Journal de la société statistique de Paris

Lucien Amy

Sur la valeur de la preuve indiciale

Journal de la société statistique de Paris, tome 92 (1951), p. 90-103
<http://www.numdam.org/item?id=JSFS_1951__92__90_0>

V

SUR LA VALEUR DE LA PREUVE INDICIALE

RÉSUMÉ

Application très générale des méthodes du calcul des probabilités à l'étude de la valeur de la preuve indiciale en criminologie. Causes d'erreurs dues à la possibilité, pour deux individus ou deux objets, de posséder les mêmes caractères d'identification : cas d'un nombre limité d'origines, conditions à réaliser pour définir et utiliser, dans le cas général, une origine moyenne. Emploi des méthodes statistiques dans le choix des méthodes d'identification à utiliser : efficacité des méthodes. Nombreux exemples.

ABSTRACT

Proof based on " indices " in criminalogy. Study of its value by general application of the calculus of probability methods. Possibility for two persons or two things of possessing the same identification characters as a cause of errors. Case of a limited number of origins. Conditions required to make possible the definition of an average origin and to make use of it in the general case. Choice of identification methods through the use of statistics. Efficiency of the methods. Many examples are given.

INTRODUCTION

Le sens de la justice est inné dans le cœur des hommes, mais pendant de nombreux siècles, les juges ne disposèrent que de moyens insuffisants pour découvrir la vérité et ce fut longtemps une figure de rhétorique justifiée de dire que la justice était boiteuse.

Rabelais, ce grand humoriste, ne pouvait manquer de se dauber des juges de son temps et je ne saurais résister au plaisir de rappeler devant les membres de la Société des Statistiques de Paris que Bridoye tirait ses sentences aux dés (1) ce qui n'aurait peut-être pas suffi à le faire admettre dans notre Société.

Pour établir la preuve, à cette époque, les enquêteurs ne connaissaient guère que la torture et les témoignages. Petit à petit, cependant, on fit de plus en plus souvent appel aux méthodes scientifiques pour trouver des méthodes moins sujettes à caution. C'est ainsi que naquit la « police scientifique ».

Les origines de cette discipline sont, en réalité, aussi vieilles que la science, toutefois, la première application qui ne relève pas de la légende, ne date que de vingt-deux siècles; elle est due à Archimède. On sait que le tyran Hélias, inquiet de l'honnêteté de son orfèvre, avait fait appel à ce grand savant pour rechercher la présence d'argent dans une couronne d'or fin sans altérer l'œuvre d'art.

Mais il fallut attendre la fin du XIX^{e} siècle pour que l'étude des traces indiciales cesse d'être un sujet de roman et devienne une réalité courante de laboratoire. De nombreux chercheurs ont contribué à transformer ainsi la police scientifique en une véritable doctrine indépendante ayant ses lois et ses méthodes qui lui sont propres.

Parmi ces chercheurs, l'un des plus connus est, à juste titre, l'illustre Bertillon et nous rappellerons, à ce propos, que la méthode de classification et d'identification des récidivistes avait été précédée d'études statistiques approfondies et résultait directement de ces études.

Les conclusions que l'on peut tirer de l'étude des traces indiciales sont-elles toujours formelles? Il n'en est rien. Si, en matière de traces papillaires, d'empreintes digitales par exemple, les chances d'erreur sont pratiquement insignifiantes, au moins pour un expert averti, il n'en est pas ainsi dans tous les domaines. C'est ainsi, qu'interrogé par un président de cour d'assises sur la confiance que l'on pouvait avoir dans une analyse graphologique, un de nos plus célèbres experts, dans cette spécialité, avouait « qu'en matière de graphologie, on avait une chance sur deux de se tromper » ce qui lui a valu de la part du président la réplique que « dans ce cas, il aurait été aussi simple de tirer le résultat aux dés ». Vous voyez que, décidément, l'emploi des dés est très à l'honneur dans la magistrature.

La première pensée qui vient à l'esprit, c'est que l'expert doit s'abstenir lorsqu'il ne possède pas de méthode sûre lui permettant de conclure en toute certitude. Mais la question n'est pas aussi simple en réalité et, pour la bien comprendre, il est nécessaire de préciser le rôle exact de l'expert.

Un crime, un délit, ont été commis : le but de la justice est double : établir s'il y a bien eu crime ou délit, rechercher, s'il y a lieu, le coupable. L'enquête préliminaire conduisant à la présomption du délit ou à soupçonner un individu, le juge s'adresse à l'expert pour obtenir de lui les preuves de la culpabilité. Peut-il poser directement les questions est-ce tel individu le coupable, a-t-il commis tel délit? Poser les questions sous cette forme, serait une abdication de la part du juge. L'expert ne doit être qu'un simple informateur technique; sa responsabilité est limitée à la constatation et à l'interprétation matérielle des faits, non à leur interprétation humaine; le *juge*, seul, peut prononcer le *jugement*. Dès lors, la question posée doit prendre une forme ainsi conçue : telle trace, tel indice, proviennent-ils ou non de l'individu suspect, y a-t-il des traces établissant l'existence d'une fraude, etc...

