

Comparaison entre les échantillonnages à poids constant et à effectif constant

par G. MATHERON

Ingénieur des Mines, Docteur ès Sciences

La Revue de l'Industrie Minérale va bientôt publier sous forme d'un ou de deux numéros spéciaux, un ouvrage sur l'Echantillonnage par Pierre GY.

Dans cet ouvrage, Pierre GY s'appuie sur une théorie de l'échantillonnage équiprobable à effectif constant, c'est-à-dire en considérant des échantillons qui sont tous constitués du même nombre de fragments de minerai ; les résultats de l'analyse théorique le conduisent toutefois à justifier l'application de la formule trouvée à un échantillonnage pratique à masse ou volume constant ; par ailleurs, des expériences américaines récentes paraissent pleinement ce point de vue.

G. MATHERON, reprenant le modèle de P. GY, a effectivement montré, par une analyse mathématique rigoureuse, que les deux types d'échantillonnage à masse et à effectif constants conduisaient à des dispersions de la teneur du lot échantillonné caractérisées par des variances équivalentes.

Le niveau mathématique de l'exposé ne sera peut-être pas familier à tous nos lecteurs. Mais il s'agit d'un point très importants de la théorie de l'échantillonnage, que beaucoup d'auteurs ont déjà abordé sans parvenir à l'élucider, et qui, dans le cas particulier de la théorie de P. GY, a soulevé des polémiques passionnées.

La Revue de l'Industrie Minérale s'honore d'apporter à cette discussion la contribution d'une autorité aussi indiscutable que celle de G. MATHERON.

S. I. M.

RESUME

Dans son traité « l'échantillonnage des minerais en vrac », qui doit être prochainement publié par les soins de la Revue de l'Industrie Minérale, P. GY donne l'expression de la variance attachée à une échantillonnage à effectif constant. Dans la pratique, on prélève plutôt des échantillons de poids donné. Nous montrons, dans les pages qui suivent, que ces types d'échantillonnages conduisent à des variances équivalentes.

**

INTRODUCTION

Dans son ouvrage fondamental (1), consacré à l'« échantillonnage des minerais en vrac », P. GY a donné l'expression asymptotique de la variance de l'erreur associée à la prise d'un échantillon d'effectif donné n (échantillon constitué d'un nombre donné n de fragments élémentaires) lorsque n est grand et que le mode de prélèvement vérifie le principe d'équiprobabilité (tous les échantillons possibles de même effectif n sont supposés avoir la même probabilité d'être tirés). La démarche adoptée par GY a fait l'objet de certaines critiques qui, pour l'essentiel, peuvent se ramener aux deux points suivants :

- en ce qui concerne le calcul de la variance elle-même, on a prétendu contester la légitimité de certains développements utilisés par GY.
- plus fondamentalement, la pratique réelle ne consiste jamais à prélever des échantillons d'effectif n donné d'avance, mais toujours des échantillons de poids (ou de volume) donné, dont l'effectif est, par suite, nécessairement aléatoire. On peut se demander, par conséquent, dans quelle mesure les conclusions tirées de l'étude de l'échantillon à effectif donné peuvent se transposer à l'échantillon de poids constant.

Nous nous proposons, dans cette étude, d'examiner ces objections, et de montrer la légitimité des résultats de GY. Or, il apparaît, d'un point de vue purement mathématique, que, si la

théorie de l'échantillon d'effectif donné ne conduit qu'à des calculs somme toute assez élémentaires, celle de l'échantillon de poids constant nécessite, au contraire, la mise en œuvre d'un arsenal mathématique assez notable. On comprend donc que GY, dans un ouvrage destiné à des praticiens, ait préféré suivre la voie la plus rapide. En ce qui nous concerne, pour éviter des complications mathématiques superflues, nous idéaliserons un peu les données du problème en supposant que le lot de minerai à échantillonner peut être considéré comme infiniment grand (ce qui est, d'ailleurs, le plus souvent le cas en pratique). Plus précisément, nous ferons les hypothèses suivantes :

1. Le poids ω d'un fragment de minerai et la quantité q de métal qu'il contient peuvent être considérés comme deux variables aléatoires (non indépendantes). Nous admettrons, de plus, à seule fin de simplifier les notations, que leur fonction de répartition $F(q, \omega)$ possède une densité de probabilité $f(q, \omega)$.
2. Le mode de prélèvement est tel que l'échantillon obtenu peut être considéré comme constitué de la réunion de fragments de minerais de même loi $f(q, \omega)$ tirés au sort successivement et indépendamment les uns des autres. Autrement dit :
 - s'il s'agit d'un échantillon d'effectif donné n , le poids Y_n de cet échantillon et la quantité de métal X_n qu'il contient sont de la forme :

$$(1) \quad \left. \begin{aligned} X_n &= \sum_{i=1}^n q_i \\ Y_n &= \sum_{i=1}^n \omega_i \end{aligned} \right\}$$

(1) P. GY : « L'échantillonnage des minerais en vrac », à paraître par les soins de la Revue de l'Industrie Minérale.

D. 151

chaque q_i étant indépendant de q_j et ω_i pour $j \neq i$ (mais q_i et ω_i ne sont pas indépendants), et chaque couple (q_i, ω_i) possédant la même loi de probabilité $f(q, \omega)$.

— S'il s'agit, au contraire, de l'échantillon de poids donné p , son effectif N , défini par la condition

$$\sum_{i=1}^N \omega_i < p \quad \text{et} \quad \sum_{i=1}^{N+1} \omega_i \geq p$$

apparaît comme aléatoire. Le poids Y_N et la quantité de métal X_N de cet échantillon sont alors définis comme des sommes d'un nombre N aléatoire de variables ω_i ou q_i .

Dans une première partie, nous calculerons l'espérance mathématique et la variance de la

teneur $\frac{X_n}{Y_n}$ de l'échantillon d'effectif n , lorsque n est grand. Dans la deuxième partie, nous

ferons le même calcul pour la teneur $\frac{X_N}{p}$ de l'échantillon de poids p donné, lorsque p est grand, et nous montrerons que les variances correspondantes à ces deux types d'échantillons sont asymptotiquement équivalentes.

