COURS DE SÉLECTIVITÉ

par

Christian LANTUÉJOUL

Cours C–140

Janvier 1994

(2ème édition)

CENTRE DE GÉOSTATISTIQUE

35, RUE SAINT–HONORÉ, 77305 FONTAINEBLEAU (France)

ECOLE DES MINES
DE PARIS

CENTRE DE GÉOSTATISTIQUE,
Ecole Nationale Supérieure des Mines de Paris,
35 rue Saint–Honoré, 77305 FONTAINEBLEAU, France

Christian LANTUÉJOUL. Cours de sélectivité. 2ème édition. Janvier 1994. 72 p.

C–140

© ENSMP 1994
Imprimé en France

Cours de Sélectivité

A l'origine de ce cours de Sélectivité figurent deux problèmes apparus dans le contexte de la Géostatistique Minière.

L'exploitation d'un gisement par exemple métallique se fait fréquemment par blocs, le contenu de chaque bloc étant envoyé soit en laverie, soit au rebut stérile. En fonction de la taille des blocs d'exploitation, la quantité de métal récupérée n'est pas la même. En effet, plus la taille des blocs augmente, plus on a de chance d'envoyer en laverie des blocs riches contenant des parties pauvres, ainsi que de mettre au stérile des blocs de faible teneur mais pouvant néanmoins renfermer quelques parties fortement minéralisées. Aussi faut-il s'attendre à récupérer de moins en moins de métal au fur et à mesure que la taille des blocs augmente. Tel est l'*Effet de Support.*

Le deuxième problème considéré dans ce cours est l'*Effet d'Information.* La décision d'orienter un bloc exploité vers la laverie ou bien vers le rebut stérile est prise en fonction de la teneur estimée du bloc et non pas en fonction de sa teneur vraie (qui est inconnue à cette étape de l'exploitation). Il s'ensuit que des blocs pauvres estimés riches sont envoyés en laverie, tandis que des blocs riches estimés pauvres sont abandonnés au stérile, ce qui contribue encore à réduire la quantité de métal effectivement récupérée.

L'objet de ce cours est de voir comment décrire et modéliser les Effets de Support et d'Information. Les chapitres 6 et 7 y sont consacrés. Plutôt que de travailler sur la quantité de métal récupéré, il est avantageux de s'intéresser à la teneur en métal des blocs. A taille de blocs donnée, l'ensemble des teneurs de blocs fournit un histogramme, et étudier l'Effet de Support consiste à voir comment évolue cet histogramme en fonction de la taille des blocs. D'où la nécessité de disposer d'outils de comparaison d'histogrammes. C'est le rôle imparti aux *Courbes de Sélectivité* qui sont introduites aux chapitres 4 et 5. L'utilisation des courbes de sélectivité pour modéliser les effets de support et d'information requiert une bonne connaissance de l'*Espérance Conditionnelle* ainsi que des *Lois Multigaussiennes* et de leur manipulation en termes de *Polynômes d'Hermite.* Ces notions sont présentées en détail aux chapitres 2 et 3.

Ce cours de Sélectivité est très largement inspiré du cours de Probabilité dis-

pensé par G. Matheron au Cycle de Formation Spécialisée en Géostatistique de 1980 à 1988. Par rapport à ce cours, signalons quelques modifications de présentation. Les courbes de sélectivité sont introduites de façon plus graphique. Les parties consacrées à l'espérance conditionnelle et aux lois multigaussiennes ont été davantage étoffées.

Quelques exercices agrémentent ce cours. Certains en sont une application immédiate, d'autres au contraire constituent de véritables compléments au cours. Ces compléments ont été en grande majorité puisés dans un ensemble de notes et de publications de G. Matheron, dont on trouvera la liste en bibliographie.

Chapitre 1

Quelques rappels probabilistes

L'objet de ce chapitre est de rappeler les notions de probabilité qui seront constamment utilisées tout au long de ce cours. Quelques exercices sont proposés pour permettre au lecteur d'apprécier son niveau d'assimilation de ces notions.

1.1 Lois monovariables

Soit X une variable aléatoire. Pour caractériser la loi de X, il est commode d'introduire la *Fonction de Répartition*

$$F(x) = P\{X < x\}$$

La connaissance de F permet le calcul de la *moyenne* de X

$$E\{X\} = \int_{-\infty}^{+\infty} x \, dF(x)$$

où chaque valeur x prise par X est pondérée par sa fréquence d'apparition $dF(x)$. Plus généralement, on peut évaluer la moyenne de n'importe quelle fonction $\phi(X)$ de X par la formule

$$E\{\phi(X)\} = \int_{-\infty}^{+\infty} \phi(x) \, dF(x)$$

Lorsque $\phi(x) = (x - E\{X\})^2$, on obtient la *Variance* de X qui s'écrit

$$Var\{X\} = \int_{-\infty}^{+\infty} (x - E\{X\})^2 \, dF(x) = E\{X^2\} - [E\{X\}]^2$$

La racine carrée de la variance est appelée *Ecart-Type.*

Du point de vue pratique, deux cas sont essentiellement à considérer pour le calcul de la moyenne de $\phi(X)$

- Si la loi de X est *continue* de densité f, c'est-à-dire que

  $$F(x) = \int_{-\infty}^{x} f(u) \, du$$

  la moyenne est l'intégrale

  $$E\{\phi(X)\} = \int_{-\infty}^{+\infty} \phi(x) \, f(x) \, dx$$

- Si la loi de X est *discrète*, c'est-à-dire que X prend les valeurs $x_1, x_2, ...$ avec les probabilités $p_1, p_2, ...$, la moyenne est la somme discrète

  $$E\{\phi(X)\} = \sum_{n \geq 1} \phi(x_n) \, p_n$$

1.2 Lois bivariables

Soit maintenant (X, Y) un couple de variables aléatoires. La fonction de répartition du couple est

$$F(x, y) = P\{X < x, Y < y\}$$

En prenant x ou y infini, on retrouve les fonctions de répartition de X et de Y

$$F(x, +\infty) = F_X(x) \qquad F(+\infty, y) = F_Y(y)$$

La loi $dF(x, y)$ de (X, Y) peut s'écrire comme le produit d'une *Loi Conditionnelle* et d'une loi marginale

$$dF(x, y) \;=\; dF(x \mid y) \, dF_Y(y) \;=\; dF(y \mid x) \, dF_X(x)$$

Les deux variables X et Y sont dites *indépendantes* si la connaissance d'une information sur l'une d'entre elles n'implique rien sur l'autre. Dans ce cas, les lois conditionnelles coïncident avec les lois marginales

$$dF(x \mid y) = dF_X(x) \qquad dF(y \mid x) = dF_Y(y)$$

de sorte que la fonction de répartition du couple (X, Y) se factorise

$$F(x, y) = F_X(x)\, F_Y(y)$$

Entre les deux variables X et Y, on peut définir la *Covariance*

$$Cov\{X, Y\} = E\Big\{[X - E\{X\}]\,[Y - E\{Y\}]\Big\} = E\{XY\} - E\{X\}\,E\{Y\}$$

ainsi que le *Coefficient de Corrélation*

$$\rho_{X,Y} = \frac{Cov\{X, Y\}}{\sqrt{Var\{X\}}\,\sqrt{Var\{Y\}}}$$

qui est une quantité comprise entre -1 et 1. Dans le cas où la covariance entre X et Y est nulle, on dit que les deux variables sont non corrélées. Deux variables non corrélées ne sont pas nécessairement indépendantes. Elles le seront seulement s'il y a absence de corrélation entre toute fonction de l'une et toute fonction de l'autre.

1.3 Fonctions caractéristiques

La *Fonction Caractéristique* d'une variable aléatoire X est la transformée de Fourier de sa loi

$$\Phi(u) = \int_{-\infty}^{+\infty} e^{iux}\, dF(x) = E\Big\{e^{iuX}\Big\}$$

Deux raisons au moins justifient l'utilisation des fonctions caractéristiques. Tout d'abord, compte tenu de ce qu'il y correspondance bijective entre une loi et sa transformée de Fourier, il est parfois bien avantageux d'identifier une loi à l'aide de sa fonction caractéristique. Ensuite et dans certains cas, le

passage aux fonctions caractéristiques simplifie largement les calculs. Pour ne citer que deux exemples, mentionnons que la somme de deux variables aléatoires indépendantes a pour fonction caractéristique le produit des fonctions caractéristiques de chaque variable. Pour déterminer la moyenne et la variance d'une variable, il est commode de considérer le développement à l'origine de sa fonction caractéristique qui s'écrit

$$\Phi(u) = 1 + iuE\{X\} - \frac{u^2}{2}E\{X^2\} + o(u^2)$$

Exercices

1) Soit X une variable aléatoire *positive* de fonction de répartition F (i.e. $F(x) = 0$ pour $x \leq 0$). Montrer que la moyenne de X peut s'écrire

$$E\{X\} = \int_0^{+\infty} [1 - F(x)]\, dx$$

Que se passe-t-il lorsque cette intégrale est infinie?

2) On considère une variable aléatoire X qui suit une loi de *Poisson* de paramètre λ

$$P\{X = n\} = e^{-\lambda} \frac{\lambda^n}{n!} \qquad n = 0, 1, 2, ...$$

a) Quelles sont la moyenne et la variance de X?

b) Donner l'expression de la fonction caractéristique de X.

c) Si Y est une autre variable de Poisson de paramètre μ et indépendante de X, quelle est la loi de $X + Y$?

d) Quelle est la loi de X conditionnellement au fait que $X + Y = n$?

3) Donner un exemple de deux variables aléatoires X et Y non corrélées et non indépendantes.

4) Soit X une variable aléatoire positive de loi F. Il sera commode d'interpréter X comme la durée de vie d'une machine. A chaque instant t, on s'intéresse à la

loi de la durée de vie résiduelle

$$1 - F_t(x) = P\{X - t \geq x \mid X \geq t\}$$

Donner l'expression de F_t en fonction de F. Pour quelles lois F a-t-on non vieillissement (i.e. la durée de vie résiduelle a la même loi que la durée de vie a priori)?

5) Soit X une variable aléatoire positive de densité exponentielle

$$f(x) = a\, e^{-ax} \qquad (a > 0)$$

On demande d'évaluer la moyenne et la variance de X, d'une part par un calcul direct, en passant par la fonction caractéristique d'autre part.

Chapitre 2

L'espérance conditionnelle

Soient X et Y deux variables aléatoires. On cherche à "approcher" X par une fonction de Y. Bien entendu, plusieurs critères d'approximation peuvent être envisagés. Dans le cas où X et Y admettent une moyenne et une variance finie, il est commode de retenir comme critère l'écart quadratique moyen. Le problème se ramène alors à trouver une fonction ϕ telle que la quantité $E\left\{[X - \phi(Y)]^2\right\}$ soit minimale. Nous verrons qu'une telle fonction existe et qu'elle est unique (à une réserve près). On lui donne le nom *d'Espérance Conditionnelle* ou de *Régression*. Dans une deuxième partie, nous étudierons les propriétés de l'espérance conditionnelle. Pour clore ce chapitre, nous présenterons une généralisation qui est l'approximation de X par une fonction de n variables $Y_1, ..., Y_n$.

2.1 Existence et unicité de l'espérance conditionnelle

Supposons que la quantité $E\left\{[X - \phi(Y)]^2\right\}$ soit minimale. Cela veut dire que pour toute variable $\psi(Y)$ de variance finie, on a

$$E\left\{[X - \phi(Y)]^2\right\} \leq E\left\{[X - \psi(Y)]^2\right\}$$

Prenons $\psi = \phi + \epsilon h$ ainsi que $\psi = \phi - \epsilon h$, où ϵ est un nombre positif quelconque. En développant l'inégalité précédente, on obtient respectivement:

$$E\Big\{[X - \phi(Y)]^2\Big\} \leq E\Big\{[X - \phi(Y)]^2\Big\} - 2\epsilon E\Big\{[X - \phi(Y)]h(Y)\Big\} + \epsilon^2 E\Big\{h^2(Y)\Big\}$$

$$E\Big\{[X - \phi(Y)]^2\Big\} \leq E\Big\{[X - \phi(Y)]^2\Big\} + 2\epsilon E\Big\{[X - \phi(Y)]h(Y)\Big\} + \epsilon^2 E\Big\{h^2(Y)\Big\}$$

ce qui donne, après simplification:

$$0 \leq -2E\Big\{[X - \phi(Y)]h(Y)\Big\} + \epsilon E\Big\{h^2(Y)\Big\}$$

$$0 \leq \ 2E\Big\{[X - \phi(Y)]h(Y)\Big\} + \epsilon E\Big\{h^2(Y)\Big\}$$

Faisons tendre maintenant tendre ϵ vers 0 par valeurs positives. Il vient:

$$0 \leq -E\Big\{[X - \phi(Y)]h(Y)\Big\}$$

$$0 \leq \ E\Big\{[X - \phi(Y)]h(Y)\Big\}$$

et par conséquent:

$$E\Big\{[X - \phi(Y)]h(Y)\Big\} = 0$$

Inversement, supposons que l'on ait $E\Big\{[X - \phi(Y)]h(Y)\Big\} = 0$ pour toute variable $h(Y)$ de variance finie. Soit $\psi(Y)$ une variable de variance finie. Calculons $E\Big\{[X - \psi(Y)]^2\Big\}$ à partir de la décomposition $\psi = \phi + (\psi - \phi)$:

$$\begin{aligned} E\Big\{[X - \psi(Y)]^2\Big\} = E\Big\{[X - \phi(Y)]^2\Big\} \ &+ \ 2E\Big\{[X - \phi(Y)]\,[\psi(Y) - \phi(Y)]\Big\} \\ &+ \ E\Big\{[\psi(Y) - \phi(Y)]^2\Big\} \end{aligned}$$

Par hypothèse, le terme médian du deuxième membre est nul. Il reste:

$$E\Big\{[X - \psi(Y)]^2\Big\} = E\Big\{[X - \phi(Y)]^2\Big\} + E\Big\{[\psi(Y) - \phi(Y)]^2\Big\}$$

et par conséquent:

$$E\Big\{[X - \psi(Y)]^2\Big\} \geq E\Big\{[X - \phi(Y)]^2\Big\}$$

Nous avons ainsi établi la propriété caractéristique suivante:

La quantité $E\{[X - \phi(Y)]^2\}$ *est minimale si et seulement si pour toute variable* $h(Y)$ *de variance finie on a*

$$E\{[X - \phi(Y)]h(Y)\} = 0$$

ou encore

$$E\{X\, h(Y)\} = E\{\phi(Y)\, h(Y)\}$$

L'unicité découle simplement de cette propriété. Supposons que X ait un écart quadratique moyen minimal avec deux variables $\phi(Y)$ et $\psi(Y)$ de variances finies. Appliquons la formule caractéristique en choisissant $h(Y) = \psi(Y) - \phi(Y)$:

$$E\{[X - \phi(Y)]\,[\psi(Y) - \phi(Y)]\} = 0$$

$$E\{[X - \psi(Y)]\,[\psi(Y) - \phi(Y)]\} = 0$$

Par soustraction, il vient:

$$E\{[\psi(Y) - \phi(Y)]^2\} = 0$$

ce qui implique $\phi(Y) = \psi(Y)$ presque sûrement.

Pour établir l'existence d'une telle fonction ϕ, nous allons en donner une construction explicite dans un certain nombre de cas particuliers, puis nous vérifierons que la fonction ainsi construite répond au critère de minimisation.

