

# THÉORIE DES PROBABILITÉS ET PRATIQUES INDUSTRIELLES

## Les débuts difficiles du contrôle statistique de qualité en France

*À la mémoire de Maurice Dumas (1898-1993)*

Le terme de « révolution probabiliste » recouvre un ensemble de méthodes, apparues dans les premières années du xx<sup>e</sup> siècle dans les pays développés et liées à la généralisation de la fabrication industrielle en grandes séries. En France, le domaine de l'organisation militaire et des industries d'armement a été un terrain d'expérimentation particulièrement fécond pour le contrôle statistique de la qualité. C'est ce domaine que deux artilleurs – le chef d'escadron d'artillerie J.E. Estienne et M. Dumas, jeune ingénieur de l'artillerie navale – ont défriché, parmi les tout premiers, en mobilisant les outils et les modes de pensée du raisonnement probabiliste. Il faut reconnaître à ces pionniers le mérite d'avoir accompli un travail remarquable sur ces questions et cela, avec les moyens théoriques relativement limités de leur époque.

PAR **Denis BAYART**, CENTRE DE RECHERCHE EN GESTION – ÉCOLE POLYTECHNIQUE (\*)

*Cet article est dédié à Maurice Dumas, pionnier des applications statistiques dans l'industrie et infatigable défenseur du raisonnement bayésien et des Lois de Lhoste, dont l'amitié m'a été un précieux stimulant. Je remercie Bernard Bru, Yves Cohen, Pierre Crépel, Alain Desrosières pour leur aide et leurs encouragements.*

(\*) [bayart@polytechnique.fr](mailto:bayart@polytechnique.fr)

De nombreux travaux ont été consacrés à la « révolution probabiliste » dans les sciences, exactes ou humaines (1), mais très peu se sont intéressés à la mise en œuvre de ces nouveaux modes de pensée dans l'organisation industrielle et dans la gestion des entreprises. Les possibilités d'application y sont nombreuses, depuis les études de marché et la prévision des ventes jusqu'à la planification des opérations et ce qu'on a appelé la recherche opérationnelle. Le domaine de l'organisation militaire et des industries d'armement a souvent été un terrain d'expérimentation de ces méthodes. C'est particulièrement le cas en France pour ce qui concerne le contrôle statistique de la qualité, domaine que deux artilleurs – un officier et un ingénieur – ont défriché, parmi les tout premiers, en mobilisant les outils et les modes de pensée du raisonnement probabiliste. Ce domaine recouvre un ensemble de méthodes, apparues simultanément, surtout dans les années 1920-1930, dans plusieurs pays industrialisés (France, Allemagne, États-Unis, Grande-Bretagne...) et liées à la généralisation de la fabrication industrielle en grandes séries ainsi qu'à des évolutions dans les conventions commerciales (apparition de l'assurance qualité aux États-Unis, au début du siècle).

Le plus ancien de ces travaux probabilistes, publié en 1903, est nettement antérieur à la vague d'intérêt, dans le monde industriel, pour le contrôle statistique. Il est dû au chef d'escadron d'artillerie J.E. Estienne, plus connu comme promoteur des chars d'assaut français pendant la première Guerre mondiale (2). Il s'agit en fait d'un exemple grâce auquel l'auteur illustre sa propre théorie probabiliste, le « calcul des cotes ». Mais cet exemple, « Application à l'analyse des conditions de réception de lots d'objets fabriqués », est développé sur plusieurs pages et témoigne d'une réelle pénétration du sujet (3).

Le second auteur est Maurice Dumas, jeune ingénieur de l'artillerie navale qui reprend le problème en 1925 et le traite de façon approfondie dans le *Mémorial de l'artillerie française* (4). La même année, un pharmacien militaire, Lucien Vallery, fait également paraître un article, plus court, sur l'utilisation du calcul des

probabilités pour le contrôle (5). Les entreprises civiles ne semblent pas, en revanche, s'intéresser à la question. Le domaine de l'armement – et notamment de l'artillerie, par suite de la formation que recevaient les officiers et ingénieurs – apparaît comme un champ propice au développement de la pensée probabiliste. Par-delà le désir de faire reconnaître les mérites de ces pionniers qui ont accompli, il faut le souligner, un travail remarquable sur la question, et cela avec des moyens théoriques relativement limités, je souhaite explorer quelques pistes afin d'expliquer pourquoi, en définitive, leurs réflexions sont restées à peu près sans conséquences pratiques jusqu'à la veille de la Seconde Guerre mondiale. C'est seulement en 1936-1937, en effet, que l'Administration militaire commence à expérimenter réellement ces méthodes ; cet intérêt est principalement déclenché, semble-t-il, par l'action militante du physicien et statisticien Georges Darmois, s'appuyant sur des travaux américains et britanniques (6) publiés entre 1930 et 1935.

#### POUR UNE HISTOIRE DES CRITÈRES D'EFFICACITÉ DANS LES TECHNIQUES DE GESTION

La problématique que j'adopte résulte d'un croisement entre la recherche en gestion, qui est mon activité principale de chercheur, et l'histoire des techniques de production et d'organisation. Le contrôle statistique de qualité est un ensemble de techniques de gestion de la production qui aident à prendre des décisions face à la variabilité tant des produits fabriqués que des conditions de fabrication. Ces techniques interviennent aussi dans la régulation des rapports client-fournisseur, comme nous le verrons plus loin, en explicitant des règles d'acceptation ou de rejet des lots de produits fabriqués, lorsqu'on en examine des échantillons. Ce sont donc des méthodes qui permettent de porter des jugements en certaines situations récurrentes dans l'activité industrielle ou commerciale, au nom de certains critères d'efficacité

(1) KRÜGER L. et al. (eds) : *The Probabilistic Revolution*, 2 vol., Cambridge : MIT Press, 1987

(2) Estienne est un personnage très intéressant de l'histoire militaire et industrielle française. Voir l'article de DETCHESSAHAR M. & LEMARCHAND Y. : « Des hommes et des projets dans l'urgence. La naissance du char d'assaut français, 1914-1918 », *Annales des Mines, Série Gérer et comprendre*, juin 2001, pp. 43-54.

