JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

J. ULMO

# Quelques applications des méthodes statistiques dans la recherche industrielle

*Journal de la société statistique de Paris*, tome 106 (1965), p. 237-248
<http://www.numdam.org/item?id=JSFS_1965__106__237_0>

© Société de statistique de Paris, 1965, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
*Article numérisé dans le cadre du programme*
*Numérisation de documents anciens mathématiques*
http://www.numdam.org/

# III

# QUELQUES APPLICATIONS DES MÉTHODES STATISTIQUES DANS LA RECHERCHE INDUSTRIELLE

### INTRODUCTION

L'objet de cet exposé n'est pas de décrire les principales applications des méthodes statistiques dans la recherche industrielle, ni même dans la recherche sidérurgique.

J'ai seulement voulu décrire dans leurs grandes lignes quelques applications que j'ai choisies en raison de leur caractère représentatif des situations que nous avons rencontrées et des techniques que nous employons pour y faire face.

## I — ESSAIS COMPARATIFS DE QUELQUE IMPORTANCE, NOTAMMENT ESSAIS COOPÉRATIFS A L'ÉCHELLE INTERNATIONALE

Dans ce type d'essais où les conditions d'exécution sont très variables en raison de la grande dispersion géographique des laboratoires ou usines chargés de les exécuter et où le nombre des participants est souvent assez élevé, il apparaît indispensable d'établir un plan des essais préalablement à leur réalisation si l'on veut pouvoir en analyser et interpréter les résultats de façon valable.

La phase de préparation du plan d'essais qui oblige les techniciens à envisager par avance et en détail l'ensemble du déroulement des opérations et à préciser leurs objectifs conduit souvent à des échanges de vue fructueux.

Voici deux exemples d'essais de ce type.

### I. A) ÉTUDE DE LA RÉSISTANCE A LA FATIGUE DE JOINTS SOUDÉS

1° *Généralités sur les essais de fatigue et leur interprétation*

L'interprétation des résultats des essais de fatigue est un problème difficile pour la solution duquel tout le monde s'accorde à penser que l'emploi de la statistique est indispensable.

Les essais de fatigue consistent à faire subir à des éprouvettes des sollicitations cycliques dont l'amplitude S varie avec l'éprouvette. Les essais sont poursuivis jusqu'à la rupture de l'éprouvette dont l'endurance est ainsi caractérisée par le nombre N des cycles subis, à condition que N ne dépasse pas une certaine limite au delà de laquelle l'essai est suspendu, sa durée devenant prohibitive. (Cette limite est en général de l'ordre de quelques millions de cycles.)

La durée de vie des éprouvettes soumises à des cycles de même amplitude S varie considérablement d'une éprouvette à l'autre et l'expérience montre que les éprouvettes non rompues au bout de $10^6$ ou $10^7$ cycles, dont la proportion augmente quand l'amplitude S des sollicitations diminue, ne rompent généralement pas, aussi loin qu'on poursuive l'essai, en sorte qu'on peut considérer leur durée de vie comme infinie.

On peut alors définir la limite d'endurance au risque $\alpha$ comme la valeur $E_\alpha$ de l'amplitude pour laquelle la proportion des éprouvettes de durée de vie infinie est égale à $1 - \alpha$, c'est-à-dire comme la valeur de l'amplitude S pour laquelle le quantile d'ordre $\alpha$ de la distribution des durées de vie est infini.

L'expérience montre que si on poursuit les essais jusqu'à un nombre donné $N_0$ de cycles, la courbe représentative des variations de la probabilité de rupture en fonction de l'amplitude S des sollicitations peut être assimilée à une courbe en S de Galton, forme qu'on rencontre fréquemment pour représenter les variations en fonction d'un paramètre (ici S) de la probabilité d'un phénomène se traduisant par une alternative (la rupture ou la non rupture).

Par ailleurs, à partir du moment où $N_0$ est assez élevé (de l'ordre de $10^6$ ou $10^7$), comme nous l'avons déjà dit, cette courbe ne dépend pas de $N_0$, en sorte qu'il suffit d'estimer les valeurs des deux paramètres « médiane $E_{50}$ et écart type $\sigma$ » de cette courbe pour en déduire une estimation de $E_\alpha$.

Pratiquement on adopte comme limite d'endurance l'endurance médiane $E_{50}$, valeur de S pour laquelle la moitié des éprouvettes ont une durée de vie infinie; c'est en effet celle que l'on peut estimer le plus facilement et avec le plus de précision.

L'estimation de la limite d'endurance et de l'écart type $\sigma$ des endurances limites relève des techniques « d'analyse des probits » que je ne décrirai pas ici [2, 3, 4].

Je ne voudrais pas terminer ces généralités sur les essais de fatigue sans ajouter que la distribution des durées de vie des éprouvettes sous une contrainte d'amplitude S donnée et son évolution avec S ont été étudiées à l'I. R. S. I. D. par M. F. Bastenaire dont les travaux [5, 6] constituent un bel exemple de construction d'un modèle mathématique cohérent pour représenter un phénomène physique complexe.