Désignons par $dF(x, y)$ la loi du couple (X, Y). On sait que cette loi est le produit de la loi conditionnelle de X à $Y = y$ fixé et de la loi marginale de Y

$$dF(x, y) = dF(x \mid y) dF_Y(y)$$

Distinguons deux cas:

- La variable Y admet une densité positive f_Y au voisinage du point y. Appliquons la formule caractéristique avec $h(Y) = 1_{Y<y}$:

$$E\{\phi(Y)\, 1_{Y<y}\} = E\{X\, 1_{Y<y}\}$$

ce qui se récrit:

$$\int_{-\infty}^{y} \phi(v)\, dF_Y(v) = \int_{-\infty}^{y} \int_{-\infty}^{+\infty} x\, dF(x \mid v)\, dF_Y(v)$$

Dérivons par rapport à y:

$$\phi(y)\, f_Y(y) = \int_{-\infty}^{+\infty} x\, dF(x \mid y)\, f_Y(y)$$

et après division par $f_Y(y)$, on obtient:

$$\phi(y) = \int_{-\infty}^{+\infty} x\, dF(x \mid y)$$

c'est-à-dire:

$$\phi(y) = E\{X \mid Y = y\}$$

- La variable Y admet un atome en y. En d'autres termes, $P\{Y = y\} \neq 0$. Prenons $h(Y) = 1_{Y=y}$. Dans ce cas, la formule caractéristique s'écrit:

$$E\Big\{\phi(Y)\, 1_{Y=y}\Big\} = E\Big\{X\, 1_{Y=y}\Big\}$$

En termes de probabilité, cela donne:

$$\phi(y)\, P\{Y = y\} = \int_{-\infty}^{+\infty} x\, dF(x \mid y)\, P\{Y = y\}$$

et l'on obtient, après division par $P\{Y = y\}$

$$\phi(y) = \int_{-\infty}^{+\infty} x\, dF(x \mid y)$$

Ainsi donc, dans les deux cas, on a

$$\phi(y) = E\{X \mid Y = y\}$$

Montrons maintenant que la fonction $\phi(y) = E\{X \mid Y = y\}$ vérifie la propriété caractéristique. A cette fin, soit $h(Y)$ une variable de variance finie.

On a:

$$\begin{aligned}
E\Big\{\phi(Y)\,h(Y)\Big\} &= \int_{-\infty}^{+\infty} \phi(y)\,h(y)\,dF_Y(y) \\
&= \int_{-\infty}^{+\infty} E\{X \mid Y = y\}\,h(y)\,dF_Y(y) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,dF(x \mid y)\,h(y)\,dF_Y(y) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,h(y)\,dF(x,y) \\
&= E\Big\{X\,h(Y)\Big\}
\end{aligned}$$

En conclusion

Théorème 2.1 *Si X et Y sont deux variables aléatoires de variance finie, il existe une fonction de Y notée $E\{X \mid Y\}$ qui minimise l'écart quadratique moyen avec X. Cette fonction est unique à une égalité presque sûre près. Elle est caractérisée par le fait que pour toute variable $h(Y)$ de variance finie, on a*

$$E\Big\{Xh(Y)\Big\} = E\Big\{E\{X \mid Y\}\,h(Y)\Big\}$$

2.2 Propriétés de l'espérance conditionnelle

2.2.1 Linéarité

$E\{aX + a'X'|Y\} = aE\{X|Y\} + a'E\{X'|Y\}$ *ps.*

Soit en effet $h(Y)$ une variable de variance finie.

$$\begin{aligned}
E\Big\{E\{aX + a'X'|Y\}h(Y)\Big\} &= E\Big\{(aX + a'X')h(Y)\Big\} \\
&= aE\Big\{Xh(Y)\Big\} + a'E\Big\{X'h(Y)\Big\} \\
&= aE\Big\{E\{X|Y\}h(Y)\Big\} + a'E\Big\{E\{X'|Y\}h(Y)\Big\} \\
&= E\Big\{\Big(aE\{X|Y\} + a'E\{X'|Y\}\Big)h(Y)\Big\}
\end{aligned}$$

D'où le résultat, par unicité de l'espérance conditionnelle.

2.2.2 Positivité

Si $X \geq 0$ *ps., alors* $E\{X|Y\} \geq 0$ *ps.*

Supposons que l'événement "$E\{X|Y\} < 0$" soit non négligeable. Appliquons la formule caractéristique avec

$$h(Y) = 1_{E\{X|Y\}<0}$$

On obtient:

$$E\Big\{E\{X|Y\}1_{E\{X|Y\}<0}\Big\} = E\Big\{X1_{E\{X|Y\}<0}\Big\}$$

Par hypothèse, le membre de droite est ≥ 0. Par construction, le membre de gauche est < 0. D'où la contradiction.

2.2.3 Monotonie

Si $X \leq X'$ *ps., alors* $E\{X \mid Y\} \leq E\{X' \mid Y\}$ *ps.*

Ecrivons $X' = (X' - X) + X$. La linéarité implique

$$E\{X'|Y\} = E\{X' - X|Y\} + E\{X|Y\} \quad ps.$$

Mais $E\{X' - X \mid Y\} \geq 0$ *ps.* puisque $X' - X \geq 0$ *ps.* (positivité), d'où le résultat.

2.2.4 Moyenne

$E\Big\{E\{X \mid Y\}\Big\} = E\{X\}$

Il suffit d'appliquer la formule caractéristique en prenant $h(Y) = 1$.

2.2.5 Cas d'indépendance

Si X *et* Y *sont indépendantes, alors* $E\{X \mid Y\} = E\{X\}$ *ps.*

En effet, l'indépendance entre X et Y implique

$$E\Big\{X\,h(Y)\Big\} = E\Big\{E\{X\}\,h(Y)\Big\}$$

pour toute variable $h(Y)$ de variance finie.

2.2.6 Formule d'invariance

$E\{Xu(Y) \mid Y\} = u(Y)\,E\{X \mid Y\} \quad ps.$

Soit $h(Y)$ une variable de variance finie.

$$E\Big\{E\{X\,u(Y) \mid Y\}h(Y)\Big\} = E\Big\{X\,u(Y)\,h(Y)\Big\} = E\Big\{E\{X \mid Y\}\,u(Y)\,h(Y)\Big\}$$

et l'on applique l'unicité. En particulier, on a

$$E\{u(Y) \mid Y\} = u(Y)$$

2.2.7 Formule des projections successives

$E\{X \mid Y \mid u(Y)\} = E\{X \mid u(Y)\} \quad ps.$

Ici, la notation abrégée $E\{X \mid Y \mid u(Y)\}$ est utilisée à la place de $E\Big\{E\{X \mid Y\} \mid u(Y)\Big\}$. Soit $h[u(Y)]$ une variable de variance finie. alors:

$$\begin{aligned} E\Big\{E\{X \mid Y \mid u(Y)\}\,h[u(Y)]\Big\} &= E\Big\{E\{X \mid Y\}\,h[u(Y)]\Big\} \\ &= E\Big\{X\,h[u(Y)]\Big\} \\ &= E\Big\{E\{X \mid u(Y)\}\,h[u(Y)]\Big\} \end{aligned}$$

et l'on applique l'unicité. La formule des projections successives est souvent utilisée de la droite vers la gauche.

2.2.8 Inégalité de Jensen

Si ϕ est une fonction convexe, alors $\phi\Big[E\{X \mid Y\}\Big] \leq E\Big\{\phi(X) \mid Y\Big\} \quad ps.$

Si ϕ est convexe, ϕ est l'enveloppe supérieure de ses droites d'appui:

$$\phi(x) = \sup_{i \in I}(a_i x + b_i)$$

En vertu de la linéarité et de la monotonie de l'espérance conditionnelle, on peut écrire:

$$a_i E\{X \mid Y\} + b_i = E\{a_i X + b_i \mid Y\} \leq E\{\phi(X) \mid Y\} \quad ps.$$

et l'inégalité est maintenue lorsque l'on passe à la borne supérieure sur I.

2.3 Conditionnement multivariable

Dans cette partie, on considère $n+1$ variables aléatoires X et $Y_1, ..., Y_n$ de moyennes et de variances finies. On cherche à approcher X par une fonction $\phi(Y_1, ..., Y_n)$ des n variables $Y_1, ..., Y_n$, de telle façon que l'écart quadratique moyen $E\Big\{[X - \phi(Y_1, ..., Y_n)]^2\Big\}$ soit minimal.

Tout comme dans le cas monovariable, il existe une fonction ϕ qui minimise cet écart quadratique. Cette fonction est unique à une égalité presque sûre près. Elle est caractérisée par le fait que l'on a

$$E\Big\{[X - \phi(Y_1, ..., Y_n)]\, h(Y_1, ..., Y_n)\Big\} = 0$$

pour toute variable $h(Y_1, ..., Y_n)$ de variance finie. Explicitement, on trouve

$$\phi(y_1, ...y_n) = E\Big\{X \mid Y_1 = y_1, ..., Y_n = y_n\Big\}$$

Toutes les propriétés de l'espérance conditionnelle monovariable se retrouvent dans le cas multivariable. Mentionnons tout particulièrement la *formule d'invariance*

$$E\Big\{X\, u(Y_{i_1}, ..., Y_{i_p}) \mid Y_1, ..., Y_n\Big\} = u(Y_{i_1}, ..., Y_{i_p})\; E\Big\{X \mid Y_1, ..., Y_n\Big\}$$

ainsi que la *formule des projections successives*

$$E\Big\{X \mid Y_1, ..., Y_n \mid Y_{i_1}, ..., Y_{i_p}\Big\} = E\Big\{X \mid Y_{i_1}, ..., Y_{i_p}\Big\}$$

pour toute sous-famille $(i_1, ..., i_p)$ de $(1, ..., n)$.

Exercices

1) Si le couple (X, Y) a pour densité

$$f(x, y) = x \exp\left\{-3y - xe^{-y}\right\} \quad (x, y \geq 0)$$

que vaut $E\{X \mid Y\}$?

2) Quelles sont la moyenne et la variance de la somme d'un nombre aléatoire N de variables $X_1, X_2, ...$ équidistribuées? On supposera que toutes les variables mises en jeu admettent moyenne et variance, et qu'elles sont de plus mutuellement indépendantes.

3) Donner la valeur de $E\left\{[X - E\{X \mid Y\}]^2\right\}$

4) Que vaut $E\left\{X \middle| E\{X \mid Y\}\right\}$?

5) Montrer à l'aide d'un contre-exemple que l'égalité $E\{X \mid Y\} = E\{X\}$ n'implique pas nécessairement l'indépendance entre X et Y.

6) Montrer que si la variable X est indépendante du *vecteur* (Y, Z), on a

$$E\{c(Z) \mid X, Y\} = E\{c(Z) \mid Y\}$$

pour toute variable $c(Z)$. En est-il ainsi si X est seulement indépendante à la fois de Y et de Z?

(Conseil: il suffit d'appliquer la formule caractéristique aux fonctions du type $h(x, y) = a(x)b(y)$).

7) On dira que les deux variables X et Z sont conditionnellement indépendantes de Y si l'on a

$$E\{a(X)c(Z) \mid Y\} = E\{a(X) \mid Y\} E\{c(Z) \mid Y\}$$

pour toutes les variables $a(X)$ et $c(Z)$. Montrer qu'il en est ainsi si et seulement si

$$E\{c(Z) \mid X, Y\} = E\{c(Z) \mid Y\}$$

8) Soit X et Y deux variables aléatoires. On cherche la *régression linéaire de Y selon X*, c'est-à-dire la droite d'équation $y = ax+b$ qui minimise $E\{(Y-aX-b)^2\}$. Montrer qu'une telle droite a pour équation

$$\frac{y - m_Y}{\sigma_Y} = \rho \frac{x - m_X}{\sigma_X}$$

ce qui met en jeu la moyenne et l'écart-type de X et de Y, ainsi que le coefficient de corrélation ρ entre ces deux variables.

9) Soit $X, Y_1, ..., Y_n$ $n+1$ variables aléatoires de moyennes et de variances finies. On cherche à approximer X non pas par une fonction de $Y_1, ..., Y_n$, mais par une combinaison linéaire $X^* = \phi_1(Y_1) + ... + \phi_n(Y_n)$ de fonctions de chaque Y_i, de telle sorte que la quantité

$$E\left\{(X - X^*)^2\right\}$$

soit minimale. Montrez que de telles fonctions ϕ_i existent et qu'elles sont uniques (presque sûrement) à une constante près. Montrez que de plus ces fonctions sont caractérisées par le fait que

$$E\{X^*|Y_i\} = E\{X|Y_i\} \quad i = 1, ..., n$$

(Cette technique d'estimation est connue sous le nom de *Krigeage Disjonctif).*

Chapitre 3

Vecteurs aléatoires Gaussiens

Ce chapitre a pour but de donner quelques propriétés des lois de Gauss à n variables. On en profite pour introduire les polynômes d'Hermite qui permettent une manipulation aisée des lois bigaussiennes.

3.1 Quelques résultats sur les lois multigaussiennes

On rappelle que la densité d'une loi de Gauss de moyenne m et de variance σ^2 est

$$g_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Sa transformée de Fourier vaut:

$$\Phi(u) = e^{ium - \frac{u^2\sigma^2}{2}}$$

Plus généralement, posons la définition suivante:

Définition 3.1 *On dit que le vecteur aléatoire* $X = \begin{pmatrix} X1 \\ \vdots \\ X_n \end{pmatrix}$ *est multigaussien si sa transformée de Fourier vaut:*

$$E\left\{e^{iu^tX}\right\} = e^{iu^tm - \frac{1}{2}u^tCu}$$

Dans cette définition, u est un vecteur colonne de composantes $u_1, ..., u_n$, m un vecteur colonne de composantes $m_1, ..., m_n$ et C désigne une matrice carrée d'ordre n. La transposition d'un vecteur ou d'une matrice est figurée par l'indice supérieur ^{t}. Ainsi, pour ce jeu de notations, on a $u^tX = u_1X_1 + \cdots + u_nX_n$. Par développement limité au voisinage de l'origine de la transformée de Fourier, on peut voir que m est le vecteur moyenne de X et que C est sa matrice de covariance. Dans le cas $n = 1$, on retrouve le cas monovariable. Si $n = 2$, on parlera de vecteur *bigaussien.*

Théorème 3.1 *Soit X un vecteur multigaussien d'ordre n, et soit A une matrice (p, n). Alors AX est un vecteur multigaussien d'ordre p.*

En d'autres termes, toute transformation linéaire préserve le caractère multigaussien.