(3) ESTIENNE J.E. : « Essai sur l'art de conjecturer », *Revue d'artillerie*, t. 61, mars 1903 : 405-449, mai 1903 : 73-117 ; t. 62, avril 1904 : 5-39, mai 1904 : 65-97. L'analyse des conditions de réception occupe les pages 439-449 (mars 1903).

(4) DUMAS M. : « Sur une interprétation des conditions de recette », *Mémorial de l'artillerie française*, t.4 (1925), fasc.2, pp. 395-438. Maurice Dumas (1898-1993) était polytechnicien de la promotion 1919S.

(5) VALLERY L. : « Le calcul des probabilités appliqué au prélèvement des échantillons en matière d'expertise », *Annales des falsifications et des fraudes*, décembre 1925, pp. 597-605. L'article traite du contrôle de la qualité des vins par des dosages ; bien qu'intéressant, il est beaucoup moins approfondi que les deux autres et nous le laisserons de côté.

(6) Les plus notoires sont : SHEWHART W.A. : *Economic Control of Quality of Manufactured Products*, New York, Van Nostrand, 1931. PEARSON E.S. : *The application of statistical methods to industrial standardization and quality control*, British Standards Institution, London, 1935.

ou de pertinence. Écrire l'histoire de ces méthodes, c'est contribuer à l'histoire des critères d'efficacité en vigueur dans les organisations industrielles. En étudiant leur genèse, on voit les acteurs de l'organisation mettre en forme de nouveaux critères d'efficacité ou faire évoluer ceux existants. Les critères d'efficacité étant un objet d'étude spécifique de la recherche en gestion, il est logique de trouver là un point de croisement avec l'approche historique. Le point de vue de la recherche en gestion peut « ouvrir de nouvelles boîtes noires », selon l'expression devenue classique en sociologie des sciences et des techniques, en interrogeant des objets historiques, dans une perspective différente.

Les recherches menées sur des organisations contemporaines montrent qu'il existe une grande diversité de ces critères d'efficacité selon les organisations, selon les types de production et selon les situations (7). Elles montrent aussi le rôle crucial que joue la traduction opérationnelle des critères sous forme de règles et de dispositifs concrets pour les collectifs de travail.

Les techniques du contrôle statistique de qualité constituent un bon objet pour étudier ces dispositifs concrets. Elles reposent en effet sur des raisonnements et concepts de statistique mathématique qui seraient inutilisables par la main d'œuvre des usines s'ils n'étaient pas traduits en moyens opérationnels tels que procédures, graphiques, feuilles de calcul, etc., qui ramènent les processus mathématiques à des opérations arithmétiques et graphiques élémentaires, selon le principe de la division du travail intellectuel qui avait tant séduit Charles Babbage. Nous appelons *outils cognitifs* ces moyens opérationnels qui donnent un corps aux idées abstraites et permettent de les intégrer dans un collectif de travail. Il est possible de les étudier sous les mêmes angles que des outils ou des machines au sens habituel, en s'interrogeant par exemple sur leur ergonomie, sur les possibilités de coordination qu'ils offrent aux agents du collectif, sur leur compatibilité avec l'organisation sociale du travail dans l'entreprise (8).

Des outils cognitifs performants constituent une première gamme d'exigences pour qu'une méthode de gestion puisse être mise en œuvre. Or, dès 1925, Maurice Dumas a conçu des outils cognitifs très

proches de ceux qui, par la suite, ont été un facteur de succès de ces méthodes ; nous les examinerons de plus près dans la suite. Si nous voulons expliquer l'absence d'applications, nous sommes dès lors incités à en chercher ailleurs la raison.

Il existe une deuxième gamme d'exigences tenant au fait qu'une opération de contrôle (le contrôle de qualité en est une) n'a pas de sens si elle n'est pas intégrée dans un système plus large. Nous verrons Estienne invoquer explicitement cette raison. L'expérience de AT&T aux États-Unis montre aussi que la vision « à l'échelle de l'entreprise tout entière » est un facteur de succès important. L'insuccès des propositions de Dumas est lié, nous le verrons, à l'absence d'une vision plus large de l'interaction client-fournisseur et des processus industriels qui l'encadrent. À ce propos, il est intéressant de renvoyer le lecteur au concept de « réseaux de métrologie » employé par Christian Licoppe (9) : dans le réseau des processus industriels et des procédés métrologiques qui établissent la qualité des biens (et même plus largement leur identité, ce qui les définit en termes scientifiques), la théorie des probabilités apporte de nouvelles possibilités de mesure, là où l'on ne voyait auparavant que du hasard. Des notions intuitives telles que le « risque du client » se voient dotées d'une signification précise et opérationnelle. Une nouvelle ère du contrôle s'ouvre, qui entraînera de nouveaux modes d'action, notamment la mise sous contrôle statistique des systèmes de production, technique cruciale, à l'heure actuelle, en microélectronique.

L'ordre de présentation sera thématique plutôt que chronologique. D'abord, nous verrons comment, avec ces nouvelles méthodes, on définit la qualité des fabrications, et selon quels critères. Le recours aux méthodes statistiques quantitatives a pour effet de forcer à définir la qualité de façon précise et tranchée, pas forcément pertinente pour l'activité des différents acteurs. Autre nouveauté : il s'agit de la qualité d'un ensemble de produits, ce que nos auteurs appellent un « lot de fabrication ». Ces redéfinitions sont indissociables des épreuves qui établissent la bonne ou la mauvaise qualité, autrement dit de ce qui fait preuve dans les modes de raisonnement et de pensée statistiques. Dans un deuxième point, nous verrons comment émergent différents critères permettant de trier, comparer, ordonner les méthodes possibles selon leurs performances, et comment on apprécie ces dernières. Enfin, nous étudierons les questions soulevées par

(7) Certains textes programmatiques de recherche en gestion définissent en effet comme objet de recherche possible « l'analyse des conditions de la performance des agents producteurs de biens et de services ». Voir par exemple les études réunies dans : CHARUE-DUBOC F. (éd) : *Des savoirs en action. Contributions de la recherche en gestion*. L'Harmattan, Paris, 1995.