Il fallait en effet que les formes proposées pour la distribution des durées de vie des éprouvettes quand S est fixée, fussent en accord avec les formes trouvées expérimentalement pour les courbes d'isoprobabilité de rupture de niveau $\alpha$ qui sont les courbes donnant, en fonction du nombre de cycles N, la valeur $S = f_\alpha(N)$ pour laquelle la probabilité de rupture avant N cycles est $\alpha$ ($E_\alpha$ est alors la limite de $f_\alpha(N)$ quand N augmente indéfiniment).

### 2° *Étude de la résistance à la fatigue de joints soudés*

Les essais de fatigue sont des essais de traction répétée (la traction minimale est invariable, tandis que la traction maximale varie) réalisés dans les conditions assez proches de celles rencontrées dans les constructions métalliques.

Chaque laboratoire participant (il y en a 6 ou 7) ne disposant que d'une machine d'essai, il a fallu se contenter d'une quinzaine d'éprouvettes par laboratoire. La limite d'endurance est alors déterminée par une méthode séquentielle, la méthode « staircase » aussi appelée « up and down » [2, 3].

Une première série d'essais a été réalisée avec des éprouvettes toutes préparées par un même laboratoire et a permis de vérifier que dans ces conditions les limites d'endurance

obtenues par les différents laboratoires n'étaient pas significativement différentes (On a mis au point à cette occasion une méthode de comparaison statistique des limites d'endurance sous l'hypothèse que le paramètre $\sigma$ des courbes en S donnant la probabilité de rupture en fonction de la charge maximale ne dépend pas du laboratoire).

On a alors réalisé une deuxième série d'essais dans laquelle les éprouvettes provenaient d'une même tôle d'acier mais étaient préparées par chacun des participants.

On a obtenu cette fois-ci des limites d'endurance différentes d'un laboratoire à l'autre. Ces différences sont vraisemblablement dues aux conditions de dépôt du cordon de soudure et une étude de l'influence de différents facteurs intervenant dans ce dépôt est en cours dans un des laboratoires.

### I. B) ÉTUDE DU DOSAGE DE L'OXYGÈNE DANS L'ACIER

L'objectif des essais réalisés est double, à savoir :

1° Vérifier l'homogénéité d'un acier « pseudo étalon » pour voir s'il peut être retenu comme étalon international.

2° Comparer les résultats obtenus par différents laboratoires ayant la pratique du dosage de l'oxygène dans l'acier, soit avec des appareils de même type, soit avec des appareils de types différents, ce qui implique l'étude de la dispersion des résultats obtenus *à des dates variées* par un même laboratoire.

On a ajouté à ce programme l'étude de la dispersion instantanée (c'est-à-dire dans un jour) des résultats obtenus par un même laboratoire.

Un plan d'expérience rigoureux a permis d'obtenir des résultats intéressants et indiscutables lors d'une première série d'essais à laquelle participèrent 44 laboratoires répartis dans toute l'Europe occidentale.

## II — PROBLÈMES POSÉS PAR LA MISE AU POINT DE NOUVEAUX PROCÉDÉS

Ces problèmes sont naturellement très variés et je me contenterai d'en donner deux exemples choisis en raison du caractère général des méthodes employées.

### II. A) ÉTUDE DE L'INFLUENCE D'UNE AUGMENTATION DE LA « QUANTITÉ DE MOUVEMENT » DU FLUIDE D'ATOMISATION DU COMBUSTIBLE DANS LES BRULEURS DES FOURS MARTIN

Les travaux de la fondation d'études internationales sur les flammes ont montré l'avantage que présentait une « quantité de mouvement » aussi grande que possible du jet de combustible dans les échanges de chaleur par convection entre la flamme et la charge.

Il restait à examiner les avantages que l'on pouvait pratiquement retirer d'une augmentation de la quantité de mouvement du fluide d'atomisation du combustible en fonction de la valeur de cette augmentation et des moyens employés pour l'obtenir.

L'amélioration des échanges de chaleur par convection entre la flamme et le four doit se traduire par une augmentation de la vitesse de fusion de la charge solide introduite et par conséquent de la production horaire d'acier.

Les premières expériences furent faites sur des fours à mazout dans lesquels le fluide d'atomisation était de l'air comprimé.

Elles consistèrent à comparer les vitesses de fusion et les productions horaires obtenues dans les conditions de référence et pour des pressions de l'air d'atomisation supérieures à la pression de référence.

Ultérieurement on fit des expériences analogues en augmentant la quantité de mouvement au nez des brûleurs par une insufflation complémentaire d'oxygène et ce procédé fut ensuite étendu à des brûleurs à gaz naturel et à gaz de gazogène.

La méthode expérimentale fut mise au point au cours des premières études et fut toujours la même. Elle devait tenir compte de l'influence de l'âge du four sur la vitesse de fusion et la production horaire et notamment du « vieillissement du four ».

