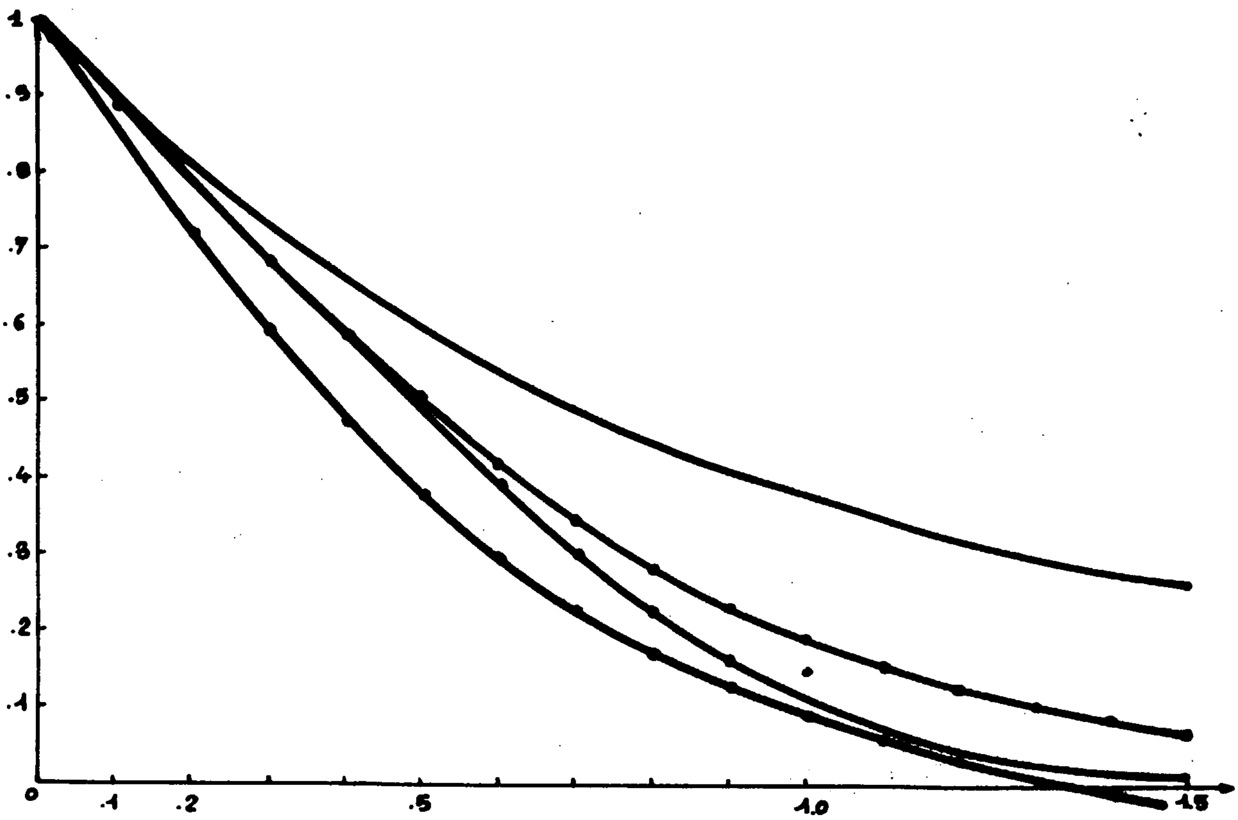


Fig.6 V_{ill}, V_{id}, V_{opt} et V_{eff}

Cas lognormal, $m_o = 1$, $\sigma_X = 1$ $\rho = \sigma_Y = 0.5$

$$T_*(y_0) = 1 - G(z_*)$$

$$Q_*(y_0) = m_0 \left[1 - G(z_* - \sigma_*) \right]$$

$$V_*(y_0) = Q_*(y_0) - y_0\, T_*(y_0)$$

(l'indice $*$ remplace opt, id et ill, et, pour les écarts types, les indices H, Y et X respectivement)

$T_{eff}(y_0)$ et $Q_{eff}(y_0)$ coïncident avec $T_{opt}(y_1)$ et $Q_{opt}(y_1)$, avec Y_1 telle que $h(y_1) = y_0$

Les Fig.5 et 6 montrent l'allure des fonctions Q et V dans le cas m = 1, $\sigma_X = 1$, $\sigma_Y = \rho = 0.5$. On notera l'apparition de valeurs négatives pour V_{eff} aux grandes valeurs de y_0 - Les indices de sélectivité correspondants sont :

$$S_{opt} = \frac{S_H}{m_0} = 0.1403$$

$$S_{id} = \frac{S_Y}{m_0} = 0.2763$$

$$S_{ill} = \frac{S_X}{m_0} = 0.5205$$

Je laisse ces chiffres à la méditation du lecteur, ainsi que les valeurs numériques suivantes des diverses fonctions T, Q, m et V pour quelques teneurs de coupure (voir Tableau 1), calculées sous les mêmes hypothèses.

TABLEAU 1

y_o		T	Q	M	V x 10 000
0.5	ill	.577	.884	1.532	5953
	Eff	.577	.671	1.164	3829
	Opt	.996	.998	1.002	5001
	Id	.872	.949	1.088	5131
0.75	Ill	.416	.785	1.886	4726
	Eff	.416	.515	1.238	2031
	Opt	.847	.899	1.061	2634
	Id	.627	.795	1.267	3248
I.00	Ill	.308	.691	2.241	3829
	Eff	.308	.401	1.301	928
	Opt	.450	.550	1.221	995
	Id	.411	.599	1.492	1974
1.25	Ill	.235	.609	2.594	3156
	Eff	.235	.318	1.355	246
	Opt	.154	.221	1.433	283
	Id	.243	.422	1.736	1183
I.50	Ill	.183	.538	2.944	2637
	Eff	.183	.256	1.402	- 178
	Opt	.040	.067	1.667	67
	Id	.143	.287	1.991	709