Or, par suite d'une dissymétrie dans la valeur de la preuve, dissymétrie sur laquelle nous reviendrons plus loin, la réponse à une même question ne peut, en général, être affirmative que dans l'une des éventualités et n'est que plus ou moins probable dans l'autre. On peut affirmer par exemple, qu'un document est faux, on ne peut jamais affirmer qu'il est authentique, mais seulement qu'on n'a rien retrouvé de suspect. De même, on peut être certain qu'une trace de pas ne provient pas d'un individu donné, mais jamais on ne peut établir avec une certitude absolue qu'elle en provient, etc...

Avec ou sans preuves décisives, le juge doit toujours conclure. Or, dans bien des questions, il y a dilemme entre deux individus : acquitter l'un, c'est condamner l'autre. Même si le sort d'un individu seul est en cause, il n'est pas possible de ne condamner qu'en cas de preuve absolue. Certes, le doute doit bénéficier à l'inculpé, mais le juge ne doit pas non plus oublier qu'acquitter systématiquement serait encourager le crime et renoncer pratiquement à toute justice : son devoir le plus strict est de protéger la société. Il y a donc le plus souvent conflit entre les deux positions.

Réciproquement, lorsque les charges contre l'inculpé sans être complètement réduites à néant sont devenues très faibles, il y a le plus grand intérêt à préciser dans quelle mesure. L'acquittement au bénéfice du doute n'est jamais un acquittement moral total et la réputation de l'accusé reste hélas définitivement ternie contre toute justice. Préciser la valeur exacte des charges qui sont retenues ou non contre l'inculpé, présente donc dans tous les cas un intérêt capital.

Dès lors, le devoir de l'expert est très net : il doit communiquer au juge tous les éléments d'appréciation dont il dispose. Si sa méthode ne conduit qu'à une simple présomption, il devra quand même en indiquer le résultat, quitte à mentionner qu'il s'agit d'une présomption et, dans ce cas, le degré de confiance que l'on doit lui accorder; vous pressentez tout de suite que la valeur de cette présomption devrait s'exprimer en langage probabiliste et par conséquent, reposer sur l'application des méthodes statistiques.

Une conclusion présentée sous la forme « rien ne s'oppose à ce que... », que l'on rencontre trop souvent, est donc inadmissible car elle ne peut jamais traduire complètement l'opinion de l'expert. Elle montre que celui-ci n'a pas osé conclure, mais laisse sous-entendre que son opinion était positive. Si réellement il ne peut rien tirer dés études effectuées, c'est-à-dire si d'une manière précise, la probabilité est de 0,5, il doit compléter la formule précédente et dire : « rien ne s'oppose à ce que..., mais rien non plus ne permet de penser le contraire » ou mieux encore, avouer plus simplement : « on ne peut tirer aucune conclusion des faits observés ». Dans tous les cas contraires, l'expert doit préciser quelle est l'hypothèse la plus probable.

Mis en contact avec les méthodes statistiques, par la fréquentation assidue des séances de notre société, nous avons, tout naturellement, pensé utiliser ses méthodes pour tirer le maximum de renseignements des examens physico-chimiques ou biologiques que nous avions l'occasion de faire. Mais une telle méthode n'est pas sans danger. Nous n'avons pas oublié qu'une théorie probabiliste sur la valeur des jugements a été édifiée par un des grands maîtres de mathémathiques, théorie que Stuart Mill (2) a qualifiée de « scandale de mathématiques ».

D'autres auteurs tels que Balthazard (3), avaient d'ailleurs déjà essayé d'appliquer ces méthodes et nous avons nous-mêmes présenté des critiques de détail sur l'emploi qu'ils en avaient fait. Ce n'est donc qu'avec la plus grande prudence, que nous avons entrepris ces recherches.

Nous avons actuellement étudié trois problèmes nettement différents :

— La recherche de la paternité par les groupes sanguins (4);

— L'identification des traces papillaires (5, 6, et 7);

— La comparaison des encres au campêche chrome (8).

Dans ces trois études, les méthodes de calcul sont aussi différentes que les sujets. C'est ainsi que, dans un cas, nous faisons appel aux probabilités continues, dans deux autres, aux probabilités discontinues. La précision des résultats obtenus n'est pas moins variable. Dans le cas des empreintes digitales, on arrive en général à une quasi-certitude — la comparaison des groupes sanguins présente tous les cas possibles — tandis que dans l'exemple choisi, relatif aux encres, on n'avait guère plus de raison d'accepter l'une ou l'autre des éventualités.

Nous avons pu cependant dégager de ces trois études, un certain nombre de caractères communs qui nous ont permis de préciser le sens exact des probabilités que nous avions calculées dans chaque cas particulier. En dehors de l'intérêt évident d'une étude générale, nous estimons indispensable d'attirer l'attention sur les difficultés rencontrées.

Dans les cas que nous avons étudiés, en effet, et *a fortiori* dans les autres, l'interprétation est très délicate et nous sommes inquiets que, s'appuyant sur nos travaux, des personnes mal averties des méthodes de statistiques, n'en tirent de fausses conclusions. En dehors des conséquences relatives aux cas particuliers, ces erreurs ne manqueraient pas de discréditer l'emploi des méthodes statistiques dans le domaine judiciaire.