I. L'ECHANTILLON D'EFFECTIF CONSTANT

L'effectif n d'un échantillon étant donné, prenons

$$\left\{ \begin{aligned} X &= \frac{1}{n} X_n = \frac{1}{n} \sum_{i=1}^n q_i \\ Y &= \frac{1}{n} Y_n = \frac{1}{n} \sum_{i=1}^n \omega_i \end{aligned} \right.$$

Lorsque n est grand, les variances de X et de Y sont en $\frac{1}{n}$, et les moments centrés

d'ordre supérieur sont au moins en $\frac{1}{n^2}$. La

teneur $\frac{X}{Y}$ de l'échantillon apparaît donc com-

me le quotient de deux variables aléatoires ayant des variances très faibles. Pour calculer son espérance et sa variance, on peut poser :

$$Y = m_y + \varepsilon \quad (m_y = E(Y))$$

et partir du développement :

$$\frac{X}{Y} = \frac{X}{m_y} \left[1 - \frac{\varepsilon}{m_y} + \frac{\varepsilon^2}{(m_y)^2} + \dots + (-1)^n \frac{\varepsilon^n}{(m_y)^n} + \dots \right]$$

utilisable si $\left| \frac{\varepsilon}{m_y} \right| < 1$. S'il y a une probabilité

unité pour que l'inégalité $\left| \frac{\varepsilon}{m_y} \right| < 1$ soit vérifiée,

on peut ensuite prendre l'espérance terme à terme, et en déduire, sous forme de série entière, l'expression de la moyenne et de la variance

$\frac{X}{Y}$. Naturellement, si l'inégalité $\left| \frac{\varepsilon}{m_y} \right| < 1$

n'est pas vérifiée avec une probabilité unité, ce mode de calcul cesse d'être valable, et d'ailleurs les séries entières que l'on obtiendrait formellement par ce procédé sont généralement divergentes. Néanmoins, sous forme de développements limités (et non plus de série entière) les résultats auxquels conduit ce procédé conservent leur validité.

Espérance $E\left(\frac{1}{Y}\right)$

Pour le voir sans développements mathématiques inutiles, intéressons-nous simplement à

l'espérance $E\left(\frac{1}{Y}\right)$ d'une variable Y possédant

une densité $f(y)$, soit

$$E\left(\frac{1}{Y}\right) = \int_{-\infty}^{+\infty} \frac{1}{y} f(y) dy$$

Si $f(0) \neq 0$, cette intégrale est divergente, de sorte que $\frac{1}{Y}$ n'a pas d'espérance mathématique :

Par exemple, si Y est une variable nor-

male, son inverse n'a jamais d'espérance mathématique. Il faut donc bien se garder d'assimiler la loi $f(y)$ à une loi de Gauss. En fait, Y représentant un poids, $f(y)$ n'est différente de 0 que pour $y \geq 0$, et l'intégrale

$$(3) \quad E\left(\frac{1}{Y}\right) = \int_0^{\infty} \frac{1}{y} f(y) dy$$

existera pourvu que la densité $f(y)$ soit un infiniment petit d'ordre $\varepsilon > 0$ en $y = 0$. On montre facilement que cette condition est toujours remplie dans le cas où Y est la somme d'au moins deux variables indépendantes admettant elles-mêmes des lois continues. Dans le problème étudié ici, $E\left(\frac{1}{Y}\right)$ existe donc toujours.

Pour évaluer $E\left(\frac{1}{Y}\right)$ il est commode d'introduire la transformée de Laplace de la loi $f(y)$:

$$\Phi(\lambda) = \int_0^{\infty} e^{-\lambda y} f(y) dy$$

En effet, $\Phi(\lambda)$ existe toujours (pour $\lambda \geq 0$), ainsi que l'intégrale :

$$(4) \quad \int_0^{\infty} \Phi(\lambda) d\lambda = \int_0^{\infty} \frac{1 - e^{-\lambda y}}{y} f(y) dy$$

En faisant tendre μ vers l'infini dans la relation (4), on voit que l'espérance de $\frac{1}{Y}$ existe en même temps que l'intégrale $\int_0^{\infty} \Phi(\lambda) d\lambda$, et que l'on a :

$$E\left(\frac{1}{Y}\right) = \int_0^{\infty} \Phi(\lambda) d\lambda$$

Désignons par m , σ^2 , et μ_n la moyenne, la variance et le moment centré d'ordre n de la variable Y (dont nous supposons l'existence), et par $\Phi_c(\lambda)$ la transformée de Laplace de la loi de la variable centrée $Y - m$:

$$\Phi_c(\lambda) = E[e^{-\lambda(Y-m)}] = e^{\lambda m} \Phi(\lambda)$$

On a par conséquent :

$$(5) \quad E\left(\frac{1}{Y}\right) = \int_0^{\infty} e^{-\lambda m} \Phi_c(\lambda) d\lambda$$

Or, il peut arriver que $\Phi_c(\lambda)$ se développe en série entière sous la forme

$$(6) \quad \Phi_c(\lambda) = 1 + \frac{\lambda^2 \sigma^2}{2!} + \dots + (-1)^k \frac{\lambda^k \mu_k}{k!} + \dots$$

Si l'on porte dans (5) et si l'on intègre terme à terme, on obtient formellement :

$$(7) \quad E\left(\frac{1}{Y}\right) = \frac{1}{m} + \frac{\sigma^2}{m^3} + \dots + (-1)^k \frac{\mu_k}{m^{k+1}} + \dots$$

c'est-à-dire, justement, le résultat que l'on pouvait prévoir à partir du développement (2). Mais :

— Il peut se faire que la série (6) ne soit pas convergente. C'est le cas, par exemple, si Y est une variable lognormale.

— Il peut également se faire que, tout en convergeant, la série (6) ne soit pas uniformément convergente, de sorte que l'intégration terme à terme ne soit pas légitime. C'est ce qui se produit, par exemple, si f est une loi gamma :

$$f(y) = \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by}$$

Pour $a > 1$, $E\left(\frac{1}{Y}\right)$ existe et (6) converge.

Mais la convergence n'est pas uniforme, et on montre facilement que le développement formel

écrit en (7) est divergent ($\frac{\mu_k}{m^k}$ tend vers l'infini avec k).

En général, donc, on devra renoncer à utiliser le développement (7) sous forme de série entière. Mais on pourra le récupérer à titre de développement limité.

