D-180

ANNALES DES MINES

TECHNIQUES D'ÉVALUATION DES GISEMENTS FONDÉES SUR LA STATISTIQUE MATHÉMATIQUE

8, rue Edouard V
Tél. Opéra 74.97 et 77.46

5, rue Jules Lefebvre
Tél. Trinité 46-38
C.C.P. 640.93 PARIS

144e ANNÉE
DÉCEMBRE 1955

Application des méthodes statistiques à l'évaluation des gisements

par Georges MATHERON,

Ingénieur au Corps des Mines
en service détaché auprès du Bureau de Recherches Minières de l'Algérie

---

On peut, à l'heure actuelle, admettre, comme une certitude expérimentale, que les teneurs, dans un gisement, se répartissent suivant la loi théorique dite lognormale. L'importance, pour la recherche minière, des méthodes statistiques fondées sur cette loi ne provient pas seulement du gain de précision qu'elles apportent, ni même du calcul d'erreur qu'elles justifient. Elle provient surtout de l'explication qu'elles donnent de l'existence d'une erreur systématique d'échantillonnage, erreur particulièrement sensible dans le cas d'une reconnaissance par sondages, et conduisant invariablement à la surestimation de la teneur moyenne, et de la correction rigoureuse qu'elles permettent d'y apporter. Certains spécialistes jugent inutile le recours à une loi théorique, puisque, disent-ils, il suffit de raisonner sur la distribution empirique des teneurs des carottes de sondage. Cela serait parfaitement exact, si, lors de l'exploitation future, le minerai devait être sorti sous forme de petits morceaux, de la taille d'une carotte, que l'on étalerait au jour, et trierait un par un. En fait, le minerai est sorti par paquets de plusieurs tonnes, et il va de soi que la distribution des teneurs de blocs de plusieurs tonnes n'est pas la même que celle de petits morceaux de la taille d'une carotte. Telle est l'origine de l'erreur systématique, erreur que permet seule d'éliminer une théorie capable de déduire, de la loi de distribution des carottes, celle des blocs de plusieurs tonnes.

L'exposé ci-après comporte trois parties principales. La première énonce les propriétés générales de la loi log normale, ou, du moins, celles qui peuvent servir à l'évaluation d'un gisement. La deuxième définit les conditions de l'application de la loi au cas d'un gisement, et tente de résoudre la question de l'erreur systématique. Dans une troisième partie enfin, est traité un exemple de calcul complet, portant sur une minéralisation en plomb argentifère. La parfaite compréhension des démonstrations mathématiques ne suppose rien de plus que la connaissance des propriétés générales de la loi normale, à une et à deux variables. Ces propriétés sont exposées dans tous les traités, même élémentaires, de calcul des probabilités ou de statistique.

SOMMAIRE

Pages.

I. *Généralités sur la loi log-normale* 51

Définition et exemples. — Les paramètres γ et σ^2; les moments. — Estimation de la moyenne. — Décomposition par tranche de teneur. — Les corrélations.

II. *Application de la loi log-normale au cas de gisements* 56

Notion de population. — Les propriétés de la variance. — Les erreurs d'échantillonnage. — Retour sur le cas où il y a corrélation entre puissance et teneur.

III. *Exemple de calcul complet* 70

ANNEXES :

1. Estimation de la moyenne 73
2. Table de la fonction G (*z*) 74
3. Table pour le coefficient de corrélation 75

RÉFÉRENCES.

I. GÉNÉRALITÉS SUR LA LOI LOG-NORMALE

CE titre a pour objet de réunir toutes les formules utiles. Les précautions à prendre dans leur emploi seront précisées dans le titre suivant :

A. Définition et exemples

L'expérience montre qu'un grand nombre de phénomènes, naturels ou économiques, se manifestant sous la forme d'une grandeur dont la mesure est obligatoirement comprise entre 0 et l'infini, obéissent, non pas à la loi normale symétrique, mais à une loi dissymétrique, présentant une queue allongée dans la direction des grandes valeurs de x (fig. 1). Citons par exemple :

— la répartition des fortunes dans un pays;

— le nombre d'habitants des différentes communes de France;

— la surface des différents pays du monde;

— la superficie des terres émergées en fonction de l'altitude;

— la granulométrie des matériaux naturels (sables, graviers) ou artificiels (concassage, broyage);

— le débit d'un fleuve;

— les teneurs dans un gisement.

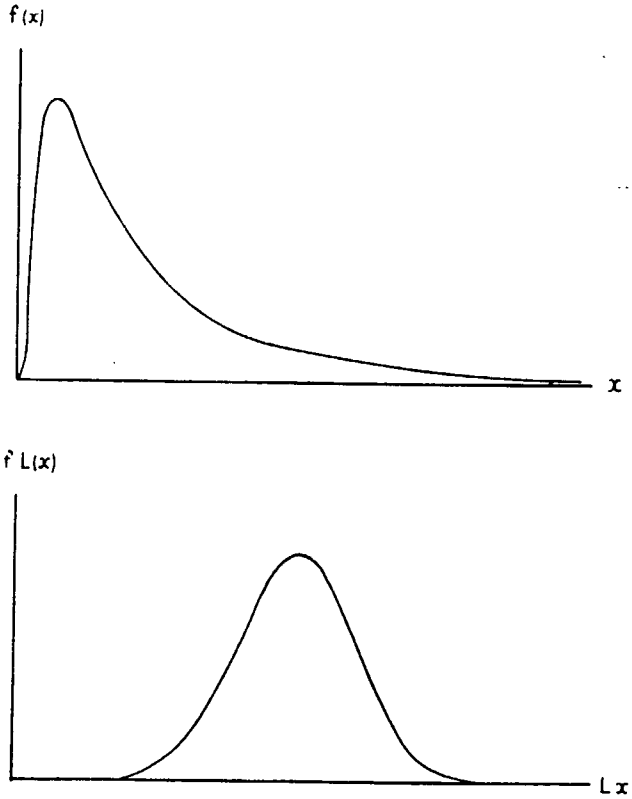

Fig. 1.

Dans tous les cas, on constate que si, au lieu de x, on porte en abscisse log x, la courbe des fréquences prend la forme familière de la courbe en cloche. Autrement dit, ce n'est pas x, mais son logarithme qui est distribué suivant la loi normale. C'est pourquoi la loi de répartition de x est dite log-normale.

Effet proportionnel. — On peut expliquer philosophiquement l'apparition de la loi log-normale grâce à la notion d'*effet proportionnel.* En ce qui concerne la répartition des fortunes dans un pays, par exemple, il est intuitif qu'il est aussi facile, à un homme possédant un milliard de gagner un million, qu'à un homme possédant un million de gagner mille francs. De même l'histoire montre qu'il est aussi facile à un pays de 5 millions de kilomètres carrés d'annexer 5 000 kilomètres carrés, qu'à un pays de 500 000 kilomètres carrés d'en annexer 500. Autrement dit, l'élément différentiel représentatif n'est pas dx, mais $\frac{dx}{x}$, c'est-à-dire la différentielle de Lx. Posons donc :

$$dy = \frac{dx}{x} \quad \text{ou} \quad y = \mathrm{L}x$$

On sait que l'addition d'un grand nombre de petits effets indépendants, ici représentés par des dy, engendre une variable normale. On doit donc s'attendre à ce que y soit une variable normale, donc x une variable log-normale.

Cas d'un gisement. — Dans le cas d'un gisement, on a bien l'intuition que le passage d'une teneur de 10 % à une teneur de 11 % est un phénomène équivalent au passage de 1 % à 1,1 %. M. Allais (1) a remarqué que, dans ce cas particulier, la loi de l'effet proportionnel pouvait se rattacher à la loi physico-chimique d'Action de Masse. En effet, la genèse d'un gisement s'explique par un ensemble d'équilibres chimiques du type :

$$\alpha \mathrm{A} + \beta \mathrm{B} = \gamma \mathrm{C} + \delta \mathrm{D}$$

Et la loi d'Action de Masse nous apprend que les variations des concentrations des différents constituants A, B, C, D... sont liées par des relations du type :

$$\alpha \frac{d\mathrm{A}}{\mathrm{A}} + \beta \frac{d\mathrm{B}}{\mathrm{B}} - \gamma \frac{d\mathrm{C}}{\mathrm{C}} - \delta \frac{d\mathrm{D}}{\mathrm{D}} = 0$$

de sorte que l'élément différentiel significatif n'est pas dA, mais $\frac{d\mathrm{A}}{\mathrm{A}}$:

Toute variation fortuite d'une des concentrations se répercute sur toutes les autres suivant la loi de l'effet proportionnel.

(1) Référence : *a*.

Si maintenant on remarque que « l'affinité », ou énergie libre, d'une réaction est une fonction linéaire des logarithmes des concentrations des différents constituants, on voit qu'admettre une répartition log-normale des concentrations revient à admettre une répartition normale des énergies libres, et une répartition normale d'une énergie apparaît comme assez satisfaisante pour l'esprit.

Définition mathématique de la loi log-normale. — Soit Lγ la valeur moyenne de Lx (log-neperien), et σ^2 sa variance.

La probabilité pour que LX soit compris entre Lx et Lx + d (Lx), donc x entre x et $x + dx$ est, puisque Lx est une variable normale :

$$f(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\mathrm{L}x - \mathrm{L}\gamma)^2}{2\sigma^2}} d(\mathrm{L}x)$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{\left(\mathrm{L}\left(\frac{x}{\gamma}\right)\right)^2}{2\sigma^2}} \frac{dx}{x}$$

La fonction de fréquence, ou densité de probabilité, est donc :

$$(1) \qquad f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\mathrm{L}x - \mathrm{L}\gamma)^2}{2\sigma^2}}$$

Par définition de la fonction de fréquence, l'expression $f(x)dx$ représente, en pourcentage du tonnage total, le tonnage du minerai dont la teneur est comprise entre x et $x + dx$. Toutefois, aussi bien pour les démonstrations théoriques que pour les calculs pratiques, il est préférable d'effectuer le changement de variable suivant :

$$(2) \qquad z = \frac{1}{\sigma} \mathrm{L} \frac{x}{\gamma}, \quad \text{soit} \quad x = \gamma e^{\sigma z}$$

En effet, z est une variable normale réduite : sa fonction de fréquence est $\frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ et les tables de statistique donnant la valeur numérique de l'intégrale :

$$(3) \qquad \mathrm{G}(z) = \int_z^{+\infty} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$$

il y a tout intérêt à pouvoir utiliser les données ainsi disponibles.

B. Les paramètres γ et σ^2; Les moments

La médiane γ. — Par définition, Lγ est la valeur moyenne de Lx. On en déduit que γ est à la fois *la*

médiane de la variable aléatoire X (il y a, en effet, une chance sur deux pour que Lx soit supérieur à Lγ, donc également une chance sur deux pour que x soit supérieur à γ), et *la moyenne géométrique* de X, puisque Lγ est la moyenne arithmétique de Lx.

La variance σ^2. — C'est la variance de Lx. Elle donne une mesure de la dispersion relative de x. Ces deux paramètres, γ et σ, suffisent à déterminer la loi. Nous indiquerons tout à l'heure comment les calculer.

Les moments. — Par définition, on appelle *moment d'ordre n*, l'expression :

$$m_n = \int_0^{\infty} x^n f(x) dx$$

Compte tenu de la formule (2), on peut écrire :

$$m_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \gamma^n e^{n\sigma z} e^{-\frac{z^2}{2}} dz$$

Cette expression s'intègre aisément. On trouve :

$$m_n = \gamma^n e^{\frac{n^2\sigma^2}{2}} \tag{4}$$

En particulier, la valeur moyenne de x ($n = 1$) a pour expression :

$$\boxed{m = \gamma e^{\frac{\sigma^2}{2}}} \tag{5}$$

σ^2 est une variance logarithmique; la variance vraie de x est égale à :

$$\Sigma^2 = m_2 - m^2 = m^2(e^{\sigma^2} - 1) \tag{6}$$

Cette expression montre bien que σ^2 mesure une dispersion relative.

Estimation de γ. — Le meilleur estimateur de Lγ est :

$$\boxed{L\gamma^* = \frac{1}{n} \sum_{i=1}^{n} Lx_i} \tag{7}$$

c'est-à-dire la moyenne arithmétique des teneurs des n échantillons. Lx étant une variable normale, il en est de même de Lγ^*. La valeur moyenne de Lγ^* est Lγ, et sa varianceσ^2/n. Par suite, γ^* est une variable log-normale, de même variance, et de médiane γ. La valeur moyenne de γ^* n'est pas γ, mais $\gamma e^{\sigma^2/2n}$. γ^* est donc un estimateur biaisé. En pratique, n est assez grand pour qu'il n'y ait pas lieu de faire une correction de biais.

En ce qui concerne la précision de γ^*, on peut dire qu'il y a environ 95 chances sur cent pour que la valeur véritable de γ soit comprise à l'intérieur de l'intervalle :

$$\gamma^* e^{-\frac{2\sigma}{\sqrt{n}}} - \gamma^* e^{\frac{2\sigma}{\sqrt{n}}}$$

En toute rigueur, il faudrait connaître la vraie valeur de σ alors que l'on n'en a qu'une estimation. Pratiquement, n est assez grand pour que l'on puisse négliger de telles subtilités.