(8) Un exemple classique d'analyse des outils cognitifs est : HUTCHINS, Edwin : *Cognition in the Wild*, Cambridge (Mass.), MIT Press, 1995. L'étude porte sur l'opération collective consistant à faire le point à bord d'un navire militaire.

(9) LICOPPE C. : *La Formation de la pratique scientifique : le discours de l'expérience en France et en Angleterre (1630-1820)*, Paris, La Découverte, 1996.

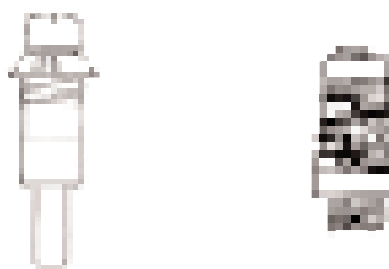
l'insuccès de ces nouvelles méthodes du point de vue de la pratique, et cela sous trois angles : les outils cognitifs permettant de les rendre concrètes ; la mise en scène des démonstrations de validité de ces nouvelles théories ; et, enfin, leur insertion dans les réseaux métrologiques existants. Nous nous appuyons pour ce faire sur des comparaisons avec les méthodes développées aux Laboratoires Bell à la demande de la direction de AT&T, la grande entreprise américaine de téléphone, méthodes qui furent conçues en étroite relation avec les pratiques.

## JUGER LA QUALITÉ SUR ÉCHANTILLON : REDÉFINIR DES PRATIQUES ANCIENNES

Le contexte des opérations envisagées par Estienne et Dumas est assez particulier et bien défini : il s'agit de la « *prise en recette* » de fabrications effectuées par un fournisseur de l'Administration militaire, qu'il s'agisse d'une entreprise privée ou d'un arsenal. C'est une opération possédant un fort caractère administratif, une routine à laquelle il est bien possible que les acteurs ne prêtent pas grande signification. Estienne et Dumas travaillent à élaborer un sens nouveau, plus riche, plus précis : d'abord, définir la qualité sous forme d'une grandeur mesurable qui sera, en l'occurrence, la proportion de défectueux dans le lot ; puis préciser l'entité par rapport à laquelle cette variable prend signification : soit le lot, soit le processus de fabrication.

Comment mesurer la qualité d'un lot de fournitures ? En premier lieu, le contrôle statistique nécessite une définition de la qualité qui conduise à des grandeurs mesurables. Estienne choisit un cas apparemment simple : les fusées d'obus, dispositifs destinés à faire exploser la charge de l'obus lorsqu'il arrive sur l'objectif, et dont le bon fonctionnement se mesure par une variable « *oui/non* ». Dumas reprend le même exemple. Cependant, il s'agit là d'une simplification qui fait peut-être perdre du réalisme au modèle. En effet, les fusées d'obus peuvent être de différents types et leurs performances ne sont pas généralement aussi simples à définir. Le principe d'action d'une fusée est la détonation soit à la percussion, soit au bout d'un laps de temps déterminé par la combustion d'une mèche (souvent réglable avant lancement). Les fusées étaient couramment à double effet. Or, si l'effet par percussion relève bien d'une variable binaire, l'effet de temporisation est mesuré par une durée, variable continue. La fusée était un élément évidemment crucial de l'obus car celui-ci n'explosait pas tout seul. Pendant la guerre de 1914-1918, la mauvaise qualité

des fusées (notamment pour les obus de 75) était un problème sérieux.



Fusées pour obus de 75 (10)

À gauche : Fusée détonateur percutante 24/31 – modèle 1899 – instantanée ou court retard

À droite : Fusée à double effet 22/31 – modèle 1897 – 24 secondes – maxi+percutante

Estienne et Dumas esquivent là un point fort difficile, celui de la mesure de ce qui fait réellement la qualité d'un produit pour les clients ou les utilisateurs. Ce thème parcourt toute l'histoire des politiques de qualité, générant périodiquement des controverses et des renouvellements importants de la pensée et des pratiques. Des travaux modernes de recherche en management (11) établissent clairement qu'il existe des conceptions très différentes de la qualité selon les services ou les fonctions de l'entreprise, entre le marketing, le bureau d'études, la fabrication, le service après-vente, etc. Dans les articles de statistique sur le contrôle de qualité, cependant, par une sorte de convention académique, la question de la mesure de la qualité est généralement considérée comme résolue ou hors sujet.

L'entreprise américaine de téléphone AT&T a développé des méthodes de contrôle statistique de qualité en adoptant une démarche d'investigation (12) qui couvrait l'ensemble de l'activité de l'entreprise, depuis la conception jusqu'à l'installation des centraux téléphoniques chez les clients (qui étaient des compagnies locales de téléphone). Cette démarche a permis de prendre en compte, dans leur diversité, les conceptions de la qualité qui étaient pertinentes pour le but recherché et d'arriver à des compromis se traduisant dans la méthode proposée. Estienne et Dumas adoptent au contraire un point de vue unilatéral, celui de l'Administration militaire. Dumas se trouve amené,

(10) Les illustrations et certaines informations proviennent du site de David DHOORY, <http://canonde75.free.fr/>, consacré au canon de 75.

(11) Voir par exemple : GARVIN D.A. : *Managing Quality : The Strategic and Competitive Edge*, New York : The Free Press, 1988

(12) Décrite en détail dans : BAYART D. : « Savoir organisationnel, savoir théorique et situation : le contrôle statistique sur échantillons », *Entreprises et Histoire*, 1996, n° 13, pp. 67-81

Estienne a certainement bénéficié de l'excellent cours de probabilités dispensé par le colonel Henry à l'École d'artillerie de Fontainebleau. Et la promotion 1919S (celle de Dumas) est la première à recevoir un enseignement de probabilités plus moderne, par le grand mathématicien Paul Lévy. (Guerre de 1914-1918 – Canon de 75mm mis en position contre un avion.)

nous le verrons, à considérer un critère qui intéresse directement le fabricant (critère du risque de rejet d'un lot de bonne qualité), mais il ne l'explique pas comme tel ; son point de vue reste celui du client. En fait, Dumas comme Estienne ne considèrent qu'une situation abstraite, un type pur.