En effet, posons

$$(8) \quad \Phi_n(\lambda) = 1 + \lambda^2 \frac{\sigma^2}{2!} + \dots + \frac{(-1)^k}{k!} \lambda^k \mu_k + R_k(\lambda)$$

Le reste $R_k(\lambda)$ de ce développement est :

$$R_k(\lambda) = \int_0^{\infty} \left[e^{-\lambda(x-m)} - 1 - \lambda^2 \frac{(x-m)^2}{2!} - \dots - \frac{(-1)^k \lambda^k}{k!} (x-m)^k \right] f(x) dx$$

Portant dans (5) l'expression (8) de $\Phi_n(\lambda)$, nous obtiendrons :

$$E\left(\frac{l}{y}\right) = \frac{l}{m} + \frac{\sigma^2}{m^2} + \dots + (-1)^k \frac{\mu_k}{m^{k+1}} + R'_k$$

le reste R'_k étant donné par :

$$R'_k = \int_0^{\infty} e^{-\lambda m} R_k(\lambda) d\lambda = \int_0^{\infty} \frac{l}{x} \left(1 - \frac{x}{m}\right)^{k+1} f(x) dx$$

Pour majorer ce reste, prenons un nombre $x > l$, que nous préciserons dans un instant, et écrivons :

$$R'_k = \int_0^{\frac{m}{x}} \frac{l}{x} \left(1 - \frac{x}{m}\right)^{k+1} f(x) dx + \int_{\frac{m}{x}}^{\infty} \frac{l}{x} \left(1 - \frac{x}{m}\right)^{k+1} f(x) dx$$

Pour $x \geq \frac{m}{x}$, on a

$$\left| \frac{l}{x} \left(1 - \frac{x}{m}\right)^{k+1} \right| \leq \frac{x}{m^{k+1}} (m-x)^{k+1}$$

de sorte que la deuxième intégrale est majorée par :

$$\frac{x}{m^{k+1}} E \left[|Y - m|^{k+1} \right]$$

Or si Y est de la forme $\frac{l}{n} \sum_{i=1}^n Y_i$, le moment absolu centré $E \left[|Y - m|^{k+1} \right]$ tend vers 0 lorsque n tend vers l'infini (c'est

généralement un infiniment petit d'ordre $h+1 < k+1$ en $\frac{l}{n}$).

Il reste à majorer la première intégrale de l'expression de R'_k . Si la densité de Y_i est majorée par un nombre B , celle de Y vérifie

$$f(y) \leq (nB)^n \frac{x^{n-1}}{(n-1)!}$$

et on a :

$$\int_0^{\frac{m}{x}} \frac{l}{x} \left(1 - \frac{x}{m}\right)^{k+1} f(x) dx \leq (nB)^n \int_0^{\frac{m}{x}} \frac{x^{n-1}}{(n-1)!} dx = \frac{(nB)^n}{(n-1)(n-1)!} \left(\frac{m}{x}\right)^{n-1}$$

Pour n suffisamment grand, la formule de Stirling permet de remplacer ce terme par un terme en $(B e)^n \left(\frac{m}{\alpha}\right)^{n-1}$, qui tend exponentiellement vers 0 lorsque $n \rightarrow \infty$ pourvu que l'on ait choisi α supérieur à $B m e$ (par exemple $\alpha = 3 B m$).

Finalement donc le reste R'_k est un infiniment petit en $\frac{1}{n}$ de l'ordre de

$\frac{1}{m^{k+2}} E \left[|Y - m|^{k+1} \right]$ (ordre généralement inférieur à $k + 1$). Il n'importe pas que R'_k tende ou non vers 0 pour $k \rightarrow \infty$. Si, pour une valeur donnée de k , on vérifie que R'_k est d'ordre $h + 1$ en $\frac{1}{n}$, on peut, à l'approximation d'ordre h (généralement $< k$) utiliser le développement (9) arrêté au terme en μ_k .

Espérance $E\left(\frac{1}{Y^k}\right)$

On montre, de la même manière, que l'espérance de $\frac{1}{Y^k}$ existe en même temps que l'intégrale $\int_0^\infty \frac{\lambda^{k-1}}{(k-1)!} \Phi(\lambda) d\lambda$ et que l'on a :

$$(10) \quad E \left[\frac{Y^k}{Y^k} \right] = (-1)^k \int_0^\infty \frac{\mu^{k-1}}{(k-1)!} \frac{\partial^k \Phi(\lambda, \mu)}{\partial \lambda^k} d\mu \quad \lambda=0$$

Désignant ensuite par Φ_c la transformée de la loi des variables centrées :

$$\begin{aligned} \Phi_c(\lambda, \mu) &= E \left[e^{-\lambda(X-m_x) - \mu(Y-m_y)} \right] \\ &= 1 + \frac{\lambda^2 \sigma_x^2 + 2 \lambda \mu \sigma_{xy} + \mu^2 \sigma_y^2}{2!} + \dots \end{aligned}$$

il suffit de remplacer Φ par $e^{-\lambda m_x - \mu m_y} \Phi_c$ dans la relation (10) pour obtenir des développements formels, en général divergents, mais récupérables sous forme de développements limités. Ainsi, de :

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

Il suffit ensuite de remplacer $\Phi(\lambda)$ par $e^{-\lambda m} \Phi_c(\lambda)$ pour obtenir, comme ci-dessus, le

développement formel de $E\left(\frac{1}{Y^k}\right)$ généralement divergent, mais récupérable sous forme de développement limité. Par exemple, pour $k = 2$, on obtient :

$$E\left(\frac{1}{Y^2}\right) = \frac{1}{m^2} \left(1 + \frac{3 \sigma^2}{m^2} - 4 \frac{\mu_3}{m^3} + \dots \right)$$

et on en déduit la partie principale de la variance de $\frac{1}{Y}$:

$$D^2 \left(\frac{1}{Y} \right) = \frac{\sigma^2}{m^4} + \dots$$

Etablissement de la formule de GY

Si X et Y sont deux variables aléatoires de loi $f(x, y)$, on introduit la transformée de Laplace

$$\Phi(\lambda, \mu) = \int_0^\infty \int_0^\infty e^{-\lambda x - \mu y} f(x, y) dx dy$$

et on établit, comme ci-dessus, la relation

$$\left(\begin{array}{l} -\frac{\partial}{\partial \lambda} \phi(\lambda, \mu) \Big|_{\lambda=0} = \left[m_x \left(1 + \frac{\mu^2}{2} \sigma_y^2 + \dots \right) - \mu \sigma_{xy} + \dots \right] e^{-\mu m_y} \\ -\frac{\partial^2}{\partial \lambda^2} \phi(\lambda, \mu) \Big|_{\lambda=0} = \left[m_x^2 \left(1 + \frac{\mu^2}{2} \sigma_y^2 \right) - 2 \mu m_x \sigma_{xy} + \sigma_x^2 + \dots \right] e^{-\mu m_y} \end{array} \right.$$

on déduit sans difficulté les premiers termes du développement de $E\left(\frac{X}{Y}\right)$ et $E\left(\frac{X^2}{Y^2}\right)$, soit :