Estimation de la variance σ^2. — La théorie de la loi normale montre que le meilleur estimateur de σ^2 est :

$$\boxed{\sigma^{*2} = \frac{1}{n-1} \sum_{i=1}^{n} (Lx_i - L\gamma^*)^2} \tag{8}$$

Il faut prendre n–1, et non pas n, parce que, en remplaçant Lγ par son estimation Lγ^*, on a diminué d'une unité le nombre des « degrés de liberté ». Dès que n est assez grand, σ^{*2} est pratiquement une variable normale, de valeur moyenne σ^2 et de variance $\frac{2\sigma^4}{n-1}$ Il y a donc 95 chances sur 100 pour que la valeur véritable de σ^2 soit comprise dans l'intervalle :

$$\sigma^{*2}\left[1 - 2\sqrt{\frac{2}{n-1}}\right] \qquad \sigma^{*2}\left[1 + 2\sqrt{\frac{2}{n-1}}\right]$$

C. Estimation de la moyenne *m*

La pratique habituelle consiste à prendre la moyenne arithmétique. Ce procédé paraît tout naturel; cependant, ce n'est pas le meilleur. On aura toujours une plus grande précision avec l'estimateur :

$$\boxed{m^* = \gamma^* e^{\frac{\sigma^{*2}}{2}}} \tag{9}$$

En effet, m, produit de deux variables log-normales, est une variable log-normale, de variance logarithmique $\frac{\sigma^2}{n} + \frac{\sigma^4}{2n}$, donc de variance brute :

$$\Sigma_1^2 = m^2\left[e^{\sigma^2/n + \sigma^4/2n} - 1\right] \simeq \frac{m^2}{n}\left(\sigma^2 + \frac{\sigma^4}{2}\right)$$

De son côté, la moyenne arithmétique ordinaire (qui n'est pas une variable log-normale) a pour variance :

$$\Sigma_2^2 = \frac{m^2}{n}\left(e^{\sigma^2} - 1\right) = \frac{m^2}{n}\left(\sigma^2 + \frac{\sigma^4}{2!} + \frac{\sigma^6}{3!} + \frac{\sigma^8}{4!} + \cdots\right)$$

expression toujours supérieure à la précédente. Dès que σ est de l'ordre de l'unité, valeur tout à fait courante,

la différence devient énorme, et Krige (*c*, *d*) a montré l'importance du gain de précision apporté par la formule (9).

L'estimateur (9) n'est pas, en fait, le meilleur estimateur possible; c'est plutôt l'estimateur de Sichel (1). En pratique, n est assez grand pour qu'il n'y ait pas de différence appréciable entre ces deux estimateurs. Il y a — approximativement — 97,5 chances sur 100 pour que la valeur réelle de m soit supérieure à Km^*, le coefficient de sécurité K étant donné par la formule :

$$(10) \qquad K = e^{-2\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^4}{2n}}}$$

D. Décomposition par tranche de teneur

Une fois γ et σ déterminés, des formules purement mathématiques permettent de déterminer, en pourcentage du tonnage total, le tonnage compris entre deux teneurs x_1 et x_2, ainsi que la teneur moyenne de cette tranche.

Posons :

$$(11) \qquad z_1 = \frac{1}{\sigma} L\left(\frac{x_1}{\gamma}\right) \qquad z_2 = \frac{1}{\sigma} L\left(\frac{x_2}{\gamma}\right)$$

Le tonnage correspondant aux teneurs supérieures à x, est donné par :

$$(12) \qquad T_1^{\infty} = G(z_1)$$

Le tonnage compris entre x_1 et x_2 est donné par :

$$(13) \qquad T_1^2 = G(z_1) - G(z_2)$$

Teneur moyenne de la tranche correspondante. Elle est égale à :

$$(m)_1^2 = \frac{\int_{x_1}^{x_2} x f(x)\, dx}{\int_{x_1}^{x_2} f(x)\, dx} = \frac{\int_{z_1}^{z_2} \gamma e^{\sigma z} e^{-\frac{z^2}{2}}\, dz}{\int_{z_1}^{z_2} e^{-\frac{z^2}{2}}\, dz}$$

On trouve :

$$(14) \qquad (m)_1^2 = m \frac{G(z_1 - \sigma) - G(z_2 - \sigma)}{G(z_1) - G(z_2)}$$

Si x_2 est infini, c'est-à-dire *si l'on considère le minerai à teneur supérieure à x_1, on trouve :*

$$(15) \qquad \boxed{m_1 = m \frac{G(z_1 - \sigma)}{G(z_1)}}$$

(1) Référence *e*. Voir aussi : annexe I.

Remarque. — Si l'on écrit cette relation sous la forme :

$$m_1 T G(z_1) = m T G(z_1 - \sigma)$$

T étant le tonnage de minerai total, le premier membre représente le tonnage métal compris dans la tranche considérée, et mT le tonnage métal total : $G(z_1 - \sigma)$ représente donc le pourcentage du tonnage métal compris dans cette tranche.

E. Les corrélations

Étant données deux variables aléatoires log-normales X et Y, qui seront ici *les teneurs en métal A et en métal B d'un même échantillon,* nous définirons le coefficient de corrélation ρ comme le coefficient de corrélation de LX et LY, variables normales. Les résultats classiques de l'étude des corrélations entre variables normales s'étendent d'eux-mêmes au cas log-normal. On sait que le coefficient ρ, toujours compris entre $+1$ et -1, donne une mesure du degré de dépendance des deux variables. $\rho = 0$ signifie l'indépendance, tandis que les valeurs extrêmes, $+1$ et -1, correspondent au cas où il existe une relation fonctionnelle entre X et Y.

Calcul du coefficient de corrélation ρ. — Soient γ_1 et γ_2, σ_1 et σ_2 les médianes et écarts-types logarithmiques de X et Y. On évalue ρ par la formule suivante :

$$(16) \qquad \boxed{r = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(Lx - L\gamma_1)}{\sigma_1} \frac{(Ly - L\gamma_2)}{\sigma_2}}$$

La variable aléatoire r n'est pas normale. Des tables donnent, en fonction de r, les valeurs extrêmes que ρ n'a pas plus de 5 chances sur 100 de dépasser. Il est très important, en particulier, d'apprécier si r est significativement différent de zéro ou non.

Seul le coefficient de corrélation permet de substituer à la notion géologique qualitative de « liaison » entre deux minéralisations, une mesure rigoureuse de leur degré de dépendance.

Calcul rapide de ρ. — On construit, en coordonnées logarithmiques, *le nuage de corrélation :* portant en abscisse Lx, et en ordonnée Ly; chaque échantillon est représenté par un point, et l'ensemble des n échantillons, par un nuage de points. Cette notion de nuage est extrêmement utile dans la pratique, parce qu'elle donne une image géométrique intuitive des lois de répartition.

Ayant construit le nuage, on trace les axes passant par le centre de gravité ($L\gamma_1$, $L\gamma_2$), et on compte le nombre des points contenus dans chacun des quadrants (fig. 2).

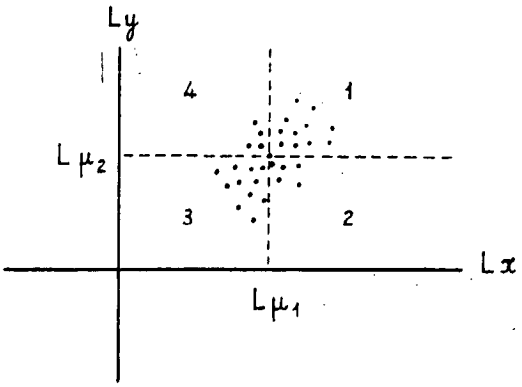

Fig. 2.

Soit N_1 le nombre des points contenus dans le 1er et le 3e.

Soit N_2 le nombre des points contenus dans le 2e et le 4e;

La formule :

$$r = \sin \frac{\pi}{2} \cdot \frac{N_1 - N_2}{N_1 + N_2} \tag{17}$$

fournit une évaluation de ρ, un peu moins précise, mais de calcul plus rapide.

Ayant calculé r, deux cas peuvent se présenter. Si r est nul, ou non significativement différent de zéro, le nuage, en coordonnées logarithmiques, présente une allure elliptique, les axes de l'ellipse étant parallèles aux axes de coordonnées, X et Y sont indépendants. Si, au contraire, r est significativement différent de zéro, x et y ne sont pas indépendants, et le nuage affecte la forme d'une ellipse, dont les axes sont obliques. La pente du grand axe a le même signe que ρ.

Loi d'action de masse. — Lorsque ρ est égal à ± 1, l'ellipse se réduit à une droite, et il y a une relation linéaire entre les logarithmes de x et de y. Si $\rho = +1$, la relation entre x et y s'écrit :

$$\left[\frac{x}{\gamma_1}\right]^{\frac{1}{\sigma_1}} = \left[\frac{y}{\gamma_2}\right]^{\frac{1}{\sigma_2}} \tag{18}$$

Pour $\rho = -1$, elle s'écrit :

$$\left[\frac{x}{\gamma_1}\right]^{\frac{1}{\sigma_1}} \left[\frac{y}{\gamma_2}\right]^{\frac{1}{\sigma_2}} = 1 \tag{19}$$

Dans les deux cas, elle est de la forme de la loi d'action de masse. On doit chercher s'il existe un équilibre chimique, géologiquement possible, entre les métaux A et B, les coefficients de A et de B étant proportionnels à $1/\sigma_1$ et $1/\sigma_2$. L'apparition de la loi d'action de masse n'a rien d'étonnant, elle est conforme à la nature profonde de la loi log-normale.

En pratique, r n'est jamais rigoureusement égal à ± 1. Le seul fait qu'il existe une erreur d'analyse chimique, laisse subsister une variance résiduelle pour y lorsque x est fixé. Mais si r est de l'ordre de $\pm 0{,}95$, on peut sans crainte admettre que ρ est égal à ± 1, et écrire la loi d'action de masse. C'est ainsi, par exemple, qu'au gisement de manganèse du Djebel-Guettara (*h*) on a trouvé, entre manganèse et arsenic, un coefficient de corrélation négatif, égal à $-0{,}97$. Les écarts-types, σ_1 et σ_2, étant dans le rapport de 1 à 8, la relation s'écrit ici :

$$(As)\,(Mn)^8 = \text{constante},$$

et doit être interprétée par une équation chimique du type :

$$As + 8Mn = \text{quelque chose}.$$

Formule des Minerais connexes. — En général, cependant, ρ, tout en ayant une valeur notable, n'est pas assez voisin de $+1$ ou -1 pour que l'on puisse admettre une relation fonctionnelle entre x et y. On est conduit à étudier la loi de probabilité de y, lorsque x est fixé (loi conditionnelle : quelle est la valeur moyenne des teneurs en zinc, lorsque la teneur en Pb est de 4 % par exemple, et quelle est leur dispersion?)

En transposant les données classiques de la loi normale à deux variables, on voit que lorsque x est fixé, Ly est une variable normale, mais :

1° Sa moyenne n'est plus $L\gamma_2$, mais :

$$L\gamma_2 + \rho \frac{\sigma_2}{\sigma_1} L\left(\frac{x}{\gamma_1}\right)$$

2° Sa variance n'est plus σ_2^2, mais $\sigma_2^2 (1 - \rho^2)$, donc plus faible.

Elle est indépendante de y.

On en conclut que, pour x fixé, y est encore une variable log-normale, mais :

1° Sa médiane n'est plus γ_2, mais :

$$\gamma_2 \left[\frac{x}{\gamma_1}\right]^{\beta} \qquad \text{avec : } \beta = \rho \frac{\sigma_2}{\sigma_1}$$

2° Sa variance n'est plus σ_2^2, mais $\sigma_2^2 (1 - \rho^2)$.

Par suite, la moyenne de y, pour x fixé, est, d'après la formule (5) :

$$m_2(x) = \gamma_2 \left[\frac{x}{\gamma_1}\right]^{\beta} \cdot e^{\frac{\sigma_2^2(1-\rho^2)}{2}} \tag{20}$$

Si nous considérons maintenant non plus une valeur fixée de x, mais une tranche de teneur définie par x_1 et x_2, la valeur moyenne de x dans cette tranche est donnée par la formule (14). Il est possible de calculer également la teneur moyenne de y pour l'ensemble des échantillons pour lesquels x est compris entre x_1 et x_2.

En effet, posons :

$$(21) \qquad z_1 = \frac{1}{\sigma_1} L\left(\frac{x_1}{\gamma_1}\right) \qquad z_2 = \frac{1}{\sigma_2} L\left(\frac{x_2}{\gamma_1}\right)$$

Le tonnage de cette tranche est donné, en pourcentage, par :

$$\frac{1}{\sqrt{2\pi}} \int_{z_1}^{z^2} e^{-z^2/2} dz = G(z_1) - G(z_2)$$

$\frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$ représente le tonnage compris entre :

$$x = \gamma_1 e^{\sigma_1 z} \quad \text{et} \quad \gamma_1 e^{\sigma_1 z}(1 + \sigma_1) dz = x + dx$$

tonnage dont la teneur moyenne en B (valeur moyenne de Y pour cette valeur de x) est d'après (20) et (21) :

$$m_2(x) = \gamma_2 \left(\frac{x}{\gamma_1}\right)^{\beta} e^{\frac{\sigma^2_2(1-\rho^2)}{2}} = \gamma_2 e^{\beta\sigma_1 z + \sigma_2^2 \frac{(1-\rho^2)}{2}}$$

$$\left(\beta = \rho \frac{\sigma_2}{\sigma_1}\right)$$

Or, la valeur moyenne de y dans la tranche $x_1 - x_2$ est :

$$m_2(x_1, x_2) = \frac{\frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} m_2(x) e^{-z^2/2} dz}{\frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-z^2/2} dz}$$

L'intégrale du numérateur est égale à :

$$\frac{1}{\sqrt{2\pi}} \gamma_2 \int_{z_1}^{z_2} e^{\beta\sigma_1 z + \sigma_2 \frac{(1-\rho^2)}{2} - z^2/2} dz$$

$$= \gamma_2 e^{\sigma^2/2} \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{(z-\rho\sigma_2)^2}{2}} dz$$

$$= m_2[G(z_1 - \rho\sigma_2) - G(z_2 - \rho\sigma_2)]$$

La valeur moyenne de y, dans la tranche $x_1 x_2$, est donc donnée par la formule suivante, dite formule des minerais connexes :

$$(22) \qquad m_2(x_1, x_2) = m_2 \frac{G(z_1 - \rho\sigma_2) - G(z_2 - \rho\sigma_2)}{G(z_1) - G(z_2)}$$

On remarquera l'analogie avec la formule (15).

Si x_2 est infini, c'est-à-dire *si l'on ne considère que le minerai à teneur $x > x_1$, la valeur moyenne de y est :*

$$(23) \qquad \boxed{m_2(x_1) = m_2 \frac{G(z_1 - \rho\sigma_2)}{G(z_1)}}$$

$G(z_1 - \rho\sigma_2)$ représente le pourcentage du tonnage total du corps B compris dans cette tranche.