Démonstration: Si u désigne un vecteur d'ordre p

$$\begin{aligned} E\left\{e^{iu^t(AX)}\right\} &= E\left\{e^{i(u^tA)X}\right\} \\ &= E\left\{e^{i(A^tu)^tX}\right\} \\ &= e^{i(A^tu)^tm - \frac{1}{2}(A^tu)^tC(A^tu)} \\ &= e^{iu^tAm - \frac{1}{2}u^tACA^tu} \end{aligned}$$

On reconnait la transformée de Fourier d'un vecteur multigaussien de moyenne Am et de matrice de covariance ACA^t. ■

Théorème 3.2 *Soit X un vecteur multigaussien. Si $Cov(X_1, X_i) = 0$ pour $i = 2, ..., n$, alors la variable X_1 est indépendante de $(X_2, ..., X_n)$.*

Démonstration: Posons $X = \begin{pmatrix} X_1 \\ Y \end{pmatrix}$ où Y est le vecteur de composantes $X_2, ..., X_n$. D'après le théorème précédent, Y est un vecteur multigaussien dont la moyenne m_Y est le vecteur colonne de composantes $m_2, ..., m_n$ et dont la covariance est la matrice C_Y de terme général $(C_Y)_{ij} = Cov\{X_i, X_j\}$ $2 \leq i, j \leq n$. Posons $u = \begin{pmatrix} u_1 \\ v \end{pmatrix}$, v étant le vecteur colonne de composantes $u_2, ..., u_n$, et calculons la tranformée de Fourier de X:

$$E\left\{e^{iu^tX}\right\} = \exp\left\{i \begin{pmatrix} u_1 & v^t \end{pmatrix} \begin{pmatrix} m_1 \\ m_Y \end{pmatrix} - \frac{1}{2} \begin{pmatrix} u_1 & v^t \end{pmatrix} \begin{pmatrix} C_{11} & 0 \\ 0 & C_Y \end{pmatrix} \begin{pmatrix} u_1 \\ v \end{pmatrix}\right\}$$

En développant, On arrive à la factorisation suivante:

$$E\left\{e^{iu^tX}\right\} = e^{iu_1m_1 - \frac{1}{2}C_{11}u_1^2} \; e^{iv^tm_Y - \frac{1}{2}v^tC_Yv}$$

Il s'agit du produit de la transformée de Fourier de X_1 et de celle de Y. Par conséquent, la variable X_1 est indépendante de $(X_2, ..., X_n)$. ■

Théorème 3.3 *Soit $(X_1, ...X_n)$ un vecteur multigaussien, et soit X_1^{KS} l'estimateur par krigeage simple de X_1 à l'aide de $X_2, ..., X_n$. Alors*

$$E\{X_1 \mid X_2, ..., X_n\} = X_1^{KS} \quad ps.$$

Démonstration: D'après le théorème 3.1, le vecteur $X_1 - X_1^{KS}, X_2, ..., X_n$ est multigaussien. Par ailleurs, on a toujours

$$\forall i = 2, ..., n \;\; Cov\{X_1 - X_1^{KS}, X_i\} = 0$$

En effet, posant $X_1^{KS} - m_1 = \sum_{j=2}^n \lambda_j(X_j - m_j)$, la covariance entre $X_1 - X_1^{KS}$ et X_i s'écrit $C_{1i} - \sum_{j=2}^n \lambda_j C_{ji}$. D'après les équations du krigeage simple, cette covariance vaut 0.

En vertu du théorème 3.2, il s'ensuit que la variable $X_1 - X_1^{KS}$ est indépendante de $X_2, ..., X_n$. Par conséquent, la covariance entre $X_1 - X_1^{KS}$ et toute fonction $h(X_2, ..., X_n)$ de variance finie est nulle. Cela s'écrit:

$$E\left\{[X_1 - X_1^{KS}]\, h(X_2, ...X_n)\right\} = 0$$

ce qui signifie précisément que

$$E\{X_1 \mid X_2, ..., X_n\} = X_1^{KS} \quad ps. \blacksquare$$

Pour finir, il est parfois plus facile de travailler en termes de densité qu'à l'aide de la transformée de Fourier. Cette densité n'existe que dans le cas où la matrice de covariance est inversible:

Théorème 3.4 *Soit X un vecteur multigaussien d'ordre n. Si sa matrice de covariance est inversible, ce vecteur admet la densité*

$$f(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C}} e^{-\frac{1}{2}(x-m)^t C^{-1}(x-m)}$$

Ce résultat est admis.

3.2 Les polynômes d'Hermite

3.2.1 Définition et aspects fonctionnels

Partons de la loi de Gauss centrée réduite

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

On pose:

$$H_n(x) = \frac{g^{(n)}(x)}{g(x)}$$

Aux premiers ordres, cela donne:

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= -x \\ H_2(x) &= x^2 - 1 \end{aligned}$$

Théorème 3.5 *H_n est un polynôme de degré n vérifiant les relations de récurrence*

$$H_{n+1}(x) + xH_n(x) + nH_{n-1}(x) = 0$$
$$H_n'(x) = -n\, H_{n-1}(x)$$

H_n est appelé polynôme d'Hermite de degré n.

Démonstration: Calculons les dérivées successives de la loi de Gauss:

$$g'(x) + xg(x) = 0$$

$$g''(x) + xg'(x) + g(x) = 0$$
$$g^{(3)}(x) + xg''(x) + 2g'(x) = 0$$

et plus généralement

$$g^{(n+1)}(x) + xg^{(n)}(x) + ng^{(n-1)}(x) = 0$$

La division par $g(x)$ fournit une première relation de récurrence

$$H_{n+1}(x) + xH_n(x) + nH_{n-1}(x) = 0$$

qui montre que H_n est un polynôme de degré n.

On a d'autre part:

$$\begin{aligned} H_{n+1}(x)g(x) = g^{(n+1)}(x) &= \left(g^{(n)}(x)\right)' \\ &= \left(H_n(x)g(x)\right)' = H_n'(x)g(x) - xH_n(x)g(x) \end{aligned}$$

ce qui fournit une seconde relation de récurrence

$$H_{n+1}(x) + xH_n(x) - H_n'(x) = 0$$

La comparaison entre ces deux relations de récurrence conduit immédiatement à la formule cherchée. ■

Théorème 3.6 *Les polynômes d'Hermite sont orthogonaux pour la loi de Gauss:*

$$\int_{-\infty}^{+\infty} H_n(x)\, H_p(x)\, g(x)\, dx = \begin{cases} n! & si\ n = p \\ 0 & si\ n \neq p \end{cases}$$

Démonstration: Pour fixer les idées, supposons $n \leq p$ et posons

$$I_{n,p} = \int_{-\infty}^{+\infty} H_n(x)\, H_p(x)\, g(x)\, dx$$

Revenons à la définition des polynomes d'Hermite

$$I_{n,p} = \int_{-\infty}^{+\infty} H_n(x)\, g^{(p)}(x)\, dx$$

et intégrons par partie. On obtient:

$$I_{n,p} = -\int_{-\infty}^{+\infty} H_n'(x)\, g^{(p-1)}(x)\, dx$$

On applique maintenant la formule de récurrence

$$I_{n,p} = n \int_{-\infty}^{+\infty} H_{n-1}(x)\, g^{(p-1)}(x)\, dx = n\, I_{n-1,p-1}$$

En itérant la même procedure n fois, on aboutit à

$$I_{n,p} = n!\, I_{0,p-n} = n! \int_{-\infty}^{+\infty} H_0(x)\, g^{(p-n)}(x)\, dx$$

Deux cas alors se présentent:

- si $n = p$, alors $I_{n,p} = n! \int_{-\infty}^{+\infty} g(x)\, dx = n!$
- si $n \neq p$, alors $I_{n,p} = n! \left[g^{(p-n-1)}(x)\right]_{-\infty}^{+\infty} = n! \left[H_{n-p-1}(x) g(x)\right]_{-\infty}^{+\infty} = 0$

■

Soit λ une valeur quelconque. Considérons le développement de Taylor de la fonction $g(x - \lambda)$ au point x:

$$g(x - \lambda) = \sum_{n=0}^{+\infty} \frac{(-\lambda)^n}{n!}\, g^{(n)}(x) = \sum_{n=0}^{+\infty} \frac{(-\lambda)^n}{n!}\, H_n(x)\, g(x)$$

Par ailleurs, on a aussi

$$g(x-\lambda) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\lambda)^2} = g(x)\, e^{x\lambda - \frac{\lambda^2}{2}}$$

Par identification, on obtient:

$$e^{\lambda x} = e^{\frac{\lambda^2}{2}} \sum_{n=0}^{+\infty} \frac{(-\lambda)^n}{n!} H_n(x)$$

Ainsi, la fonction exponentielle peut se développer en polynomes d'Hermite. Cela n'est au reste qu'un cas particulier du résultat général suivant:

Théorème 3.7 *Toute fonction ϕ qui est de carré intégrable pour g, c'est-à dire telle que $\int_{-\infty}^{+\infty} \phi^2(x)\, g(x)\, dx < +\infty$ peut être développée en polynômes d'Hermite*

$$\phi(x) = \sum_{n=0}^{+\infty} \frac{\phi_n}{n!} H_n(x)$$

Les coefficients ϕ_n sont donnés par la formule

$$\phi_n = \int_{-\infty}^{+\infty} \phi(x)\, H_n(x)\, g(x)\, dx$$

et de plus

$$\int_{-\infty}^{+\infty} \phi^2(x)\, g(x)\, dx = \sum_{n=0}^{+\infty} \frac{\phi_n^2}{n!}$$

Plus généralement, si ϕ et ψ sont de carré intégrable pour g, alors

$$\int_{-\infty}^{+\infty} \phi(x)\, \psi(x)\, g(x)\, dx = \sum_{n=0}^{+\infty} \frac{\phi_n\, \psi_n}{n!}$$

La décomposition de ϕ en polynômes d'Hermite est de démonstration délicate. Nous l'admettrons. Les trois autres formules découlent directement de l'orthogonalité des polynômes d'Hermite. Leur établissement est laissé à titre d'exercice.

Pour la suite, et conformément à l'habitude, nous noterons $L^2(I\!R, g)$ l'ensemble des fonctions de carré intégrable pour g.

3.2.2 Quelques aspects probabilistes

Soit X une variable aléatoire suivant une loi de Gauss centrée réduite. On s'intéresse aux caractéristiques probabilistes des variables $H_n(X)$.

Partons de la formule d'orthogonalité des polynômes d'Hermite. En termes probabilistes, elle s'écrit

$$E\{H_n(X)\, H_p(X)\} = \begin{cases} n! & \text{si } n = p \\ 0 & \text{si } n \neq p \end{cases}$$

D'où la moyenne de $H_n(X)$

$$E\{H_n(X)\} = E\{H_n(X)\, H_0(X)\} = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \neq 0 \end{cases}$$

ainsi que sa variance

$$Var\{H_n(X)\} = \begin{cases} n! & \text{si } n \neq 0 \\ 0 & \text{si } n = 0 \end{cases}$$

La covariance entre $H_n(X)$ et $H_p(X)$ vaut

$$Cov\{H_n(X), H_p(X)\} = \begin{cases} n! & \text{si } n = p \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Soit maintenant $\phi \in L^2(\mathbb{R}, g)$. Alors ϕ admet un développement en polynômes d'Hermite

$$\phi(x) = \sum_{n=0}^{+\infty} \frac{\phi_n}{n!} H_n(x)$$

La moyenne de $\phi(X)$ s'obtient par linéarité

$$E\{\phi(X)\} = \sum_{n=0}^{+\infty} \frac{\phi_n}{n!} E\{H_n(X)\} = \phi_0$$

Quant à sa variance, elle vaut:

$$Var\{\phi(X)\} = \sum_{n,p=0}^{+\infty} \frac{\phi_n}{n!} \frac{\phi_p}{p!} Cov\{H_n(X), H_p(X)\} = \sum_{n=1}^{+\infty} \frac{\phi_n^2}{n!}$$

Plus généralement, si ϕ et ψ sont dans $L^2(I\!R, g)$, alors

$$Cov\{\phi(X), \psi(X)\} = \sum_{n,p=0}^{+\infty} \frac{\phi_n}{n!} \frac{\psi_p}{p!} Cov\{H_n(X), H_p(X)\} = \sum_{n=1}^{+\infty} \frac{\phi_n \psi_n}{n!}$$

Considérons maintenant un couple (X, Y) de variables centrées réduites bigaussiennes de coefficient de corrélation ρ. Tout l'intérêt des polynômes d'Hermite apparaît dans la propriété suivante:

Théorème 3.8 *La loi bigaussienne admet un développement en polynômes d'Hermite de la forme:*

$$g_\rho(x, y) = \sum_{n=0}^{+\infty} \frac{\rho^n}{n!} H_n(x) H_n(y) g(x) g(y)$$

Démonstration: La transformée de Fourier de la loi bigaussienne

$$\hat{g}_\rho(u, v) = e^{-\frac{1}{2}(u^2 + v^2 + 2\rho uv)}$$

s'écrit après développement en série

$$\hat{g}_\rho(u, v) = e^{-\frac{1}{2}(u^2 + v^2)} \sum_{n=0}^{+\infty} \frac{\rho^n}{n!} (iu)^n (iv)^n$$

Or on a

$$e^{-\frac{u^2}{2}} (iu)^n = (-1)^n \int_{-\infty}^{+\infty} e^{iux} H_n(x) g(x) dx$$

ce qui s'établit au moyen de n intégrations par partie, ou bien à partir du développement en polynômes d'Hermite de l'exponentielle complexe. De même

$$e^{-\frac{v^2}{2}} (iv)^n = (-1)^n \int_{-\infty}^{+\infty} e^{ivy} H_n(y) g(y) dy$$

et par conséquent

$$\hat{g}_\rho(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(ux + vy)} \sum_{n=0}^{+\infty} \frac{\rho^n}{n!} H_n(x) H_n(y) g(x) g(y) dx dy$$

ce qui est le résultat annoncé. ■

Théorème 3.9 *Si (X, Y) suit une loi bigaussienne de coefficient de corrélation ρ, on a*

$$E\{H_n(X) \mid Y\} = \rho^n H_n(Y) \quad n = 0, 1, ...$$

Démonstration: En effet

$$E\{H_n(X) \mid Y = y\} = \int_{-\infty}^{+\infty} H_n(x)\, g_\rho(x \mid y)\, dx$$

où $g_\rho(x \mid y)\, dx$ désigne la loi de X à $Y = y$ fixé. D'après le théorème précédent, cette loi peut être développée en polynômes d'Hermite

$$g_\rho(x \mid y) = \sum_{m=0}^{+\infty} \frac{\rho^m}{m!} H_m(x)\, H_m(y)\, g(x)$$

Il suffit alors d'utiliser la propriété d'orthogonalité. ■

Donnons ici deux conséquences importantes de ce résultat:

La première est que si $\phi \in L^2(\mathbb{R}, g)$

$$E\{\phi(X) \mid Y\} = \sum_{n=0}^{+\infty} \frac{\phi_n}{n!} E\{H_n(X) \mid Y\} = \sum_{n=0}^{+\infty} \frac{\phi_n}{n!} \rho^n H_n(Y)$$

On obtient donc une fonction de Y qui ne dépend que des coefficients ϕ_n et du coefficient de corrélation ρ. On la note traditionnellement $\phi_\rho(Y)$.

Une deuxième conséquence est que si $\phi, \psi \in L^2(\mathbb{R}, g)$, alors

$$Cov\big\{\phi(X), \psi(Y)\big\} = \sum_{n=1}^{+\infty} \frac{\phi_n \psi_n}{n!} \rho^n$$

Cela résulte de ce que

$$\begin{aligned} E\{\phi(X)\,\psi(Y)\} &= E\Big\{E\{\phi(X)\,\psi(Y) \mid X\}\Big\} \\ &= E\Big\{\phi(X)E\{\psi(Y) \mid X\}\Big\} \\ &= \sum_{n=0}^{+\infty} \frac{\psi_n}{n!}\,\rho^n E\Big\{\phi(X)\,H_n(X)\Big\} \\ &= \sum_{n=0}^{+\infty} \frac{\phi_n\psi_n}{n!}\,\rho^n \end{aligned}$$

Exercices

1) Donner l'expression de la densité d'un couple bigaussien (X, Y) centré normé, de coefficient de corrélation ρ. Quelle est la densité de X à $Y = y$ fixé?

2) On dira qu'une variable Z est *lognormale* si $\ln(Z)$ suit une loi Gaussienne. Une telle variable se met sous la forme

$$Z = m\exp\Big\{\sigma Y - \frac{\sigma^2}{2}\Big\}$$

où Y est une variable Gaussienne centrée réduite, et où m et σ sont deux valeurs positives. Quelles sont les moyennes et les variances de Z et de $\ln(Z)$?

3) Soit Z et Z' deux variables lognormales de la forme

$$Z = m\exp\Big\{\sigma Y - \frac{\sigma^2}{2}\Big\} \qquad Z' = m'\exp\Big\{\sigma' Y' - \frac{\sigma'^2}{2}\Big\}$$

On suppose que le couple (Y, Y') est bigaussien de coefficient de corrélation ρ. Quelle est la covariance entre Z et Z'? Que vaut $E\{Z \mid Z'\}$?