La première solution qui vint à l'esprit pour éliminer cette influence fut de réaliser des alternances aussi courtes que possible de coulées normales et de coulées d'essai.

Malheureusement il semble que les opérateurs eurent du mal à s'adapter aux changements de rythme dans le chargement et la conduite du four entraînés par le changement des conditions de combustion aux brûleurs et, peut-être aussi en raison de l'inertie thermique du four, la production horaire et la vitesse de fusion furent peu influencées par les modifications essayées.

On dut donc revenir à l'exécution dans les mêmes conditions de séries de coulées assez importantes, c'est-à-dire durant au moins une semaine.

Il devenait alors indispensable de tenir compte de l'âge du four dans les comparaisons de résultats.

La méthode consista à étudier, pour une série de coulées élaborées dans les mêmes conditions, les régressions de la durée de fusion et de la production horaire sur l'âge du four mesuré en nombre de coulées.

L'expérience a montré que, comme on pouvait s'y attendre, après une période de mise en régime d'environ 40 coulées, ces régressions sont sensiblement linéaires et ont des

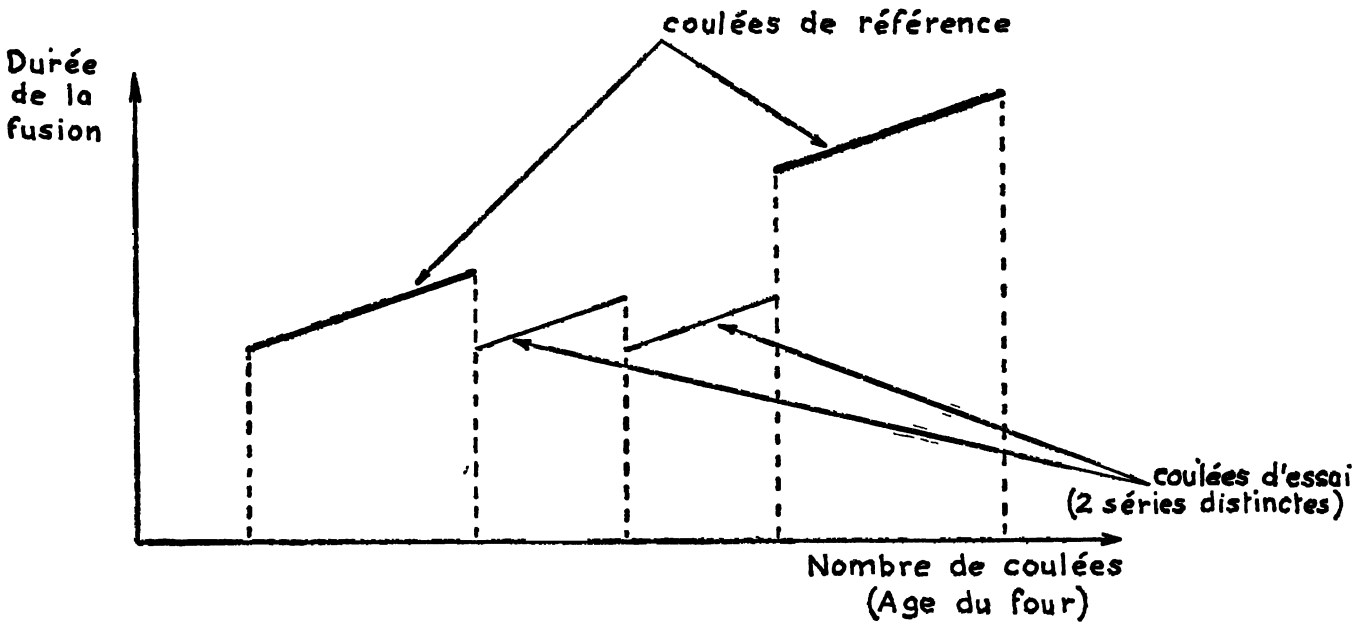


Figure 1

pentes indépendantes des conditions d'atomisation du combustible. La « vitesse de vieillissement » du four ne dépend donc pas des conditions de la combustion, tout au moins dans le domaine de variation de ces conditions qui a été exploré. Le parallélisme des lignes de régression correspondant aux différentes conditions d'atomisation fut testé statistiquement.

La comparaison des conditions d'atomisation se ramène alors (*cf.* figure 1) à la comparaison des ordonnées à l'origine ou en un point donné des droites de régression sur l'âge du four correspondantes, ces droites étant supposées parallèles, et l'estimation de l'amélioration apportée par une des conditions essayées se ramène à une estimation de la différence des ordonnées à l'origine ou en un point donné entre la droite correspondant aux conditions nouvelles et celle correspondant aux conditions normales.

En fait l'âge du four n'est pas le seul facteur variable d'une série d'essais à l'autre et on a étudié, conjointement à l'âge du four, l'influence d'autres facteurs tels que la quantité de fonte liquide enfournée dont l'influence n'a pas été trouvée sensible.