II. APPLICATION DE LA LOI LOG-NORMALE AU CAS DES GISEMENTS

APRÈS avoir exposé les généralités, passons à l'application au cas des gisements. En fait, toutes les formules utiles ont déjà été données, et un exemple pratique nous montrera la manière de s'en servir. Toutefois, et afin d'éviter un emploi mécanique de ces formules, pouvant conduire à de graves erreurs, il est nécessaire de préciser quelques points délicats. Avant de faire des calculs, il faut réfléchir, sous peine de n'obtenir que des résultats dénués de signification. Au contraire, une étude critique des données met souvent en lumière des faits nouveaux, même des faits géologiques. Il convient de parler, d'une part, de la notion de population, et en particulier de son homogénéité ou de son hétérogénéité; de l'autre, de certaines propriétés de la variance, importantes pour les calculs pratiques.

A. Notion de population

S'il s'agit d'étudier, par exemple, la loi de répartition de la taille des français, la population, au sens statistique, est définie sans ambiguïté, comme l'*ensemble des tailles, mesurées en centimètres, de chacun des français.* Dans le cas d'un gisement, la population est constituée par l'ensemble de toutes les teneurs possibles. Mais il est essentiel de préciser qu'il s'agit des teneurs de morceaux de minerai de quelques kilogrammes ou de blocs de plusieurs tonnes. La teneur moyenne sera bien la même, mais non pas la dispersion. Dans le cas d'un filon, il est plus représentatif de considérer, non pas les teneurs, mais les produits des teneurs par la hauteur minéralisée. Mais il faut alors préciser la surface de référence. Il y a

donc deux façons possibles de définir la population :

1° *On considère le gisement comme formé par la juxtaposition d'un grand nombre de blocs jointifs de même volume* v *et de même forme géométrique.* v peut être quelconque, mais doit être fixé une fois pour toutes. *La population est alors constituée de l'ensemble des teneurs de chacun des volumes* v. Lorsqu'on cherchera à évaluer le gisement, les échantillons prélevés devront tous être des volumes v.

2° Ou bien *on considère la surface S correspondant à l'extension du gisement comme formée de la juxtaposition d'un grand nombre d'aires jointives, de même surface* s *et de même forme géométrique.* A chaque aire correspond une hauteur h de minéralisation. *La population est alors constituée de l'ensemble des produits* hx *de la puissance minéralisée* h *par la teneur correspondante pour chacune des surfaces* s. Au lieu de hx, il est plus rigoureux de considérer ρhx, ρ étant la densité; ρ, en effet, est en général d'autant plus élevée que la teneur x est plus forte. ρhx est un *poids métal/mètre carré.* En fait, la variation de ρ est presque toujours négligeable, et il suffit de considérer les hx. Lorsqu'on cherche à évaluer le gisement, les carottes, ou les saignées, devront toutes avoir une section égale à s.

Donc, il y a *deux méthodes d'attaque : les volumes* v et *les surfaces* s.

Méthode des volumes v. — Elle conduit directement à la teneur moyenne (valeur moyenne de x) et à la dispersion. Il faut l'employer chaque fois que toutes les dimensions du gisement sont du même ordre de grandeur, c'est-à-dire *lorsqu'il s'agit d'un amas.* Lorsque le gisement se présente sous forme de *lentilles* plus ou moins superposées, et réparties irrégulièrement, il peut y avoir avantage à utiliser la méthode des volumes v, qui permet une analyse plus fine. Lorsque les volumes v sont représentés par des volées, ou des carottes de longueurs irrégulières, il est tout à fait légitime de se servir des teneurs non pondérées à condition qu'il n'y ait pas de corrélation entre la longueur de la volée ou de la carotte et la teneur correspondante.

Méthodes des surfaces s. — Elle s'impose dans le cas *d'un gisement filonien ou en couches minces;* il y a, en effet, presque toujours corrélation entre h et x, x étant par exemple d'autant plus faible que h est plus grand. Les hx comme les x, suivent la loi log-normale. On évaluera le hx moyen par la formule (5). En divisant par la valeur moyenne de h, on aura la teneur moyenne. S'il s'agit d'un filon ou d'une couche mince, et que les conditions d'exploitation obligent à dépiler une hauteur h' supérieure à la puissance h du filon ou de la couche, la teneur est égale à $\frac{hx}{h'}$. h' étant constant, hx doit être alors considéré comme une véritable teneur, et sa dispersion comme celle des teneurs elles-mêmes. Nous reviendrons plus loin sur le cas où h n'est pas toujours inférieur à h'.

Homogénéité et hétérogénéité. — Ayant judicieusement choisi la méthode, il peut arriver que les teneurs d'échantillonnages ne suivent pas la loi log-normale. Or l'échantillonnage donne une image d'autant meilleure de la distribution réelle qu'il y a davantage d'échantillons. Si donc le nombre des échantillons est déjà important (une trentaine au moins), et *si la répartition n'est manifestement pas log-normale, on doit suspecter une hétérogénéité.*

Hétérogénéité, cela signifie que la minéralisation n'appartient pas à un type unique, mais se répartit entre deux ou plusieurs types géologiquement distincts.

La façon la plus simple de mettre en évidence une hétérogénéité consiste, à condition que les échantillons aient été analysés pour deux constituants au moins, *à construire le nuage de corrélation en coordonnées logarithmiques.* S'il y a hétérogénéité, les points se répartiront entre deux ou plusieurs groupes elliptiques qui apparaîtront de façon intuitive. Si les ellipses en question ne se chevauchent pas trop, ce procédé très simple permettra même de séparer les différents types. Même s'il y a chevauchement, une étude géologique, orientée sur l'idée de l'existence de plusieurs types, donnera une interprétation de cette distinction, et fournira un critère permettant de séparer les types composants.

Une fois cette séparation faite, on s'apercevra généralement que chacun des types constitue une population log-normale homogène, avec ses paramètres propres. *L'analyse statistique permet ainsi de fonder des distinctions géologiques non évidentes* a priori.

Frange minéralisée. — Dans tous les cas, ou presque, on s'apercevra que le nuage elliptique de points bien groupés est accompagné, vers les faibles teneurs, d'un nuage de points beaucoup moins dense et plus dispersé. Ces points constituent *la frange faiblement minéralisée du gisement.* Les couches correspondantes ne font plus partie du gisement, mais le prolongent à la manière d'un halo. Le plus souvent, mais pas toujours, *un simple tronquage* des faibles teneurs, effectué à un seuil convenablement choisi, permet d'éliminer cette frange. Il importe de bien préciser que le *seuil, ou teneur limite du tronquage, est absolument indépendant de toute idée d'exploitabilité.* Il est généralement inférieur à la limite d'exploitabilité, mais il peut aussi lui être supérieur. Les calculs doivent porter sur tou-

ce qui est supérieur à ce seuil, et non pas sur ce qui est supérieur à la limite d'exploitabilité. C'est seulement une fois γ et σ^2 déterminées, pour l'ensemble de la population ainsi définie, que la teneur limite doit être introduite, par le biais du calcul des tranches de teneurs. Nous verrons plus loin qu'en procédant autrement on risque de commettre de graves erreurs.

B. Les propriétés de la variance

Avant de justifier ce qui a été dit plus haut, sur la nécessité de connaître les paramètres avant de procéder à un tronquage, démontrons d'abord une propriété purement théorique, mais assez instructive, de la variance.

1° *Existence d'une limite supérieure de la variance* (h)

En effet, étant donné un gisement de teneur moyenne m, on conçoit très bien que la plus grande dispersion possible, donc la plus grande variance brute Σ^2, est obtenue dans l'hypothèse où chacun des volumes v est constitué soit par du minerai pur, de teneur x_0, soit par du stérile pur, de teneur 0.

Soit p la proportion des volumes v de teneurs x_0. On a :

$$m = px_0.$$

La variance, est dans cette hypothèse, égale à :

$$\Sigma_0^2 = p(x_0 - m)^2 + (1-p)m^2$$
$$= \frac{m}{x_0}(x_0 - m)^2 + \left(\frac{x_0 - m}{x_0}\right)m^2 = m(x_0 - m)$$

Telle est la plus grande variance possible. Pour la distribution réelle, log-normale, la variance Σ^2 est :

$$\Sigma^2 = m^2(e^{\sigma^2} - 1)$$

On doit nécessairement avoir :

$$\Sigma^2 < \Sigma_0^2$$

On en déduit immédiatement :

$$\boxed{\sigma^2 < 1 + \frac{x_0}{m}} \qquad (24)$$

Cette inégalité très simple montre que, *plus* m *est voisin de* x_0, *plus la limite supérieure de la variance est basse. Telle est l'explication du fait bien connu que les gisements à haute teneur moyenne sont généralement peu dispersés,* et inversement.

2° *Permanence de la log-normalité par rapport à v et V* (ou s et S)

Dans ce qui suit, nous utiliserons la première définition de la population (gisement de volume V considéré comme la juxtaposition des volumes v). Les conclusions seront également valables pour la deuxième définition, à condition de remplacer v et V par s et S.

Cette définition comprend deux paramètres susceptibles de variation. D'une part le volume v : que se passe-t-il si je définis la population à partir d'un autre volume v', quel rapport y a-t-il entre les populations des v et des v', dont chacune représente le même gisement de volume V? D'autre part le volume V lui-même : que se passe-t-il, en effet si au lieu de considérer le gisement V entier je n'en considère qu'une portion de volume V' ? Quel rapport y a-t-il entre la population des v dans V, et la sous-population, extraite de la précédente, des v dans V' ?

Intuitivement, on se rend compte que *la dispersion doit être d'autant plus faible que* v *est plus grand et que* V' *est plus petit.* C'est bien ce que l'on observe : les échantillons représentant des volées sont moins dispersés que des carottes, et d'autre part, lorsqu'on découvre une extension d'un gisement, la dispersion des résultats antérieurs à cette découverte est plus faible que celle des résultats d'ensemble. On peut aller plus loin, et se demander si la loi reste log-normale lorsque v et V varient.

En fait, l'expérience montre qu'elle le reste, mais ses paramètres γ et σ se modifient. Cette *permanence de la log-normalité,* vis-à-vis de v et de V, n'est pas du tout aussi naturelle qu'elle le paraît. En effet, un échantillon de volume $2v$ peut être regardé comme la juxtaposition de deux volumes v contigus. Sa teneur x est la moyenne arithmétique des teneurs x_1 et x_2 des deux volumes v voisins. On sait que la moyenne arithmétique de deux variables normales indépendantes est toujours une variable normale. Dans le cas de la loi log-normale, il n'en est pas de même, parce que $\log(x_1 + x_2)$ n'est pas égal à $\log(x_1) + \log(x_2)$. On démontre, en fait, que la moyenne arithmétique de deux variables log-normales indépendantes n'est pas une variable log-normale.

Comment concilier ce fait avec la permanence de la log-normalité? On doit conclure que les teneurs x_1 et x_2 de deux volumes v voisins *ne sont pas indépendantes.* Cette notion de corrélation entre teneurs de volumes voisins exprime très bien *la constatation empirique de l'existence d'une certaine continuité de la minéralisation.* Il est évident pour tout le monde qu'une zone contiguë à une zone riche a plus de chances *a priori* qu'une autre d'être riche également. L'étude des corrélations entre les tronçons de carottes successifs d'un même sondage conduit toujours, en effet,

à des coefficients ρ positifs, significativement différents de 0.

3° *Hypothèse et formule de Krige* (c)

Ce sont ces corrélations qui assurent la permanence de la log-normalité, mais par quel mécanisme analytique, nous n'en savons rien actuellement. Toutefois, cette permanence est un fait expérimental; nous exposerons tout à l'heure la théorie approximative de de Wijs, qui permet d'en comprendre la raison profonde, et nous tenterons d'en donner une démonstration plus générale.

Pour l'instant, *nous admettons cette permanence à titre de postulat.* Voyons quelles conséquences en découlent.

Soit un gisement de volume V et de teneur moyenne m. Il peut être considéré comme la juxtaposition d'un grand nombre de panneaux de même volume Vi. Chacun de ces volumes Vi a une teneur m_i, variable.

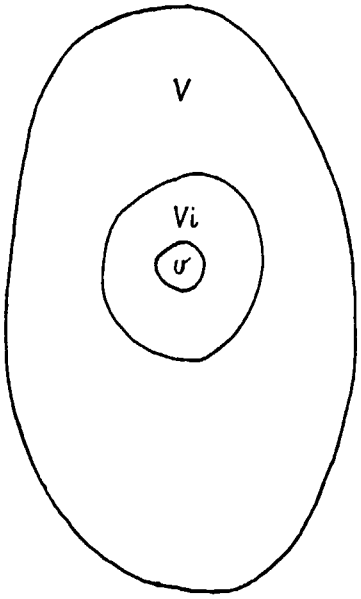

Fig. 3.

A son tour chaque Vi, de teneur m_i peut être considéré comme la juxtaposition d'un grand nombre de petits volumes v, tous égaux entre eux. Soit x la teneur d'un volume v (fig. 3).

Du postulat de la permanence découlent les conséquences suivantes :

a. Les m_i forment une population log-normale dans V; soient γ' leur médiane et σ^2_γ leur variance.

b. Dans chaque Vi, les x forment une population log-normale, de moyenne m_i par définition. Soit σ^2_i leur variance. Rien ne permet, *a priori*, d'affirmer que σ^2_i soit le même dans tous les Vi.

c. Dans V tout entier, les x forment une population log-normale; soient γ leur médiane et σ^2 leur variance.

Étudions la corrélation entre la teneur m$_i$ *d'un volume* Vi, *et les teneurs* x *des* v *contenus dans ce* Vi. Soit ρ le coefficient de corrélation entre x et m_i. Lorsque m_i est fixé, la variance de x est égale à $\sigma^2(1-\rho^2)$. Elle est indépendante de m_i. Mais la variance de x à m_i fixé n'est pas autre chose que la variance σ^2_i des x dans le Vi correspondant, donc :

$$\boxed{\sigma_i^2 = \sigma^2(1-\rho^2)} \tag{25}$$

La variance σ^2_i *de* x *est la même pour tous les* Vi. — Ce résultat, posé par Krige à titre d'hypothèse, apparaît ainsi comme une première conséquence de la permanence de la log-normalité.