$$(11) \quad \left(\begin{array}{l} E\left(\frac{X}{Y}\right) = \frac{m_x}{m_y} \left[1 + \frac{\sigma_y^2}{(m_y)^2} - \frac{\sigma_{xy}}{m_x m_y} + \dots \right] \\ E\left(\frac{X^2}{Y^2}\right) = \left(\frac{m_x}{m_y}\right)^2 \left[1 + 3 \frac{\sigma_y^2}{(m_y)^2} - 4 \frac{\sigma_{xy}}{m_x m_y} + \frac{\sigma_x^2}{(m_x)^2} + \dots \right] \end{array} \right.$$

et la partie principale de la variance de $\frac{X}{Y}$:

$$D^2\left(\frac{X}{Y}\right) = \left(\frac{m_x}{m_y}\right)^2 \left[\frac{\sigma_x^2}{(m_x)^2} - \frac{2 \sigma_{xy}}{m_x m_y} + \frac{\sigma_y^2}{(m_y)^2} + \dots \right]$$

qui se met également sous la forme :

$$(12) \quad D^2\left(\frac{X}{Y}\right) = \left(\frac{m_x}{m_y}\right)^2 E\left[\left(\frac{X}{m_x} - \frac{Y}{m_y}\right)^2\right]$$

Revenant aux notations initiales (1), on voit que la teneur

$$\frac{X}{Y} = \frac{Y_n}{X_n}$$

de l'échantillon d'effectif n possède une espérance et une variance qui peuvent s'écrire, en négligeant les termes en $\frac{1}{n^2}$, et en posant

$$x_0 = E(q) \text{ et } y_0 = E(m)$$

$$(13) \quad \left(\begin{array}{l} E\left(\frac{X}{Y}\right) = \frac{x_0}{y_0} + \frac{1}{n} \frac{x_0}{y_0} E\left[\frac{m}{y_0} \left(\frac{m}{y_0} - \frac{q}{x_0}\right)\right] \\ D^2\left(\frac{X}{Y}\right) = \frac{1}{n} \left(\frac{x_0}{y_0}\right)^2 E\left[\left(\frac{q}{x_0} - \frac{m}{y_0}\right)^2\right] \end{array} \right.$$

Telles sont bien, en effet, les formules obtenues par GY, et, en particulier, la formule 4-49 de son *chapitre IV*, dont la validité est ainsi établie de manière incontestable. On remarquera que l'espérance $E\left(\frac{X}{Y}\right)$ de la teneur de l'échantillon d'effectif n ne coïncide pas avec la véritable teneur moyenne du lot, qui est $\frac{x_0}{y_0}$, mais en diffère d'une quantité, d'ailleurs très

petite en $\frac{1}{n}$. Il y a toujours un léger biais.

La variance est, comme il se doit, en raison inverse de l'effectif n .

II. L'ÉCHANTILLON DE POIDS DONNE

Dans la pratique, les échantillons effectivement prélevés correspondent à un poids (ou à

un volume) donné d'avance plutôt qu'à un effectif n imposé. Il n'est pas évident que les variances d'échantillonnages soient les mêmes pour ces deux types de prélèvement. Du fait

que la variable $\frac{1}{n} \sum_{i=1}^n \omega_i$ converge presque

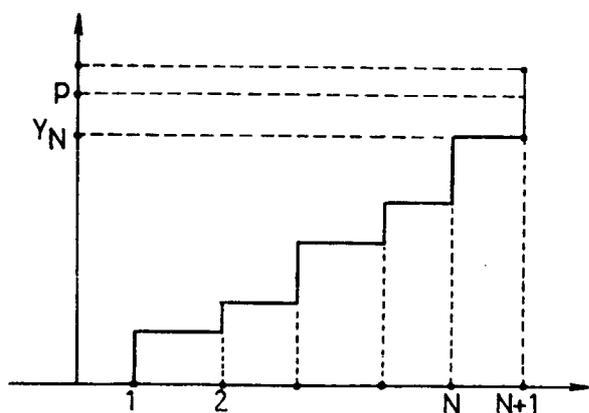
sûrement vers $E(\omega)$ lorsque n tend vers l'infini, on soupçonne bien, comme l'indique GY, qu'il doit en être ainsi, mais il faut le montrer de manière plus précise.

Nous admettons que tout se passe, dans le prélèvement, comme si l'on tirait au sort, indépendamment les uns des autres, des fragments successifs dont les poids ω_i et les quantités de métal q_i possèdent la même loi de probabilité. Ainsi, au tirage numéro n on obtient l'échantillon de caractéristiques

$$\left\{ \begin{array}{l} X_n = \sum_{i=1}^n q_i \\ Y_n = \sum_{i=1}^n \omega_i \end{array} \right.$$

Lorsque n varie, le vecteur (X_n, Y_n) constitue un *processus stochastique* (processus vectoriel, à deux composantes X_n et Y_n , défini sur l'ensemble discret des entiers n positifs. Du fait de l'indépendance des tirages successifs, il s'agit d'un processus markovien à accroissements indépendants et stationnaires).

Chacun des deux composantes X_n et Y_n peut se représenter sous la forme d'un *escalier aléatoire*. L'échantillon d'effectif k donné d'avance, étudié au paragraphe précédent, est défini par



REVUE DE L'INDUSTRIE MINERALE

(X_k, Y_k) , qui est la valeur en $n = k$ du processus (X_n, Y_n) .

L'échantillon de poids p donné d'avance peut être défini de deux manières différentes, soit par défaut, soit par excès. En effet, soit N le temps (la valeur de n) aléatoire pour lequel on a :

$$(14) \quad Y_N < p, \quad Y_{N+1} \geq p$$

Nous dirons que N est l'*effectif* (aléatoire) de l'échantillon de poids donné p . Les inégalités (14) signifient, en effet, que les N premiers fragments constituent un échantillon de poids inférieur à p , et que le poids total des $N+1$ premiers fragments atteint ou dépasse p . L'échantillon lui-même peut être défini soit par défaut — ses caractéristiques sont alors X_N et Y_N — soit par excès — elles sont alors X_{N+1} et Y_{N+1} . Ces deux définitions peuvent être regardées comme pratiquement équivalentes. Les deux échantillons correspondants, en effet, ne diffèrent que par un fragment unique, le fragment tiré à la $N+1$ ème épreuve.