Un travail d'explicitation important est en revanche fait par nos deux auteurs sur ce qui constitue la qualité d'un lot. Ils font certes l'impasse sur les problèmes que soulève la mesure de la qualité d'un élément individuel, mais ils réfléchissent longuement sur ce qu'est un lot, sur ce qu'est un échantillon pris dans le lot, et sur ce que peut signifier une mesure du taux d'éléments défectueux dans l'échantillon et dans le lot.

Dumas définit un lot de mauvaise qualité comme un lot où la proportion d'éléments défectueux est supérieure à un nombre **b** (« *supposé non nul* », ajoute-t-il en puriste !). C'est clair et net. Mais Estienne prend le soin d'articuler un long raisonnement, dont le caractère laborieux reflète sans doute la difficulté du sujet : « *Considérons donc un lot de vingt mille fusées, présenté en recette par un fabricant, et supposons que la conjecture*

*porte sur le nombre de ratés que pourront donner ces vingt mille fusées, une fois en service. Concevons qu'un procédé ignoré ou quelque être infiniment perspicace ait fait connaître que sur les vingt mille fusées, il y en a cent qui rateront, mais sans indiquer lesquelles ; prononcera-t-on la recette ? On peut dire oui, on peut dire non. C'est une question technique préalable à trancher ; le calcul des cotes n'y peut rien, pas plus que l'arithmétique ne peut dire combien de francs vaut un mètre de drap d'Elbeuf. Il faut donc convenir d'un certain nombre que nous laisserons sous sa forme littérale pour n'être engagés à rien, indiquant le pour-cent d'objets mauvais qui, si on le savait certainement dépassé, entraînerait le rebut du lot »* (p. 440).

Estienne souligne ainsi que nous n'avons aucun moyen de connaître exactement le taux de défectueux dans le lot. C'est une idée tout à fait centrale dans sa théorie, à tel point que ce n'est pas cette variable qu'il pose comme objet direct de la conjecture, mais l'idée qu'on s'en fait. La nuance se comprend mieux avec la déclaration suivante qu'il adresse à Maurice Dumas, en réponse à l'article que celui-ci lui a fait parvenir :

« Puisque vous avez bien voulu me lire, vous savez que mon idée dominante est que l'épreuve expérimentale ne fait que modifier une conviction a priori » (13). En termes modernes, nous dirions qu'Estienne a une approche bayésienne du problème. Cette façon de poser le problème du contrôle de qualité est totalement originale par rapport aux problématiques qui se développeront ultérieurement, qui toutes relèvent des probabilités dites « directes », dont Dumas va nous fournir un exemple.

Maurice Dumas, quant à lui, tente de suivre chacune des deux approches, la bayésienne et la directe, mais c'est seulement la seconde qui l'amène à des conclusions utilisables dans la pratique. L'approche bayésienne suppose de choisir, comme l'indique Estienne, une distribution de probabilités a priori qui reflète l'idée qu'on se fait de la qualité du lot, distribution de probabilités qui va être modifiée selon les résultats de l'examen de l'échantillon. Or Dumas recule devant la difficulté de choisir une distribution a priori. C'est pourtant cette approche qu'il poursuivra ensuite, tout au long de sa vie, écrivant de nombreux articles sur les lois de probabilité a priori et les thèmes apparentés, notamment les lois « impropres », telles celles proposées par Ernest Lhoste, autre artilleur probabiliste (14).

## LE TAUX DE DÉFECTUEUX : PROPRIÉTÉ DU PROCESS OU PROPRIÉTÉ DU LOT ?

Voyons maintenant comment nos auteurs proposent de « se faire une idée » de la qualité du lot à partir d'un échantillon, selon l'expression de Dumas. Il existe une différence importante entre les deux au sujet de ce que représente la variable « taux de défautueux » : faut-il la considérer comme une propriété du seul lot proposé en recette, ou comme une propriété du processus de fabrication, ce qui donne une portée plus large au jugement qu'elle supporte ?

Estienne considère que le taux de défautueux est une caractéristique du procédé de fabrication du fournisseur, au moment de la fabrication. Si le lot présente ce taux de défautueux, c'est parce que le lot est prélevé au hasard dans le flot des pièces fabriquées : « On assimile plus ou moins consciemment la production des fusées sortant de l'usine, bonnes ou mauvaises, suivant un degré

à peu près constant de perfection de la fabrication pendant la courte durée considérée, à l'écoulement d'un fleuve qui roulerait des fusées bonnes et mauvaises dans une proportion à peu près constante pendant un court espace de temps, et on puise au hasard dans ce fleuve vingt mille fusées pour former un lot. (...) Voilà les raisons qui justifient la conviction qu'il y a une fréquence a priori de bonnes fusées et tout raisonnement a sa base dans cette conviction » (p. 441).

Cette idée de constance des conditions de fabrication, clairement exposée par Estienne, se révélera d'une très grande importance pour la théorie du contrôle de qualité. On la retrouve en effet au fondement de toute la théorie américaine du contrôle statistique de processus, développée par W.A. Shewhart pour AT&T aux Laboratoires Bell. Les techniques statistiques, telles que les cartes de contrôle de Shewhart, servent essentiellement à amener le processus de production dans un état de stationnarité où les caractéristiques des produits suivent des distributions statistiques stables. De son côté, Estienne a une vision très claire de cette propriété de stationnarité : « Le problème se pose ainsi : un établissement a fabriqué, en série, avec des procédés identiques, avec des matières premières de même provenance, avec un personnel permanent, ou à peu près, etc., un nombre plus ou moins considérable d'objets ; ces objets ont été réunis en lots, c'est-à-dire qu'on a groupé un nombre déterminé d'objets fabriqués dans un espace de temps aussi court que possible » (p. 439). Les termes qu'il emploie ressemblent beaucoup à ceux qu'emploiera Shewhart vingt-cinq ans plus tard en théorisant les « constant systems of chance causes » (systèmes stationnaires, en termes modernes).

Dumas ne reprend aucunement ce raisonnement et limite ses conjectures à la qualité du lot présenté en recette (15). À vrai dire, rien dans son raisonnement ne nécessite le recours à l'hypothèse de constance du taux de défautueux dans la fabrication, mais nous verrons que cela limite considérablement la portée pratique de sa méthode.