La valeur moyenne de x, à m_i fixé, est donnée par la formule (20).

$$m_x(m_i) = \gamma\left[\frac{m_i}{\gamma'}\right]^{\rho\frac{\sigma_\gamma}{\sigma}} e^{\frac{\sigma^2(1-\rho^2)}{2}} \tag{26}$$

Par définition, cette moyenne doit être égale à m. — On en déduit immédiatement :

$$\boxed{\rho = \frac{\sigma_\gamma}{\sigma}} \tag{27}$$

Portons cette valeur ρ dans la formule (25). Il vient :

$$\boxed{\sigma^2 = \sigma^2 + \sigma_\gamma^2} \tag{28}$$

Telle est la formule de Krige. *La variance des* v *dans V est égale à la somme de la variance des* v *dans* Vi *et de la variance des* Vi *dans V. Elle est donc supérieure à chacune de ces deux variances,* comme annoncé au 2° ci-dessus.

Conservation de la moyenne m. — Il faut naturellement vérifier que les x et les m_i ont même valeur moyenne m dans V. Les x ont pour teneur moyenne, dans V :

$$m_1 = \gamma e^{\sigma^2/2}$$

et les m_i :

$$m_2 = \gamma' e^{\sigma^2_\gamma/2}$$

Mais γ' est la médiane des m_i dans V. Or chaque m_i est égale à :

$$m_i = \gamma_i \; e^{\sigma_i^2/2}$$

σ^2_i étant constante, et γ_i étant la médiane des x dans le Vi correspondant. La médiane γ' de m_i dans V est donc égale à la médiane de γ_i multipliée par $e^{\sigma^2/2}$ Mais la médiane de γ_i dans V n'est pas autre chose que γ (son logarithme est en effet, la moyenne arithmétique des logarithmes des x dans V). Donc :

$$m_2 = \gamma e^{\frac{\sigma^2_i + \sigma_\gamma^2}{2}}$$

et la formule (28) nous montre que l'on a bien :

$$m_1 = m_2$$

Conséquence. — Si l'on admet que la variance σ^2 *ne dépend que de V et* v :

(29) $$\sigma^2 = \mathrm{F}(\mathrm{V}, v)$$

On déduit de la formule de Krige que *la fonction F est de la forme* f (V) — f(v). En effet, la formule de Krige, en termes de fonction F, s'écrit :

$$\mathrm{F}(\mathrm{V}, v) = \mathrm{F}(\mathrm{V}i, v) + \mathrm{F}(\mathrm{V}, \mathrm{V}i)$$

Supposons Vi fixé, et faisons varier V et v. On voit que F est de la forme :

$$\mathrm{F}(\mathrm{V}, v) = g_1(v) + g_2(\mathrm{V})$$

Appliquons une seconde fois la formule de Krige. Il vient :

$$g_1(v) + g_2(\mathrm{V}) = g_1(v) + g_2(\mathrm{V}i) + g_1(\mathrm{V}i) + g_2(\mathrm{V})$$

Ceci devant être vrai *quel que soit* Vi, on en déduit :

$$g_1(x) = \mathrm{K} - g_2(x)$$

K est une constante indéterminée pour l'instant. On a donc :

$$\mathrm{F}(\mathrm{V}, v) = g_2(\mathrm{V}) + \mathrm{K} - g_2(v)$$

Mais σ^2 est nul lorsque V = v, donc K = 0, et σ^2 est de la forme :

(30) $$\boxed{\sigma^2 = f(\mathrm{V}) - f(v)}$$

4° *Conséquence pratique de la théorie de Krige*

Erreur systématique d'échantillonnage. Nous avons étudié la corrélation entre x et m_i, en supposant m_i fixé et en déterminant la loi liée de x. Dans la pratique, le problème posé est exactement inverse. On a prélevé dans Vi un échantillon v, de teneur x (donc x est fixé), et on voudrait en déduire m_i, teneur de Vi, ou du moins connaître la loi de m_i liée par x.

x étant fixé, la valeur moyenne de mi, soit m (x) est, d'après la formule (20) :

$$m(x) = \gamma' \left(\frac{x}{\gamma}\right)^{\rho \frac{\sigma_\gamma}{\sigma}} \cdot e^{\frac{\sigma^2\gamma(1-\rho^2)}{2}}$$

mais :

$$\gamma' = \gamma \, e^{-\sigma_\gamma^2/2} \qquad \gamma = m \, e^{-\sigma^2/2}$$

$$\rho = \frac{\sigma_\gamma}{\sigma} \qquad \sigma^2 = \sigma^2_i - \sigma^2_\gamma$$

On en déduit :

(31) $$m(x) = m \left[\frac{x}{m}\right]^{\frac{\sigma^2_\gamma}{\sigma^2}} e^{\frac{\sigma^2_i \, \sigma^2_\gamma}{2\sigma^2}}$$

On peut écrire cette relation sous la forme :

(32) $$\boxed{m(x) = \mathrm{K}x}$$

avec :

(33) $$\boxed{\mathrm{K} = \left[\frac{x}{m}\right]^{-\frac{\sigma^2_i}{\sigma^2}} e^{\frac{\sigma^2_i \, \sigma^2_\gamma}{2\sigma^2}}}$$

A proprement parler, cette relation signifie que si N échantillons prélevés sur N blocs ont tous donné une teneur x, la valeur moyenne des teneurs des N blocs n'est pas x, mais Kx.

Autrement dit *si un échantillon prélevé sur un bloc a une teneur* x, *la teneur probable du bloc*, c'est-à-dire celle sur laquelle il est raisonnable de compter *n'est pas* x *mais* Kx. Un résultat d'apparence aussi surprenante mérite que l'on s'y arrête un moment. Le coefficient K est inférieur à 1 lorsque x est supérieur à $me^{\sigma^2_\gamma/2}$, il est supérieur à 1 dans le cas contraire. Autrement dit, si x est nettement plus faible que la valeur moyenne du gisement, la valeur probable de m_i est supérieure à x et inversement.

Or les blocs Vi ont une variance σ^2_γ plus faible que la variance σ^2 des échantillons v. Globalement, les v sont plus dispersés que les Vi. Autrement dit, *si l'on admet, sans plus, que la teneur* x *de chaque* v *représente convenablement la teneur* m *de chaque* Vi, *on a une estimation correcte de* m, *mais on surestime à la fois le tonnage des basses teneurs et celui des hautes teneurs.* Si x est inférieur à la moyenne m, on doit s'attendre, en moyenne à ce que la teneur m_i du Vi correspondant, soit supérieure à x, et inversement.

Le coefficient de Krige, K n'exprime rien d'autre. Il a pour effet, lorsque l'on veut estimer la teneur d'un panneau déterminé, de corriger l'influence de la surestimation de la dispersion, surestimation résultant

elle-même de ce que l'on a analysé non pas V*i*, mais *v*. Précisons bien qu'*il s'agit d'une corrélation locale*, c'est-à-dire de la correction à appliquer à l'évaluation *x*

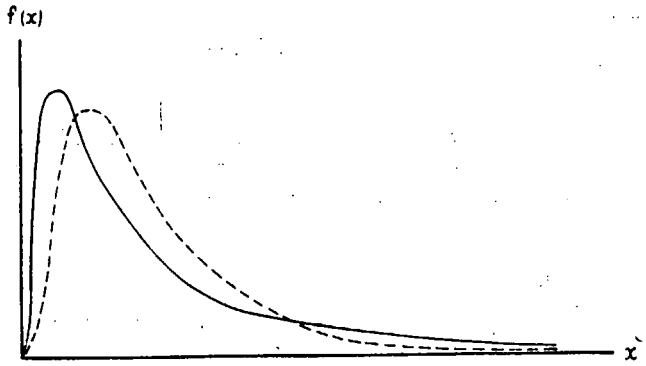

Fig. 4.

de la teneur d'un panneau déterminé pour obtenir une meilleure évaluation. Mais si, à la population des *x*, on substitue la population des K*x*, on commettra une erreur inverse (sous-estimation de la variance). En effet, on voit aisément que la variable aléatoire K*x* est, pour l'ensemble du gisement V, une variable log-normale de valeur moyenne *m*, mais de variance σ^2_γ/σ^2 toujours inférieure à σ^2_γ. La sous-estimation de la variance provient du fait que l'on a négligé la variance $\frac{\sigma^2_i\sigma^2_\gamma}{\sigma^2}$ de m_i à *x* fixé, en identifiant *mi* à sa valeur moyenne K*x*. La somme de ces deux variances est naturellement égale à la variance véritable σ^2_γ de m_i.

Remarques. — 1. Si au lieu d'un échantillon *v* on en prélève plusieurs dans V*i*, la théorie précédente reste valable. La teneur moyenne des *v*, estimée par la formule (5), est, en effet, une variable log-normale, de variance $\frac{\sigma^2_i}{n}+\frac{\sigma^2_i}{2n}$. Il suffit de remplacer σ^2_i par cette valeur dans les formules précédentes, σ^2_γ n'étant pas modifiée.

2. Si au lieu de la valeur probable K*x*, on cherche une valeur m'_i telle qu'il y ait, par exemple, 97,5/100 chances que la valeur véritable lui soit supérieure, il convient de remplacer le coefficient de Krige par le coefficient de sécurité :

$$(34) \qquad K^1 = \left[\frac{x}{m}\right]^{-\frac{\sigma^2_i}{\sigma^2}} e^{-\frac{2\,\sigma_i\,\sigma_\gamma}{\sigma}}$$

En effet, *x* étant fixé, m_i est une variable log-normale de variance $\frac{\sigma^2_i\sigma^2_\gamma}{\sigma^2}$ et de médiane $m\left(\frac{x}{m}\right)^{\frac{\sigma^2_\gamma}{\sigma^2}}$, donc a 97,5 chances sur cent d'être supérieure à

$$m\left(\frac{x}{m}\right)^{\frac{\sigma^2_\gamma}{\sigma^2}} e^{\frac{2\sigma_i\sigma_\gamma}{\sigma}}.$$

Si l'on se contente d'une sécurité de 84/100, on prendra :

$$(35) \qquad K^{11} = \left[\frac{x}{m}\right]^{-\frac{\sigma^2_i}{\sigma_\gamma^2}} e^{-\frac{\sigma_i\,\sigma_\gamma}{\sigma}}$$

Cette correction n'est négligeable que si σ^2_i est très petit, c'est-à-dire si l'échantillonnage est excellent.

5° *Théorie de de Wijs* [1]

Cependant, dans la marche normale de l'échantillonnage on ne connaît pas σ^2_γ. Il est possible de le déterminer, ou plutôt de déterminer σ^2_i, en prélevant un grand nombre d'échantillons normaux *v* dans V*i*. Il en faut au moins une trentaine. Dans le cas de sondages, il n'en est pas question. Il est pourtant possible d'avoir une évaluation *a priori* de σ^2_γ, par le biais de la théorie de de Wijs. De plus, cette théorie nous permettra d'expliquer la permanence de la loi, permanence admise jusqu'ici à titre d'hypothèse.

Le raisonnement de de Wijs est le suivant : considérons un bloc de minerai de teneur *m*, divisons-le par la pensée en deux blocs égaux de teneur x_1 et x_2.

Posons :

$$\delta = \frac{(x_2 - x_1)}{x_2 + x_1} \qquad \text{ou} \qquad \begin{cases} x_1 = (1-\delta)m \\ x_2 = (1+\delta)m \end{cases} \quad (x_2 > x_1)$$

δ *écart relatif* de x_2 et x_1, donne une mesure de la dispersion de x_1 et x_2. C'est un gradient relatif de teneur. Il y a une infinité de manières de couper le bloc initial en deux parties ; à chacune d'elles correspond une valeur de δ. *δ est une variable aléatoire.* Soit *d* sa valeur moyenne. *De Wijs admet que* d *a la même valeur quelle que soit la taille du bloc initial.* En moyenne, l'écart relatif des teneurs d'un bloc et de son voisin dans le gisement est le même que ces blocs fassent 1 kilogramme ou 1 tonne. Cette hypothèse est très importante, et contient en germe le principe de similitude.

Pour aller plus loin, de Wijs admet que l'écart relatif est toujours égal à sa valeur moyenne *d*. Moyennant cette approximation, on obtient, en coupant en deux le bloc initial, deux blocs de teneur :

$$(1+d)m \text{ et } (1-d)m$$

En coupant de même chacun d'eux, on obtient quatre blocs de teneur :

$$(1+d)^2m,\ (1+d)(1-d)m,\ (1-d)(1+d)m,\ (1-d)^2m$$

(1) Références : *f* et *g*.

Après avoir répété k fois cette opération, on obtient 2^k blocs, dont :

1 à la teneur $(1+d)^k m$
k à la teneur $(1+d)^{k-1}(1-d)m$

.....

$C_k^r = \frac{k!}{r!(k-r)!}$ à la teneur $(1+d)^{k-r}(1-d)^r m$

.....

1..... à la teneur $(1-d)^k m$

Par suite, la probabilité pour qu'un bloc ait la teneur :

$$x = (1+d)^{k-r}(1-d)^r m$$

est égale à :

$$\frac{C_k^r}{2^k} = \frac{k!}{2^k r!(k-r)!}$$

Cette probabilité est égale à la probabilité pour tirer r fois pile en jetant k fois de suite une pièce de monnaie (loi dite binomiale). On démontre que, lorsque k est grand, la loi de probabilité de r converge vers une loi normale de variance égale à $\frac{k}{4}$.

Or le logarithme de x est fonction linéaire de r :

$$\mathrm{L}x = \mathrm{L}m + k\mathrm{L}(1+d) + r\mathrm{L}\frac{1-d}{1+d}$$

Lx est donc une variable normale de variance :

$$\sigma^2 = \frac{k}{4}\left[\mathrm{L}\frac{1-d}{1+d}\right]^2 \tag{36}$$

Ainsi, lorsque k est grand, le schéma de de Wijs montre que x est une variable log-normale, dont la variance est donnée par (36).

Permanence de la loi. — Deux blocs v voisins ont, dans le schéma de de Wijs, des teneurs :

$$x_1 = (1+d)^{k-r}(1-d)^r m$$
$$x_2 = (1+d)^{k-r-1}(1-d)^{r+1} m$$

Le bloc $2v$, formé de la réunion des deux précédents, a donc la teneur :

$$x = (1+d)^{k-r-1}(1-d)^r m$$

On voit que les blocs $2v$ suivent la même loi que les blocs v, *à condition de remplacer* k *par* k–1. k étant grand, ces deux lois sont toutes deux des lois log-normales. La permanence est ainsi assurée.