Dans ce qui suit, nous adopteront la définition par excès (X_{N+1}, Y_{N+1}) . Nous examinerons, en premier lieu, la loi de l'effectif aléatoire N de l'échantillon de poids p donné, puis, en deuxième lieu, la loi du métal contenu dans cet échantillon, c'est-à-dire la loi de X_{N+1} . Le poids p donné étant supposé grand (vis-à-vis du poids moyen $y_n = E(\omega)$ des fragments élémentaires) nous chercherons surtout à déterminer les expressions asymptotiques de la moyenne et de la variance de ces différentes variables.

a/ Loi de l'effectif N de l'échantillon de poids donné p

Soit $P_n(p)$ la probabilité d'avoir $N = n$, c'est-à-dire un effectif n . L'événement « $N = n$ » coïncide par définition avec l'événement

$$\text{« } Y_n < p \text{ et } Y_{n+1} \geq p \text{ »}.$$

Désignons par $f_n(y)$ et $F_n(y)$ la densité de probabilité et la fonction de répartition de Y_n . La probabilité de « $Y_n < p$ » est donc $F_n(p)$, et celle de « $Y_{n+1} < p$ » est $F_{n+1}(p)$. Comme l'événement « $Y_n < p$ » est somme logique des événements incompatibles « $N = n$ » et « $Y_{n+1} < p$ », on a :

$$F_n(p) = P_n(p) + F_{n+1}(p)$$

D'où l'on déduit :

$$(15) \quad \begin{cases} P_n(p) = F_n(p) - F_{n+1}(p) \\ P_n(p) = 1 - F_n(p) \end{cases}$$

Il est commode d'introduire la fonction génératrice $G(s; p)$ des $P_n(p)$, qui, d'après (15), se met sous la forme :

$$(16) \quad G(s; p) \equiv \sum_{n=0}^{\infty} s^n P_n(p) = 1 + \sum_{n=1}^{\infty} s^{n-1} (s-1) F_n(p)$$

Il suffit, comme on sait de dériver la fonction génératrice et de faire $s = 1$ pour obtenir les moments successifs de la loi discrète des $P_n(p)$. Les deux premiers moments, en particulier, sont :

$$\begin{cases} E(N) = G'(1) \\ E[N(N-1)] = G''(1) \end{cases}$$

Compte tenu de l'expression (16) de la fonction génératrice, on obtient :

$$(17) \quad \begin{cases} E(N) = \sum_{n=1}^{\infty} F_n(p) \\ E[N(N-1)] = 2 \sum_{n=1}^{\infty} (n-1) F_n(p) \end{cases}$$

Nous sommes ainsi conduits à évaluer, lorsque p est grand, les deux sommes $\sum F_n(p)$ et $\sum n F_n(p)$. Nous y parviendrons grâce à la transformation de Fourier, et en utilisant la règle suivante :

Règle. — Si $h(x)$ est une fonction, et $H(u)$ sa transformée de Fourier (prise, en général, au sens des distributions), on sait que les propriétés de continuité de $H(u)$ donnent une image de la régularité de $h(x)$ à l'infini. En particulier, si la distribution $H(u)$ s'identifie à une fonction continue à croissance lente, $h(x)$ tend vers 0 lorsque x tend vers l'infini.

Soit alors $h(x)$ une fonction identiquement nulle pour $x < 0$, et $H(u)$ sa transformée de Fourier, qui est en général une distribution. Si la distribution

$$H(u) + \frac{a_0}{iu} - \frac{a_1}{(iu)^2} + \frac{a_2}{(iu)^3}$$

s'identifie avec une fonction continue à croissance lente, alors

$$\lim_{x \rightarrow +\infty} \left[h(x) - a_0 - a_1 x - \frac{a_2}{2} x^2 \right] = 0$$

autrement dit, $h(x)$ est asymptotiquement égal au polynôme $a_0 + a_1 x + \frac{a_2}{2} x^2$.

Dans ce qui suit, nous poserons $\lambda = -iu$, ce qui reviendra (formellement) à utiliser la transformation de Laplace. Nous passerons également sous silence certaines difficultés de nature purement mathématique (notamment dans la sommation des séries géométriques du type $\sum [\Phi(u)]^n$, où $\Phi(u)$ est une fonction caractéristique : il est nécessaire, en fait, de supposer que l'inégalité $|\Phi(u)| < 1$ est stricte dès que u n'est pas nulle. Cette condition est d'ailleurs vérifiée par toutes les lois usuelles, à l'exception des lois discrètes, comme la loi de Poisson, où la variable aléatoire n'admet pas d'autres valeurs que des multiples entiers d'une même quantité : ces lois, en effet, possèdent des fonctions caractéristiques périodiques et l'égalité $\Phi(u) = 1$ est possible pour $u \neq 0$).

Pour trouver l'expression asymptotique

$$h(x) \simeq a_0 + a_1 x + \frac{a_2}{2} x^2$$

d'une fonction $h(x)$ (identiquement nulle pour $x < 0$), nous prendrons sa transformée de Laplace $\Phi(\lambda)$, et nous déterminerons les constantes a_0, a_1 et a_2 telles que

$$\Phi(\lambda) - \frac{a_0}{\lambda} - \frac{a_1}{\lambda^2} - \frac{a_2}{\lambda^3}$$

soit une fonction continue en $\lambda = 0$.

Calcul de $E(N)$ et $E(N^2)$

Soit alors $F_1(y)$ la fonction de répartition de Y_1 (poids d'un fragment) et $\Phi(\mu)$ sa transformée

$$\Phi(\mu) = \int_0^{\infty} e^{-\mu y} dF_1(y)$$

La variable Y_n , somme de n variables indépendantes de loi F_1 , obéit à une loi dont la transformée de Laplace est $[\Phi(\mu)]^n$. La transformée

de $F_n(p)$ est donc $\frac{1}{\mu} [\Phi(\mu)]^n$, et la somme

$\sum_{n=0}^{\infty} F_n(p)$ admet la transformée :

$$\frac{1}{\mu} \sum_{n=0}^{\infty} [\Phi(\mu)]^n = \frac{1}{\mu[1 - \Phi(\mu)]}$$

Désignons par $y_0 = E(Y_1)$ et σ_{y^2} la moyenne et la variance du poids Y_1 d'un fragment. Du développement limité

$$\Phi(\mu) = 1 - y_0 \mu + \frac{1}{2} \mu^2 (y_0^2 + \sigma_{y^2}) + \dots$$

on déduit

$$\begin{aligned} & \frac{1}{\mu[1 - \Phi(\mu)]} \\ &= \frac{1}{y_0 \mu^2} + \frac{1}{2 \mu} \left(1 + \frac{\sigma_{y^2}}{y_0^2} \right) + \dots \end{aligned}$$

Par application de la règle énoncée ci-dessus on en tire l'expression asymptotique suivante :

$$\sum_{n=0}^{\infty} F_n(p) \simeq \frac{p}{y_0} + \frac{1}{2} \left(1 + \frac{\sigma_{y^2}}{y_0^2} \right)$$

Reportons-nous à (17), en remarquant que $F_n(p) = 1$. L'espérance $E(N; p)$ de l'effectif de l'échantillon de poids p , qui est évidemment une fonction de p , admet donc l'expression asymptotique :

$$E(N; p) = \frac{p}{y_0} + \frac{1}{2} \left(\frac{\sigma_{y^2}}{y_0^2} - 1 \right)$$

Sa partie principale coïncide, comme il se doit, avec le rapport du poids p donné de l'échantillon au poids moyen y_0 des fragments individuels.