PROBABILITÉ D'ERREUR

Quelle que soit la question posée par le juge à l'expert, quatre groupes de causes principales interviennent pour limiter la valeur de l'identification par l'étude des indices :

a) Les erreurs d'expérience et, en particulier, celles dues à l'inattention de l'expert et de ses aides.

Pour identifier, par exemple, un récidiviste, il faut d'abord retrouver les fiches établies et classées antérieurement. Or, dans l'établissement et la notation des fiches antérieures et de la nouvelle fiche, comme dans le classement et la recherche de ces fiches, l'expérimentateur peut se tromper. La réponse ne devrait donc pas être limitée à : « l'individu recherché n'a jamais été bertillonné », elle devrait être complétée par « *sous réserve d'erreurs commises par les employés chargés du classement des fiches* ». Suivant la valeur professionnelle des employés, cette cause d'erreur peut être de une sur mille, une sur dix mille..., mais on ne pourra jamais avoir une certitude absolue.

b) La théorie sur laquelle on s'appuie, est plus ou moins parfaite et toute imperfection introduit une cause d'erreur que l'expérimentateur le plus habile et le plus consciencieux ne peut éviter.

Par exemple, dans la recherche de la paternité, on admet un certain nombre de lois relatives à l'hérédité des caractères d'agglutination des globules sanguins. Mais il existe des cas où les lois de l'hérédité cessent de s'appliquer; c'est ce que l'on appelle les mutations. L'existence d'une mutation peut fausser complètement les conclusions tirées de l'examen des caractères sanguins.

Pour qu'une théorie soit applicable à un problème d'identification, il est indispensable qu'elle ait été soumise à des vérifications nombreuses. Mais si nombreuses que soient ces dernières, elles ne sauraient déceler des cas très rares. Pour déceler, par exemple, avec une probabilité raisonnable, une mutation sur 100.000, il faudrait examiner au moins 100.000 cas. Or, il est impossible dans l'étude des lois de l'hérédité humaine de procéder à des expériences présentant toutes les garanties. On ne peut jamais être certain, en particulier, que le père officiel d'un enfant est bien son père. Certains auteurs estiment que la cause d'erreur due à cette incertitude est d'environ 1 %, ce qui rend pratiquement impossible la mise en évidence de mutations ayant une fréquence inférieure à 1 %.

c) On peut imaginer *a priori* que les traces ont été volontairement altérées soit par les auteurs du délit, soit par les premières personnes qui sont venues sur le lieu du crime dans le but de détourner les soupçons sur des innocents. Cette cause d'erreur est une des péripéties de choix des auteurs de romans policiers, mais en fait, elle est purement théorique et nous ne l'avons jamais rencontrée au cours des très nombreuses expertises auxquelles nous avons participé, sauf, dans les cas où des falsifications étaient l'objet même du délit. A l'occasion de trois expertises, nous avons pu en effet établir que des faussaires, certainement au courant des techniques de police scientifique, avaient altéré des documents en conséquence, ce qui normalement aurait dû conduire à des conclusions erronées. Les conclusions ont pu être rectifiées par l'utilisation de nouvelles techniques plus fines que les précédentes et non encore publiées mais il n'est pas possible d'affirmer que des erreurs n'ont jamais été commises de ce fait.

d) Reste, enfin, les causes d'erreurs dues à la possibilité pour deux individus, pour deux objets, etc... de posséder les mêmes caractères d'identification.

La première cause d'erreur rentre dans le cadre général des erreurs d'observation. Elle a été longtemps étudiée au point de vue théorique par de nombreux auteurs. Mais il est souvent très difficile de chiffrer pratiquement l'importance exacte des erreurs commises, en particulier des erreurs d'inattention.

Les causes d'erreur des 2e et 3e catégories sont encore plus difficiles à déceler et, *a fortiori*, à chiffrer.

Nous n'étudierons ni les unes ni les autres dans ce travail et nous les supposerons négligeables pour nous attacher seulement à celles de la quatrième catégorie : erreurs dues au fait que plusieurs individus, objets, etc... possèdent les mêmes caractères d'identification.

***

En réalité, il ne nous a pas été possible de calculer directement la probabilité d'une telle erreur. Tout ce que nous avons pu faire, dans le cas le plus favorable, c'est l'inverse, c'est-à-dire dégager les éléments permettant de calculer la probabilité pour qu'un ensemble de traces ait une origine déterminée. Dans ce but il faut que nous puissions :

1° Dresser la liste de toutes les origines possibles.

2° Calculer la probabilité pour qu'un ensemble de traces identiques à celles retrouvées sur les lieux, aient pu être produites par chacune des origines considérées. Si nous appelons alors p_i celle d'une origine i nous admettrons que cette probabilité d'origine i est :

$$\varpi_i = \frac{p_i}{\Sigma\, p_i} \tag{1}$$

la sommation étant étendue à toutes les origines possibles.

Ce qui rend un tel problème rarement soluble et toujours très délicat, c'est la difficulté de calculer ces probabilités d'origine. Celles-ci sont en effet très complexes.