Formule de de Wijs. — La variance est donnée par la formule (36). Par définition de k, le volume V du bloc initial est égal à 2^k fois le volume v, donc :

$$k = \frac{\mathrm{L}\,\mathrm{V}/v}{\mathrm{L}2} \tag{37}$$

Et la variance est égale à :

$$\boxed{\sigma^2 = \alpha \mathrm{L}\,\mathrm{V}/v} \tag{38}$$

avec :

$$\alpha = \frac{1}{4\mathrm{L}2}\left[\mathrm{L}\frac{1-d}{1+d}\right]^2 \tag{39}$$

Le paramètre α — ou dispersion absolue — ne dépend que de d *: c'est donc, comme* d, *une constante caractéristique du gisement,* indépendante à la fois de V et de v.

Remarque. — La théorie précédente suppose que l'on assimile δ à d. Le procédé, qui consiste à remplacer une variable aléatoire par sa valeur moyenne, ne peut conduire qu'à une approximation. Cependant, cette approximation apparaît comme d'autant plus justifiée que la variance propre de δ est plus petite vis-à-vis de la variance σ^2, donc d'autant plus justifiée que k est plus grand. Or c'est seulement pour les grandes valeurs de k que la loi binomiale de de Wijs devient la loi log-normale. Donc, dans la mesure même où la loi est log-normale (et l'expérience montre qu'elle l'est), sa permanence est assurée, et la formule de de Wijs est valable. *Cette formule peut d'ailleurs se déduire directement du principe de similitude,* sans passer par la théorie approximative de de Wijs.

Le principe de similitude (h). — De Wijs a admis que l'écart relatif moyen d des teneurs de deux blocs égaux contigus était le même, quelle que soit la taille de ces blocs. Généralisant cette hypothèse énonçons le principe de similitude :

« La variance logarithmique (c'est-à-dire la dispersion relative) des teneurs des volumes v dans un volume V est la même que celle des teneurs des volumes λv dans un volume λV du même gisement. »

De ce principe résulte que σ^2 est une fonction de la forme $g(\mathrm{V}/v)$. Or la formule de Krige nous a montré que σ^2 est également de la forme $f(\mathrm{V}) - f(v)$. Seule la fonction logarithmique peut se mettre à la fois sous ces deux formes. On déduit donc de ce principe que σ^2 est de la forme :

$$\boxed{\sigma^2 = \alpha \mathrm{L}\frac{\mathrm{V}}{v}} \tag{38}$$

On retrouve bien la formule de de Wijs.

Remarque. — En fait, σ^2 dépend aussi de *la forme* des volumes v et V, et de leur disposition relative. Une même carotte prise parallèlement ou perpendiculairement aux épontes, une carotte allongée et un bloc ramassé de même volume, deux portions de gisement

de même volume V, mais l'une en forme d'amas, l'autre en chevelus ramifiés, ne donneront pas la même variance. Le principe de similitude, au sens strict, doit s'énoncer ainsi :

« Si les figures formées par le volume v dans le volume V d'une part, et le volume v' dans le volume V' d'autre part, sont *géométriquement semblables*, à une translation près de v ou de v', les teneurs v dans V et de v dans V' ont même variance logarithmique. »

En pratique, *des échantillons de forme ramassée prélevés dans un gisement en forme d'amas* (si l'on raisonne sur des volumes), *ou des carottes prélevées dans un gisement étendu dans les deux dimensions* (si l'on raisonne en termes de surface), *vérifieront approximativement la formule de de Wijs.*

Limite de la formule de de Wijs. — Pour v très petit, elle conduit à une variance infinie. Or nous avons démontré l'existence d'une limite supérieure. En fait, lorsque v est petit, la loi cesse d'être log-normale. A la limite, en effet, tous les volumes v sont ou bien du stérile pur, ou bien du minerai pur. En pratique, les échantillons usuels sont assez grands pour que cet effet perturbateur, dit de *noyau métallique*, ne se manifeste pas.

La dispersion absolue α. — Dans la formule (38), ou formule de de Wijs, $L\frac{V}{v}$ est un facteur purement géométrique : α, au contraire, est un coefficient intrinsèque, caractéristique du gisement. C'est la dispersion absolue. Il est possible qu'à chaque type de minéralisation corresponde une valeur déterminée de la dispersion absolue. Il ne sera permis, cependant, de l'affirmer qu'après un énorme travail de compilation statistico-géologique, portant sur les types métallogéniques les plus variés. On conçoit l'intérêt que présenterait une classification des gisements, fondée sur un critère aussi précis. Les minéralisations sulfurées, en Pb, Zn, Cu, montrent des dispersions absolues généralement comprises entre 3/100 et 8/100. Pour des gisements de fer, ou de manganèse, elles sont de l'ordre de 1/1 000, ou 1/10 000.

La dispersion absolue, α, est toujours inférieure à 1. En effet, la variance brute, d'après les formules (6), et (38), est égale à :

$$\Sigma^2 = m^2 (e^{\sigma^2} - 1) = m^2 \left[\left(\frac{V}{v}\right)^{\alpha} - 1 \right]$$

$\alpha = 1$ conduirait à une variance brute en $\frac{1}{v}$, comme dans l'hypothèse où les teneurs des volumes v, voisins dans le gisement, seraient indépendantes. En fait, elles ne peuvent pas être indépendantes; elles sont toujours liées par une corrélation positive, de sorte que α décroît moins vite que $\frac{1}{v}$. On en conclut que l'on a nécessairement :

$$\boxed{\alpha < 1} \tag{40}$$

6° *Démonstration de la permanence de la log-normalité* (*b*)

Il est possible en effet, de démontrer la permanence de la log-normalité, sans passer par l'intermédiaire de la théorie approchée de de Wijs, à condition d'énoncer le Principe de Similitude sous une forme indépendante de toute hypothèse sur la loi théorique. Indroduisons, d'abord, une notion nouvelle, la *teneur relative*. Soit x la teneur d'un volume v prélevé dans un volume V de teneur m. On appelle teneur relative de v par rapport à V le quotient :

$$y = \frac{x}{m} \tag{41}$$

Formulons alors de la façon suivante le principe de similitude : « La teneur relative du volume v par rapport au volume V, et la teneur relative y' du volume v' par rapport au volume V', obéissent à la même loi de probabilité, pourvu que l'on ait :

$$\frac{V}{v} = \frac{V'}{v'}$$

« De plus, la valeur moyenne de y est toujours égale à l'unité, quels que soient v et V ».

Aucune hypothèse n'a été avancée, en ce qui concerne la loi théorique de y : on a simplement supposé que sa valeur moyenne était toujours égale à l'unité, et que sa variance ne dépendait que du rapport $\frac{V}{v}$.

Dans ces conditions, considérons un volume v_1 intérieur à un volume v_2, lui-même intérieur à un volume v_3, etc., tous intérieurs à un volume v_n, les valeurs de ces volumes étant telles que l'on ait :

$$\frac{v_2}{v_1} = \frac{v_3}{v_2} = \cdots = \frac{v_p}{v_{p-1}} = \cdots = \frac{v_n}{v_{n-1}} = \lambda \tag{42}$$

λ étant une constante, on en déduit :

$$v_n = \lambda^{n-1} v_1 \tag{43}$$

Soient y_1 la teneur relative de v_1 par rapport à v_2.
Soient y_2 la teneur relative de v_2 par rapport à v_3.
Soient y_p la teneur relative de v_p par rapport à v_{p+1}.
Soient y_{n-1} la teneur relative de v_{n-1} par rapport à v_n.
Et $x_1, x_2, \ldots x_n$ les teneurs absolues de $v_1, v_2, \ldots v_n$.

Du principe de similitude, et des relations (42), on déduit que les y_p peuvent être considérées comme

autant de valeurs d'une même variable aléatoire Y, de valeur moyenne égale à l'unité. Considérons alors la teneur relative de v_1 par rapport à v_n :

$$y' = \frac{x_1}{x_n} = y_1 y_2 \cdots y_{n-1}$$

Prenons son logarithme :

$$\text{(44)} \qquad \mathrm{L}y' = \sum_{i=1}^{n-1} \mathrm{L}y_i$$

Les volumes v_1 et v_n étant fixés, faisons tendre n vers l'infini, et λ vers 1. Le rapport $v_1/v_2 = \lambda$ tendant vers 1, la variance de la teneur relative y_1 tend vers zéro, car elle s'annule pour $v_1 = v_2$. L'expression (44) est la somme d'un grand nombre de variables aléatoires de variance très petite. La loi de $\mathrm{L}y'$ converge donc vers une loi normale (loi des grands nombres), lorsque n tend vers l'infini. En fait, cette loi, étant indépendante de n, ne peut être que normale. Il en résulte que y' est une variable log-normale.

Si h est la variance, très petite, de $\mathrm{L}y$, la variance logarithmique de y' est :

$$\sigma^2 = (n-1)\, h$$

Mais, d'après la formule (43), on a :

$$n - 1 = \frac{\mathrm{L}v_n/v_1}{\mathrm{L}\lambda}$$

Donc, on peut écrire :

$$\text{(45)} \qquad \sigma^2 = \alpha\, \mathrm{L}\frac{v_n}{v_1} \qquad \alpha = \frac{h}{\mathrm{L}\lambda}$$

On retrouve la formule de de Wijs. Aucune hypothèse n'ayant été formulée, ni sur les volumes v, ni sur la loi des y, on voit que le principe de similitude entraîne les conséquences suivantes :

a. La teneur relative y' de n'importe quel volume v dans n'importe quel volume V est une variable log-normale. Il en est de même de la teneur réelle, $x = my$, puisqu'ici m est une constante. La permanence de la log-normalité est donc démontrée;

b. La variance logarithmique de y, donc aussi celle de x, est donnée par la formule (45), c'est-à-dire par la formule de de Wijs.

En fait, le principe de similitude peut être pris comme point de départ unique d'une théorie purement déductive.

7° *Importance pratique de la formule de de Wijs*

La formule de de Wijs n'est pas seulement une formule théorique, elle a aussi une très grande importance pratique. En effet, de tous les paramètres de la loi log-normale, deux seulement intéressent l'exploitant : d'abord, la teneur moyenne du gisement, ensuite, la dispersion, dans la mesure où cette dispersion donne effectivement une idée de la variation de la teneur d'une zone à l'autre. Par zone, il faut naturellement entendre ici une unité qui soit à l'échelle du mode d'exploitation adopté. Suivant les cas, cette unité peut être de quelques tonnes ou de quelques centaines de tonnes. Or, après une reconnaissance par sondages, la variance calculée représente, non pas la dispersion de semblables unités, mais celles de carottes de quelques kilogrammes. On connaît σ^2, mais non pas σ^2_γ, et σ^2 est toujours plus grande que σ^2_γ, et la différence peut être importante. Si donc on se fonde sur σ^2, on surestimera le tonnage des hautes et des basses teneurs, on sous-estimera les teneurs moyennes. Raisonner sur σ^2 revient à admettre implicitement que le mineur sortira et étalera sur le carreau des petits morceaux de quelques kilogrammes, et qu'il les triera un par un, séparant les morceaux riches et les morceaux pauvres. Le scheidage à main est tout de même un cas exceptionnel. Dans le cas général, on traite plusieurs tonnes à la fois. La confusion de σ^2 et de σ^2_γ — et la pratique habituelle, qui consiste à calculer directement les moyennes pondérées revient aussi à confondre σ^2 et σ^2_γ puisqu'il s'agit des moyennes des teneurs des carottes — conduit à deux sortes d'erreurs :

a. *Une erreur systématique d'échantillonnage,* dont il a été parlé plus haut, et qui conduit invariablement *à la surestimation de la teneur des panneaux riches, et à la sous-estimation de celle des panneaux pauvres.* Cette erreur était bien connue des mineurs du Rand, qui avaient mis au point empiriquement des coefficients correcteurs numériques, dont les valeurs coïncident remarquablement avec celles des coefficients théoriques de Krige, lorsque l'on donne aux paramètres γ et σ les valeurs du Rand.

L'erreur en question porte sur l'estimation de la valeur locale de la teneur dans une zone déterminée. La correction correspondante, ou correction de Krige, donnée par la formule (32) est *une correction locale. Elle nécessite la connaissance de la variance vraie* σ^2_γ.

b. *Une erreur globale* dans le calcul de la répartition du tonnage par tranches de teneurs. On est sûr, pour l'ensemble du gisement, de surestimer le tonnage des hautes et des basses teneurs, et de sous-estimer celui des teneurs intermédiaires. L'erreur est particulièrement grave lorsque la teneur limite d'exploitabilité se situe à l'intérieur de l'intervalle des teneurs les mieux représentées. La figure 5 montre, en traits pleins, la répartition réelle du tonnage (σ^2_γ), et en pointillés la répartition déduite de σ^2. On distingue trois cas possibles :

1° La teneur d'exploitabilité, x_1, est plus faible que

les teneurs courantes du gisement (tout le gisement est exploitable). L'erreur globale porte sur la répartition du tonnage, mais on a une estimation correcte à la fois du tonnage total et de la teneur moyenne. C'est le cas le moins grave;

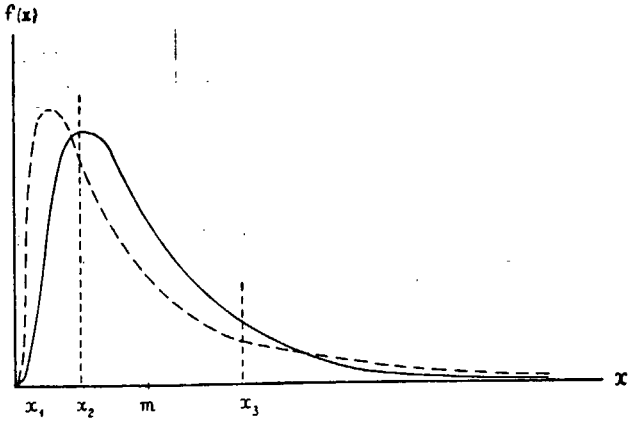

Fig. 5.

2° La teneur d'exploitabilité, x_2, plus faible que la teneur moyenne, se situe cependant à l'intérieur de l'intervalle des teneurs courantes (une partie non négligeable du gisement est en dessous de la teneur limite). Seule est alors intéressante la partie de la courbe située à droite de x_2. L'erreur est double :

— on sous-estime le tonnage exploitable;

— on surestime la teneur moyenne du tonnage exploitable;

3° La teneur d'exploitabilité, x_3, est supérieure à la teneur moyenne (seules les parties très riches du gisement sont exploitables). C'est là le cas le plus grave. En effet :

— on surestime le tonnage exploitable;

— on surestime la teneur moyenne du tonnage exploi table.