L'exemple que nous venons de citer n'en est qu'un parmi beaucoup d'autres de l'emploi des techniques « d'analyse de la covariance » auxquelles nous faisons très fréquemment appel.

Ces techniques consistent, pour tenir compte de l'influence de facteurs mesurables dont il n'a pas été possible de neutraliser l'intervention dans l'expérience, à remplacer la classique comparaison des moyennes par une comparaison des droites ou des plans de régression des caractéristiques à comparer sur les facteurs perturbateurs.

Elles permettent également de voir, en testant le parallélisme des surfaces de régression, si les influences des facteurs perturbateurs sont les mêmes dans différentes conditions.

## II. B) PRÉCISION DES PROCÉDÉS DE MESURE D'UNE CARACTÉRISTIQUE BASÉS SUR L'EMPLOI D'UNE DROITE D'ÉTALONNAGE

Les méthodes de détermination de teneurs chimiques par des procédés physiques telles que les méthodes d'analyse spectrographique ou spectrophotométrique reposent sur un étalonnage préalable de l'appareil qui se fait en observant les indications $y_i$ qu'il donne pour des teneurs bien connues $x_i$, $i = 1, 2 \ldots n$ de l'élément auquel on s'intéresse.

On doit ensuite, ayant fait sur un produit de teneur inconnue une mesure qui donne le résultat $y$, en déduire une estimation de la teneur $x$ de ce produit et la précision de cette estimation.

Comme pour une valeur déterminée de $x$, $y$ varie par suite des erreurs de mesure, on adopte pour courbe d'étalonnage la ligne de régression de $y$ sur $x$.

On se limite généralement à un domaine de variation de $x$ pour lequel la régression de $y$ sur $x$ ou sur $\log x$ peut être considérée comme linéaire, cette linéarité pouvant être testée statistiquement.

On détermine ensuite les lignes limites de l'intervalle de probabilité au risque $\alpha$ pour les valeurs observables $y$, qui dans le cas où la régression de $y$ sur $x$ est linéaire, sont les deux branches :

$$(\mathrm{H}_1)\quad y_1 = a + b\,x - t_{1-\frac{\alpha}{2}}\, s'_{y/x} \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}$$

$$(\mathrm{H}_2)\quad y_2 = a + b\,x + t_{1-\frac{\alpha}{2}}\, s'_{y/x} \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

d'une hyperbole dont la droite de régression estimée (D) $y' = a + b\,x$ constitue le diamètre conjugué des cordes verticales (*cf.* figure 2).

Dans ces formules $\bar{x} = \frac{\Sigma\, x_i}{n}$ est la moyenne arithmétique des teneurs utilisées pour estimer la droite d'étalonnage, $s'_{y/x}$ est une estimation « du $\chi^2$ » de l'écart type $\sigma_{y/x}$ de $y$ quand $x$ est fixé basée sur $n - 2$ degrés de liberté et $t_{1-\frac{\alpha}{2}}$ est la limite fiducielle au risque $\frac{\alpha}{2}$ de la variable $t$ de Student basée sur $n - 2$ degrés de liberté.

Si donc on a observé la valeur Y, il y a une probabilité $P = 1 - \alpha$ pour que la valeur X correspondante soit contenue dans l'intervalle $(X_2, X_1)$ délimité par les abscisses des points d'intersection de l'horizontale d'ordonnée Y avec les deux courbes $H_2$ et $H_1$.

L'intervalle de confiance au niveau $1 - \alpha$ pour X a donc une longueur variable avec Y et n'est pas symétrique par rapport à l'abscisse $X' = \frac{Y - a}{b}$ du point d'ordonnée Y de la droite d'étalonnage qu'on adopte généralement pour estimation de X.