Considérons, par exemple, le cas d'une empreinte digitale :

Pour calculer la probabilité qu'un individu donné ait pu produire une trace identique, il faut :

1° que l'individu puisse être sur les lieux,

2° qu'il ait produit une trace,

3° avec la région convenable du doigt qui porte le groupement caractéristique,

4° que la trace ne soit pas détruite ou oblitérée par une autre trace, etc...

En vertu du principe des probabilités composées p_i est alors le produit des probabilités de chacune de ces éventualités.

Une certaine simplification peut toutefois s'introduire du fait qu'un de ces facteurs sera le même pour toutes les origines et pourra ainsi être supprimé.

Dans une usine de torréfaction de café nous avons trouvé le moulage d'un doigt dans de la poussière impalpable de café sur un rebord de fenêtre par où avaient pénétré et ressorti cinq ou six cambrioleurs. La probabilité pour qu'une telle trace ait pu se produire et se soit conservée intacte est certainement très faible, mais nous n'en pouvons pas moins la négliger. En fait, la trace existe et le facteur de probabilité dû à ce fait exceptionnel est le même pour toute origine.

La probabilité de présence (probabilité de se trouver sur les lieux) est, au contraire, très variable. Il est bien évident que ce facteur est nul pour un grand nombre d'individus dont il n'y a dès lors pas lieu de tenir compte; par exemple, il est nul pour tous les Chinois s'il s'agit d'un crime commis en France. Mais il est très rare que l'expert dispose des éléments susceptibles de se faire une opinion sur la valeur de ces facteurs. Du reste, il s'agit dans bien des cas, uniquement de témoignages, et il paraît bien difficile d'attribuer à ces derniers, une grandeur numérique indispensable pour un calcul de probabilité.

Il est des cas, cependant, où le calcul peut être poussé jusqu'au bout, c'est celui où l'enquête préalable a formellement éliminé toutes les origines, sauf un très petit nombre d'entre elles et où les charges autres que les traces soumises à l'expert restent les mêmes pour les origines retenues. Ces cas peuvent paraître purement théoriques, mais on peut cependant trouver un certain nombre de circonstances où ces conditions sont réalisées :

a) Sur un document manuscrit, existe une correction qui en altère le sens. La correction n'a pu être effectuée que par le signataire ou le propriétaire du document.

b) Une lettre anonyme a été nécessairement tapée dans un établissement qui ne possède que quatre machines car il ne s'est écoulé que quelques instants entre le moment où un certain détail a pu être connu et le dépôt de la lettre.

c) Deux individus ont commis un vol dans une maison et ont laissé une trace de pas à l'intérieur. L'un d'eux a fait le guet à l'extérieur et l'autre a blessé grièvement une personne à l'intérieur de la maison. Les deux malfaiteurs sont arrêtés en flagrant délit. Le coupable est donc sûrement l'un d'entre eux.

d) Une bande d'individus poursuivis a tiré sur les gendarmes : l'un a tué des poursuivants, les autres ont tiré en l'air. La balle extraite du corps de la victime provient donc nécessairement d'un des revolvers à l'exclusion de tout autre (1).

Dans ces différents cas, on voit immédiatement que, par définition, la probabilité de présence est nulle pour tous les individus autres que les inculpés, et égale pour chacun d'eux. Si l'on appelle alors t_1, t_2... t_i la probabilité pour que les individus, 1, 2... i puisse produire des traces identiques à celles observées sachant qu'ils étaient sur les lieux et ont effectivement produit une trace de même nature, calcul que l'expert peut souvent effectuer, on aura :

$$\frac{\varpi_1}{\varpi_2} = \frac{t_1}{t_2} \tag{2}$$

et comme

$$\varpi_1 + \varpi_2 + \ldots\ldots + \varpi_i = 1, \tag{3}$$

la probabilité d'origine i devient :

$$\varpi_i = \frac{t_i}{\Sigma\, t_i}. \tag{4}$$

(1) L'identification des balles tirées par une arme à feu est un problème qu'il serait extrêmement intéressant de traiter par les méthodes de calcul des probabilités, mais qui est très difficile à mettre en équation : sur une balle tirée par un revolver ou un pistolet automatique, on observe deux catégories de stries : des sillons larges, profonds et réguliers, dont le nombre varie de quatre à six, dont l'orientation, le pas et la largeur sont également variables. Il serait très facile, mais peu intéressant, d'utiliser ces données dans un calcul, il suffirait d'établir la statistique des différentes armes de chaque fabrication. Il existe également sur les balles des petites stries dont la position, la profondeur, l'allure paraissent distribuées au hasard et qui se retrouvent sur différentes balles tirées par la même arme. Ces stries sont donc beaucoup plus intéressantes et permettent de départager deux armes provenant d'un même lot de fabrication. Malheureusement, il est très difficile de les définir numériquement. Les balles n'ayant jamais rigoureusement le même calibre, les stries seront plus ou moins profondes, leur position même variera légèrement. D'autre part, ces stries sont dues à des défauts de polissage de l'âme de l'arme. Le moindre grain de poussière ajoutera des stries supplémentaires, tandis que le tir d'une balle peut supprimer l'aspérité qui produit une strie et celle-ci ne réapparaîtra pas au tir suivant.

Puisque, par hypothèses, nous avons tenu compte de tous les éléments d'appréciation dont on disposait, on peut dire que dans de tels cas, qui restent relativement rares, les valeurs ainsi calculées correspondent exactement à la probabilité d'origine des traces.