Dans ce cas, comme dans le précédent, *la méthode usuelle, consistant à ne retenir que les carottes ayant donné une teneur supérieure à la teneur limite,* et à calculer des moyennes, pondérées ou non, *est incapable de fournir une évaluation correcte du gisement : elle surestime toujours la teneur moyenne, et peut, suivant le cas, sousestimer ou surestimer le tonnage exploitable.*

Cette erreur est toujours très grave dans le cas de la reconnaissance par sondages. Si l'échantillonnage se fait par prélèvements à la volée, et à condition que l'échantillon moyen soit bien représentatif (c'est pratiquement toujours le cas), l'erreur est beaucoup moins grave : une volée, en effet, représente une dizaine de tonnes, c'est-à-dire quelque chose qui est déjà presque à l'échelle de l'exploitation des mines.

Il est extrêmement facile d'éviter cette erreur. La marche à suivre est la suivante :

a. *Ne pas éliminer les carottes à teneur inférieure à la teneur limite.* Prendre, au contraire, toutes les teneurs, ou plus exactement toutes les teneurs constituant la population homogène intéressante (il faut éliminer la frange minéralisée; mais la distinction entre gisement et frange est de nature géologico-statistique, elle est indépendante de toute considération économique d'exploitabilité). On calcule ainsi γ et σ^2, on en déduit la teneur moyenne m du *gisement géologique*;

b. Sachant que la carotte pèse le poids p, que le gisement géologique représente un poids total P, on évalue α par la formule de de Wijs :

$$\alpha = \frac{\sigma^2}{\mathrm{L}\, \mathrm{P}/p} \tag{46}$$

c. Connaissant maintenant la dispersion absolue σ, et sachant que l'unité d'exploitation représente un poids p', on calcule σ^2_γ par la formule de de Wijs :

$$\sigma^2_\gamma = \alpha\, \mathrm{L}\, \frac{\mathrm{P}}{p'} = \sigma^2\, \frac{\mathrm{L}\, \mathrm{P}/p'}{\mathrm{L}\, \mathrm{P}/p} \tag{47}$$

d. C'est maintenant seulement que s'introduit la notion de teneur limite d'exploitabilité. On utilise les formules (15), en remplaçant σ par sa valeur vraie σ_γ, déduite de la formule (47) ci-dessus, et γ par :

$$\gamma' = m\, e^{-\frac{\sigma^2_\gamma}{2}} \tag{48}$$

Si x_0 est la teneur limite, et z_0 la variable réduite correspondante :

$$z_0 = \frac{1}{\sigma_\gamma}\, \mathrm{L}\, \frac{x_0}{\gamma'}$$

le tonnage exploitable est :

$$\mathrm{T} = \mathrm{P}\, \mathrm{G}\, (z_0) \tag{49}$$

et la teneur moyenne du *gisement économique* est :

$$m_0 = m\, \frac{\mathrm{G}\, (z_0 - \sigma_\gamma)}{\mathrm{G}\, (z_0)} \tag{50}$$

C. Les erreurs d'échantillonage

C'est là un sujet classique. En premier lieu, pour illustrer la loi normale, on cite toujours les erreurs d'analyse chimique et les erreurs dues à un prélèvement par pelletées. C'est un premier type d'erreur d'échantillonnage, qui obéit à la loi normale, ou à la loi de Poisson.

En second lieu, et à propos de la théorie de Krige, on a vu qu'il existait une erreur systématique d'échantillonnage, chaque fois que l'on cherchait à évaluer la teneur d'un panneau au moyen d'un, ou de plusieurs échantillons prélevés à l'intérieur de ce panneau. Ce second type d'erreur obéit à la loi log-normale.

Pourquoi deux types d'erreurs, l'un normal, l'autre log-normal? Parce qu'il existe deux grands types d'échantillonnage :

a. Ou bien on prélève, sur la roche ou le minerai en place, un échantillon du type géologique, c'est-à-dire un morceau massif de roche ou de minerai;

b. Ou bien on prélève, sur un tas de minerai ayant subi diverses opérations de concassage, broyage, etc., un échantillon d'un poids déterminé, constitué d'un grand nombre de grains ou de morceaux. Ces grains et ces morceaux n'étaient nullement contigus dans le gisement originel. Il s'agit d'un échantillon moyen.

Dans le premier cas, *le prélèvement respecte la structure naturelle du gisement,* l'échantillon est constitué de particules contiguës, dont les teneurs ne sont pas indépendantes. La loi est log-normale, comme c'est presque toujours le cas lorsque le phénomène présente une organisation : la loi log-normale est la loi des ensembles structurés. Dans le second cas, au contraire, la population à échantillonner a subi diverses opérations de broyage, quartage, etc., qui ont eu pour résultat de séparer irrémédiablement les morceaux et les particules initialement voisins, donc d'effacer à tout jamais les corrélations. *La structure a disparu,* tout morceau est indépendant des autres, et la loi est normale, comme pour tous les phénomènes purement aléatoires (addition d'un grand nombre de petits effets indépendants).

Quand on parle d'erreur d'échantillonnage, il faut bien préciser ce que l'on entend par là, c'est-à-dire préciser quelle est la grandeur réelle que l'on cherche à évaluer approximativement. Dans le premier cas, la carotte de sondage, ou l'échantillon ramassé par le géologue, ont pour but d'évaluer la teneur d'un gisement, ou du moins de la zone, ou portion de gisement, où l'échantillon a été prélevé. Les structures naturelles sont respectées, l'erreur est log-normale, et la formule de de Wijs est applicable. Dans le second cas, sur une volée de volume $v = 5$ mètres cubes par exemple, on a prélevé des pelletées et constitué un échantillon moyen de teneur x. x est une évaluation de la teneur y du volume $v = 5$ mètres cubes, évaluation normale, puisque les structures ont été détruites. Mais, par contre, y, teneur *réelle* du volume v, peut à son tour être considéré comme une évaluation de la zone entourant le volume v, évaluation log-normale cette fois. Examinons ces deux types d'erreurs :

1° ***Échantillon « moyen »*** (normal)

De ce type sont :

Les prélèvements par pelletées à la volée;

Tous les prélèvements effectués, en laboratoire, après broyage des échantillons à analyser.

La teneur x d'un échantillon moyen est une variable normale ayant comme valeur moyenne la teneur réelle y du volume v (teneur moyenne de la volée elle-même, teneur de l'échantillon total reçu par le laboratoire). Sa variance doit être toujours très faible, de façon à ce qu'il n'y ait pratiquement jamais de différence importante entre x et y. Cette variance est en raison inverse du nombre n de particules prélevées, ou de volumes v' de l'échantillon moyen :

$$\Sigma^2 = \frac{A}{n} = \frac{B}{v'}$$

La précision d'un échantillon moyen est proportionnelle à la racine carrée de son volume. En pratique, elle est toujours excellente, de sorte que l'on peut sans crainte assimiler x à y.

2° ***Échantillon « géologique »*** (log-normal)

De ce type sont :

Les échantillons ramenés par les géologues;

Les carottes de sondage;

La teneur réelle y de tout volume v. En pratique, on connaît, non pas y, mais la teneur de l'échantillon moyen x. Cependant, si l'échantillon moyen est bon, on peut admettre que x et y sont égaux.

Cette teneur y est une variable log-normale dans la mesure où elle est considérée comme une évaluation de la zone V*i* entourant le volume v, dans le gisement en place.

Les dispersions de ces échantillons ne sont pas en $1/v$, mais en log v, conformément à la formule de de Wijs. C'est la différence fondamentale avec le cas précédent. *Dix échantillons de 1 kilogramme, prélevés au hasard dans le volume* v (*échantillon moyen*) *ne sont pas du tout équivalents à un échantillon unique de 10 kilogrammes* (*échantillon géologique*) *prélevé dans le même volume. Ils donneront toujours une évaluation plus précise que l'échantillon géologique.* L'échantillon de 10 kilogrammes est un peu plus précis qu'un seul échantillon de 1 kilogramme, mais toujours beaucoup moins que dix échantillons de 1 kilogramme.

3° *Comparaison de différents modes d'échantillonnage*

Un gisement de volume V a été échantillonné au moyen de n_1 échantillons géologiques de volume v_1 (par exemple n_1 carottes de sondage). L'évaluation de la teneur moyenne par la formule :

$$(51) \qquad m = \gamma\, e^{\frac{\sigma^2}{2}}$$

introduit une erreur, de variance logarithmique :

$$(52) \qquad \sigma_1^2 = \frac{\sigma^2}{n_1} + \frac{\sigma^4}{2n_1} = \frac{1}{n_1}\left[\alpha \operatorname{L}\frac{V}{v_1} + \frac{\alpha^2}{2}\left(\operatorname{L}\frac{V}{v_1}\right)^2\right]$$

Si on avait prélevé n_2 échantillons de volume v_2, la variance aurait été :

$$(53) \qquad \sigma_2^2 = \frac{1}{n_2}\left[\alpha \operatorname{L}\frac{V}{v_2} + \frac{\alpha^2}{2}\left(\operatorname{L}\frac{V}{v_2}\right)^2\right]$$

Le meilleur de ces deux échantillonnages est, naturellement, celui qui conduit à la plus petite variance, à prix de revient égal.

a. *Cas où le prix de revient est conditionné essentiellement par le volume total* $n_1 v_1$ *d'échantillons à traiter.* — Raisonnons sur un exemple numérique. Soit un gisement d'un million de tonnes. Est-il préférable de prendre n échantillons de 100 kilogrammes, ou 10 n échantillons de 10 kilogrammes. Dans le premier cas, la variance d'échantillonnage est :

$$\sigma_1^2 = \frac{1}{n}\left[\alpha \log 10^7 + \frac{\alpha^2}{2}(\log 10^7)^2\right] = \frac{1}{n}(7\alpha + 24{,}5\,\alpha^2)$$

dans le second :

$$\sigma_2^2 = \frac{1}{10n}\left[\alpha \log. 10^8 + \frac{\alpha^2}{2}(\log 10^8)^2\right] = \frac{1}{n}(0{,}8\,\alpha + 3{,}2\,\alpha^2)$$

Le second mode est de trois à quatre fois plus précis. Il est toujours préférable de prendre 10 n échantillons de 10 kilogrammes.

b. *Comparaison des puits et des sondages.* — Admettons que le prix au mètre de sondage soit le tiers du prix au mètre de puits. Une carotte de 50 centimètres pèse, disons 5 kilogrammes, une volée de 50 centimètres représente, disons 5 tonnes. Vaut-il mieux, à prix de revient égal, prélever n échantillons de 5 tonnes, ou 3 n échantillons de 5 kilogrammes. En ce qui concerne la volée de 5 tonnes, on considère que l'échantillon moyen qui la représente est très bon, de sorte que l'on connaît la teneur réelle de la volée, et que la volée peut être assimilée à un échantillon « géologique » de 5 tonnes. Supposons que le gisement fasse 50 millions de tonnes.

On trouve, pour les puits :

$$\sigma_P^2 = \frac{1}{n}\left[\alpha \log. \frac{5.10^7}{5} + \frac{\alpha^2}{2}\left(\log. \frac{5.10^7}{5}\right)^2\right] = \frac{1}{n}(7\,\alpha + 24{,}5\,\alpha^2)$$

et pour les sondages :

$$\sigma_S^2 = \frac{1}{3n}\left[\alpha \log. \frac{5.10^{10}}{5} + \frac{\alpha^2}{2}\left(\log. \frac{5.10^{10}}{5}\right)^2\right] = \frac{1}{n}(3{,}3\,\alpha + 16{,}6\,\alpha^2)$$

Quel que soit n et quel que soit α (donc quel que soit le type de la minéralisation) les sondages donneront toujours, à prix égal, une meilleure évaluation : la précision sera de 1,3 à 1,5 fois plus grande avec les sondages.

D. Retour sur le cas où il y a corrélation entre puissance et teneur (b)

Le cas où il existe de telles corrélations n'a pas encore été traité. Il est cependant de la plus haute importance, puisqu'il renferme la quasi-totalité des gîtes filoniens ou alluvionnaires. La pratique habituelle, en pareil cas, consiste à raisonner sur les teneurs pondérées, hx, h étant la puissance minéralisée et x la teneur, ou bien sur les poids métal au mètre carré. Les hx et les h suivent la loi log-normale, et la teneur moyenne du gisement est le rapport du hx moyen au h moyen. Mais ce procédé ne donne pas la dispersion des teneurs elles-mêmes. En fait, on peut distinguer deux cas bien différents du point de vue pratique :

a. *Cas d'un filon très mince.* — La puissance h est presque partout inférieure à la hauteur minimale h' de dépilage. Dans ce cas, la teneur pratique, c'est-à-dire la teneur que l'on observera à l'exploitation, est hx/h'. h' étant une constante, le produit hx doit être considéré comme une véritable teneur. La dispersion de hx, en particulier, n'est pas différente de la dispersion des teneurs et, par suite, l'étude des hx suffit à épuiser le problème;

b. *Cas d'un filon épais.* — La puissance h de la minéralisation est à peu près partout supérieure à la hauteur minimale h' de dépilage, telle qu'elle résulte des conditions techniques de l'exploitation. Il est absolument nécessaire de connaître la loi de répartition de la teneur x

en fonction du tonnage, parce que c'est ici cette teneur x qui fixe la rentabilité du gisement. Si le gisement est reconnu par travaux miniers, on atteint directement cette loi. Dans le cas, étudié ici, de la reconnaissance par sondages, ce que l'on connaît, c'est la loi de répartition de x en fonction de la surface horizontale, et non pas du tonnage.

Comment passer d'une loi à l'autre, et comment appliquer la correction de Krige, telle est la question à laquelle nous allons maintenant tenter de répondre, en transposant en langage log-normal la pratique séculaire des moyennes pondérées.

1° *Loi de répartition de la teneur* x *en fonction du tonnage.*

Le gisement a été reconnu par sondages. Chaque carotte, prise d'une éponte à l'autre, a donné :

— une puissance h;

— une teneur x;

— une teneur « pondérée » $hx = y$.