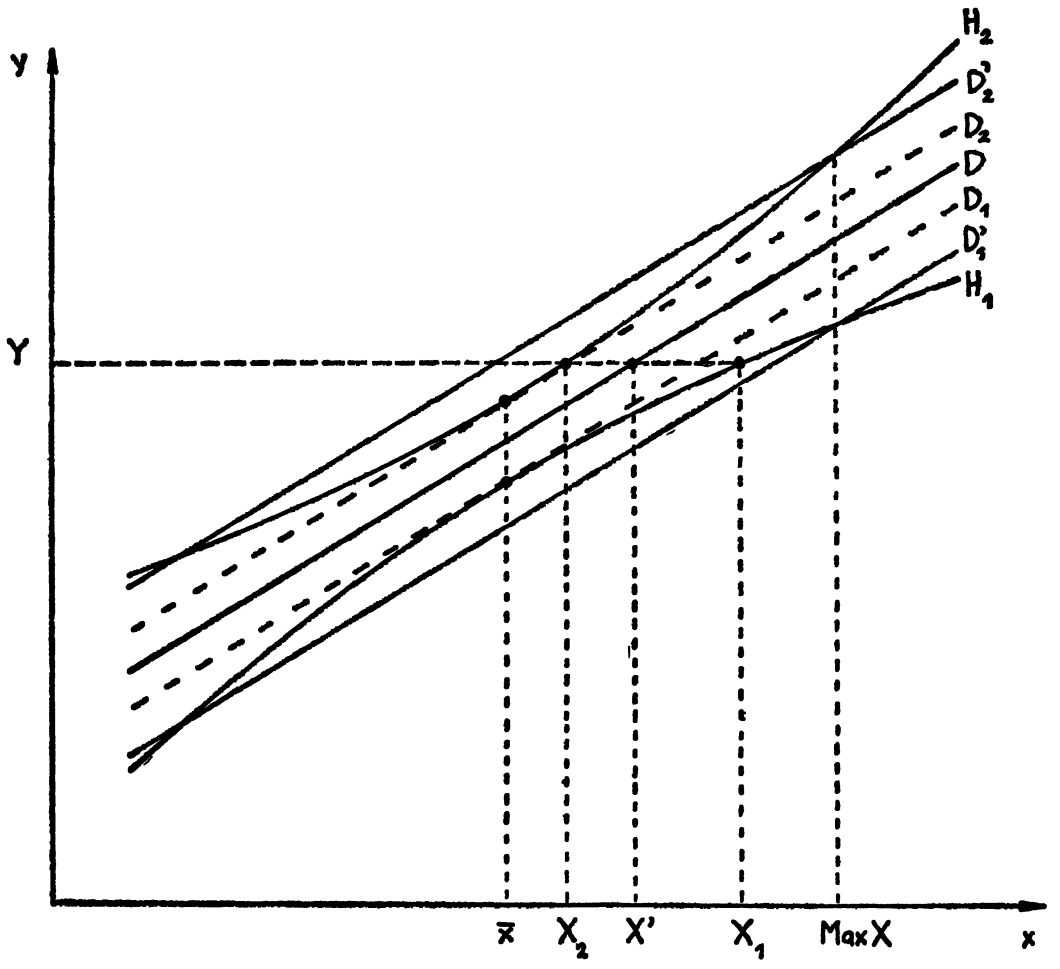


Figure 2

Comme il n'est pas commode d'avoir un intervalle de longueur variable et qui ne soit pas symétrique par rapport à X', nous avons proposé de remplacer les deux branches d'hyperbole par les deux droites $D'_1$ et $D'_2$ parallèles à D passant par les points de $H_1$ et $H_2$ qui correspondent à la limite la plus éloignée de $\bar{x}$ du domaine des valeurs de $x$ considéré (*cf* figure 2) (1). Leur distance à (D) comptée parallèlement à l'axe des $y$ est donc :

$$t_{1-\frac{\alpha}{2}}\, s'_{y/x} \sqrt{1 + \frac{1}{n} + \frac{Max\,(X - \bar{x})^2}{\sum\limits_i (x_i - \bar{x})^2}}$$

Ces deux droites définissent, dans l'intervalle de variation considéré pour X, un intervalle de confiance à un niveau variable supérieur à $1 - \alpha$, mais symétrique par rapport à X' et de longueur constante égale à :

$$2\, t_{1-\frac{\alpha}{2}} \frac{s'_{y/x}}{b} \sqrt{1 + \frac{1}{n} + \frac{Max\,(X - \bar{x})^2}{\sum\limits_i (x_i - \bar{x})^2}}$$

On voit que cet intervalle de confiance dépend d'une part du rapport $\frac{s'_{y/x}}{b}$ (il ne sert à rien d'augmenter la pente de la droite d'étalonnage si la dispersion de $y$ augmente dans la même proportion) et d'autre part de $\frac{1}{n} + \frac{Max\,(X - \bar{x})^2}{\sum\limits_i (x_i - \bar{x})^2}$ que l'on a intérêt à rendre petit par rapport à 1 en faisant un nombre suffisant de mesures sur étalons et en choisissant des étalons dont la teneur moyenne $\bar{x}$ soit au centre du domaine intéressant pour X et qui soient régulièrement répartis dans ce domaine. (En toute rigueur on aurait intérêt à prendre seulement deux étalons dont les teneurs soient aux extrémités du domaine intéressant pour X mais on perdrait alors toute possibilité de contrôler la linéarité de l'étalonnage dans l'ensemble du domaine considéré).