Théoriquement, on voit que les quantités t_1, t_2... t_i n'interviennent que par leur rapport. Mais pratiquement, il est indispensable de ne pas oublier que les calculs qui permettent d'obtenir ces quantités résultent plus ou moins de considérations théoriques. En particulier, les valeurs extrêmement petites des probabilités résultent toujours d'extrapolations. Or, il faut bien le dire, rien ne permet d'affirmer que les lois de probabilité sont encore valables en dehors de la base expérimentale qui a permis de les établir. Considérons, par exemple, le cas de l'identification d'un plomb de chasse isolé à partir de la réserve de plomb utilisée par un chasseur. L'expérience montre que le poids des plombs de chasse d'une même série se répartit sensiblement suivant une loi de Laplace-Gauss. Supposons alors que, pour deux chasseurs, les écarts réduits soient respectivement de 3,3 et 6,1, les probabilités correspondantes seraient respectivement 10^{-3} et 10^{-9}, et le rapport t_1/t_2 serait 10^{6}; l'un des chasseurs serait formellement éliminé. En vérité, rien ne permet une telle extrapolation. La vérification de la loi de Laplace-Gauss a été effectuée en dénombrant un millier de plombs; on ne peut pas savoir si cette loi reste valable au delà de cette limite. Il suffirait, par exemple, que l'un des chasseurs ait mélangé dans un fond de tiroir quelques plombs d'un calibre différent et que le mélange n'aît pas été homogène, pour fausser tous les calculs.

Dans un tel cas, l'expert devra avertir le juge que les deux origines sont peu probables et qu'il importe de rechercher la possibilité d'une troisième. Dans le cas où l'enquête confirmerait l'impossibilité d'une troisième origine, on ne pourrait guère utiliser le premier résultat que comme une indication provisoire qu'il faudrait confirmer par d'autres méthodes. Pour les plombs de chasse, par exemple, on pourrait essayer de rechercher et doser les impuretés du plomb par une méthode spectrographique.

Dans le cas plus général où le nombre de suspects n'est pas strictement limité, tout ce que peut faire l'expert c'est comparer deux probabilités d'origine car on a évidemment :

$$\frac{\varpi_1}{\varpi_2} = \frac{p_1}{p_2}. \tag{5}$$

Deux cas peuvent alors se présenter :

a) les facteurs qui ne dépendent pas des traces, les facteurs de présence en particulier, sont identiques. On peut les éliminer et on retrouve l'équation (2).

On ne sait pas calculer ϖ_1, et ϖ_2, mais on peut évaluer leur rapport.

b) les facteurs qui ne dépendent pas des traces sont eux aussi différents, l'expert ne peut plus terminer son calcul. Il pourra encore essayer d'évaluer le rapport t_1/t_2, mais il devra préciser que ce rapport n'est égal au rapport de probabilité d'origine que dans la mesure où le rapport des autres facteurs est égal à 1, ce qui peut se résumer par l'expression « toutes choses étant égales d'ailleurs ».

Dans certains cas, du reste très fréquents, l'expert n'est consulté que sur une seule origine. Puisque l'on consulte l'expert, c'est évidemment que les autres origines n'ont pas une probabilité nulle. Or, l'expert, comme nous venons de le voir, ne peut établir que des rapports de probabilité. Il est tout naturellement amené à comparer l'origine sur laquelle il est consulté avec une origine arbitrairement choisie, ce qui nous conduit à étudier la notion d'individus moyens.

Nous allons maintenant rechercher les conditions à réaliser pour utiliser une telle notion que nous avons introduite pour la première fois dans notre étude sur les groupes sanguins.

ORIGINE MOYENNE

Supposons que, pour tous les individus d'un certain groupe, nous ayons déterminé la loi de répartition des fréquences des différentes valeurs t d'un des facteurs ou d'un groupe de facteurs, nous pourrons calculer une moyenne t_m et, en supposant que les valeurs des autres facteurs et en particulier que les probabilités de présence soient les mêmes pour tous les individus du groupe, nous pourrons dire que t_m est proportionnel à la probabilité pour un individu pris au hasard, d'être à l'origine de la trace. On peut alors considérer que le rapport t_i/t_m est caractéristique de l'individu i. S'il est égal à 100, i a cent fois plus de chances qu'un individu du groupe pris au hasard, d'être l'auteur de la trace.

L'introduction de cette notion peut rendre d'appréciables services dans l'application des méthodes statistiques à ce genre de question, mais il convient d'être extrêmement prudent dans la définition du groupe choisi pour calculer t_m si l'on veut éviter des erreurs grossières susceptibles de fausser complètement toutes les conclusions.

Considérons le cas où l'on recherche dans un village l'auteur d'une lettre anonyme. Choisir comme élément de statistique l'ensemble de toutes les encres utilisées en France, pourrait masquer le fait, qu'il n'y a qu'un seul revendeur d'encre dans le village; si l'encre identifiée est celle vendue par celui-ci, tous les habitants du village posséderont probablement un flacon d'encre identique. La présence d'un élément très rare dans cette encre n'aura alors aucune valeur probante.