La démarche de Dumas est d'une rigueur logique sans faille. Elle relève de ce qu'on appelle maintenant l'approche par les probabilités « directes » (alors que lui-même la qualifie du terme « a posteriori » afin de l'opposer aux démarches « a priori » ou bayésiennes). Soit  $a$  la proportion inconnue de fusées mauvaises dans le lot ; soit  $N$  la taille du lot et  $n$  la taille de l'échantillon. On sait calculer, par des méthodes de dénombrement,

(13) Lettre à Maurice DUMAS du 8 avril 1925

(14) LHOSTE E. : « Le calcul des probabilités appliqué à l'artillerie », *Revue d'artillerie*, t. 91, mai-août 1923. Lhoste (1880-1948) fut directeur de l'AFNOR à partir de 1938 ; il y créa en 1939 une commission pour l'étude des essais par prélèvement (autre nom pour le contrôle de qualité par échantillonnage).

(15) Dodge et Romig ont introduit ultérieurement la classification des plans en « type A » et « type B » selon qu'ils se réfèrent à un lot isolé (cas Dumas) ou à un lot issu d'une production régulière poursuivie sur la longue durée (hypothèse d'Estienne). Cf DODGE H.F., ROMIG H.G. : *Sampling Inspection Tables*, New York : Wiley & Sons, 1944

la probabilité d'observer dans l'échantillon 0, 1, 2, ...,  $c$  éléments défectueux. Par exemple, pour un lot de mille pièces et un échantillon de vingt pièces, la probabilité de n'observer aucun défaut dans l'échantillon varie selon le pourcentage de défectueux dans le lot :

Tableau 1

proportion de défauts dans le lot	2 %	4 %	6 %	8 %
nombre de défauts dans le lot	20	40	60	80
probabilité de zéro défaut dans un tirage de vingt pièces	0,68	0,50	0,29	0,19

L'observation « *Zéro défectueux* » est peu probable dans le cas d'un lot à 8 % de défectueux ; dans le cas d'un lot à 2 %, elle se produit dans 68 % des cas en moyenne. Cette démarche de calcul toute simple permet à Dumas d'établir précisément les propriétés des « *conditions de recette* », c'est-à-dire des règles de jugement sur échantillon en vigueur dans l'Administration.

Soit par exemple la règle prescrivant de tirer un échantillon de  $p$  fusées et de rebuter le lot si on observe au moins un mauvais fonctionnement. Moyennant une approximation algébrique légitime, Dumas exprime très simplement par un polynôme la probabilité de rebuter le lot en fonction de  $a$ . On peut ainsi mesurer, en probabilité, le pouvoir qu'a la condition de recette en question de rebuter un lot de qualité insuffisante, défini comme un lot qui comporte plus de  $b$  % de défectueux.

À cette occasion, Dumas formule deux observations qui, vues depuis notre époque, paraissent évidentes mais qui étaient certainement des propositions très peu intuitives en 1925.

La première est que les conditions de recette conduisent à éliminer des lots mauvais, et non à accepter des lots bons. Il le démontre avec un raisonnement simple : en admettant, par hypothèse, que tous les lots présentés soient mauvais, un certain nombre d'entre eux sera néanmoins admis car certains échantillons n'auront pas fait apparaître de tir défectueux. Mais, cela n'empêchera pas tous les lots d'être néanmoins mauvais. Un tel résultat, aussi clairement explicité, n'était sans doute pas de nature à réjouir les opérationnels !

La deuxième est que le pouvoir d'une condition de recette d'éliminer des lots mauvais ne dépend pas de

la taille du lot (à condition que la proportion de défectueux soit faible) et, en particulier, que l'usage consistant à définir la taille de l'échantillon en pourcentage de la taille du lot n'a pas de fondement théorique acceptable. C'est là une découverte sur laquelle Dumas insiste longuement. Il en conclut notamment qu'il est beaucoup plus intéressant, lorsque le contrôle détruit les pièces comme c'est le cas ici, de présenter des lots de très grande taille. Le prélèvement lui paraît d'ailleurs tellement énorme qu'il estime la méthode impraticable dans le cas où l'essai détruit l'élément, à moins que les lots ne soient eux-mêmes très grands. Il calcule par exemple que si l'on souhaite une probabilité 0,95 de rebuter les lots contenant plus de 1 % de défectueux, une solution consiste à essayer deux cent trente-sept fusées en n'acceptant aucun échec. À titre de comparaison, Estienne indique, en 1903, que les essais courants portent sur trente fusées pour un lot de vingt mille (16).

Ces indications de volume des prélèvements nous font mesurer le coût important des méthodes de contrôle statistique, relativement aux pratiques en cours, et donc l'obstacle que cela pouvait représenter pour leur adoption. Tout au long des siècles, on a pris des échantillons pour juger de la qualité des produits fabriqués (17) et voilà que ces pratiques traditionnelles sont remises en cause au nom de raisonnements abstraits. Les méthodes statistiques conduisent à expliciter des notions telles que le risque, proposant même de le mesurer. Que des réactions incroyables se soient produites, nous pouvons aisément le comprendre. Néanmoins, le mouvement était lancé : l'étude comparative des conditions de recette allait conduire à encore plus d'explicitation et de formalisation.

#### LA RECHERCHE DU « MEILLEUR » TEST ET LES PREMIÈRES COURBES D'EFFICACITÉ (DUMAS 1925)

Egon S. Pearson, co-auteur avec Jerzy Neyman d'une théorie des tests d'hypothèse devenue classique, note dans ses souvenirs (18) que c'est surtout à partir de

(16) À certains moments de la guerre de 1914-1918, la consommation d'obus de 75 était supérieure à cent mille unités par jour.

(17) Voir STIGLER S.M. : « Eight Centuries of Sampling Inspection : the Trial of the Pyx », *Journal of the American Statistical Association*, 72 (1977) : 493-500. Repris dans Stigler SM : *Statistics on the Table*, Cambridge (Mass.) : Harvard University Press, 1999.