4) Donner le développement en polynômes d'Hermite de

$$\phi(y) = 1_{y \geq y_0}$$

5) Soit $\phi \in L^2(\mathbb{R}, g)$. Montrer qu'il existe un et un seul polynôme P_n de degré n tel que

$$\int_{-\infty}^{+\infty} \Big[\phi(y) - P_n(y)\Big]^2 g(y)\, dy$$

soit minimum. Donner la forme explicite de P_n.

6) Soit X une variable Gaussienne centrée réduite. Montrer que si les 3 nombres m, n et p sont les longueurs des cotés d'un triangle de périmètre $2s$, on a

$$E\{H_m(X)\, H_n(X)\, H_p(X)\} = \frac{m!}{(s-m)!}\, \frac{n!}{(s-n)!}\, \frac{p!}{(s-p)!}$$

et que dans le cas contraire, on a

$$E\{H_m(X)\, H_n(X)\, H_p(X)\} = 0$$

(Conseil: on cherchera à développer $\exp\Big\{(\lambda + \mu + \nu)x\Big\}$ en polynômes d'Hermite de deux façons différentes).

Chapitre 4

Les courbes de sélectivité

L'objet de ce chapitre est de proposer un certain nombre d'outils pour caractériser des lois de probabilité et leur dispersion. Ces outils, à la fois simples et commodes, sont particulièrement adaptés à l'étude des effets de support et d'information que nous aborderons aux prochains chapitres. Pour ne pas brusquer les traditions, nous avons décidé de garder la terminologie de la géostatistique minière.

Dans ce chapitre, nous désignerons par Z une variable aléatoire *positive*, de loi F, de moyenne m finie et de variance σ^2, finie ou non.

4.1 Tonnage et Quantité de Métal

Définition 4.1 *On appelle Tonnage la fonction qui vaut*

$$T(z) = \int_z^{+\infty} dF(u)$$

pour chaque teneur de coupure z.

Le tonnage peut aussi être écrit:

$$T(z) = P\{Z \geq z\} = 1 - F(z)$$

ou encore

$$T(z) = E\left\{1_{Z \geq z}\right\}$$

ce qui met en jeu la fonction indicatrice

$$1_{Z \geq z} = \begin{cases} 1 & \text{si } Z \geq z \\ 0 & \text{sinon} \end{cases}$$

Le tonnage est une fonction *décroissante* entre les valeurs $T(0) = 1$ et $T(+\infty) = 0$. Le tonnage caractérise la loi F.

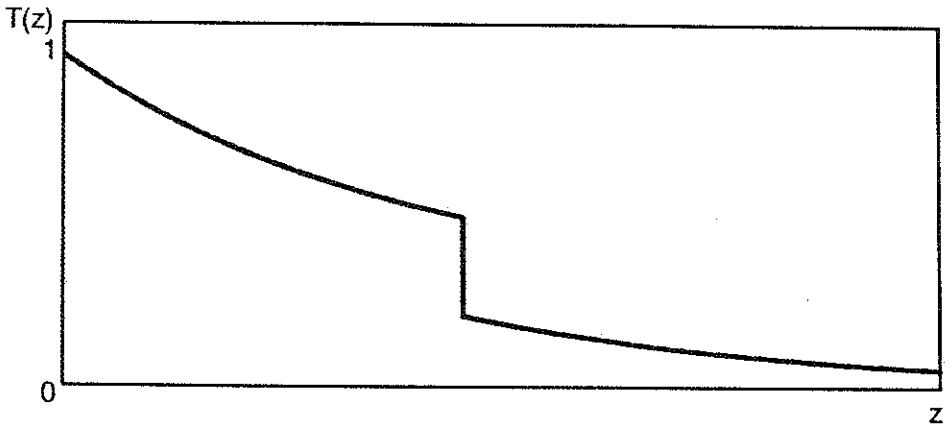

Figure 4.1 : Exemple d'une courbe de Tonnage. La loi de probabilité qui conduit à une telle courbe est le mélange d'une densité exponentielle et d'une masse de Dirac.

Définition 4.2 *On appelle Quantité de Métal la fonction qui vaut*

$$Q(z) = \int_z^{+\infty} u dF(u)$$

pour chaque teneur de coupure z.

La quantité de métal peut aussi s'écrire:

$$Q(z) = E\left\{ Z\, 1_{Z \geq z} \right\}$$

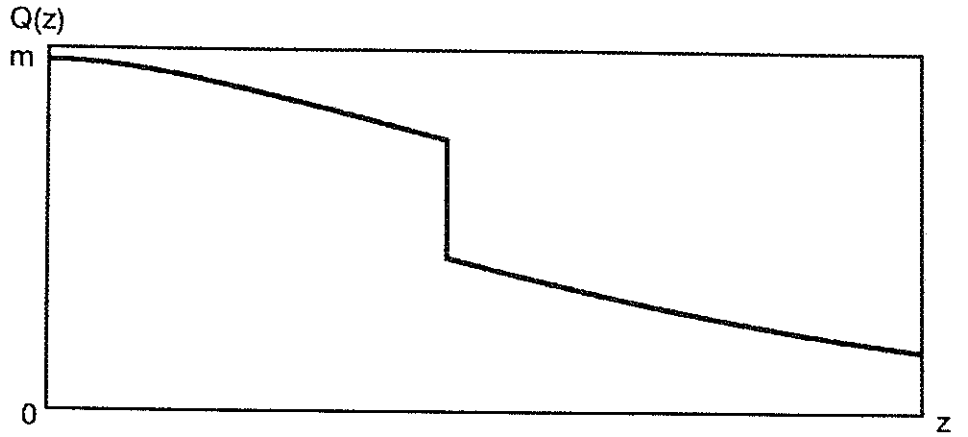

Figure 4.2 : Courbe de Quantité de Métal

Il s'agit là encore d'une fonction décroissante. Elle vaut m en 0 et s'annule à l'infini. Tout comme le tonnage, la quantité de métal caractérise la loi F.

Par intégration par parties, on a

$$Q(z) = \Big[-u\,T(u)\Big]_z^{+\infty} + \int_z^{+\infty} T(u)\,du = z\,T(z) + \int_z^{+\infty} T(u)\,du$$

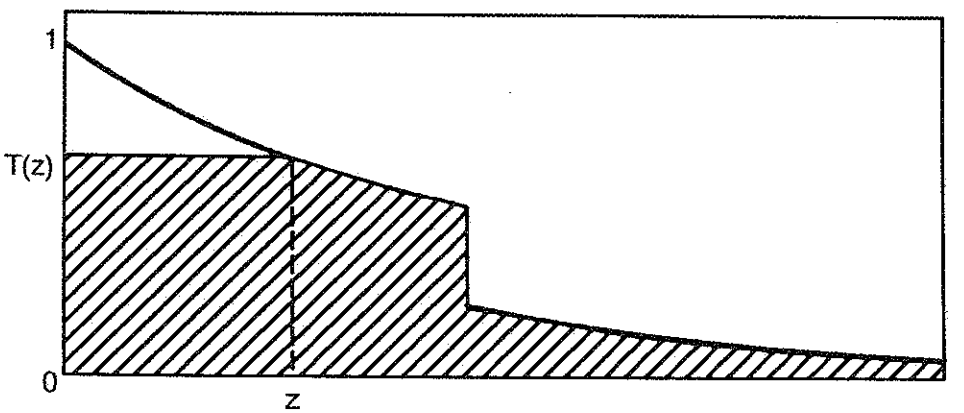

Figure 4.3 : Représentation graphique de $Q(z)$ sur la courbe de tonnage

Graphiquement, $Q(z)$ est l'aire du domaine contenu dans le sous-graphe du tonnage et constitué des points d'ordonnée inférieure à $T(z)$.

4.2 La fonction $B(z)$

Définition 4.3 *On appelle $B(z)$ la fonction qui vaut*

$$B(z) = \int_z^{+\infty} T(u)\, du$$

pour chaque teneur de coupure z.

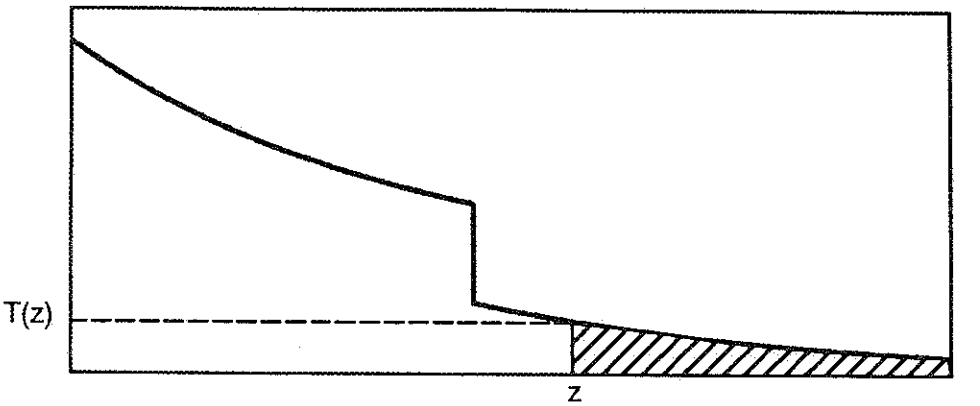

Figure 4.4 : Définition graphique de $B(z)$ à l'aide de la courbe de tonnage.

D'après le paragraphe précédent, $B(z)$ s'écrit aussi

$$B(z) = Q(z) - zT(z)$$

Ainsi, $B(z)$ est la différence entre la quantité de métal $Q(z)$ et la quantité minimale $zT(z)$ que l'on est sûr de récupérer à la teneur de coupure z. En Géostatistique minière, la fonction $B(z)$ est appelée *Bénéfice Conventionnel.*

D'autres expressions possibles pour $B(z)$ sont:

$$B(z) = \int_z^{+\infty} (u - z)\, dF(u) = E\Big\{(Z - z)1_{Z \geq z}\Big\}$$

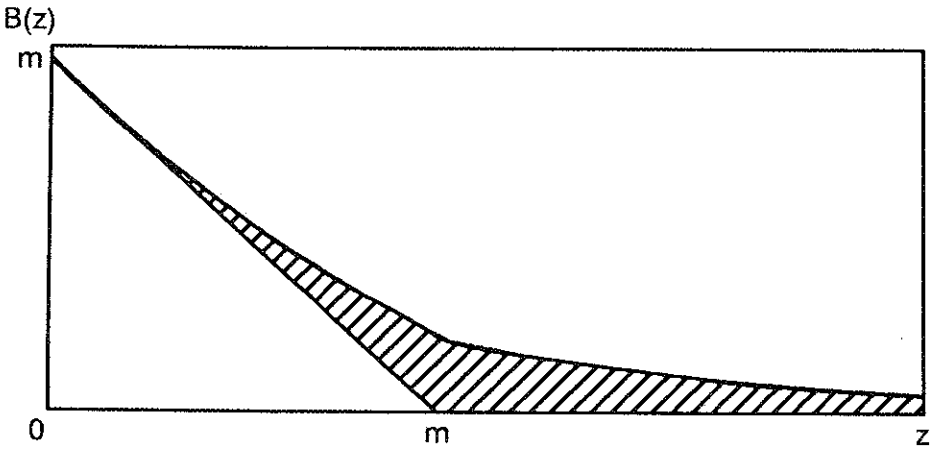

Figure 4.5 : Forme de la courbe $B(z)$. Le domaine hachuré a pour surface la moitié de la variance.

La fonction $z \longmapsto (Z - z)1_{Z \geq z}$ est décroissante et convexe. Le passage à l'espérance préserve ces propriétés et par conséquent la fonction $B(z)$ est *décroissante* et *convexe* . Elle prend la valeur m en 0 et s'annule à l'infini.

De par sa définition, la fonction $B(z)$ caractérise le tonnage et par suite la loi F. De sorte que toutes les caractéristiques de la loi F peuvent être retrouvées sur $B(z)$. Vérifions qu'il est ainsi pour la variance:

$$\int_0^{+\infty} B(z)\, dz = \int_0^{+\infty} \int_z^{+\infty} (u - z)\, dF(u)\, dz$$

On intervertit les intégrales

$$\int_0^{+\infty} B(z)\, dz = \int_0^{+\infty} \int_0^{u} (u - z)\, dz\, dF(u)$$

ce qui donne

$$\int_0^{+\infty} B(z)\, dz = \int_0^{+\infty} \frac{u^2}{2}\, dF(u) = \frac{1}{2}\left(m^2 + \sigma^2\right)$$

Ainsi, l'aire hachurée de la Figure 4.5 est égale à la moitié de la variance.

Autre exemple: la proportion q de valeurs nulles de la loi F s'obtient en

étudiant la pente de $B(z)$ au voisinage de l'origine

$$\lim_{z\to 0} \frac{B(0) - B(z)}{z} = \lim_{z\to 0} \frac{1}{z} \int_0^z T(u)\, du = T(0+) = 1 - q$$

4.3 La fonction $Q(T)$

Plutôt que de sélectionner les teneurs supérieures à la coupure z, il peut être commode de sélectionner celles qui assurent un tonnage T.

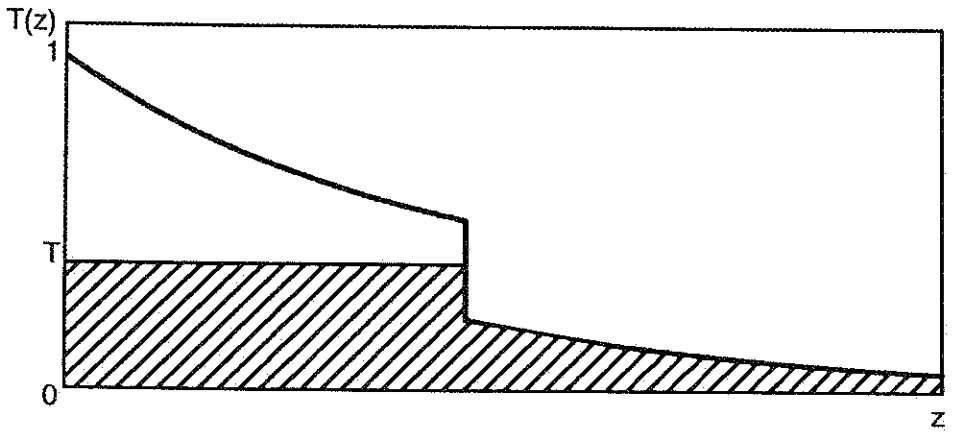

Figure 4.6 : Définition de $Q(T)$

Partons de l'expression de la quantité de métal établie plus haut

$$Q(z) = z\,T(z) + \int_z^{+\infty} T(u)\, du$$

Comme le tonnage est décroissant, on peut aussi écrire:

$$Q(z) = \int_0^{+\infty} T(u) \wedge T(z)\, du$$

où $T(u) \wedge T(z)$ désigne le minimum des deux valeurs $T(u)$ et $T(z)$. Plus généralement, introduisons la définition suivante:

Définition 4.4 *On appelle $Q(T)$ la fonction qui vaut*

$$Q(T) = \int_0^{+\infty} T(u) \wedge T\, du$$

pour toute valeur $0 \leq T \leq 1$.

Dans le cas où $T = T(z)$, on retrouve la quantité de métal associée à la teneur de coupure z.

$Q(T)$ est une fonction *croissante* entre les valeurs $Q(0) = 0$ et $Q(1) = m$. Elle est de plus *concave*. En effet, il en est ainsi de la fonction $T \longmapsto T(u) \wedge T$, et cette propriété subsiste par intégration sur u.