Voici maintenant un exemple contraire :

Un cadavre est repêché dans un canal de la région parisienne, la tête enveloppée dans un mouchoir à carreaux. L'enquête conduit chez un individu chez lequel on trouve des mouchoirs similaires. L'examen détaillé de ces linges révèle l'existence de défauts de fabrication dans la trame du tissu, défauts qui permettent de dire qu'ils proviennent d'un même coupon. On retrouve le fabricant, ce qui permet de préciser que ce coupon comportait 12 douzaines de mouchoirs et qu'il a été débité par douzaines, dans la région de Lille. Or, l'inculpé avait sa mère dans l'agglomération où se trouvait le détaillant et des mouchoirs présentant ce même défaut ont été retrouvés chez cette femme.

On voit que, dans les deux cas, la notion de valeur moyenne n'aurait pas grand sens puisqu'elle ne tient pas compte des facteurs géographiques qui jouent un rôle essentiel. En fait, cela tient à ce que l'on n'aurait pas respecté les hypothèses à la base du calcul, c'est-à-dire l'égalité au moins approximative des facteurs autres que pour tous les éléments du groupe.

Si l'on veut, en particulier, que le calcul utilise correctement toutes les informations, il faut que ces autres facteurs soient nuls pour tous les éléments étrangers au groupe. Or, dans le premier cas envisagé, limiter le groupe aux seuls habitants du village, serait encore insuffisant, car l'auteur de la lettre anonyme peut appartenir à un village voisin ou rédiger sa lettre hors de chez lui. Les calculs de probabilité, dans de tels cas, doivent être considérés comme n'ayant aucun sens.

Comme nous l'avons déjà rappelé, nous avons déjà introduit cette notion d'individu moyen dans notre premier mémoire sur les groupes sanguins présenté ici-même (4); cette notion intervenait encore implicitement dans notre travail sur la valeur de la preuve en dactyloscopie (5) (6) et (7) et de nouveau directement dans notre troisième application des méthodes statistiques aux problèmes de police scientifique (8), mais à cette époque nous n'avions pas clairement perçu les bases précises sur lesquelles doit s'appuyer cette notion. Dans les deux derniers travaux il n'y a cependant aucune modification à apporter à nos conclusions. Dans le cas des encres, nous avions en effet établi notre statistique conformément aux principes énoncés plus haut; dans le cas des empreintes digitales, il n'existe actuellement aucune relation connue entre le détail des dessins utilisés pour l'identification et les races humaines il n'y a donc aucune précaution particulière à prendre pour établir les statistiques utiles.

Il n'en est pas toujours de même avec les groupes sanguins. Il est en effet bien connu que la proportion des individus de chaque groupe varie très sensiblement avec les races humaines. Supposons, par exemple, que le père présumé soit du groupe B relativement rare en France : si la mère ne possède pas le gène B et que l'enfant le possède, il s'agit là d'un élément de grande valeur. Mais s'il existe dans le voisinage un groupe d'asiatiques pour lesquels la fréquence du groupe B est beaucoup plus élevée, la valeur de la preuve se trouve nettement modifiée; dans ce cas particulier, la notion d'individu moyen perdrait encore toute signification précise.

Lorsque la notion d'origine moyenne ne s'applique pas, la seule ressource de l'expert reste donc actuellement de soumettre au juge l'ensemble des éléments d'appréciation qu'il a pu rassembler et de lui laisser le soin de les interpréter.

En résumé, l'application des méthodes statistiques à la preuve indiciale se heurte souvent à des difficultés très importantes. Ajoutons que lorsque l'on est obligé d'établir une statistique particulière à l'objet étudié, il s'agit presque toujours d'un travail considérable. Il serait donc particulièrement intéressant avant d'entreprendre une telle étude, non de savoir si elle aboutira, ce qui supposerait que l'on connaisse d'avance le résultat, mais quelle probabilité l'on a d'arriver à un résultat significatif, autrement dit, de connaître à l'avance l'efficacité de la méthode que l'on se propose d'appliquer. Il est d'ailleurs une autre raison particulièrement importante qui rend impérieuse cette notion d'efficacité, c'est la limitation du nombre possible d'expériences.

EFFICACITÉ DES MÉTHODES UTILISÉES

Dans les sciences expérimentales, en général, on peut, au moins théoriquement, répéter les mêmes expériences d'une manière illimitée, faire varier les différents facteurs séparément, etc... Il n'en est pas de même en police scientifique. Si certaines expériences physiques et plus spécialement optiques n'altèrent pas la pièce à conviction, les analyses chimiques la détruisent plus ou moins. Il est donc indispensable d'effectuer à l'avance un choix dans la nature des essais que l'on aura à effectuer. Si le choix a été heureux, on arrive à une conclusion significative; s'il a été malheureux, non seulement on a perdu tout le bénéfice de l'effort effectué, mais encore on a anéanti toute possibilité ultérieure.

On conçoit que les méthodes statistiques soient particulièrement indiquées pour prévoir à l'avance l'efficacité éventuelle des méthodes. Cependant, à notre connaissance, il n'y a guère que pour les groupes sanguins que ces méthodes ont été appliquées (9); les auteurs ont en effet dans ce cas étudié systématiquement la probabilité d'une élimination formelle de la paternité, lorsque l'on suppose que le père présumé ne l'est pas, que l'on connaît le groupe sanguin de la mère et celui de l'enfant.