(18) PEARSON E.S. : « The Neyman-Pearson Story : 1926-1934 », in : Kendall & Pearson (eds) : *Studies in the History of Statistics and Probability*, Griffin, London, 1970

M. Marque de fabrique (Cote p. 100 a priori que le lot est bon.)	$\rho = \frac{1}{100}$ c'est à dire le lot est réputé bon si la proportion des ratés est inférieure à 1 %								
	10	20	30	40	50	60	70	80	90
$C_{30}^0$ . Cote que le lot est bon après l'essai de 30 fusées sans aucun raté.	33	41	48	56	63	70	78	85	93
$C_{30}^1$ . Cote que le lot est bon après l'essai de 30 fusées dont une seule a raté.	4	10	14	20	26	34	45	57	72
$C_{60}^1$ . Cote que le lot est bon après l'essai de 60 fusées une seule a raté.	16	20	25	31	37	45	54	65	79
Nombre de fusées qu'il faudrait essayer sans un seul raté pour avoir le droit de donner la cote 90 p. 100 au fait que le lot est bon.	217	205	192	177	159	137	109	69	0

**Tableau. 2 : gamme de résultats proposés par Estienne (1903)**

1924 que les statisticiens britanniques commencèrent à s'interroger sur la comparaison des méthodes de tests d'hypothèses. Lui-même a commencé à y travailler en 1926, à la recherche de « *principes généraux, parlant à l'intuition, qui pourrait fournir un guide pour choisir parmi les tests* ». En collaboration avec Neyman, il publia en 1928 le premier article (19) d'une longue série, qui définissait notamment les concepts d'erreurs de première et de deuxième espèce (qu'on retrouve en contrôle statistique de qualité sous les appellations de « *risque du fournisseur* » et de « *risque du client* » : rejeter à tort un lot bon, accepter à tort un lot mauvais). Les artilleurs français rencontrent ce problème de classement et de comparaison des conditions de recette. La coïncidence de dates entre l'article de Dumas et les travaux anglais est d'autant plus frappante que Dumas était isolé du milieu des statisticiens.

Estienne, en 1903, se rend bien compte du fait que le principe de l'échantillonnage permet d'expérimenter de différentes manières en variant la taille de l'échantillon et le nombre d'erreurs tolérées. Il ne cherche pas systématiquement à comparer les « méthodes pour conjecturer », il présente la sienne. Mais celle-ci se prête à des variations importantes sur le nombre d'es-

sais, sur le nombre d'échecs tolérés. Il explore les possibilités sans méthode bien explicite. Différents calculs le conduisent aux résultats de la figure 2.

Chaque colonne correspond à un niveau de confiance a priori dans la qualité du lot (faible à gauche, forte à droite, en pour cent). La ligne suivante (2) donne la « *cote que le lot est bon* » après un essai de trente fusées sans raté. Les lignes suivantes (3 et 4) donnent les cotes dans des conditions différentes. La dernière ligne donne la taille de l'échantillon pour obtenir une cote de 90 % que le lot est bon : deux cent dix-sept fusées pour le fournisseur ayant une cote a priori de 10 %, le lot étant accepté sans essai pour celui dont la cote est déjà de 90 %.

Maurice Dumas va nettement plus loin dans la réflexion théorique, en explicitant plusieurs critères de comparaison et en menant la démarche de mathématisation sur des pratiques existantes. Il introduit ainsi des concepts nouveaux. En particulier, il est certainement l'un des premiers au monde à établir et publier ce qu'on a appelé ultérieurement la « *courbe d'efficacité* » (« *efficiency curve* » en anglais), un outil de comparaison des tests qui se révélera d'une grande utilité pour les statisticiens. Il perçoit bien l'usage que l'on peut faire de ces courbes, même s'il n'en mesure pas toute l'importance.

Dumas met en place, comme nous l'avons vu plus haut, un mode de calcul qui permet de déterminer, pour une condition de recette donnée, la probabilité

(19) NEYMAN J. & PEARSON E.S. : « On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference », *Biometrika*, XX (1928), parts I & II.



de refuser un lot en fonction de son taux de défectueux. Ce mode de calcul reste applicable lorsque l'on fait varier la taille de l'échantillon et le nombre d'échecs tolérés à l'essai. Cela permet d'établir les courbes qui correspondent à trois « genres » de conditions de recette, que Dumas appelle A, B et C :

- genre A : échantillon de taille  $p$ , rebut si un seul échec ;
- genre B : échantillon de taille  $p$ , rebut si plus de  $n$  échecs ;
- genre C : échantillon de taille  $p$ , rebut si plus d'un échec ; acceptation si aucun échec ; si un seul échec, on tire un nouvel échantillon de taille  $q$  ; si cet échantillon contient un échec ou plus, le lot est rebuté ; sinon, il est accepté.

Le genre C, dénommé par Dumas « *épreuve et contre-épreuve* », est déjà évoqué par Estienne. Dans les termes de la théorie moderne, on l'appellerait « *échantillonnage double* », les genres A et B relevant de l'échantillonnage simple, comme le montrent les courbes de la page suivante..

Dumas remarque que, si l'on impose la probabilité d'éliminer un lot mauvais, il existe un grand nombre de solutions pouvant appartenir à tous les genres. Il a l'idée de rechercher quelles conditions de recette seraient « *les plus avantageuses* » et essaye d'abord de préciser ce que l'on peut entendre par là.

Il propose la définition suivante : est plus avantageuse une condition de recette qui donne une plus grande probabilité d'accepter un lot bon, tout en donnant la même probabilité de refuser un lot mauvais. Il exprime ainsi une propriété des plans d'échantillonnage qui sera reconnue plus tard comme tout à fait importante, qui est qu'un plan est défini par la fixation d'un « *risque du client* » et d'un « *risque du fournisseur* », ce qui revient à imposer à la courbe d'efficacité de passer par deux points déterminés.

Pour bien saisir le problème, il faut se référer à la théorie moderne. Le client veut être protégé contre l'acceptation de lots mauvais mais le fournisseur veut aussi, de son côté, être protégé contre le rejet de lots qui sont bons. Le test idéal devrait ainsi rejeter avec certitude les lots contenant plus de  $b$  % de défectueux, et accepter avec certitude ceux qui en contiennent moins de  $b$  %. Ce test idéal n'existe pas, ou alors c'est l'inspection du lot en totalité. Le calcul montre que, pour se rapprocher de cet idéal, il faut augmenter la taille de l'échantillon.