Passons maintenant à la somme $\sum n F_n(p)$ dont la transformée de Laplace relativement à p est :

$$\frac{1}{\mu} \sum_{n=1}^{\infty} n [\Phi(\mu)]^n = \frac{1}{\mu} \frac{\Phi(\mu)}{[1 - \Phi(\mu)]^2}$$

Prenons le développement limité, poussé au troisième ordre, et désignons par a_3 le moment d'ordre 3 de la loi Φ :

$$\Phi(\mu) = 1 - y_0 \mu + \frac{1}{2} \mu^2 [y_0^2 + \sigma_{y^2}] - \frac{1}{3!} a_3 \mu^3 + \dots$$

On obtient sans peine :

$$\frac{1}{\mu} \frac{\Phi(\mu)}{[1 - \Phi(\mu)]^2} = \frac{1}{y_0^2 \mu^3} + \frac{1}{\mu^2} \frac{\sigma_{y^2}}{y_0^3} + \frac{1}{\mu} \left[\frac{1}{4} + \frac{\sigma_{y^2}}{y_0^2} + \frac{3}{4} \frac{\sigma_{y^4}}{y_0^4} - \frac{1}{3} \frac{a_3}{y_0^3} \right] + \dots$$

En appliquant la règle déjà utilisée, on obtient l'expression asymptotique suivante :

$$\sum_{n=1}^{\infty} n F_n(p) \simeq \frac{1}{2} \frac{p^2}{y_0^2} + p \frac{\sigma_{y^2}}{y_0^3} + \frac{1}{4} + \frac{\sigma_{y^2}}{y_0^2} + \frac{3}{4} \frac{\sigma_{y^4}}{y_0^4} - \frac{1}{3} \frac{a_3}{y_0^3}$$

En fait, nous n'aurons besoin que des termes en p^2 et p et nous laisserons de côté le terme constant. Portant ce résultat dans (17), et compte tenu de l'expression déjà trouvée en (18) pour $E(N)$, on obtient facilement :

$$E(N^2) = \frac{p^2}{y_0^2} + \left(2 \frac{\sigma_{y^2}}{y_0^2} - 1 \right) \frac{p}{y_0}$$

Elevant ensuite (18) au carré et retranchant de l'expression ci-dessus, on fait apparaître

l'expression asymptotique de la variance de l'effectif de l'échantillon de poids p :

$$(19) \quad D^2(N; p) = \frac{p}{y_0} \frac{\sigma_{y^2}}{y_0^2}$$

Elle est proportionnelle à $\frac{p}{y_0}$, c'est-à-dire, au premier ordre, à $E(N; p)$.

b) Loi de la quantité de métal X_{N+1} de l'échantillon de poids p

Désignons maintenant par $f(x,y)$ la densité de probabilité des caractéristiques (X_i, Y_i) d'un fragment individuel, et par $\Phi(\lambda,\mu)$ sa transformée de Laplace

$$\Phi(\lambda,\mu) = \int_0^\infty \int_0^\infty e^{-\lambda x - \mu y} f(x,y) dx dy$$

Nous devons déterminer la densité $g(x; p)$ de la quantité de métal X_{N+1} de l'échantillon (par excès) de poids donné p et d'effectif aléatoire N .

Pour expliciter cette loi, et, en particulier, pour trouver l'expression asymptotique de la moyenne et de la variance de X_{N+1} , nous utiliserons la transformée de Laplace $\Gamma(\lambda,\mu)$ de la fonction $g(x,p)$, prise relativement aux deux variables x et p :

$$\Gamma(\lambda,\mu) = \int_0^\infty \int_0^\infty g(x;p) e^{-\lambda x - \mu p} dx dp$$

A cette fin, nous profiterons du fait que le processus (X_n, Y_n) est à accroissements indépendants et stationnaires, de telle sorte que, pour tout $n > 1$, le vecteur $(X_n - X_1, Y_n - Y_1)$ est indépendant de (X_1, Y_1) et possède la même loi de probabilité que (X_{n-1}, Y_{n-1}) . L'échantillon de poids donné p possède l'effectif $N = 0$ si le premier fragment possède un poids $Y_1 \geq p$. Si au contraire le premier fragment possède un poids $Y_1 = \eta < p$, et une quantité de métal $X_1 = \varepsilon$, la loi conditionnelle de probabilité de l'échantillon de poids p (liée par $X_1 = \varepsilon$ et $Y_1 = \eta$) possède la densité $g(x - \varepsilon; p - \eta)$. On en déduit l'équation intégrale :

$$(20) \quad g(x, p) = \int_p^\infty f(x, y) dy + \int_0^x d\varepsilon \int_0^{p-\varepsilon} g(x - \varepsilon; p - \eta) f(\varepsilon, \eta) d\eta$$

Si nous appliquons la transformation de Laplace (en x et p) aux deux membres de l'équation (20), les produits de convolution qui y figurent sont remplacés par des produits multiplicatifs ordinaires, et on obtient

$$(21) \quad \Gamma(\lambda,\mu) = \frac{1}{\mu} [\Phi(\lambda,0) - \Phi(\lambda,\mu)] + \Gamma(\lambda,\mu) \Phi(\lambda,\mu)$$

De (21), on déduit immédiatement l'expression de la transformée de Laplace $\Gamma(\lambda,\mu)$ de la densité $g(x, p)$:

$$(22) \quad \Gamma(\lambda,\mu) = \frac{1}{\mu} \frac{\Phi(\lambda,0) - \Phi(\lambda,\mu)}{1 - \Phi(\lambda,\mu)}$$

En dérivant en λ cette transformée et en faisant $\lambda = 0$, nous obtiendrons les transformées de Laplace des espérances $E(X; p)$ et $E(X^2; p)$, qui sont des fonctions de la seule variable p . Pour abrégier l'écriture, nous désignerons par X (au lieu de X_{N+1}), la quantité de métal de l'échantillon par excès de poids p . La règle déjà utilisée plus haut va donc nous permettre, à partir de leurs transformées de Laplace, de former les expressions asymptotiques de ces deux espérances, et d'en déduire la variance de l'échantillon de poids p .