Il faut cependant remarquer que, même sans avoir effectué aucun calcul, un expert averti sait par expérience que telle méthode donne généralement de bons résultats, telle autre est à rejeter. Il fait en somme appel à des méthodes non chiffrées que l'on pourrait appeler « préstatistique ».

Que l'on se limite au cas de conclusion formelle ou que l'on exploite complètement les possibilités offertes par le calcul des probabilités, il est particulièrement utile de savoir qu'il existe une dissymétrie fondamentale dans la probabilité d'une conclusion significative.

DISSYMÉTRIE FONDAMENTALE

Lorsqu'il s'agit, par exemple, de démontrer la présence d'un individu sur un lieu de crime, on montre l'identité des indices trouvés sur les lieux avec les objets propres à l'inculpé. Il suffit, alors, qu'un objet présente un caractère incompatible avec la trace pour être immédiatement éliminé, tandis que les objets qu'il y a lieu de retenir ne le sont que par l'absence de caractères dissemblables. Si, par exemple, on compare une balle de revolver avec l'arme trouvée sur l'inculpé, il suffit d'un nombre différent de rayures fondamentales pour éliminer l'arme. Au contraire, l'identité est délicate à établir et résulte de la coïncidence de tous les éléments fondamentaux et de très nombreuses petites rayures accessoires. On voit que, dans ce cas, les difficultés d'identifications jouent en faveur de l'inculpé.

Au contraire, s'il s'agit d'un document falsifié, le moindre caractère différentiel suffit à démontrer la fraude, tandis que l'authenticité n'est jamais qu'une présomption. L'expert n'a pas trouvé de différences, c'est peut-être parce qu'il n'a pas su les trouver. On conçoit, par exemple que, si dans un texte lavé on reconstitue le texte effacé par fluorescence, il ne saurait y avoir le moindre doute sur l'altération par lavage chimique. Au contraire, si le faussaire est habile, il n'est pas impossible qu'il falsifie le document sans laisser de trace. Il y a encore dissymétrie dans la valeur de la preuve, mais celle-ci joue alors contre l'inculpé.

Cette même dissymétrie n'existe pas seulement dans le cas d'élimination formelle et on la retrouverait dans les exemples traités en détail dans nos mémoires précédents. Prenons, à titre d'exemple, la comparaison des encres au campêche-chrome et supposons, pour simplifier, que nous nous limitions à un seul caractère, par exemple la concentration du fer dans l'encre. Rappelons très rapidement les résultats déjà obtenus en conservant les mêmes notations. Posons :

$$tg\,\xi = \frac{\text{moyenne des densités optiques de l'encre}}{\text{moyenne des densités optiques du fer.}} \qquad (6)$$

Considérons deux textes à comparer, soient ξ_1 *et* ξ_2 les valeurs des angles définis par la formule (6), σ_{ξ_1} et σ_{ξ_2} les valeurs des écarts quadratiques sur ξ_1 et ξ_2. On sait que la loi de répartition des écarts se rapproche d'autant plus d'une loi de Laplace-Gauss, que l'on opère sur des moyennes provenant d'un plus grand nombre de mesures. Il en est de même *a fortiori* de la différence entre les deux séries de mesures et, de plus, le carré de l'écart type σ_ξ de la différence $\lambda = \xi_1 - \xi_2$ est la somme des carrés des écarts types de chaque série.

$$\sigma_\xi^2 = \sigma_{\xi_1}^2 + \sigma_{\xi_2}^2.$$

La probabilité de la différence λ est alors :

$$\frac{1}{\sigma_\xi \sqrt{2\,\pi}}\, e^{-\frac{\lambda^2}{2\sigma^2_\xi}}\, d\,\lambda. \qquad (7)$$

La valeur moyenne de cette probabilité étant $\dfrac{d\,\lambda}{2\,\sigma_\xi\sqrt{2\,\pi}}$ le rapport entre la probabilité trouvée et la valeur moyenne est

$$\rho = \sqrt{2}\, e^{-\frac{\lambda^2}{2\,\sigma_\xi^2}}. \qquad (8)$$

Si l'on suppose que λ est essentiellement dû à une différence réelle entre les deux encres, on peut trouver encore une valeur λ sinon identique, du moins voisine, en multipliant par k le nombre de mesures dans chaque série, mais $\sigma^2{}_{\xi_1}$ et $\sigma^2{}_{\xi_2}$ étant tous deux inversement proportionnels au nombre de mesure, $\sigma^2{}_\xi$ sera aussi divisé par k et le rapport précédent deviendra :

$$\rho_k = \sqrt{2}\; e^{-\frac{\lambda^2 k}{2\,\sigma_\xi^2}} = \sqrt{2}\left(e^{-\frac{\lambda^2}{2\,\sigma_\xi^2}}\right)^k$$

$$\rho_k = \sqrt{2}\left(\frac{\rho_1}{\sqrt{2}}\right)^k \qquad (9)$$

même si ρ_1 n'a pas une valeur très petite, il suffira d'une valeur relativement faible de k pour que ρ_k ait une valeur hautement significative.