Pour Dumas, les courbes constituent le moyen le plus rapide et le plus parlant de comparer les conditions de recette. Il suffit de superposer les graphiques A et B pour constater que le fait d'accepter un nombre de défauts non nul dans l'échantillon permet d'obtenir

des courbes en « S », correspondant à des conditions de recette qui discriminent mieux entre les lots bons et mauvais. Si aucun défaut n'est accepté, les courbes sont toutes des exponentielles dont la concavité est tournée vers le haut (graphique A), et qui, si elles correspondent à un faible risque du client, entraînent un fort risque du fournisseur.

Le deuxième sens de « *solution avantageuse* » qu'évoque Dumas est la minimisation du nombre d'éléments examinés, tout en conservant la protection ( $b$ ), mais il ne développe pas cette voie. Soulignons que c'est elle qu'ont choisie les ingénieurs américains des Laboratoires Bell, Dodge et Romig (20). Ce critère de minimisation du nombre de pièces examinées (et donc du coût de l'inspection) a compté pour beaucoup dans le succès qu'a connu leur méthode dans les milieux industriels. Le contexte industriel fait ainsi émerger, dans le cas américain, un critère de comparaison des tests d'une nature différente de ceux qu'envisagent les statisticiens travaillant en milieu à dominante scientifique (tels E.S. Pearson) : plutôt que la minimisation d'un risque d'erreur, c'est la minimisation d'une quantité économique qui est formalisée. Dumas envisage les deux possibilités mais, n'ayant pas affaire à une demande effective qui pourrait l'orienter vers le critère le plus intéressant pour les milieux industriels, il reste au plus près du critère administratif traditionnel.

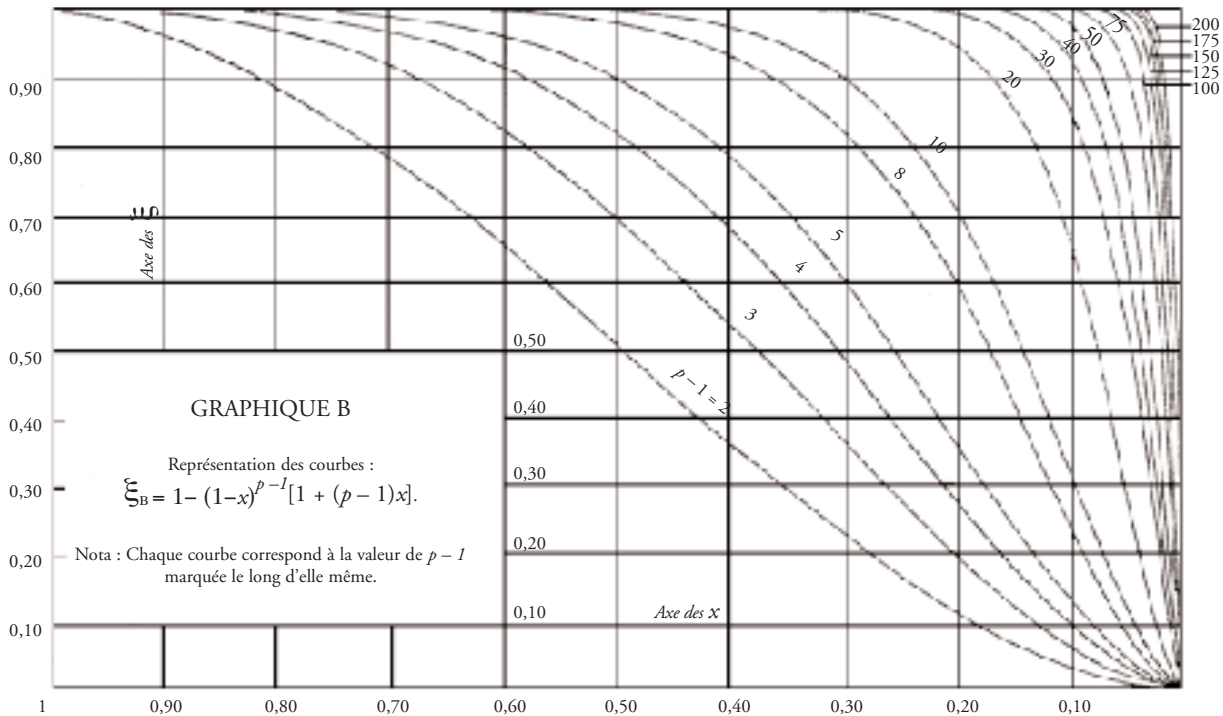
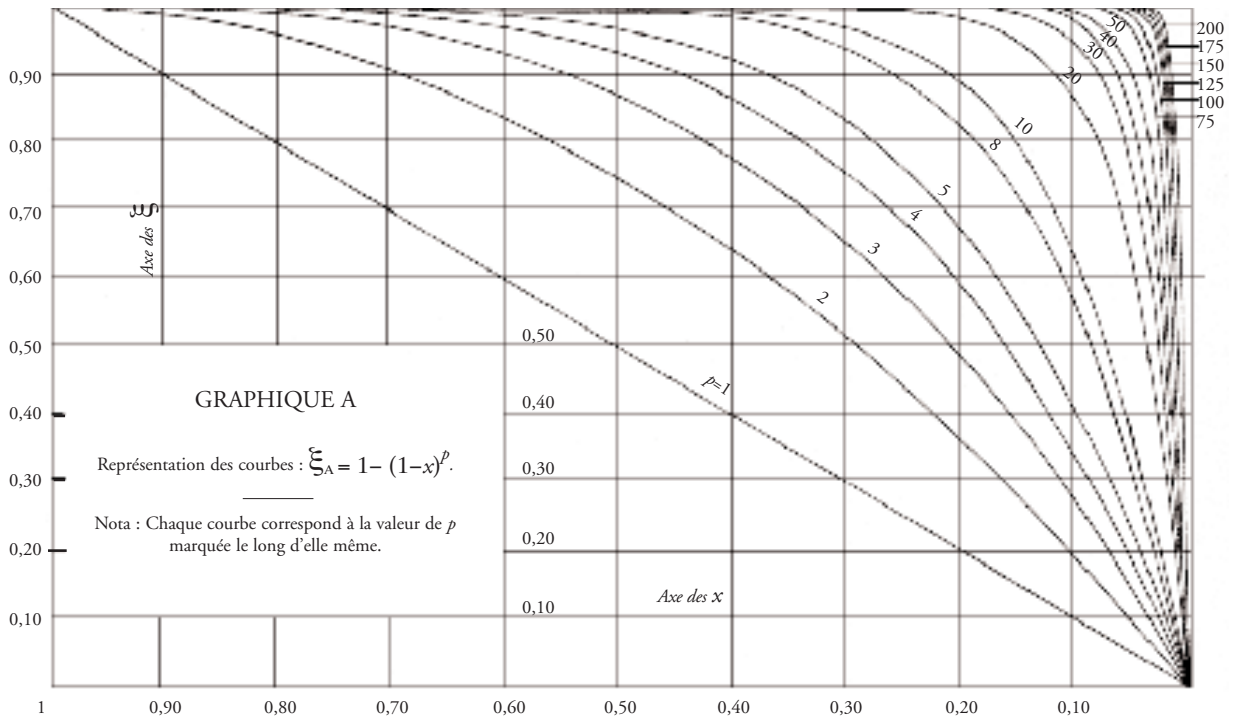
---