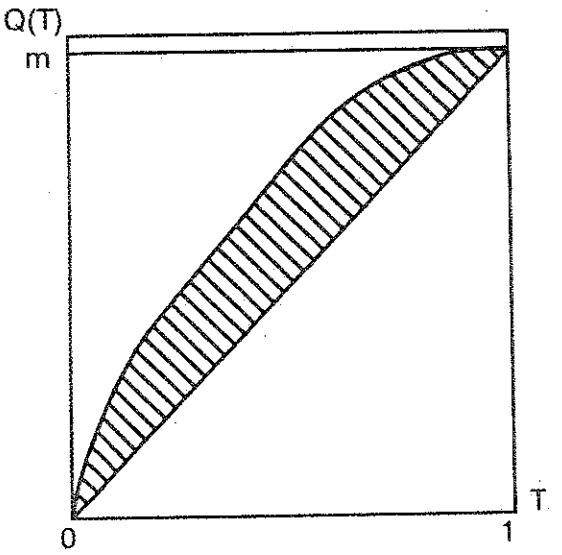

Figure 4.7 : Forme de la courbe $Q(T)$. Le domaine hachuré a pour surface la moitié de la sélectivité de la loi.

Nous verrons dans le paragraphe consacré à la dualité que la fonction $Q(T)$ caractérise la loi F. Par exemple, si F admet une proportion de valeurs nulles égale à q, alors $Q(T) = m$ sur $[1-q, 1]$. En effet, si $T \geq 1-q$, on a

$$Q(T) = \int_0^{+\infty} T(u) \wedge T\, du = \int_0^{+\infty} T(u)\, du = m$$

Signalons toutefois que certaines caractéristiques de la loi F ne se lisent pas aussi facilement sur $Q(T)$ (par exemple la variance de F).

4.4 La Sélectivité

Etant concave, la courbe $Q(T)$ est située au dessus de la droite $Q = mT$. Intéressons nous au domaine délimité par ces deux courbes.

$$\int_0^1 \left[Q(T) - mT\right] dT = \int_0^1 \int_0^{+\infty} T(u) \wedge T \, du \, dT - \frac{m}{2}$$

On intervertit les intégrales, et compte tenu de

$$\int_0^1 T(u) \wedge T \, dT = \int_0^{T(u)} T \, dT + \int_{T(u)}^1 T(u) \, dT = T(u) - \frac{T^2(u)}{2}$$

et de

$$m = \int_0^{+\infty} T(u) \, du$$

on obtient:

$$\int_0^1 \left[Q(T) - mT\right] dT = \frac{1}{2} \int_0^{+\infty} T(u) \left[1 - T(u)\right] du$$

On remarque en passant que cette quantité est nulle si et seulement si la loi est concentrée en un seul point.

Poursuivons le développement:

$$\int_0^1 \left[Q(T) - mT\right] dT = \frac{1}{2} \int_0^{+\infty} T(u) \int_0^u dF(z) \, du$$

ce qui donne après permutation des intégrales

$$\int_0^1 \left[Q(T) - mT\right] dT = \frac{1}{2} \int_0^{+\infty} \int_z^{+\infty} T(u) \, du \, dF(z) = \frac{1}{2} \int_0^{+\infty} B(z) \, dF(z)$$

Remplaçons maintenant $B(z)$ par l'une de ses expressions:

$$\int_0^1 \left[Q(T) - mT\right] dT = \frac{1}{2} \int_0^{+\infty} \int_z^{+\infty} (u - z) \, dF(u) \, dF(z)$$

Comme $u \geq z$, on peut remplacer $u - z$ par $|u - z|$, ce qui permet de faire jouer un rôle symétrique à u et z:

$$\int_0^1 \Big[Q(T) - mT\Big]\, dT = \frac{1}{4} \int_0^{+\infty} \int_0^{+\infty} |u - z|\, dF(u)\, dF(z)$$

et l'on reconnaît, à un facteur multiplicatif près, la valeur moyenne de $|X - Y|$ lorsque les deux variables X et Y suivent la même loi F indépendamment l'une de l'autre. On est ainsi conduit à poser la définition suivante:

Définition 4.5 *On appelle Sélectivité de la loi F la quantité*

$$S = \frac{1}{2} E\Big\{|X - Y|\Big\}$$

où X et Y sont deux variables indépendantes de même loi F.

Une telle définition doit être rapprochée de la formule suivante qui donne la variance de la loi F:

$$\sigma^2 = \frac{1}{2} E\Big\{(X - Y)^2\Big\}$$

Ainsi, tout comme la variance, la sélectivité mesure la dispersion de la loi. Notons toutefois que la sélectivité présente un certain nombre d'avantages sur la variance. Tout d'abord, la sélectivité existe dès que la moyenne existe ($S \leq m$), ce qui n'est pas le cas de la variance. Par ailleurs, du point de vue de l'inférence statistique, le calcul d'une sélectivité est nettement moins sensible à l'influence des fortes valeurs. On peut montrer (cf. exercices) que la sélectivité et la variance vérifient l'inégalité

$$S\sqrt{3} \leq \sigma$$

l'égalité n'ayant lieu que pour une loi F uniforme sur un intervalle.

Pour conclure ce paragraphe, nous rassemblons ci-dessous les quelques formules obtenues de la sélectivité:

Théorème 4.1 *La sélectivité de la loi F peut être calculée à l'aide de l'une des formules suivantes:*

$$S = 2 \int_0^{+\infty} \Big[Q(T) - mT\Big]\, dT$$

$$S = \int_0^{+\infty} T(u)\left[1 - T(u)\right] du$$

$$S = \frac{1}{2}\int_0^{+\infty}\int_0^{+\infty} |u - z|\, dF(u)\, dF(z)$$

4.5 La Dualité entre $B(z)$ et $Q(T)$

Soit T un tonnage, z une teneur de coupure choisie indépendamment, et $T(z)$ le tonnage correspondant. On peut avoir $T \geq T(z)$ ou $T \leq T(z)$.

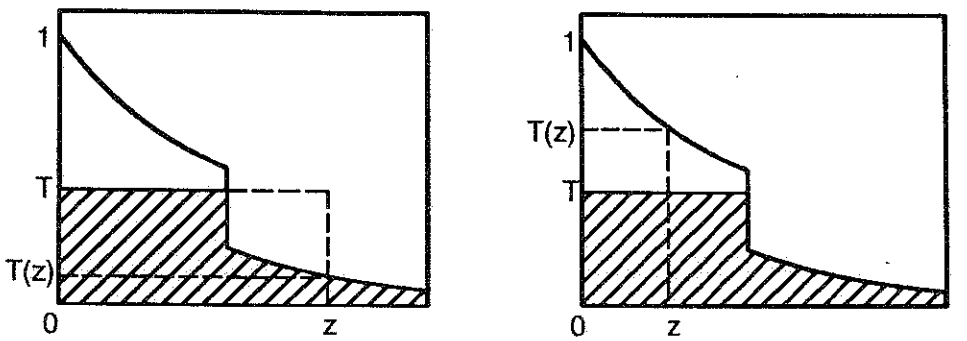

Figure 4.8 : Mise en évidence des formules de dualité entre $B(z)$ et $Q(T)$.

Dans les deux cas, on constate que l'on a

$$Q(T) \leq B(z) + zT$$

ou bien de façon équivalente

$$B(z) \geq Q(T) - zT$$

L'égalité est atteinte pour $T = T(z)$ et plus généralement lorsque l'on a $T(z+) \leq T \leq T(z)$. Nous avons ainsi établi:

Théorème 4.2 *Les fonctions $B(z)$ et $Q(T)$ sont liées par les deux formules de dualité*

$$B(z) = \sup_T \Big\{ Q(T) - zT \Big\}$$

$$Q(T) = \inf_z \Big\{ B(z) + zT \Big\}$$

Une conséquence immédiate est que la fonction $Q(T)$, tout comme la fonction $B(z)$, caractérise la loi F.

Exercices

1) Quelles sont les courbes de sélectivité de la loi δ_m (loi associée à une variable aléatoire constante égale à m)?

2) Même question pour la loi de Bernoulli $q\,\delta_0 + p\,\delta_1$. Que vaut la sélectivité d'une telle loi?

3) Quelle est la loi qui vérifie

$$\frac{Q(T)}{T} = a + b \ln T$$

4) On suppose que la loi F admet une densité positive f.

a) Etablir directement les propriétés de monotonie et de convexité des fonctions $Q(T)$ et de $B(z)$.

b) Quelle est l'équation de la tangente à la courbe $Q(T)$ passant par le point $(T(z), Q(z))$?

c) En déduire une nouvelle interprétation des formules de dualité.

5) Soit Z une variable positive de loi F.

a) Montrer que la sélectivité de Z vaut deux fois la covariance de Z et de son rang:

$$S = 2\,Cov\{Z, F(Z)\}$$

b) En déduire, par application de l'inégalité de Cauchy- Schwarz, que l'on a l'inégalité $S \leq \sigma\sqrt{3}$. Pour quelles lois l'égalité est-elle réalisée?

6) Soit $Z = \phi(Y)$ une variable positive qui est une fonction strictement croissante d'une variable Gaussienne Y.

a) Montrer que la sélectivité de Z peut s'écrire

$$S = \int_{-\infty}^{+\infty} \phi(y)\,[2G(y) - 1]\,g(y)\,dy$$

où g et G désignent respectivement la densité et la fonction de répartition de la loi de Gauss. (conseil: utiliser le résultat a) de l'exercice précédent).

b) Montrer que G admet le développement en polynômes d'Hermite

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

c) Notant ϕ_n les coefficients du développement en polynômes d'Hermite de ϕ, établir que la sélectivité de Z vaut

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

d) En déduire que la sélectivité de la loi lognormale

$$Z = m \exp\Big\{\sigma\,Y - \frac{\sigma^2}{2}\Big\}$$

a pour valeur

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

7) Soit Z une variable aléatoire positive, et soit $\Phi(u) = E\{e^{iuZ}\}$ sa transformée de Fourier.

a) Partant de la formule

$$|h| = \frac{2}{\pi} \int_0^{+\infty} \frac{1 - \cos(u\,h)}{u^2}\,du$$

montrer que la sélectivité de Z vaut

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

b) Appliquer cette formule pour évaluer la sélectivité de la loi de densité

$$f(z) = \frac{2b}{\sqrt{\pi}} e^{-bz} \sqrt{bz} \qquad (b > 0)$$

Chapitre 5

Comparaison de deux lois du point de vue de la sélectivité

Soient F_1 et F_2 deux lois de probabilité. On désigne par $Q_1(T)$ et $Q_2(T)$ les fonctions métal/tonnage, ainsi que par $B_1(z)$ et $B_2(z)$ les bénéfices conventionnels qui leur sont associées.

5.1 Quelques critères de comparaison

Lemme 5.1 *$Q_1(T) \leq Q_2(T)$ pour tout T si et seulement si $B_1(z) \leq B_2(z)$ pour tout z.*

Démonstration: En effet, supposons $Q_1(T) \leq Q_2(T)$ pour tout T. Alors

$$Q_1(T) - zT \leq Q_2(T) - zT$$

Passons à la borne supérieure sur les T. La formule de dualité qui permet de passer du Métal/Tonnage au Bénéfice Conventionnel implique

$$B_1(z) = \sup_T \{Q_1(T) - zT\} \leq \sup_T \{Q_2(T) - zT\} = B_2(z)$$

En ce qui concerne l'implication dans l'autre sens, il suffit d'appliquer l'autre formule de dualité. ■

Définition 5.1 *On dira que la loi F_1 est moins sélective que la loi F_2 si l'on a*

$$\int_0^{+\infty} \phi(z)dF_1(z) \leq \int_0^{+\infty} \phi(z)dF_2(z)$$

pour toute fonction convexe ϕ.

Pour comprendre le sens d'une telle définition, particularisons le choix de la fonction ϕ:

- $\phi(z) = z$ entraine $m_1 \leq m_2$
- $\phi(z) = -z$ entraine $-m_1 \leq -m_2$. En combinant ces deux dernières inégalités, il ressort que $m_1 = m_2$, c'est à dire que les deux lois admettent la même moyenne. Pour la suite, on notera m cette valeur commune.
- $\phi(z) = z^2 - m^2$ entraine $\sigma_1^2 \leq \sigma_2^2$. Ainsi la loi F_1 apparaît-elle moins dispersée que la loi F_2.

Mais bien d'autres inégalités sont possibles. Par exemple, on pourra prendre $\phi(z) = |z - m|^\alpha$ avec $\alpha \geq 1$. Il en résulte une inégalité sur les moments absolus d'ordre α.

Théorème 5.1 *Les trois propriétés suivantes sont équivalentes:*

1. *F_1 est moins sélective que F_2.*
2. *$B_1(z) \leq B_2(z)$ pour tout z et $m_1 = m_2$.*
3. *$Q_1(T) \leq Q_2(T)$ pour tout T et $m_1 = m_2$.*

Démonstration: Pour montrer que $1 \Longrightarrow 2$, il suffit d'appliquer la définition de la sélectivité en prenant $\phi(u) = (u - z)1_{u \geq z}$ comme fonction convexe.

Pour établir que $2 \Longrightarrow 1$, on se contentera de faire la démonstration dans le cas où ϕ est une fonction deux fois continûment dérivable. Partons de la formule de Taylor avec reste intégral

$$\phi(z) = \phi(0) + z\phi'(0) + \int_0^z (z - u)\phi''(u)du$$

Par intégration par rapport à F_1, on obtient:

$$\int_0^{+\infty} \phi(z)dF_1(z) = \phi(0) + m_1\phi'(0) + \int_0^{+\infty}\int_0^z (z-u)\phi''(u)dudF_1(z)$$

ce qui donne, après permutation des intégrales

$$\int_0^{+\infty} \phi(z)dF_1(z) = \phi(0) + m_1\phi'(0) + \int_0^{+\infty} \phi''(u)\int_u^{+\infty}(z-u)dF_1(z)du$$

c'est-à dire

$$\int_0^{+\infty} \phi(z)dF_1(z) = \phi(0) + m_1\phi'(0) + \int_0^{+\infty} \phi''(u)B_1(u)du$$

De même

$$\int_0^{+\infty} \phi(z)dF_2(z) = \phi(0) + m_2\phi'(0) + \int_0^{+\infty} \phi''(u)B_2(u)du$$

Utilisons alors le fait que $m_1 = m_2$ et que $B_1(u) \leq B_2(u)$. Comme $\phi''(u) \geq 0$ puisque ϕ est convexe, l'inégalité

$$\int_0^{+\infty} \phi(z)dF_1(z) \leq \int_0^{+\infty} \phi(z)dF_2(z)$$

s'en déduit immédiatement.

On a déjà établi $2 \Longleftrightarrow 3$ (cf. Lemme 5.1). ■

5.2 La relation de Cartier

Définition 5.2 *Soit X et Y deux variables aléatoires. On dit que X est en relation de Cartier avec Y si $E\{X|Y\} = Y$*

Théorème 5.2 *Si X est en relation de Cartier avec Y, alors la loi de Y est moins sélective que celle de X.*

Démonstration: En effet, si ϕ désigne une fonction convexe, l'inégalité de Jensen (cf. Chapitre 2) donne

$$\phi[E\{X|Y\}] \leq E\{\phi(X)|Y\}$$

Appliquons maintenant la relation de Cartier. On obtient:

$$\phi(Y) \leq E\{\phi(X)|Y\}$$

puis passons à l'espérance

$$E\{\phi(Y)\} \leq E\{E\{\phi(X)|Y\}\} = E\{\phi(X)\}$$

ce qui dit exactement que la loi de Y est moins sélective que la loi de X. ■

Cet énoncé n'aurait qu'une valeur d'exemple s'il n'était assorti de la réciproque que voici:

Théorème 5.3 (Cartier) *Si la loi F_1 est moins sélective que la loi F_2, il existe une loi bivariable F de lois marginales F_1 et F_2, et telle que si le couple de variables aléatoires X_1, X_2 suit la loi F, alors $E\{X_2|X_1\} = X_1$.*

Ce théorème est admis. En conclusion de ce chapitre, on dispose de trois critères pour comparer deux lois du point de vue de la sélectivité:

- On peut appliquer la définition.
- On peut tester l'inégalité entre les courbes $Q(T)$, ou bien de façon équivalente entre les courbes $B(z)$.
- On peut exhiber une relation de Cartier entre les deux lois.