Calcul de $E(X; p)$

Dérivant (22) une fois en λ et faisant $\lambda = 0$, nous obtenons la transformée de $E(X; p)$ sous la forme :

$$-\frac{1}{\mu} \frac{\Phi'_{\lambda}(0,0)}{1 - \Phi(0,\mu)} = \frac{x_0}{\mu} \sum_{n=0}^{\infty} [\Phi(0,\mu)]^n$$

x_0 désigne l'espérance $E(X_1)$ de la quantité de métal contenue dans un fragment. De même $\sigma_{x_1}^2$ désigne la variance de X_1 . Comme

$$\frac{1}{\mu} \sum_{n=0}^{\infty} [\Phi(0,\mu)]^n$$

est la transformée de $F_n(p)$, nous obtenons ainsi, d'après (17) :

$$(23) \quad E(X; p) = x_0 [1 + E(N; p)]$$

Ainsi, l'échantillon par excès (dont l'effectif est $N+1$ et non pas N) contient, en moyenne, une quantité de métal proportionnelle à son effectif. Vis-à-vis des espérances, la condition de liaison que l'on impose à l'échantillon en fixant son poids p ne se manifeste par aucun effet : tout se passe comme si l'on avait pré-

levé un effectif donné $n = 1 + E(N; p)$.
Compte tenu de (18), on obtient également :

$$(24) \quad E(X; p) = \frac{x_0}{y_0} p + \frac{1}{2} x_0 \left(1 + \frac{\sigma_y^2}{y_0^2} \right)$$

L'espérance $\frac{1}{p} E(X; p)$ de la teneur $\frac{X}{p}$ de l'échantillon de poids p diffère de la vraie teneur $\frac{x_0}{y_0}$ d'une quantité toujours positive, d'ailleurs petite (en $\frac{x_0}{p}$) :

Ce léger biais positif s'explique aisément si l'on se souvient qu'il s'agit de l'échantillon par excès, dont le poids réel Y_{N+1} est toujours légèrement supérieur à p .

Calcul de $E(X^2; p)$

On obtient la transformée de Laplace du moment d'ordre 2 en dérivant deux fois $\Gamma(\lambda, \mu)$ en λ avant de faire $\lambda = 0$, ce qui donne d'après (22) :

$$\frac{1}{\mu} \frac{\Phi''_{\lambda} 2(0,0)}{1 - \Phi(0,\mu)} + \frac{2}{\mu} \frac{\Phi'_{\lambda}(0,0) \Phi'_{\lambda}(0,\mu)}{[1 - \Phi(0,\mu)]^2} = \frac{1}{\mu} \frac{x_0^2 + \sigma_x^2}{1 - \Phi(0,\mu)} - \frac{2 x_0}{\mu} \frac{\Phi'_{\lambda}(0,\mu)}{[1 - \Phi(0,\mu)]^2}$$

Le premier terme ne diffère que par un facteur constant de celui qui nous a servi au calcul de $E(X; p)$. Il correspond donc, dans l'expression de $E(X^2; p)$ à un terme dont la partie principale en p est :

$$(x_0^2 + \sigma_x^2) \frac{p}{y_0}$$

Développons le deuxième terme. On trouve sans peine :

$$- \frac{2 x_0}{\mu} \frac{\Phi'_{\lambda}(0,\mu)}{[1 - \Phi(0,\mu)]^2} = \frac{2 x_0^2}{\mu^2 y_0^2} \left[1 + \mu \left(\frac{\sigma_y^2}{y_0} - \frac{\sigma_{xy}}{x_0} \right) + \dots \right]$$

L'expression asymptotique correspondante est donc (nous nous limitons aux deux termes principaux en p^2 et p)

$$\frac{x_0^2}{y_0^2} p^2 + 2 \frac{x_0^2}{y_0^2} \left(\frac{\sigma_y^2}{y_0} - \frac{\sigma_{xy}}{x_0} \right) p$$

Regroupant les deux termes ainsi obtenus, nous trouvons :

$$(25) \quad E(X^2; p) = \frac{x_0^2}{y_0^2} p^2 + x_0^2 \frac{p}{y_0} \left(1 + \frac{\sigma_x^2}{x_0^2} + 2 \frac{\sigma_y^2}{y_0^2} - 2 \frac{\sigma_{xy}}{x_0 y_0} \right)$$

Calcul de la variance de l'échantillon de poids p

Il suffit d'élever (24) au carré (en se limitant aux termes en p^2 et p) et de retrancher de (25) l'expression ainsi obtenue pour faire apparaître la partie principale de la variance :

$$D^2(X; p) = p \frac{x_0^2}{y_0} \left[\frac{\sigma_x^2}{x_0^2} + \frac{\sigma_y^2}{y_0^2} - 2 \frac{\sigma_{xy}}{x_0 y_0} \right]$$

L'expression entre crochets peut être remplacée par :

$$E \left[\left(\frac{X_i}{x_0} - \frac{Y_i}{y_0} \right)^2 \right] = E \left[\left(\frac{q}{x_0} - \frac{\sigma}{y_0} \right)^2 \right]$$

puisque $Y_i = \sigma$ est le poids d'un fragment élémentaire et $X_i = q$ la quantité de métal qu'il contient. Ainsi la teneur $\frac{X}{p}$ de l'échantillon de poids donné p possède la variance :

$$\frac{1}{p^2} D^2(X; p) = \frac{y_0}{p} \frac{x_0^2}{y_0^2} E \left[\left(\frac{q}{x_0} - \frac{\sigma}{y_0} \right)^2 \right]$$

Enfin, au premier ordre en $\frac{1}{p}$, on peut,

d'après (18), remplacer $\frac{y_0}{p}$ par $\frac{1}{E(N; p)}$.

Il vient donc, finalement :

$$(26) \quad \frac{1}{p^2} D'(X; p) = \frac{1}{E(N; p)} \left(\frac{x_0}{y_0} \right)^2 E \left[\left(\frac{q}{x_0} - \frac{\sigma}{y_0} \right)^2 \right]$$

Reportons-nous à la formule (13). Nous constatons immédiatement que l'échantillon de poids donné p possède la même variance que l'échantillon d'effectif donné n , à condition de prendre pour n l'espérance $E(N; p)$ de l'effectif (aléatoire) de l'échantillon de poids p . Ce résultat est valable si p ou $E(N)$ sont suffisamment grands pour que l'on puisse négliger

les termes en $\frac{1}{p^2}$ ou $\frac{1}{[E(N)]^2}$.