Quand on a $\frac{Max\,(X - \bar{x})^2}{\sum\limits_i (x_i - \bar{x})^2} \leqslant 0{,}20$, comme on a alors :

$$\sqrt{1 + \frac{1}{n} + \frac{Max\,(X - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \simeq 1 + \frac{1}{2\,n} + \frac{Max\,(X - \bar{x})^2}{2 \sum (x_i - \bar{x})^2}$$

la précision de la mesure est pratiquement indépendante de la valeur à mesurer et des teneurs des étalons utilisés pour tracer la droite d'étalonnage (les parties utiles des droites $D'_1$ et $D'_2$ et des courbes $H_1$ et $H_2$ sont pratiquement confondues avec les droites $D_1$ et $D_2$ d'équations :

$$y = a + bx \pm t_{1-\frac{\alpha}{2}}\, s'_{y/x} \sqrt{\left(1 + \frac{1}{n}\right)}$$

(1) Il est facile de voir qu'en négligeant $t^2_{1-\frac{\alpha}{2}} \frac{s'^2_{y/x}}{\sum\limits_i (x_i - \bar{x})^2} = t^2_{1-\frac{\alpha}{2}}\, s'^2_b$ devant $b^2$ (ce qu'on peut pratiquement faire dès qu'on a $\frac{b}{s'_b} > 3\, t_{1-\frac{\alpha}{2}}$) on obtient $\left.\begin{matrix} X_1 \\ X_2 \end{matrix}\right\} = \frac{Y - a}{b} \pm t_{1-\frac{\alpha}{2}} \frac{s_{y/x}}{b} \sqrt{1 + \frac{1}{n} + \frac{(Y - \bar{y})^2}{b^2 \sum\limits_i (x_i - \bar{x})^2}}$ ce qui définit un intervalle de confiance symetrique par rapport à X' et plus court que celui que nous proposons mais dont la longueur varie avec Y $\left(\text{ou plutôt avec } \left|\frac{Y - \bar{y}}{b}\right|\right)$. En bornant $\left|\frac{Y - \bar{y}}{b}\right|$ par Max $\left|X - \bar{x}\right|$ on retrouve l'intervalle de confiance que nous proposons.

et l'erreur au risque $\alpha$ est sensiblement égale à $\pm\, t_{1-\frac{\alpha}{2}} \frac{s'_{y/x}}{b}\left(1 + \frac{1}{2n}\right)$; c'est celle que l'on obtiendrait si la pente de la droite d'étalonnage était connue avec certitude.

La précision de la mesure est en outre pratiquement indépendante du nombre $n$ des étalons utilisés dès que ce nombre atteint 20. L'erreur au risque $\alpha$ est alors sensiblement égale à $\pm\, u_{1-\frac{\alpha}{2}} \frac{s'_{y/x}}{b}$ où $u$ est une variable normale réduite. C'est celle que l'on obtiendrait si la droite d'étalonnage était connue avec certitude.

## APPLICATIONS :

### 1° *Analyse spectrographique d'éléments métalliques d'aciers faiblement alliés.*

Nous avons appliqué ce qui précède à l'analyse spectrographique de divers éléments métalliques tels que chrome, manganèse, nickel, d'aciers faiblement alliés, pour déterminer la précision que donnait la méthode employée dans notre laboratoire, en nous basant sur l'existence, dans un domaine limité de variation de $x$, d'une régression linéaire de $y$ sur $\log x$, ce qui conduit à déterminer une valeur maximale de l'erreur relative au risque $\alpha$ sur la teneur.

Nous avons notamment testé la linéarité de la courbe d'étalonnage et sa stabilité dans le temps et regardé si les teneurs des étalons utilisés et le nombre des mesures effectuées sur chacun d'eux permettait d'obtenir l'erreur minimale possible compte tenu de la pente de la droite d'étalonnage et de la dispersion des mesures de $y$.

### 2° *Étude de la précision de méthodes d'analyse spectrophotométrique* [7].

Ces méthodes qui sont d'une très grande sensibilité reposent sur la mise en solution d'un poids constant du corps dont on veut connaître la teneur en un certain élément dans un volume fixé une fois pour toutes d'un réactif approprié, généralement incolore, qui en présence de l'élément à doser donne une coloration. On mesure alors la densité optique $y$ de la solution placée dans une cuve à faces parfaitement parallèles, pour une longueur d'onde déterminée. Pour une position donnée de la cuve par rapport à la direction des rayons lumineux la densité optique est, en raison de la loi de Lambert Beer, proportionnelle à la concentration $z$ de la solution en l'élément à doser.

Mais $z$ n'est pas proportionnelle à la teneur $x$ de l'élément à doser dans le corps, car même les meilleurs réactifs contiennent des traces de l'élément à doser auxquelles les méthodes utilisées sont sensibles. Il en résulte un « blanc » $z_0$ variable d'une série de dosages à une autre (le lot de réactif utilisé change) en sorte que $x = k(z - z_0)$ et finalement, aux erreurs de mesure près, on a $y = \alpha + \beta x$ où $\beta$ est constant d'une série de dosages à une autre tandis qu'$\alpha$ qui est proportionnel au blanc varie.

Dans ces conditions on doit, pour chaque série de dosages, déterminer la droite de dosage ou droite de régression de $y$ sur $x$ à l'aide d'étalons à teneurs connues, puis ayant observé une densité optique Y en déduire une estimation X' et un intervalle de confiance à un niveau donné pour la teneur inconnue X.