Ces trois critères sont également utilisés. Nous en verrons des exemples d'application lors de l'étude des effets de support et d'information.

Exercices

1) Si $E\{X \mid Y\} = Y$, que vaut $Cov\{X, Y\}$?

2) Soit U, X deux variables aléatoires positives indépendantes. On suppose $E\{U\} = 1$. Montrer que XU est plus sélective que X.

3) Montrer que la loi d'un estimateur sans biais conditionnel est moins sélective que la loi a priori.

4) Montrer que la loi d'un estimateur par K.O. à coefficients positifs est moins sélective que la loi a priori.

5) Soit N_X une variable de Poisson de paramètre aléatoire X. Montrer que la loi de N_X est plus sélective que celle de X.

Chapitre 6

L'effet de support

Soit $Z = \big(Z(x), x \in I\!R^d\big)$ une fonction aléatoire stationnaire de loi ponctuelle F et de fonction de covariance C. Etant donné $v \subset I\!R^d$, on cherche à prévoir la loi F_v de

$$Z(v) = \frac{1}{|v|} \int_v Z(x)\, dx$$

6.1 Quelques observations expérimentales

Commençons par une approche heuristique du problème à l'aide de deux exemples.

Pour le premier exemple, on considère un processus de Poisson de paramè-

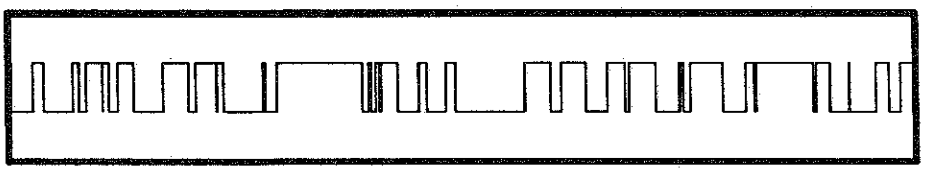

Figure 6.1 : Découpage de $I\!R$ en segments exponentiels indépendants alternativement mis à 1 et à 0

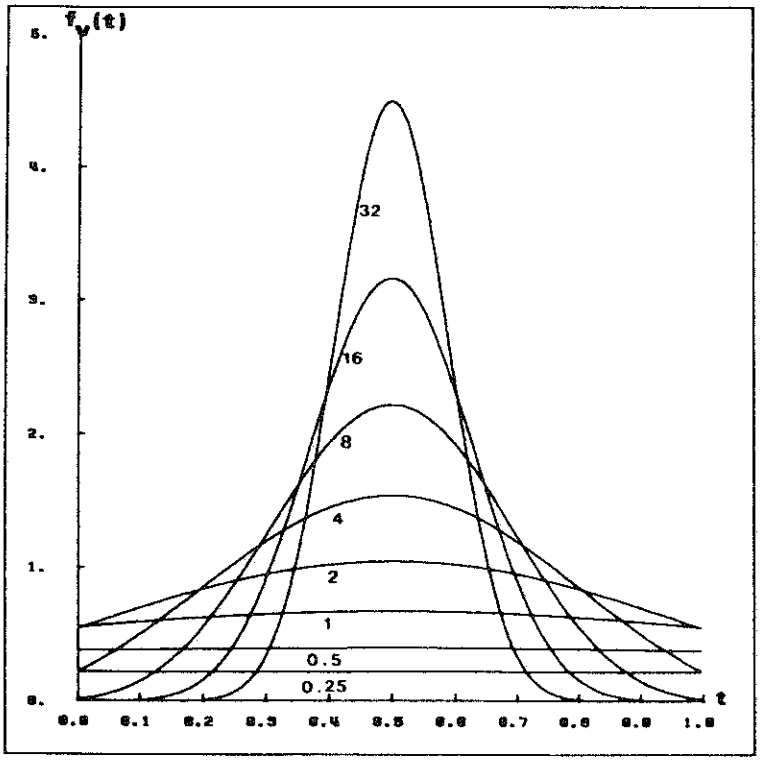

Figure 6.2 : Densité de $Z(v)$ sur $]0,1[$ pour différentes tailles de v

tre θ qui découpe $I\!R$ en intervalles exponentiels indépendants de longueur moyenne θ^{-1}. Les intervalles sont alternativement mis à 0 et 1.

Posons $v = [0,h]$, où $h > 0$. $Z(v)$ prend des valeurs comprises entre 0 et 1. $Z(v)$ vaut 0 ou 1 si v est totalement contenu dans l'un des intervalles exponentiels, ce qui se produit avec la probabilité

$$P\{Z(v) = 0\} = P\{Z(v) = 1\} = \frac{1}{2}\, e^{-\theta h}$$

Sur $]0,1[$, on peut montrer que $Z(v)$ admet une densité f_v qui est représentée Figure 6.2 pour différentes tailles de support. Pour v très petit, cette densité est pratiquement constante. Lorsque v croît, la densité s'incurve pour se transformer en une courbe d'allure Gaussienne centrée en $\frac{1}{2}$.

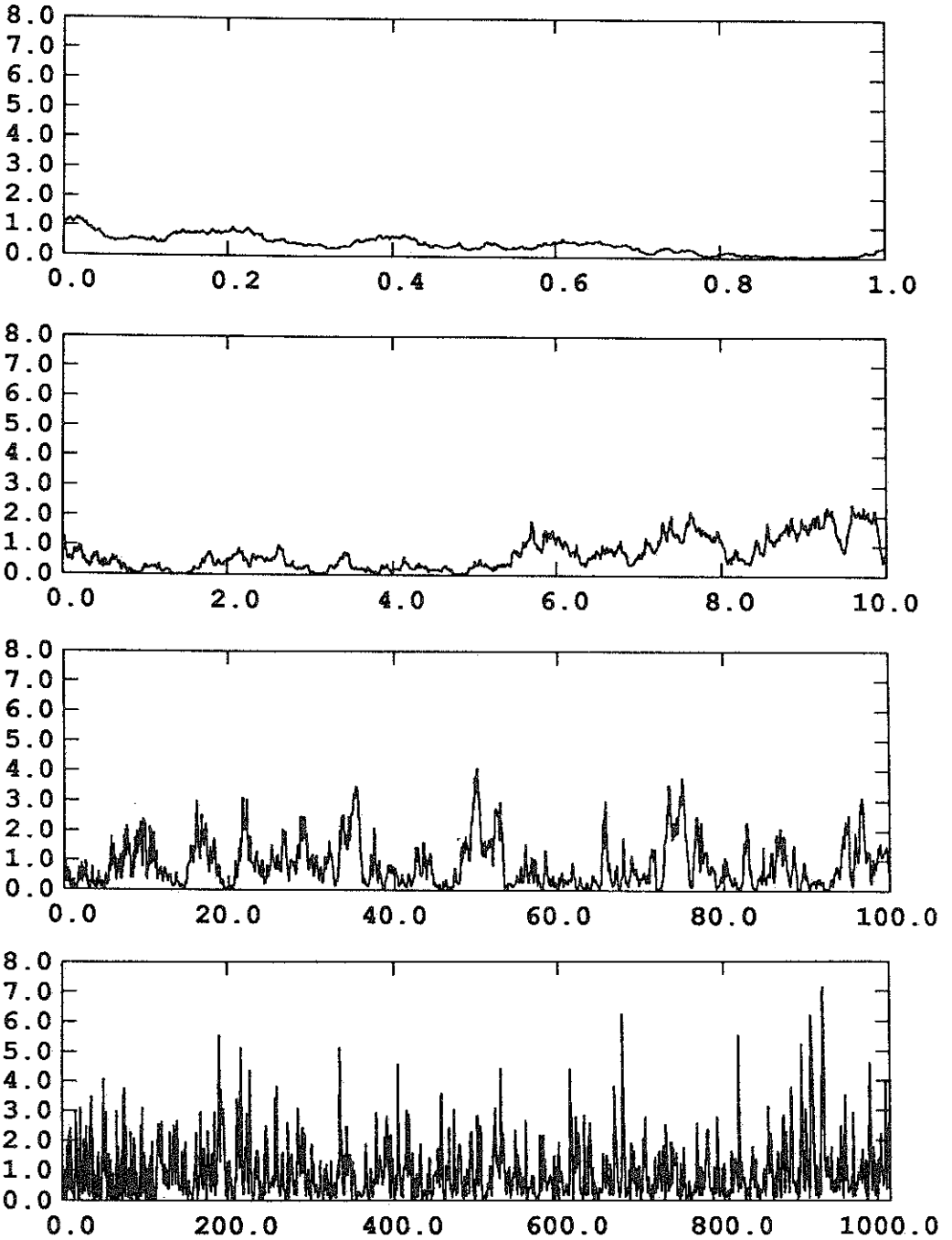

Figure 6.3 : Observation d'un processus de diffusion exponentielle à 4 échelles différentes.

Le deuxième exemple considéré est un processus de diffusion exponentielle. Il s'agit d'un processus de Markov de moyenne 1 qui satisfait les conditions infinitésimales suivantes:

i) Les valeurs prises par deux points voisins ne sont pas significativement différentes:

$$P\Big\{ |Z(x+dx) - Z(x)| > \epsilon \Big| Z(x) = z \Big\} = o(dx)$$

ii) En moyenne, tout se passe comme s'il existait une force d'attraction vers la constante 1:

$$E\Big\{ Z(x+dx) - Z(x) \Big| Z(x) = z \Big\} = (1-z)dx + o(dx)$$

iii) Plus $Z(x)$ est grand, plus grandes peuvent être les fluctuations entre les points x et $x + dx$:

$$Var\Big\{ Z(x+dx) - Z(x) \Big| Z(x) = z \Big\} = 2z\, dx + o(dx)$$

La conséquence de iii) est que le processus peut prendre des valeurs très grandes. Notons toutefois que ces grandes valeurs sont isolées en raison de ii).

La loi ponctuelle de Z est exponentielle. Au fur et à mesure que la taille du support v augmente, la densité de $Z(v)$ se symétrise pour tendre encore vers une loi d'allure Gaussienne, centrée cette fois en 1.

De ces deux exemples se dégagent un certain nombre de constatations expérimentales: tout d'abord, la loi F_v devient de plus en plus reserrée autour de sa moyenne lorsque la taille du support v augmente. Corrélativement, on observe de moins en moins de valeurs extrêmes. Aux très grandes tailles de support, on obtient des lois très concentrées autour de la moyenne. La loi devient symétrique et prend la forme d'une Gaussienne. Cela n'est pas sans rappeler les deux résultats bien connus des probabilités que sont *la Loi*

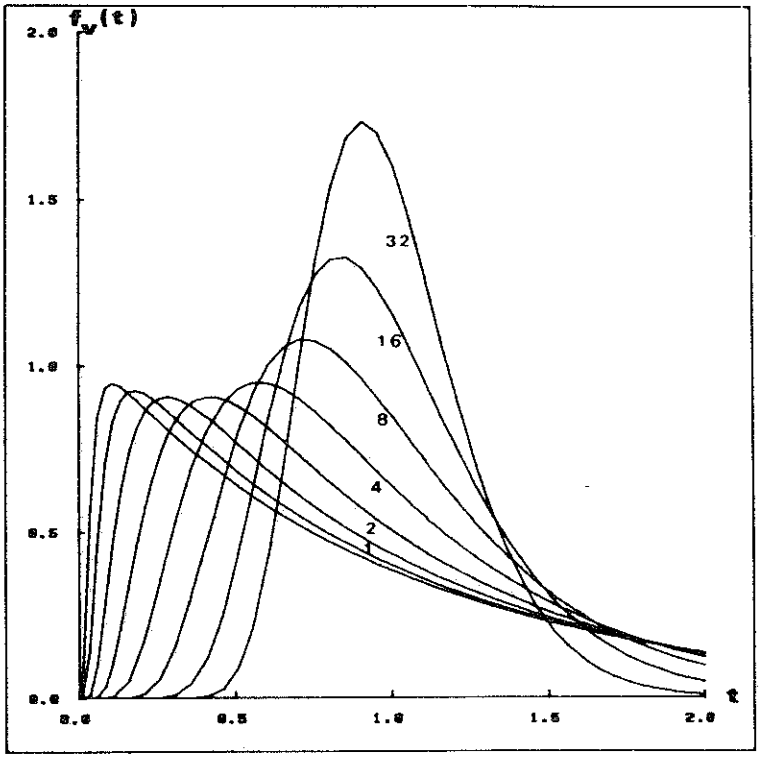

Figure 6.4 : Densité de $Z(v)$ pour différentes tailles de v.

Forte des Grands Nombres ainsi que *le Théorème de la Limite Centrale.* Ils s'énoncent usuellement de la façon suivante: si $(X_n), n \in I\!N$ est une suite de variables aléatoires mutuellement indépendantes et de même loi, de moyenne m et de variance σ^2, alors la variable

$$\frac{X_1 + \cdots + X_n}{n}$$

converge presque sûrement vers m lorque n tend vers l'infini. La loi de cette même variable, une fois centrée et normée

$$\frac{\frac{X_1 + \cdots + X_n}{n} - m}{\frac{\sigma}{\sqrt{n}}}$$

tend à devenir Gaussienne toujours lorsque n devient infini. Ces deux résultats ont une portée bien plus générale que les hypothèses peuvent laisser

croire.

6.2 Que connaît-on de $Z(v)$?

Revenons au problème de départ. On dispose d'un certain nombre de données expérimentales ponctuelles, et l'on suppose que ces données sont suffisantes pour que la loi ponctuelle F de Z (et en particulier sa moyenne m, et sa variance σ^2) ainsi que sa fonction de covariance C puissent être considérées comme connues. Que connaissons-nous de $Z(v)$?

Tout d'abord, les deux premiers moments de $Z(v)$ sont accessibles. Il en est ainsi de sa *moyenne*

$$E\left\{\frac{1}{|v|}\int_v Z(x)\,dx\right\} = \frac{1}{|v|}\int_v E\{Z(x)\}\,dx = m$$

ainsi que de sa *variance*:

$$\begin{aligned} Var\left\{\frac{1}{|v|}\int_v Z(x)\,dx\right\} &= Cov\left\{\frac{1}{|v|}\int_v Z(x)\,dx\,,\,\frac{1}{|v|}\int_v Z(y)\,dy\right\} \\ &= \frac{1}{|v|^2}\int_v\int_v Cov\{Z(x),Z(y)\}\,dx\,dy \\ &= \frac{1}{|v|^2}\int_v\int_v C(x-y)\,dx\,dy \end{aligned}$$

Mais il y a un peu plus. Si $\underline{x}$ désigne un point uniforme dans v, alors:

$$\begin{aligned} E\Big\{Z(\underline{x}) \mid Z(v)\Big\} &= \frac{1}{|v|}\int_v E\Big\{Z(x) \mid Z(v)\Big\}\,dx \\ &= E\left\{\frac{1}{|v|}\int_v Z(x)\,dx \;\Big|\; Z(v)\right\} \\ &= E\Big\{Z(v) \mid Z(v)\Big\} = Z(v) \end{aligned}$$

Il s'agit de la relation de Cartier, ce qui implique que la loi de $Z(v)$ est moins sélective que celle de $Z(\underline{x})$. Or la loi de $Z(\underline{x})$ n'est autre que F. En effet:

$$P\{Z(\underline{x}) < z\} = \frac{1}{|v|}\int_v P\{Z(x) < z\}\,dx = \frac{1}{|v|}\int_v F(z)\,dx = F(z)$$

Ainsi, la loi F_v est *moins sélective* que F. Ce phénomène de reserrement des lois en fonction de la taille de support a été clairement observé dans les exemples introductifs.