Ainsi se trouve complètement achevée la justification du mode de calcul adopté par P. Gy. Pour notre part, nous sommes heureux d'avoir pu dans cette étude, si aride soit-elle, apporter notre contribution à ce travail fondamental, et réfuter certaines critiques dont il a été l'objet.

Comptes rendus des ouvrages

OFFERTS A LA BIBLIOTHEQUE DE LA SOCIETE

La Société publie un compte rendu de chacun des ouvrages qui lui sont adressés pour sa bibliothèque en double exemplaire. Les ouvrages qui ne lui parviennent qu'en un seul exemplaire sont simplement mentionnés, à moins qu'ils ne soient accompagnés d'un compte rendu de l'éditeur qui, dans ce cas, peut être reproduit.

BERG-UND HUTTENMANNISCHER TAG VOM 20 BIS 23 MAI 1964 IN FREIBERG/SA (Journées Minières et Métallurgiques de mai 1964, à Freiberg). 177 pages, 95 figures, 32 tableaux. Prix : 30,50 MDN.

Théorie de la flottation des minéraux sulfurés et oxydés.

Base de la préparation des sels solubles par flottation et champ électrique.

Cynétique de la flottation.

Différents problèmes de la préparation des pierres et des terres.

ATLAS DES PRINCIPALES PARAGENESSES MINÉRALES VUES AU MICROSCOPE, par O. CELSNER. Traduit par R. MIGNON. Gauthier-Villars, éditeur. 309 pages, avec planches et photographies. Prix : 150 F.

Les indications d'un texte serré et des tableaux schématiques complètent une illustration très abondante. Une grille transparente permet de retrouver un minéral déterminé dans une section polie à l'aide de coordonnées fournies par une feuille de travail.

Les regroupements systématiques de données et les critères de reconnaissance seront appréciés aussi bien par l'utilisateur que par l'autodidacte et l'étudiant par correspondance, auquel l'atlas s'adresse en premier lieu.

GÉOLOGIE DU CHARBON ET DES BASSINS HOUILLERS par L. BARRABE et P. FEYS. Masson, éditeur. 220 pages, 55 figures. Prix : 42 F.

La géologie joue un rôle de premier plan dans la connaissance des gisements, la marche de leur exploitation et la recherche de leurs prolongements.

Mais cette branche de la géologie était devenue difficilement accessible aux non spécialistes, en raison même de ses récents progrès et du foisonnement bibliographique. Cet ouvrage essaie de combler une lacune des bibliothèques françaises. Il se veut d'abord un précis clair et à jour, à l'usage des étudiants, des naturalistes, des ingénieurs.

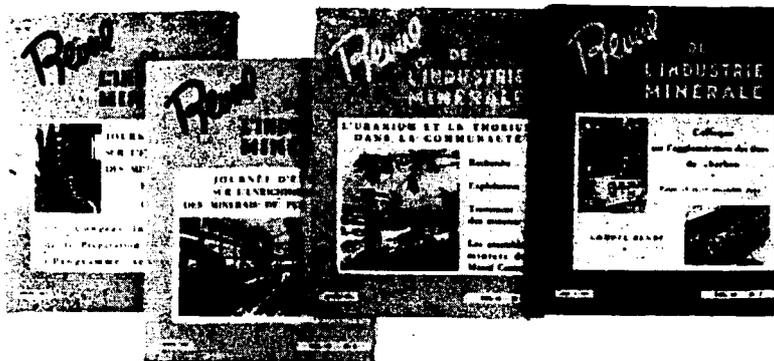
GRANDES DIVISIONS DE L'OUVRAGE

Classification des charbons. — Tourbes et tourbières. — Les houilles. — Les lignites. — Charbons particuliers riches en matières volatiles. — Evolution des charbons. — Répartition des charbons dans le monde. — Description des bassins houillers. — Description des gisements de lignite. — Bogheads. — Schistes bitumineux. — Relations entre les charbons, les schistes bitumineux et les pétroles. — Le charbon dans l'économie mondiale.

Bibliographie.

Avril 1966

NUMEROS SPECIALISES DE LA REVUE DE L'INDUSTRIE MINERALE



- Colloque sur la pétrologie des charbons** (15 juillet 1958), 224 pages, 9 F + 1 F de port.
- Colloque sur l'agglomération des fines de charbon** (janvier 1960), 168 pages, 9 F + 1 F de port.
- L'Uranium et le thorium dans la Communauté** (octobre 1959), 132 pages, 9 F + 1 F de port.
- Journées d'études sur l'enrichissement des minerais de fer français** (janvier et février 1963), 164 pages, 18 F + 2 F de port.
- Méthodes et matériel d'exploitation des Houillères en U.R.S.S.** (septembre 1959), 88 pages, 9 F + 1 F de port.
- Le creusement des voies au rocher et au charbon** (mai 1958), 132 pages, 9 F + 1 F de port.
- Compte-rendu de la Troisième Conférence internationale sur les pressions de terrains** (mars 1961), 728 pages reliées pleine toile pagamoïde, 55 F + 4 F de port.
- Journée Européenne d'information sur le grisou** (février 1964), 94 pages, 9 F + 1 F de port.
- L'Électronique et les techniques de gestion dans l'Industrie minière** (septembre et novembre 1964), 214 pages, 18 F + 2 F de port.
- Compte rendu de la réunion de novembre 1963 de la Section de préparation des minerais** (janvier 1965), 104 pages, 9 F + 1 F de port.
- Compte rendu de la réunion de novembre 1964 de la Section de préparation des minerais** (mai et juin 1965), 102 pages, 18 F + 1 F de port.
- Compte rendu de la réunion de décembre 1965 de la Section de préparation des minerais** (mai et juillet 1966), 216 pages, 18 F + 2 F de port.

Demandez ces ouvrages à l'adresse ci-après :

SOCIÉTÉ DE L'INDUSTRIE MINÉRALE

19, rue du Grand-Moulin, 42 SAINT-ETIENNE (Loire) — Tél. (77) 32-46-13 — C.C.P. : LYON 54-08

Les règlements peuvent être faits :

en espèces

par virement postal C/C Lyon 54-08

par chèque bancaire

par virement bancaire, Compte N° 90 100 B à l'agence de St-Etienne du Crédit Lyonnais.

REVUE DE L'INDUSTRIE MINÉRALE