Pour des raisons de commodité on ne peut introduire plus de deux étalons dans une série de dosages mais on peut disposer d'une gamme d'étalons synthétiques suffisamment variés pour pouvoir les choisir judicieusement lors de chaque série de dosages.

On procède alors de la façon suivante :

La pente $\beta$ de la droite de dosage ainsi que la dispersion $\sigma_{y/x}$ des résultats des mesures au spectrophotomètre sont estimées une fois pour toutes à partir de plusieurs gammes d'étalonnage.

Dans chaque série de dosages on introduit deux étalons qui servent à l'estimation de la position de la droite de dosage et au contrôle du bon fonctionnement de l'appareillage et de la correction de la manipulation.

Nos études ont porté sur les points suivants :

*a*) Analyse des résultats obtenus sur les gammes d'étalonnage. Vérification de la constance de la dispersion $\sigma_{y/x}$ pour les différentes gammes et estimation de cette dispersion. Vérification de la constance de la pente des droites d'étalonnage. Estimation de cette pente et de l'erreur attachée à cette estimation. Comparaison des blancs des droites d'étalonnage pour voir qu'ils diffèrent significativement les uns des autres.

*b*) Estimation avec deux étalons du blanc associé à une série de dosages et contrôle du bon fonctionnement de l'appareillage et de la manipulation.

*c*) Étude de la précision de la méthode et notamment choix des teneurs des deux étalons à associer à une série de dosages, et des dosages à faire dans l'étude préliminaire destinée à estimer la pente et la dispersion des mesures.

Nous avons pu définir à l'intention des chimistes usagers des méthodes spectrophotométriques un mode de détermination des droites de dosage et de la précision des analyses spécialement adapté à leur problème.

## III — QUELQUES PROBLÈMES NE RELEVANT PAS DES MÉTHODES STATISTIQUES CLASSIQUES

Nous avons déjà signalé en I. A quelques uns des problèmes posés par l'interprétation des résultats des essais de fatigue.

L'intérêt porté par les techniciens à ce qu'ils appellent la courbe de répartition granulométrique d'un produit en morceaux (minerai, charbon, coke etc...) et le fait que cette courbe a souvent la forme de la courbe de répartition d'une variable aléatoire dont la distribution est logarithmico normale, ont conduit à rechercher la liaison que pouvait avoir cette courbe avec une courbe de répartition statistique et à généraliser la notion de courbe de répartition granulométrique [8].

La courbe de répartition granulométrique d'un produit n'est en effet pas une courbe de répartition statistique puisque c'est la courbe, donnant en fonction du diamètre $d$, la fraction pondérale du produit constituée par les morceaux dont le plus grand diamètre est inférieur à $d$ (morceaux passant à travers un tamis de maille égale à $d$). Sa détermination pose un problème d'échantillonnage sur lequel on s'est également penché.

Nous avons d'ailleurs fréquemment à étudier des problèmes d'échantillonnage de produits en vrac ou massifs tels que coke, minerai, fonte et acier liquides, lingots d'acier etc...

Les méthodes statistiques permettent d'estimer, grâce à un plan d'expérience convenable les influences des différents facteurs systématiques (tels que le choix d'un emplacement de prélèvement dans un produit massif : lingot, tôle, etc...) et aléatoires intervenant dans l'échantillonnage.

Il nous arrive aussi d'avoir à étudier des problèmes de caractère moins général posés

par nos collègues expérimentateurs et, comme toujours, un premier travail non négligeable consiste à traduire ces problèmes sous forme mathématique et à poser des hypothèses convenables, compte tenu du problème concret considéré.

## CONCLUSIONS

Il semble que ce soit au stade de la recherche technique, plus qu'à celui de la recherche fondamentale, que les méthodes statistiques sont susceptibles d'apporter une aide précieuse. Cependant une étude statistique et logique de phénomènes complexes permet parfois de les dégrossir et ses conclusions servent alors assez souvent de point de départ à des recherches fondamentales.

Dans l'application des méthodes statistiques à la recherche industrielle il importe de tenir compte des faits suivants :

Il y a généralement possibilité d'établir un plan d'essais et de répéter les essais, mais comme ceux-ci sont généralement assez longs et coûteux il importe d'en limiter le nombre surtout si on veut que ceux-ci soient conformes au plan établi.

Pour notre part, nous préférons de beaucoup un petit nombre d'essais bien organisés et bien exécutés à un grand nombre de résultats « tout venant ».

Par ailleurs, tout au moins dans le domaine où nous opérons, qui est étudié depuis longtemps, il ne faut généralement pas espérer des résultats spectaculaires, mais chercher plutôt à compléter ou à préciser les connaissances des techniciens.

On doit donc s'efforcer d'obtenir des résultats quantitatifs plutôt que qualitatifs, c'est-à-dire faire de la statistique fine et travailler en étroite liaison avec les techniciens, voire même devenir soi-même un peu technicien, afin de tenir compte, dans l'organisation des essais et dans l'analyse de leurs résultats, des connaissances techniques qu'on a sur le sujet, voire même des idées à priori dont il est alors prudent de tester la validité.