Voilà tout ce qui est connu de $Z(v)$ à partir des données ponctuelles disponibles. On est par conséquent loin d'avoir déterminé la loi F_v. Il demeure une assez large indétermination.

Pour lever l'indétermination, il faut introduire une information supplémentaire. Cela va se faire au moyen d'un *Modèle de Changement de Support.* Recourir à un tel modèle reviendra en gros à spécifier la loi bivariable entre $Z(\underline{x})$ et $Z(v)$ en cohérence avec les données disponibles. Bien entendu, de nombreux modèles peuvent être envisagés. A chaque modèle sera attaché une prédiction, et la qualité de la prédiction vaudra ce que vaut l'adéquation du modèle à la réalité.

6.3 Trois modèles de changement de support

Dans cette partie, nous allons passer en revue les trois modèles suivants: La Correction Affine, le Modèle Mosaïque, ainsi que le Modèle Gaussien Discret. Nous donnerons les avantages et les inconvénients de chacun de ces modèles afin de mieux pouvoir comparer leur efficacité.

6.3.1 La Correction Affine

Dans ce modèle, la forme des lois reste inchangée par changement de support. Pour tenir compte de la préservation de la moyenne et de la réduction de variance, on pose par hypothèse:

$$\frac{Z(\underline{x}) - m}{\sigma} \stackrel{\mathcal{L}}{=} \frac{Z(v) - m}{\sigma_v}$$

Il se trouve que la courbe métal/tonnage associée à $Z(v)$

$$Q_v(T) = \int_0^{+\infty} T_v(z) \wedge T\, dz$$

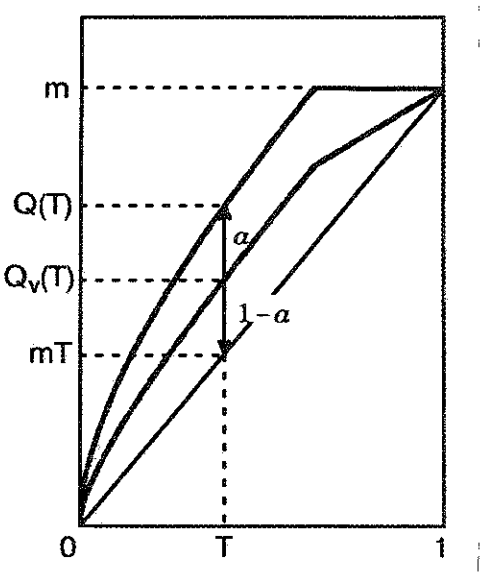

Figure 6.5 : Construction de la Correction Affine.

a une forme particulièrement simple. Si l'on introduit y tel que

$$\frac{y-m}{\sigma} = \frac{z-m}{\sigma_v}$$

alors $T_v(z) = T(y)$ de sorte que

$$\begin{aligned} Q_v(T) &= \frac{\sigma_v}{\sigma}\int_{m\left(1-\frac{\sigma}{\sigma_v}\right)}^{+\infty} T(y) \wedge T\,dy \\ &= \frac{\sigma_v}{\sigma}\int_{m\left(1-\frac{\sigma}{\sigma_v}\right)}^{0} T(y) \wedge T\,dy + \frac{\sigma_v}{\sigma}\int_{0}^{+\infty} T(y) \wedge T\,dy \end{aligned}$$

La deuxième intégrale vaut $Q(T)$. Pour le calcul de la première, il suffit de remarquer que $T(y) = P\{Z(\underline{x}) \geq y\} = 1$ pour $y \leq 0$, et finalement on obtient:

$$Q_v(T) = mT\left(1 - \frac{\sigma_v}{\sigma}\right) + \frac{\sigma_v}{\sigma}\,Q(T)$$

Ainsi, la courbe $Q_v(T)$ obtenue est une moyenne pondérée entre la courbe ponctuelle $Q(T)$ et la courbe mT qui correspond à un support infini.

Avantage:

Puisque $Q(T)$ est concave, on a $mT \leq Q(T)$, et par suite

$$Q_v(T) \leq Q(T)\left(1 - \frac{\sigma_v}{\sigma} + \frac{\sigma_v}{\sigma}\right) = Q(T)$$

Il s'agit donc bien d'un modèle cohérent de changement de support.

Inconvénients:

i) Si $T < 1$, $Q_v(T) < m$. La correction affine gomme tout effet 0.

ii) De par sa construction même, la correction affine n'est pas compatible avec le théorème de la limite centrale aux grandes tailles de support.

6.3.2 Le changement de support Mosaïque

La construction du modèle mosaïque de changement de support repose sur l'hypothèse que le bénéfice conventionnel de $Z(v)$ est une moyenne pondérée horizontale entre deux autres courbes de bénéfice conventionnel, à savoir celle associée à support ponctuel, et celle associée à un support infini. Plus précisément, étant donné le paramètre α, $0 < \alpha < 1$, on pose (cf. Figure 6.6)

$$z_v = (1-\alpha)z + \alpha\big[m - B(z)\big]$$

ainsi que

$$B_v(z_v) = B(z)$$

Par passage aux fonctions réciproques, on vérifie que B_v est décroissante et convexe. B_v a donc bien le statut d'une courbe de bénéfice conventionnel. Par ailleurs $B_v(0) = B(0) = m$.

Pour avoir une interprétation physique du paramètre α, partons de l'égalité obtenue par le changement de variable $z_v \longmapsto z$:

$$\int_0^{+\infty} B_v(z_v)\,dz_v = \int_0^{+\infty} B(z)\left[1 - \alpha + \alpha T(z)\right] dz$$

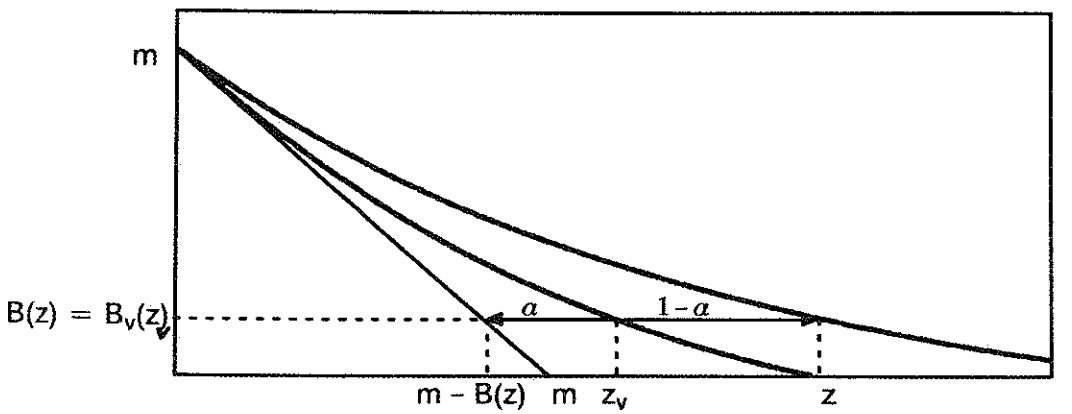

Figure 6.6 : Construction du Modèle Mosaïque

L'intégrale de gauche vaut (cf. Chapitre 4)

$$\int_0^{+\infty} B_v(z_v)\, dz_v = \frac{1}{2}\left(\sigma_v^2 + m^2\right)$$

De son coté, compte tenu de

$$\int_0^{+\infty} B(z)\, T(z)\, dz = \int_0^{m} B\, dB = \frac{m^2}{2}$$

l'intégrale de droite vaut

$$\int_0^{+\infty} B(z)\left[1 - \alpha + \alpha T(z)\right] dz = (1-\alpha)\,\frac{1}{2}\left(\sigma^2 + m^2\right) + \alpha\,\frac{m^2}{2}$$

Par conséquent

$$\frac{1}{2}\left(\sigma_v^2 + m^2\right) = (1-\alpha)\,\frac{1}{2}\left(\sigma^2 + m^2\right) + \alpha\,\frac{m^2}{2}$$

ce qui montre que α est la réduction de variance

$$\alpha = 1 - \frac{\sigma_v^2}{\sigma^2}$$

Pour évaluer la courbe $Q_v(T)$ associée à $Z(v)$, le plus simple est de recourir à la formule de dualité

$$Q_v(T) = \inf_{z_v}[B_v(z_v) + z_v T]$$

$$B_v(z_v) + z_v T = B(z) + (1-\alpha)zT + \alpha\Big[m - B(z)\Big]\,T$$

ce qui se récrit

$$B_v(z_v) + z_v T = (1-\alpha T)\left[B(z) + z\,\frac{(1-\alpha)T}{1-\alpha T}\right] + \alpha m T$$

Passons maintenant à la borne inférieure sur z_v. Comme z et z_v sont en bijection croissante, on peut passer à la borne inférieure sur z au second membre, et l'on trouve finalement

$$Q_v(T) = (1-\alpha T)\,Q\left[\frac{(1-\alpha)T}{1-\alpha T}\right] + \alpha m T$$

Avantages:

i) Le modèle mosaïque est un modèle cohérent de changement de support. En effet, écrivant $m = Q(1)$, la concavité de Q_v implique

$$Q_v(T) \leq Q\left[(1-\alpha T)\,\frac{(1-\alpha)T}{1-\alpha T} + \alpha T\,1\right] = Q(T)$$

ii) Le modèle mosaïque peut prendre un effet 0. Si $Q_v(T) = m$, alors nécessairement

$$Q\left[\frac{(1-\alpha)T}{1-\alpha T}\right] = m$$

et donc

$$\frac{(1-\alpha)T}{1-\alpha T} \geq 1 - q$$

où q désigne la proportion de valeurs nulles de la loi ponctuelle. Cela donne:

$$T \geq \frac{1-q}{1-\alpha q}$$

et par conséquent

$$q_v = 1 - \frac{1-q}{1-\alpha q} = \frac{q(1-\alpha)}{1-\alpha q}$$

Inconvénient:

Le modèle mosaïque n'est pas compatible avec le théorème de la limite centrale aux grandes tailles de support.

6.3.3 Le Modèle Gaussien Discret

Soit $\underline{x}$ un point uniforme dans v. Si $Z(\underline{x})$ admet une densité, alors il existe une fonction bijective croissante ϕ ainsi qu'une variable aléatoire Gaussienne centrée réduite Y telles que

$$Z(\underline{x}) = \phi(Y)$$

ϕ est appelée *fonction d'anamorphose.* Elle est donnée par la formule

$$\phi(y) = F^{-1}\Big[G(y)\Big]$$

De la même façon, Si $Z(v)$ admet une densité, il existe une fonction d'anamorphose ϕ_v et une variable Gaussienne centrée réduite Y_v telles que

$$Z(v) = \phi_v(Y_v)$$

Supposons que ces conditions soient satisfaites. Le modèle Gaussien Discret repose sur l'hypothèse que le couple (Y, Y_v) suit une loi bigaussienne.

Pour assurer une complète détermination de ce modèle, il importe de savoir évaluer la fonction d'anamorphose ϕ_v de $Z(v)$, ainsi que le coefficient de corrélation r entre Y et Y_v.

La relation de Cartier

$$E\{Z(\underline{x}) \mid Z(v)\} = Z(v)$$

conduit à

$$E\{\phi(Y) \mid Y_v\} = \phi_v(Y_v)$$

ce qui permet d'exprimer ϕ_v à partir de ϕ. Considérant le développement de ϕ en polynomes d'Hermite

$$\phi(y) = \sum_{n=0}^{+\infty} \frac{\phi_n}{n!} H_n(y)$$

on trouve (cf. Chapitre 3)

$$\phi_v(y) = \sum_{n=0}^{+\infty} \frac{\phi_n r^n}{n!} H_n(y)$$

La formule

$$Var\{Z(v)\} = Var\Big\{\phi_v(Y_v)\Big\} = \sum_{n=1}^{+\infty} \frac{\phi_n^2 r^{2n}}{n!}$$

permet d'identifier r qui doit être positif (cf. exercices).

Pour évaluer la courbe métal/tonnage associée à $Z(v)$, le plus simple sans doute est d'écrire $Q_v(T) = Q_v(z)$ avec $T = T_v(z)$. Alors:

$$\begin{aligned} Q_v(T) &= E\Big\{Z(v)\, 1_{Z(v) \geq z}\Big\} \\ &= E\Big\{\phi_v(Y_v)\, 1_{Y_v \geq G^{-1}(1-T)}\Big\} \\ &= \int_{G^{-1}(1-T)}^{+\infty} \phi_v(y)\, g(y)\, dy \end{aligned}$$

Avantages:

i) Le modèle Gaussien discret est un modèle cohérent de changement de support. Sa construction repose en effet sur la relation de Cartier.

ii) Le modèle Gaussien discret est compatible avec le théorème de la limite centrale. Supposons en effet que les dimensions de v deviennent très grandes. Alors la variance de $Z(v)$ tend vers 0, et par suite r tend également vers 0. Développons ϕ_v au premier ordre en r:

$$\phi_v(y) = \Big[\phi_v(y)\Big]_{r=0} + r\left[\frac{\partial \phi_v(y)}{\partial r}\right]_{r=0} + o(r)$$

Un simple calcul montre que

$$\Big[\phi_v(y)\Big]_{r=0} = m$$

$$\left[\frac{\partial \phi_v(y)}{\partial r}\right]_{r=0} = y \int_{-\infty}^{+\infty} \phi'(u)\, g(u)\, du$$

Par conséquent $\phi_v(y)$ devient une fonction affine en y lorsque r tend vers 0. De cela découle que $Z(v)$ devient Gaussienne.

Inconvénient:

Le modèle Gaussien discret s'accommode mal d'un effet 0.

Exercices

1) Soit Z la fonction aléatoire définie sur $I\!R$ par

$$Z(x) = \cos(x + 2\pi U)$$

où U est une variable uniforme sur $(0,1)$. Pour tout $h > 0$, on pose

$$Z_h = \frac{1}{h}\int_0^h Z(x)\, dx$$

Que vaut la variance de Z_h en fonction de h?

2) Dans le modèle Gaussien discret, donner l'expression de ϕ_v en fonction de ϕ. En déduire que le coefficient r qui caractérise le changement de support est positif.

3) Montrer que le modèle Gaussien discret assure la permanence de la loi lognormale.

4) Soit $Q(T)$ une courbe quantité de métal/tonnage associée à une loi F de moyenne m. Etant donné $0 < \alpha < 1$, on construit la courbe $Q_\alpha(T)$ comme moyenne pondérée horizontale dans la proportion α entre la courbe $Q(T)$ et la courbe mT:

$$Q_\alpha(T_\alpha) = Q(T) \quad \text{avec} \quad T_\alpha = (1-\alpha)T + \alpha\frac{Q(T)}{m}$$

a) Montrer que la courbe $Q_\alpha(T)$ a toutes les caractéristiques d'une courbe quantité de métal/tonnage, et que le passage de $Q(T)$ à $Q_\alpha(T)$ constitue un modèle cohérent de changement de support. Si la loi F a une proportion q de valeurs nulles, qu'en est-il de la loi F_α associée à Q_α?

b) Quelle est la pente maximale que peut prendre la courbe $Q_\alpha(T)$? Quelle est la conséquence qui en résulte pour la loi F_α?

c) En utilisant les formules de dualité, donner l'expression du bénéfice conventionnel attaché à F_α.

d) Montrer que α n'est autre que la réduction de sélectivité entre les lois F et F_α.

Chapitre 7

L'effet d'information

En de nombreux domaines, il arrive fréquemment qu'une décision doive être prise dans un contexte d'incertitude. Faute de pouvoir disposer de valeurs exactes non observables, la décision s'appuie sur des valeurs incertaines ou entâchées d'erreur.

Tel est le cas par exemple dans l'industrie minière. La décision d'orienter le contenu d'un bloc exploité vers la laverie ou bien vers le rebut stérile s'effectue à partir de la teneur estimée du bloc et non pas à partir de sa teneur vraie.