Si l'on suppose au contraire que λ n'est dû qu'aux écarts de mesure, il faut faire une hypothèse sur la loi de répartition des valeurs de ξ et ceci introduit une première incertitude. Supposons, pour fixer les idées, que cette loi soit une loi de Laplace-Gauss dont l'écart quadratique moyen est Θ_ξ. La probabilité de chaque valeur ξ est alors :

$$d\,P_\xi = \frac{1}{\Theta_\xi \sqrt{2\,\pi}}\, e^{-\frac{\xi^2}{2\,\Theta_\xi^2}}\, d\,\xi. \qquad (10)$$

S'il s'agit du même tracé, la probabilité du couple ξ_1, ξ_2 est alors :

$$d\,P_1 = \frac{1}{2\,\pi\,\Theta\;\sigma_\xi}\, e^{-\frac{\xi^2}{2\,\Theta_\xi^2}-\frac{\lambda^2}{2\,\sigma_\xi^2}}\, d\xi\; d\lambda.$$

S'il s'agit de deux tracés n'ayant aucune corrélation,

$$d\,P_2 = \frac{1}{2\,\pi\,\Theta_\xi^2}\, e^{-\frac{\xi_1^2+\xi_2^2}{2\,\Theta_\xi^2}}\, d\xi_1\, d\xi_2$$

et l'on a :

$$r = \frac{d\,P_1}{d\,P_2} = \frac{\Theta_\xi}{\sigma_\xi}\; e^{-\frac{\xi_1^2+\xi_2^2-\xi^2}{2\,\Theta_\xi^2}+\frac{\lambda^2}{2\,\sigma_\xi^2}}. \qquad (11)$$

Si, maintenant, nous multiplions par k le nombre de prélèvements dans chaque tracé, ξ_1 et ξ_2 tendent vers ξ. σ^2_ξ est inversement proportionnel à k, λ naturellement varie, mais le rapport $\frac{\lambda^2}{2\,\sigma^2_\xi}$ garde une valeur moyenne constante. On aura donc :

$$r_k = \frac{\Theta_\xi\,\sqrt{k}}{\sigma_\xi}\, e^{-\frac{\xi^2}{2\,\Theta_\xi^2}+\frac{\lambda^2}{2\,\sigma_\xi^2}}$$

et en moyenne :

$$r_k = r\,\sqrt{k}. \qquad (12)$$

Il faudra multiplier par un nombre considérable de mesures pour augmenter sensiblement la valeur de r.

Si donc les premiers résultats permettent de penser que les deux tracés n'ont pas la même origine, l'expert aura intérêt à multiplier les mesures du même caractère pour établir avec une haute probabilité que la différence est significative. Si, au contraire, les premiers résultats permettent de penser à une origine unique, l'expert aura intérêt à choisir d'autres caractères pour renforcer la valeur de la probabilité en faveur de l'identité d'origine.

CONCLUSIONS

On voit, en résumé, que le calcul des probabilités peut apporter une aide précieuse à la justice. Mais son emploi est très délicat et ce serait une grave erreur de croire qu'il suffit d'une connaissance superficielle de nos méthodes de calcul à un spécialiste de la seule police scientifique. Nous estimons qu'il est indispensable qu'une collaboration étroite s'établisse entre les spécialistes des deux disciplines et que chacun d'eux se pénètre intimement des méthodes de l'autre, car ils ne doivent jamais oublier les conséquences graves d'une conclusion auxquelles leurs calculs les conduira, conséquences pouvant se traduire, en cas d'erreur, par la condamnation d'un innocent ou, au contraire, par un grave préjudice moral et matériel pour la victime si l'on acquitte à tort un coupable.

Lucien AMY.
Laboratoire Municipal
Paris.

BIBLIOGRAPHIE

(1) RABELAIS — *Pantagruel,* livre 1, chapitre 39.
(2) J. BERTRAND — *Calcul des probabilités,* Paris, Gauthier-Villars, 1889, pages 319 à 327.
(3) V. BALTHAZARD — *De l'identification par les empreintes digitales* — Comptes rendus Académie des Sciences. — **152,** 1911, page 1862.
(4) L. AMY. — *Probabilité, groupes sanguins et paternité.* Journal de la Société de Statistique de Paris, t. **86,** pages 102 à 112 (1945).
(5) L. AMY. — *Valeur de la preuve en dactyloscopie.* Journal de la Société de Statistique de Paris (1re partie) t. **87,** pages 80 à 87 (1947).
(6) L. AMY. — *Valeur de la preuve en dactyloscopie.* — Journal de la Société de Statistique de Paris (2e partie) t. **87,** pages 183 à 195 (1947).
(7) L. AMY. — *Recherches sur l'identification des traces papillaires.* — Annales de Médecine Légale. — Mars 1948.
(8) L. AMY. — *Recherches statistiques sur l'identification des encres au campêche-chrome.*— Journal de la Société de Statistique de Paris. t. **91,** pages 188 à 202 (1950).
(9) Voir la bibliographie dans la thèse présentée par Mme BARGE-NICOUX en 1944, devant la Faculté de Pharmacie de Paris, pour l'obtention du grade de Docteur en pharmacie.