Nous pensons qu'il doit en être de même dans les autres domaines industriels, car on ne fait bien souvent appel au statisticien que quand on n'a pas pu résoudre le problème sans lui, ou tout au moins, quand les études préliminaires laissent penser que le problème posé sera difficile à résoudre. Nous pensons d'ailleurs qu'il est souvent souhaitable de procéder à quelques essais ou études préliminaires pour dégrossir ou préciser les données d'un problème avant de le soumettre au statisticien.

J. Ulmo

## BIBLIOGRAPHIE

[1] Ulmo J., St Guilhem R., Loison R. — Sur les variations en fonction d'un paramètre de la probabilité d'un phénomène se traduisant par une alternative. *Annales des Mines* (1950) IV, p. 3-9.

[2] Ulmo J. — Détermination expérimentale de la courbe représentant les variations en fonction d'un paramètre de la probabilité d'un phénomène se traduisant par une alternative quand celle-ci est une « courbe en S de Galton ». *Annales des Mines* (1952) III-IV, p. 93-130.

[3] Finney D. J. — « Probit Analysis ». Cambridge University Press.

[4] Bastenaire F. — Étude de la précision d'une méthode d'approximations successives pour la détermination d'un paramètre. *Revue de Statistique Appliquée*, vol. 1, n° 3-4 (1953).

[5] Bastenaire F. — Étude statistique et physique de la dispersion des résistances et des endurances à la fatigue. Thèse de Doctorat d'État, Paris (29 janvier 1960).

[6] Bastenaire F., Bastien M., Pomey G. — L'analyse statistique des résultats d'essais de fatigue et son application à de nouvelles séries d'observations. Conférence de l'Académie Hongroise des Sciences sur la résistance des matériaux (Budapest, octobre 1961). *Acta Technica Academia Scientiarum Hungaricae*, t. 35-36, p. 7-26 (1961).

[7] Ulmo J. — Étude de la précision de méthodes d'analyse spectrophotométrique. *Revue de Statistique Appliquée* (1960), vol. VIII, n° 3, p. 53-73.

[8] Bastenaire F. — Étude théorique du mode de génération des distributions granulométriques. *Revue de Statistique Appliquée*, n° 1, vol. XI (1963).

## DISCUSSION

M. Brichler déclare qu'il a été très intéressé par les exemples concrets d'utilisation de la méthode statistique présentés avec beaucoup de clarté par M^lle^ Ulmo.

La Conférencière, à l'occasion du premier exemple, ayant indiqué que les essais avaient été effectués dans différents pays, M. Brichler voudrait savoir quels sont ces pays et sous l'égide de quels organismes s'est déroulée l'expérience.

M. Maurice Dumas. — Dans les organismes jouant dans les pays étrangers un rôle analogue à celui de l'I. R. S. I. D. en France, existe-t-il des spécialistes en statistique et, si oui, M^lle^ Ulmo a-t-elle des occasions de s'entretenir avec ses collègues étrangers?

M. Bastenaire. — L'Institut International de la Soudure est un groupement réellement international puisqu'il est ouvert à toutes les nations, sans considération de régime politique. En font partie, par exemple, la France, les États-Unis, le Royaume-Uni mais aussi, l'U. R. S. S., la Tchécoslovaquie, etc... La Chine populaire qui n'a envoyé jusqu'à présent que des observateurs aura probablement une délégation officielle au prochain Congrès de l'I. I. S. Cet organisme est purement technique et s'intéresse à tous les problèmes posés par le soudage et les techniques connexes.

M. Morice signale qu'une autre méthode d'inversion de l'équation de régression linéaire a été proposée par Brownlee (Statistical Theory and Methodology in Science and Engineering, 1960, p. 284).

Applicable lorsque l'écart type du coefficient de régression est petit, par rapport au coefficient, elle conduit aussi à remplacer l'hyperbole par deux droites parallèles à la droite de régression et d'équations

$$Y = a + b\,X \pm t\, s'_{y/x} \sqrt{1 + \frac{1}{n} + \frac{(Y - \bar{y})^2}{b^2 \Sigma\,(x_i - \bar{x})^2}}$$

On obtient les deux solutions :

$$X = \frac{Y - a}{b} \pm \frac{t}{b}\, s'_{y/x} \sqrt{1 + \frac{1}{n} + \frac{(Y - \bar{y})^2}{b^2 \Sigma\,(x_i - \bar{x})^2}}$$

symétriques par rapport à l'abscisse $X' = \dfrac{Y - a}{b}$ du point d'ordonnée Y de la droite d'étalonnage.

Pour un risque donné, qui définit la valeur de $t$, l'intervalle de confiance est d'autant plus réduit que la valeur observée Y est plus voisine de la moyenne $\bar{y}$ des observations d'étalonnage.

Par contre cet intervalle est variable avec Y et doit être calculé à chaque fois ou préalablement tabulé en fonction de Y.