L'objet de ce chapitre est d'étudier les conséquences résultant de ce contexte d'incertitude. Compte tenu du fait que l'effet d'information peut surgir dans les situations les plus disparates, une présentation générale n'est guère envisageable. Dans ce qui suit, nous nous bornerons à regarder en détail l'exemple minier évoqué ci-dessus. Nous verrons que l'effet d'information se traduit en pratique par une perte de sélectivité. Il agit donc exactement dans le même sens que l'effet de support. Un exemple numérique donné en exercice montre l'importance relative de ces deux effets. Il apparaît qu'aucun des deux ne peut être négligé.

7.1 L'effet d'information

Replaçons nous dans le contexte minier introduit ci-dessus. Si l'on décide d'exploiter les blocs de teneur *estimée* supérieure à z, le tonnage dégagé est

$$T_v^*(z) = E\Big\{1_{Z^*(v)} \geq z\Big\}$$

Dans ces blocs estimés riches, la quantité de métal *effectivement* récupérée a pour valeur

$$Q_v^*(z) = E\Big\{Z(v)\, 1_{Z^*(v)} \geq z\Big\}$$

Commençons par évaluer la courbe $Q_v^*(T)$ résultant d'une telle sélection. A cette fin, il est commode d'introduire la régression entre la teneur vraie du bloc et sa teneur estimée

$$h(z) = E\Big\{Z(v) \;\Big|\; Z^*(v) = z\Big\}$$

Supposant cette fonction croissante, la quantité de métal se récrit

$$\begin{aligned} Q_v^*(z) &= E\Big\{E\Big\{Z(v)\, 1_{Z^*(v)} \geq z \;\Big|\; Z^*(v)\Big\}\Big\} \\ &= E\Big\{h[Z^*(v)]\; 1_{Z^*(v)} \geq z\Big\} \\ &= E\Big\{h[Z^*(v)]\; 1_{h[Z^*(v)]} \geq h(z)\Big\} \end{aligned}$$

De son coté, le tonnage vaut quant à lui

$$T_v^*(z) = E\Big\{1_{h[Z^*(v)]} \geq h(z)\Big\}$$

Ainsi, à condition de remplacer la teneur z par $h(z)$, les courbes de tonnage et de quantité de métal obtenues sont identiques à celles de $h[Z^*(v)]$. Cela suggère que $Q_v^*(T)$ n'est autre que la courbe métal/tonnage associé à $h[Z^*(v)]$. Il en est effectivement ainsi, comme nous allons le voir maintenant.

Soient $\bar{T}(z)$ et $\bar{Q}(z)$ les fonctions de récupération associées à $h[Z^*(v)]$. Nous venons d'établir

$$T_v^*(z) = \bar{T}[h(z)]$$

$$Q_v^*(z) = \bar{Q}[h(z)]$$

Or nous savons que

$$\bar{Q}[h(z)] = \int_0^{+\infty} \bar{T}(u) \wedge \bar{T}[h(z)]\, du$$

Par conséquent

$$Q_v^*(z) = \int_0^{+\infty} \bar{T}(u) \wedge T_v^*(z)\, du$$

Tout comme pour la définition des courbes métal/tonnage, cette fonction se généralise en une fonction $Q_v^*(T)$ définie pour tout tonnage T par

$$Q_v^*(T) = \int_0^{+\infty} \bar{T}(u) \wedge T\, du$$

Ainsi donc

$$Q_v^*(T) = \bar{Q}(T)$$

ce qui est le résultat anoncé. Sous une forme imagée, on peut dire avec G. Matheron que "l'on exploite non pas des teneurs, mais des espérances conditionnelles de teneur".

Remarque: dans le cas où l'estimateur est *sans Biais Conditionnel*, i.e.

$$E\Big\{Z(v) \,\Big|\, Z^*(v)\Big\} = Z^*(v)$$

alors $h(z) = z$ et les courbes de sélectivité effectivement récupérées sont celles de l'estimateur.

7.2 Comparaison entre les effets de support et d'information

Il importe maintenant de remarquer que la loi de $h[Z^*(v)]$ est moins sélective que celle de $Z(v)$. En effet, ces deux variables sont en relation de Cartier:

$$\begin{aligned} E\Big\{Z(v) \,\Big|\, h[Z^*(v)]\Big\} &= E\Big\{Z(v) \,\Big|\, Z^*(v) \,\Big|\, h[Z^*(v)]\Big\} \\ &= E\Big\{h[Z^*(v)] \,\Big|\, h[Z^*(v)]\Big\} = h[Z^*(v)] \end{aligned}$$

De cette relation de Cartier, il découle immédiatemment

$$Q_v^*(T) \leq Q_v(T)$$

En rassemblant ce résultat et ceux qui ont été obtenus lors de l'étude du changement de support, nous arrivons à la conclusion suivante:

Théorème 7.1 *L'effet d'information, tout comme l'effet de support, agit dans le sens d'une perte de sélectivité. On a toujours*

$$Q(T) \geq Q_v(T) \geq Q_v^*(T)$$

la première inégalité étant la conséquence de l'effet de support, la deuxième de l'effet d'information.

Pour évaluer $Q_v^*(T)$, il faut connaître la loi de $h[Z^*(v)] = E\Big\{Z(v)|Z^*(v)\Big\}$, ce qui revient à connaître la loi bivariable entre $Z(v)$ et $Z^*(v)$. Cette fois encore, cette loi n'est pas accessible à partir des données expérimentales disponibles, et il nous faudra recourir à un modèle. Dans ce qui suit, nous allons reprendre le modèle Gaussien discret et voir comment l'effet d'information peut être pris en compte dans ce modèle.

7.3 Le modèle Gaussien discret

Supposons que l'estimateur soit une combinaison linéaire des valeurs prises par les échantillons

$$Z^*(v) = \sum_i \lambda_i Z(x_i)$$

où les coefficients λ_i sont *positifs* et de *somme* 1. Il pourra s'agir par exemple d'un krigeage ordinaire à poids positifs ou même d'une simple moyenne des échantillons. Dans un tel cas, la loi de l'estimateur est moins sélective que celle des échantillons. En effet, si I désigne le numéro aléatoire d'échantillon

qui est égal à i avec la probabilité λ_i, alors $Z(x_I)$ suit la loi des échantillons et de plus

$$\begin{aligned} E\Big\{Z(x_I) \mid Z^*(v)\Big\} &= \sum_i \lambda_i \, E\Big\{Z(x_i) \mid Z^*(v)\Big\} \\ &= E\Big\{\sum_i \lambda_i Z(x_i) \;\Big|\; Z^*(v)\Big\} \\ &= E\Big\{Z^*(v) \mid Z^*(v)\Big\} = Z^*(v) \end{aligned}$$

Commençons par modéliser la loi de $Z^*(v)$. Pour ce faire, nous adoptons un procédé tout à fait analogue à celui qui a permis de modéliser la loi de $Z(v)$ en modèle Gaussien discret. Posons $Z(x_I) = \phi(Y_I)$ ainsi que $Z^*(v) = \phi_v^*(Y_v^*)$, où Y_I et Y_v^* sont deux variables Gaussiennes centrées réduites, et faisons l'hypothèse que le couple (Y_I, Y_v^*) suit une loi bigaussienne. Alors, la relation de Cartier entre $Z(x_I)$ et $Z^*(v)$ implique

$$\phi_v^*(y) = \sum_{n=0}^{+\infty} \frac{\phi_n s^n}{n!} H_n(y) = \phi_s(y)$$

où s est le coefficient de corrélation entre Y_I et Y_v^*. Ce coefficient, nécessairement positif, est donné par la variance de l'estimateur

$$Var\{Z^*(v)\} = \sum_{n=1}^{+\infty} \frac{\phi_n^2 s^{2n}}{n!}$$

Il reste à modéliser la loi de $Z(v)$ à $Z^*(v)$ fixé. Ici, il est bon de rappeler que la construction du modèle Gaussien discret de changement de support conduit à poser $Z(v) = \phi_r(Y_v)$ où Y_v est une variable Gaussienne et où r est un coefficient de corrélation dont la valeur explicite peut être obtenue grâce à la variance de $Z(v)$.

L'hypothèse retenue ici pour l'effet d'information est que le couple (Y_v, Y_v^*) suit une loi bigaussienne de coefficient de corrélation ρ. Ce coefficient ρ peut être évalué expérimentalement grâce à la covariance entre $Z(v)$ et $Z^*(v)$ qui est calculable à partir de la covariance ponctuelle

$$\sum_i \frac{\lambda_i}{|v|} \int_v C(x_i - x)\, dx = Cov\Big\{Z(v), Z^*(v)\Big\} = \sum_{n \geq 1} \frac{\phi_n^2 (rs\rho)^n}{n!}$$

Alors:

$$E\Big\{Z(v) \mid Z^*(v)\Big\} = E\Big\{\phi_r(Y_v) \mid Y_v^*\Big\} = \phi_{r\rho}(Y_v^*)$$

de sorte que la courbe métal/Tonnage effective est

$$Q_v^*(T) = \int_{G^{-1}(1-T)}^{+\infty} \phi_{r\rho}(y)\, g(y)\, dy$$

Exercices

1) L'objet de cet exercice est de comparer les pertes de variance et de sélectivité dues aux effets de support et d'information lorsque l'on cherche à estimer la teneur $Z(v)$ à partir d'un échantillon uniforme $\underline{x}$ dans v

$$Z^*(v) = E\Big\{Z(v) \;\Big|\; Z(\underline{x})\Big\}$$

Pour mener les calculs à bien, on se placera dans le cadre du modèle bilognormal pour lequel

$$\begin{aligned} Z(\underline{x}) &= m \exp\Big\{\sigma Y - \frac{\sigma^2}{2}\Big\} \\ Z(v) &= m \exp\Big\{\sigma_v Y_v - \frac{\sigma_v^2}{2}\Big\} \end{aligned}$$

et où le couple (Y, Y_v) suit une loi bigaussienne de coefficient de corrélation r. Application numérique $m = \sigma = 1$ et $r = 0.5$.

Bibliographie

DEMANGE, C. et al. (1987) Global recoverable reserves: testing various change of support models on uranium data. In MATHERON G. and ARMSTRONG M. (eds.) *Geostatistical Case Studies* pp. 187-208, Reidel Pub. Co., Dordrecht.

GUIBAL, D. et TOUFFAIT, Y. (1982) Grade-tonnage relationships - their use in predicting future reserves and estimating the global recoverable reserves in a deposit. Proc. 17th APCOM Symp., pp. 535-543, Golden (Colorado).

HU, L.Y. (1988) *Mise en oeuvre du modèle gamma pour l'estimation de distributions spatiales.* Thèse de Doctorat en Géostatistique, Ecole des Mines de Paris.

KLEINGELD, W.J. (1987) *Geostatistics and the discrete particles*, Thèse de Doctorat en Géostatistique, Ecole des Mines de Paris.

KLEINGELD, W.J. and MATHERON, G. (1987) The evolution of geostatistics. Proc. 20th APCOM, Vol. 3, pp. 9-12, Johannesburg.

LAJAUNIE, Ch. & LANTUEJOUL, Ch. (1986) Setting up the general methodology for discrete isofactorial models. In ARMSTRONG, M. ed. *Geostatistics* Vol. 1, pp. 323-334, Kluwer Ac. Publ., Dordrecht.

LANTUEJOUL, Ch. (1984) Trois modèles de changement de support à partir d'une distribution ponctuelle gamma. *Sciences de la Terre*, Série Informatique Géologique, Vol. 21 pp. 139-173, Nancy.

LANTUEJOUL, Ch. et RIVOIRARD, J. (1984) Une méthode de détermination d'anamorphose. N-916, Ecole des Mines de Paris.

MATHERON, G. (1971) *La Théorie des Variables Régionalisées et ses Applications.* Ecole des Mines de Paris.

MATHERON, G. (1976) Forecasting block grade distributions: the transfer functions. In M. GUARASCIO et al. (eds.) *Advanced Geostatistics in the Mining Industry*, pp. 237-251, NATO ASI Series C, Vol. 24, Reidel, Dordrecht.

MATHERON, G. (1978) L'estimation globale des réserves récupérables. C-

75, Ecole des Mines de Paris.

MATHERON, G. (1982) La destucturation des hautes teneurs et le krigeage des indicatrices. N-761, Ecole des Mines de Paris.

MATHERON, G. (1984) The selectivity of the distributions and "the second principle of geostatistics". In VERLY, G. et al. (eds.) *Geostatistics for natural ressources characterization*, pp. 421-423, NATO ASI Series C, Vol 122, Reidel, Dordrecht.

MATHERON, G. (1984) Isofactorial models and change of support. In VERLY, G. et al. (eds.) *Geostatistics for natural ressources characterization*, pp. 449-467, NATO ASI Series C, Vol 122, Reidel, Dordrecht.

MATHERON, G. (1984) Changement de support en modèle mosaïque. *Sciences de la Terre*, Vol. 20, pp. 435-454, Nancy.

MATHERON, G. (1984) Une méthodologie générale pour les modèles isofactoriels discrets. *Sciences de la Terre*, Vol. 21, pp. 1-64, Nancy.

MATHERON, G. (1985) Comparaison de quelques distributions du point de vue de la sélectivité. *Sciences de la Terre*, Vol. 24, pp. 1-21, Nancy.

MATHERON, G. (1985) Change of support for diffusion-type random functions. *Mathematical Geology*, Vol. 17-2, pp. 137-165.

MATHERON, G. (1988) Two classes of isofactorial models. In ARMSTRONG, M. ed. *Geostatistics* Vol. 1, pp. 309-322, Kluwer Ac. Publ., Dordrecht.

RIVOIRARD, J. (1984) Convergence des développements en polynômes d'Hermite. *Sciences de la Terre*, Vol. 24, pp. 129-159, Nancy.

RIVOIRARD, J. (1989) *Introduction au Krigeage Disjonctif et à la Géostatistique non linéaire.* C-139, Ecole des Mines de Paris.

Table des matières

1 Quelques rappels probabilistes **3**

1.1 Lois monovariables 3

1.2 Lois bivariables 4

1.3 Fonctions caractéristiques 5

2 L'espérance conditionnelle **8**

2.1 Existence et unicité de l'espérance conditionnelle 8

2.2 Propriétés de l'espérance conditionnelle 12

2.2.1 Linéarité 12

2.2.2 Positivité 13

2.2.3 Monotonie 13

2.2.4 Moyenne 13

2.2.5 Cas d'indépendance 13

2.2.6 Formule d'invariance 14

2.2.7 Formule des projections successives 14

2.2.8 Inégalité de Jensen 14

2.3 Conditionnement multivariable 15

3 Vecteurs aléatoires Gaussiens **18**

3.1 Quelques résultats sur les lois multigaussiennes 18

3.2 Les polynômes d'Hermite 21
3.2.1 Définition et aspects fonctionnels 21
3.2.2 Quelques aspects probabilistes 25

4 Les courbes de sélectivité 30
4.1 Tonnage et Quantité de Métal 30
4.2 La fonction $B(z)$ 33
4.3 La fonction $Q(T)$ 35
4.4 La Sélectivité 37
4.5 La Dualité entre $B(z)$ et $Q(T)$ 39

5 Comparaison de deux lois du point de vue de la sélectivité 43
5.1 Quelques critères de comparaison 43
5.2 La relation de Cartier 45

6 L'effet de support 48
6.1 Quelques observations expérimentales 48
6.2 Que connaît-on de $Z(v)$? 53
6.3 Trois modèles de changement de support 54
6.3.1 La Correction Affine 54
6.3.2 Le changement de support Mosaïque 56
6.3.3 Le Modèle Gaussien Discret 59

7 L'effet d'information 63
7.1 L'effet d'information 64
7.2 Comparaison entre les effets de support et d'information . . . 65
7.3 Le modèle Gaussien discret 66