La loi de répartition de x, c'est-à-dire des teneurs de petits cylindres de section constante, s, mais de hauteur variable, h, est une loi fonction, non du tonnage, mais de la surface horizontale. y, h et x sont lognormaux. Soient γ_y γ_h, γ_x, σ_y, σ_h et σ_x leurs paramètres. Ce ne sont pas des variables indépendantes. Désignons par ρ le coefficient de corrélation, généralement négatif, de h et x. Considérons la tranche de teneur x à $x + dx$, et posons :

$$z = \frac{1}{\sigma_x} \mathrm{L} \frac{x}{\gamma_x} \tag{54}$$

La surface S du gisement étant divisée en un grand nombre de surfaces s (section d'une carotte), la proportion des surfaces s donnant des teneurs comprises entre x et $x + dx$ est :

$$\frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

La puissance moyenne des surfaces s ayant donné cette teneur x d'après la formule (20) :

$$h(x) = \gamma_h \left[\frac{x}{\gamma_x}\right]^{\rho \frac{\sigma_h}{\sigma_x}} e^{\frac{\sigma_h^2 (1-\rho^2)}{2}} = m_h e^{\rho \sigma_h z - \rho^2 \cdot \frac{\sigma_h^2}{2}}$$

m_h n'est autre que la valeur moyenne de h, ou puissance moyenne du gisement. Par suite, le gisement étant divisé en cylindres de section s, le pourcentage, en tonnage, de ces petits cylindres donnant une teneur comprise entre x et $x + dx$, est :

$$f(z)dz = \frac{h(x)}{m_h} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

Soit

$$f(z)dz = \frac{1}{\sqrt{2\pi}} e^{\frac{-(z-\rho\sigma_h)^2}{2}} dz \tag{55}$$

Autrement dit, la loi de x en fonction du tonnage est également log-normale, la variable réduite étant, non plus z, mais :

$$z' = z - \rho\sigma_h = \frac{1}{\sigma_x} \mathrm{L} \left[\frac{x}{\gamma_x e^{\rho \sigma_h \sigma_x}}\right] \tag{56}$$

La formule (56) montre :

— que la variance de x, en fonction du tonnage, est la même que la variance σ^2_x de x en fonction de la surface;

— que la médiane de x, en fonction du tonnage, n'est plus σ_x mais :

$$\gamma'_x = \gamma_x e^{\rho\sigma_x\sigma_h} \tag{57}$$

— que, par suite, la teneur moyenne, m', du gisement, en fonction du tonnage, c'est-à-dire la moyenne pondérée, se déduit de la moyenne non pondérée, m_x, par la formule :

$$m' = m_x e^{\rho\sigma_x\sigma_h} \tag{58}$$

ρ étant généralement négatif, cette moyenne est inférieure à m_x.

Vérification. — m' doit être égale au rapport m_y/m_h (moyenne pondérée) des valeurs moyennes de $y = hx$ et de h. Il en est bien ainsi. En effet, on a d'abord :

$$m_h = \gamma_h e^{\sigma_h^2/2}$$

Quant à $y = hx$, c'est une variable log-normale de paramètres :

$$\begin{cases} \gamma_y = \gamma_h \gamma_x \\ \sigma^2_y = \sigma^2_x + \sigma^2_h = 2\rho\sigma_x\sigma_h \end{cases} \tag{59}$$

donc, de moyenne :

$$m_y = \gamma_y e^{\sigma_y^2/2} = \gamma_h e^{\sigma_h^2/2} \gamma_x e^{\sigma_x^2/2} e^{\rho\sigma_x\sigma_h} = m_x m_h e^{\rho\sigma_h\sigma_x}$$

Par suite, on a bien :

$$\frac{m_y}{m_h} = m_x e^{\rho\sigma_x\sigma_h} = m' \tag{60}$$

Règle pratique. — On calculera d'abord les paramètres log-normaux des trois variables x, h et $y = hx$. La formule (59) permet d'en déduire, sans calcul supplémentaire, le coefficient ρ. Si ce coefficient est nul, il est tout à fait légitime de raisonner sur les teneurs non pondérées, x. S'il n'est pas nul, on calculera, par la formule (5), les valeurs moyennes de hx et de h, soient $m(hx)$ et $m(h)$. La teneur moyenne du gisement est donnée par :

$$m' = \frac{m(hx)}{m(h)} \tag{61}$$

Les paramètres de la loi pondérée se déduisent alors des paramètres de la loi apparente de x, γ_x et σ'_x, par les formules :

$$(62) \qquad \begin{cases} \gamma' = m' e^{-\sigma'^2/2} \\ \sigma'^2 = \sigma_x^2 \end{cases}$$

Ces paramètres doivent ensuite subir l'habituelle correction de de Wijs (voir § E), avant l'application de la formule des tranches de teneurs.

Exemple. — La reconnaissance, par sondage, d'un horizon minéralisé en plomb a conduit aux paramètres suivants :

$$h \begin{cases} \gamma = 1{,}78 \\ \sigma^2 = 0{,}64 \\ \sigma = 0{,}80 \\ m = 2{,}48 \end{cases} \quad hx \begin{cases} \gamma = 6{,}485 \\ \sigma^2 = 0{,}41 \\ \sigma = 0{,}64 \\ m = 7{,}95 \end{cases} \quad x \begin{cases} \gamma = 3{,}51 \\ \sigma^2 = 0{,}107 \\ \sigma = 0{,}328 \\ m = 3{,}70 \end{cases}$$

Il y a, entre x et h, une forte corrélation négative, que la formule (59) permet de calculer :

$$0{,}41 = 0{,}64 + 0{,}107 + 2\,\rho\, 0{,}80.0{,}328$$
$$\rho = -0{,}64$$

La teneur moyenne réelle est égale au rapport du hx moyen au h moyen, c'est-à-dire à :

$$m' = \frac{7{,}95}{2{,}48} = 3{,}20\ \%$$

Les paramètres « pondérés » s'en déduisent ensuite :

$$m' = 3{,}20\ \%\ ,\ \gamma' = m' e^{-\sigma'^2/2} = 3{,}04\ ,\ \sigma'^2 = \sigma_x^2 = 0{,}107$$

Ces paramètres pondérés doivent ensuite subir la correction de de Wijs. Appliquée sous la forme standard, qui consiste à prendre une variance égale à la moitié de la variance des teneurs des carottes, cette correction donne les valeurs définitives suivantes des paramètres de la loi de x :

$$m'' = 3{,}20\ \%, \sigma''^2 = \frac{0{,}107}{2} = 0{,}0535, \gamma'' = m'' e^{-\sigma''^2/2} = 3{,}11$$

2° *Application de la correction de Krige pour les teneurs*

On peut démontrer qu'on a le droit d'appliquer la correction de Krige ou formule (33), aux teneurs x des carottes, à condition de prendre comme valeur de m la teneur pondérée, donnée par la formule (61). Le raisonnement, que nous ne reproduisons pas ici (voir réf. *b*), est calqué sur celui du paragraphe B, 4°, avec quatre variables au lieu de deux, et fait explicitement intervenir le principe de similitude.

La nécessité d'une démonstration rigoureuse est bien mise en évidence par le fait que la valeur probable de x n'est pas égale au rapport des valeurs probables de hx et de h, cette égalité n'ayant lieu que si le coefficient de corrélation de h et de x est nul. Il n'est pas du tout indifférent d'appliquer la correction locale de Krige à x ou à hx. On doit, en fait, l'appliquer à celle de ces deux variables dont l'importance économique est prépondérante; à x, si le filon est épais, et à hx, si le filon est mince.

A titre documentaire, le tableau suivant indique, pour l'exemple étudié au paragraphe précédent, la correspondance entre la teneur x d'une carotte, et la valeur probable, x', de la teneur du panneau centré sur cette carotte, telle qu'elle résulte de l'application, sous la forme standard, de la formule de Krige :

x	x'	x	x'	x	x'
2	2,56	5	4,05	7	4,79
3	3,14	6	4,43	9	5,41
4	3,61	7	4,79	10	5,70

Cas du filon d'épaisseur moyenne. — Le cas du filon d'épaisseur moyenne, c'est-à-dire du filon dont la puissance h est, tantôt inférieure, et tantôt supérieure à la hauteur minimale, h', de dépilage, pose un problème beaucoup plus difficile. En effet, la loi log-normale doit être cassée en deux, la teneur étant hx/h' lorsque h est inférieur à h', et x, lorsqu'il lui est supérieur. Il est possible de calculer la teneur moyenne de l'ensemble, mais non pas la teneur moyenne du tonnage exploitable, c'est-à-dire à teneur supérieure à la limite d'exploitabilité, x_0. Si, en effet, ρ n'est pas nul, la formule des tranches de teneurs n'est pas intégrable.

En ce cas, on effectuera la correction de de Wijs pour les lois de x et de hx, et seul un calcul numérique, arithmétique, effectué sur ces lois corrigées, pourra donner les tonnages et les teneurs moyennes des fractions définies par les inégalités :

$$\begin{cases} h > h' \\ x > x_0 \end{cases} \text{ ou } \begin{cases} h < h' \\ hx > h'x_0 \end{cases}$$

La correction de Krige reste applicable aux x et aux hx individuels. Ce procédé de calcul est très long, mais c'est le seul, dans le cas du filon d'épaisseur moyenne, qui permette d'éviter l'erreur systématique.

III. EXEMPLE DE CALCUL COMPLET

L'EXEMPLE choisi est un gisement de plomb argentifère, reconnu par sondages. Les carottes de sondages ont été coupées en tronçons égaux, de 5 kilogrammes chacun. Chaque carotte a été analysée pour plomb et pour argent. La figure 6 représente le nuage Ag-Pb correspondant. Le tonnage total de minerai a été évalué à un million de tonnes.

1° *Séparation de la frange minéralisée*

Le nuage fait apparaître deux groupes de points : vers les hautes teneurs, une ellipse allongée, marquée G sur la figure 6, et, vers les basses teneurs, un nuage vaguement circulaire, marqué F. F représente la frange minéralisée, G, le gisement géologique. F et G ont en

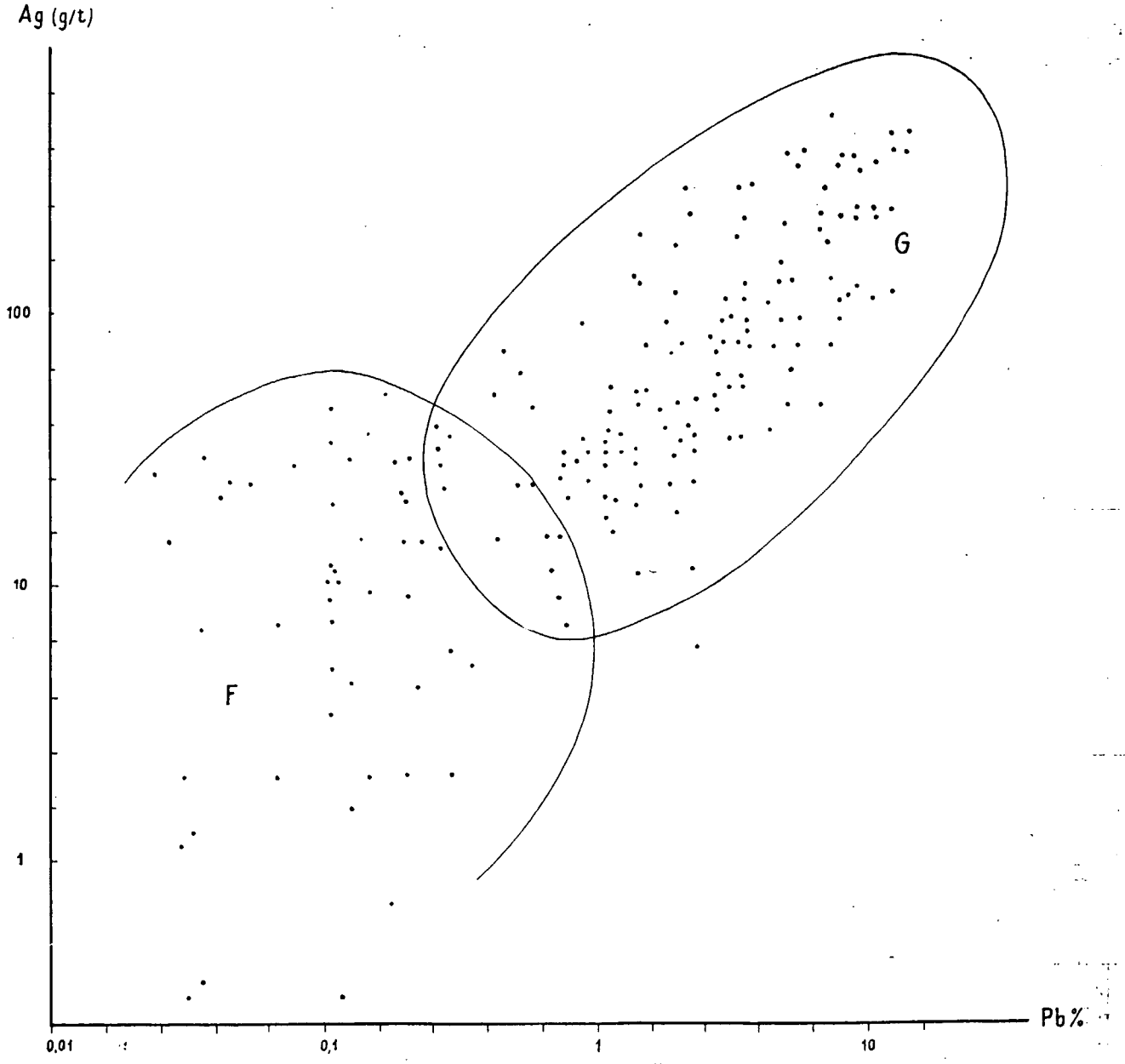

Fig. 6.

commun une petite zone; l'attribution des points correspondant à F ou à G introduit une part d'arbitraire, assez faible d'ailleurs. Le gisement G a été défini par 136 points.

2° *Calcul des paramètres du plomb et de l'argent*

Ce calcul doit être effectué sur les 136 points définissant G. En abscisse et en ordonnée, les divisions de la figure 6 représentent 0,2. en logarithme décimal. Le quadrillage correspondant définit les classes de teneurs, à la fois pour le plomb et pour l'argent. On prend, comme valeur commune de la teneur de chaque classe, la teneur centrale : cette pratique, courante en statistique appliquée, n'introduit pas d'erreur appréciable. On calcule γ, σ et m par les formules (7), (8) et (9) pour le plomb et l'argent et ρ par la formule (16). Les écarts types sont exprimés en logarithmes népériens. On a trouvé ainsi :

$$Pb \begin{cases} \gamma_1 = 2{,}565 \\ \sigma_1^2 = 0{,}82 \\ \sigma_1 = 0{,}91 \\ m_1 = 3{,}87\ \% \end{cases} \qquad Ag \begin{cases} \gamma_2 = 63{,}52 \\ \sigma_2^2 = 0{,}838 \\ \sigma_2 = 0{,}915 \\ m_2 = 91{,}5\ g/t \end{cases}$$

Coefficient de corrélation $\rho = +\ 0{,}817$.

Précision sur m_1 et m_2. — Calculons le coefficient de sécurité de la formule (10) :

$$k_1 = e^{-2\sqrt{\frac{\sigma_1^2}{136} + \frac{\sigma_1^2}{272}}} = e^{-2\sqrt{\frac{0{,}6 + 0{,}25}{100}}} = 0{,}83$$

$$k_2 = e^{-2\sqrt{\frac{\sigma_2^2}{136} + \frac{\sigma_2^2}{272}}} = e^{-2\sqrt{\frac{0{,}61 + 0{,}26}{100}}} = 0{,}83$$

Il y a 97,5 chances sur cent pour que la teneur moyenne en plomb soit supérieure à 3,87 × 0,83 = 3,21 % et autant pour que la teneur en argent soit supérieure à 91,5 × 0,83 = 76 g/t.

Ces évaluations sont sévères. Si l'on admet 16 chances sur cent de se tromper, au lieu de 2,5, le coefficient K est remplacé par :

$$K' = e^{-\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{2n}}}$$

On trouve ainsi qu'il y a 84 chances sur cent pour que les teneurs en Pb et en Ag soient supérieures à 3,50 % et 83,5 g/t respectivement.

3° *Correction de de Wijs*

Ces paramètres sont les paramètres de la loi des carottes. Prenons comme unité d'exploitation la moyenne géométrique des poids de la carotte et du gisement,

$$U = \sqrt{5.10^{-3}10^6} = 71 \text{ tonnes}$$

Les variances des teneurs en plomb et en argent de ces blocs de taille U = 71 tonnes sont égales à la moitié des variances calculées plus haut. On dit que l'on applique la correction de de Wijs *sous la forme standard.* Le coefficient de corrélation, ρ, n'est pas modifié (c'est là une conséquence du principe de similitude). Les paramètres, après correction, deviennent :

$$Pb \begin{cases} \gamma'_1 = m_1 e^{-\frac{\sigma_1^2}{2}} = 3{,}15 \\ \sigma'^2_1 = \frac{\sigma_1^2}{2} = 0{,}41 \\ \sigma'_1 = 0{,}64 \\ m'_1 = 3{,}87\ \% \end{cases} \qquad Ag \begin{cases} \gamma'_2 = m_2 e^{-\frac{\sigma'^2_2}{2}} = 70{,}6 \\ \sigma'^2_2 = \frac{\sigma_2^2}{2} = 0{,}519 \\ \sigma'_2 = 0{,}646 \\ m'_2 = 91{,}5\ g/t \end{cases}$$

$$\rho = +\ 0{,}817$$

4° *Répartition par tranches de teneurs*

Supposons que la teneur limite d'exploitabilité soit de 2,5 %, ce qui signifie que l'on décide de laisser en place, ou, du moins, de mettre au stérile, tout bloc de 71 tonnes dont la teneur est inférieure à 2,5 % en plomb.

La formule des tranches de teneurs, formule (15), permet de calculer la teneur moyenne, en plomb, du minerai conservé et la formule des minerais connexes, formule (23), la teneur moyenne en argent de ce même minerai. Mais les valeurs numériques des paramètres doivent être celles des blocs de 71 tonnes, c'est-à-dire les valeurs corrigées du paragraphe précédent.

La teneur limite étant x_0 = 2,5 %, calculons la variable réduite correspondante, donnée par la formule (2) :

$$z_0 = \frac{1}{\sigma'_1} L \frac{2{,}5}{\gamma'_1} = \frac{2{,}3}{0{,}64} \log \frac{2{,}5}{3{,}15} = -0{,}361$$

On en déduit :

$$z_0 - \sigma'_1 = -0{,}361 - 0{,}64 = -1{,}001$$

Les tables de la loi normale nous donnent :

$$G(-0{,}361) = 0{,}641$$
$$G(-1{,}001) = 0{,}841$$

Le tonnage de minerai est donc :

$$10^6.G(-0,361) = 641\,000 \text{ tonnes}$$

et la teneur moyenne en plomb, d'après la formule (15) :

$$m = 3,87 \frac{0,841}{0,641} = 5,07\ \%$$

Pour avoir la teneur en argent, calculons :

$$z_0 - \rho\sigma_2 = -0,361 - 0,817.0,646 = -0,887$$

Les tables donnent : G (— 0,887) = 0,812. La teneur moyenne en argent est donc, d'après la formule (23) :

$$m = 91,5 \frac{0,812}{0,641} = 116\ g/t$$

En recommençant le calcul pour d'autres valeurs de la teneur limite x_0, puisque celle-ci peut varier en fonction de la conjoncture économique, on obtient le tableau suivant :

x_0	T	Pb %	Ag g/t	T Pb	T Ag
2	761 000	4,62	107	35 200	81,3
2,5	641 000	5,07	116	32 500	74,3
3	530 000	5,50	125	29 200	76,3
3,5	434 000	6,08	135	27 400	58,5

En fonction de la teneur limite x_0, on lit successivement le tonnage de minerai, la teneur en plomb, la teneur en argent, le tonnage plomb métal, et le tonnage argent métal. Pour montrer l'ordre de grandeur de l'erreur systématique d'échantillonnage, on a calculé le même tableau avec les paramètres non corrigés, c'est-à-dire en confondant les lois de répartition des blocs et des carottes.

x_0	T	Pb %	Ag g/t	T Pb	T Ag
2	608 000	5,61	127	34 200	77
2,5	510 000	6,25	141	31 900	72
3	432 000	6,88	152	29 700	65.6
3,5	366 000	7,55	167	27 600	61

On voit l'ordre de grandeur de l'erreur commise. Si la teneur limite est de 2,5 %, les procédés habituels (moyennes pondérées des teneurs supérieures à 2,5 %) laissent espérer une teneur moyenne de 6,25 % en plomb et de 141 g/t en argent. A l'exploitation, on ne trouvera que 5,07 % en plomb et 116 g/t en argent. On trouvera, il est vrai, un tonnage minerai supérieur (641 000 au lieu des 510 000), et par suite à peu près le même tonnage métal. Mais du point de vue économique, c'est surtout la teneur qui fixe la rentabilité du gisement, et tomber de 6,25 % à 5,07 % peut constituer une mauvaise surprise. La correction de de Wijs n'est donc pas seulement une vue de l'esprit.

5° *Correction de Krige*

Cette correction, ou correction locale, a pour but de calculer, en fonction de la teneur x d'une carotte la teneur probable, x', du panneau entourant cette carotte. Si nous fixons à 71 tonnes, comme plus haut, la taille du panneau, de façon à ce que la variance des panneaux soit égale à la moitié de la variance des carottes, la formule de Krige (31) se met sous la forme standard très simple :

$$x' = e^{\sigma^2/8} \sqrt{mx}$$

Ici, on trouve :

$$x' = e^{\frac{0,82}{8}} \sqrt{3,87\,x} = 2,18 \sqrt{x}$$

D'où le tableau suivant :

x	x'	x	x'	x	x'	x	x'
1	2,18	4	4,35	7	5,75	10	6,86
1,5	2,66	4,5	4,60	7,5	5,95	11	7,20
2	3,08	5	4,85	8	6,15	12	7,52
2,5	3,44	5,5	5,10	8,5	6,35	13	7,85
3	3,76	6	5,31	9	6,50	14	8,13
3,5	4,06	6,5	5,54	9,5	6,70	15	8,42

Ce tableau montre qu'on ne saurait trop se méfier des belles teneurs, puisque à une carotte titrant 15 % ne correspond, en moyenne, qu'un panneau de 8,42 %. Par contre, les teneurs faibles ne sont pas aussi mauvaises qu'elles le paraissent, puisque à une carotte ayant donné 2 % correspond, en moyenne, un panneau de 3,08 %.

Alger, mai 1955.

ANNEXE I

ESTIMATION DE LA MOYENNE

Nous avons indiqué, formule (9), l'estimateur :

$$m^* = \gamma^* e^{\frac{\sigma^{*2}}{2}}$$

m^* est une véritable log-normale, de médiane :

$$\gamma' = \gamma e^{\frac{\sigma^2}{2}} = m$$

et de variance :

$$\frac{\sigma^2}{n} + 2\frac{\sigma^4}{(n-1)}$$

Donc, de valeur probable :

$$E(m^*) = m\, e^{\frac{\sigma^2}{2n} + \frac{\sigma^4}{4(n-1)}}$$

différente de la valeur réelle de m : c'est un estimateur biaisé. Sichel (e) donne un estimateur non biaisé assez compliqué. Si l'on pose :

$$V = \frac{n-1}{n}\sigma^{*2}$$

(V est l'estimateur biaisé de la variance), l'estimateur de Sichel a pour expression :

$$t = \gamma^* \left[1 + \frac{V}{2} + \frac{(n-1)V^2}{2^2\, 2!\,(n+1)} + \cdots + \frac{(n-1)^{p-1} V^p}{p!\, 2^p (n+1)(n+3)\ldots(n+2p-1)} \right]$$

Si n est supérieur à 10, on peut négliger les termes en $1/n^2$, et t se met sous la forme suivante :

$$t' = \gamma^* e^{\frac{V}{2}}\left(1 - \frac{V^2}{4n}\right) = m^* e^{-\frac{\sigma^{*2}}{2n} - \frac{\sigma^{*4}}{4(n-1)}}$$

Dès que n est supérieur à 30, il n'y a pratiquement pas de différence entre l'estimateur (9) et t ou t'.

Remarque.

Pour des compléments sur l'estimateur de Sichel, se reporter aux tables et commentaires des pages 25, 26 et 27, du présent numéro.

ANNEXE II

TABLE DE LA FONCTION G(z)

Nous donnons, ci-joint, une table de la fonction :

$$G(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-\frac{z^2}{2}} dz$$

Table peu précise, mais suffisante pour une évaluation sommaire.

Lorsque z est négatif, la valeur de G(z) est donnée par :

$$G(z) = 1 - G(-z)$$

z	G (z)	z	G (z)	z	G (z)
0.00	0.50000	1.00	0.15865	2.00	0.022750
0.05	0.49006	1.05	0.14686	2.10	0.017864
0.10	0.46017	1.10	0.13566	2.20	0.013904
0.15	0.44038	1.15	0.12507	2.30	0.010724
0.20	0.42074	1.20	0.11507	2.40	0.008197
0.25	0.40129	1.25	0.10565	2.50	0.006300
0.30	0.38209	1.30	0.09680	2.60	0.004661
0.35	0.36317	1.35	0.08851	2.70	0.003467
0.40	0.34458	1.40	0.08076	2.80	0.002555
0.45	0.32635	1.45	0.07353	2.90	0.001866
0.50	0.30854	1.50	0.06681	3.00	0.001350
0.55	0.29116	1.55	0.06057	3.10	0.000968
0.60	0.27425	1.60	0.05480	3.20	0.000687
0.65	0.25785	1.65	0.04947	3.30	0.000483
0.70	0.24196	1.70	0.04457	3.40	0.000337
0.75	0.22663	1.75	0.04006	3.50	0.000233
0.80	0.21185	1.80	0.03593	3.60	0.009159
0.85	0.19766	1.85	0.03216	3.70	0.000108
0.90	0.18406	1.90	0.02872	3.80	0.000072
0.95	0.17106	1.95	0.02559	3.90	0.000048
1.00	0.15865	2.00	0.02275	4.00	0.000042

ANNEXE III

TABLE POUR LE COEFFICIENT DE CORRÉLATION

Cette table permet de déterminer, en fonction du nombre n d'échantillons, si l'évaluation r du coefficient de corrélation ρ est, ou non, significativement différente de zéro. Elle donne, *dans l'hypothèse où la valeur réelle de ρ est nulle*, la valeur que r a 95 chances sur cent de ne pas dépasser, c'est-à-dire la valeur au-dessus de laquelle r est significativement différent de zéro, au seuil 95 %.

n	r	n	r	n	r
3	0.997	15	0.514	47	0.288
4	0.950	16	0.497	52	0.273
5	0.878	17	0.482	62	0.250
6	0.811	18	0.468	72	0.232
7	0.754	19	0.456	82	0.217
8	0.707	20	0.444	92	0.205
9	0.666	21	0.433	102	0.195
10	0.632	22	0.423	200	0.140
11	0.602	27	0.381	400	0.095
12	0.576	32	0.349	600	0.082
13	0.553	37	0.325	800	0.071
14	0.532	42	0.304	1.000	0.063

Références

a. M. Allais. — Notes de travail (non publiées).

b. B.R.M.A. — Notes et rapports statistiques divers (non publiés).

c. D. G. Krige. — *A statistical approach to some Mine valuation and allied problems on the Witwatersrand.* Thèse présentée à l'Université du Witwatersrand-Johannesburg, 1951.

d. D. G. Krige. — A statistical analysis of some of the boreholes values in the Orange Free State Goldfield. *Journal of the chemical metallurgical and Mining Society of South Africa,* septembre 1952.

e. H. S. Sichel. — *Mine valuation, and maximum likelihood.* Thèse présentée à l'Université du Witwatersrand-Johannesburg, 1949.

f. De Wijs. — *Statistics of ore distribution,* part I, Geologie in Mijnbouw, novembre 1951.

g. De Wijs. — *Statistics of ore distribution,* part II, id. janv. 1953,

h. G. Matheron. — *Utilité des méthodes statistiques dans la recherche minière.* Congrès du Centenaire de la Société de l'Industrie minérale, Paris, juin 1955.