## L'INSERTION DU CONTRÔLE STATISTIQUE DANS LES PRATIQUES

Les apports théoriques et conceptuels d'Estienne et Dumas, que nous avons étudiés jusqu'à présent, sont d'un niveau comparable à ceux des ingénieurs américains des Laboratoires Bell. Il semble ne leur avoir manqué qu'un passage à la pratique, qui aurait conduit à sélectionner et à développer les éléments les plus pertinents pour les acteurs du monde industriel. Cependant, les liens entre théories et applications restent généralement énigmatiques : telle condition qui semble un facteur de succès ici se révèle souvent sans effet ailleurs... Il est donc toujours intéressant de saisir les occasions d'étudier en détail les liens qui peuvent être créés entre les théories et le monde pratique.

---

(20) DODGE H.F. & ROMIG H.G. : « A Method of Sampling Inspection », *Bell System Technical Journal*, vol. 8 (1929), pp. 613-631. Dodge et Romig postulent que le lot peut être inspecté en totalité si l'essai n'est pas satisfaisant, les pièces mauvaises étant remplacées par des bonnes ; leur méthode minimise le nombre total de pièces inspectées, incluant l'échantillon et le reste du lot en cas de rejet (c'est une grandeur aléatoire). Cette méthode n'est évidemment pas applicable aux essais destructifs.



### Commentaire des courbes.

Axe des  $x$  : proportion de défectueux dans le lot (1 étant à gauche, 0 à droite)

Axe des  $y$  : probabilité de refuser le lot, en bas : 0 = acceptation, en haut : 1 = refus

Chacune des courbes A (graphique supérieur) correspond à une taille d'échantillon (valeurs 1, 2, 3, 4, 5, 8, 10, 20, etc.) et indique la probabilité de refuser le lot en fonction de la proportion de défectueux, en n'acceptant aucun échec.

De la même manière, chaque courbe B (graphique inférieur), indique, pour la taille d'échantillon indiquée le long de la courbe, la probabilité de refuser le lot en fonction de la proportion de défectueux, en tolérant un seul échec.

On constate que les courbes B sont en forme de "S", ce qui est dû au fait que l'on accepte au moins un défectueux. Elles donnent une plus grande probabilité d'accepter un lot bon que les courbes A. ]

Dans notre introduction, nous avons fait référence à deux types d'insertion : la facilité d'usage des méthodes en milieu industriel, dans des conditions d'exploitation courante ; la connexion avec d'autres réseaux de métrologie de façon à créer une valeur pour le mode de mesure nouvellement proposé. Nous ajouterons un troisième point : la démonstration pédagogique, autrement dit, la mise en scène de la validation des raisonnements théoriques.

Le premier point peut être traité rapidement : aucune mise en pratique réelle n'a en effet été envisagée (21), du moins avant 1936. Les courbes de Dumas, notamment, sont restées dans les cartons de l'auteur et les pages du *Mémorial* (22). Les ingénieurs américains de AT&T disent avoir utilisé très tôt des courbes similaires (23), pour le fonctionnement interne des équipes de contrôle, mais le fait est qu'ils ne les ont pas publiées. Cela tendrait à prouver qu'ils n'en avaient pas non plus perçu toute la portée théorique, laquelle n'est apparue que bien plus récemment.

#### Des expériences sous forme de simulations

Examinons ensuite le troisième point, qui est souvent de grande importance lorsque le promoteur d'une nouvelle théorie souhaite en démontrer la validité. Je ne m'attacherai pas aux thèmes classiques de l'épistémologie scientifique mais plutôt aux dispositifs qui permettent aux acteurs de mettre la nouvelle théorie à l'épreuve, de la faire cadrer avec leur intuition, de se confronter sensoriellement avec elle. Ceci est particulièrement important dans le champ du calcul des probabilités, où les conclusions théoriques sont souvent très différentes de l'intuition spontanée. Ainsi, on a souvent jugé, intuitivement, qu'un échantillon devait être de taille proportionnelle à celle du lot, conclusion que la théorie dément (lorsque la probabilité de l'événement considéré est faible). Rappelons aussi le rôle important des « *quinconces* » – ces planches portant des clous plantés en quinconce entre lesquels on laissait tomber des billes de métal, ce qui produisait des

effets aléatoires – pour montrer la loi des grands nombres en action (24). L'artillerie possède d'ailleurs une certaine pratique des simulations, notamment celle appelée « *tir simulé* » : pour entraîner les artilleurs au réglage du tir, on tirait dans une urne des boules réparties suivant la loi de Gauss, qui donnaient les positions des obus tirés, et le candidat artilleur devait allonger ou diminuer la hausse pour tenter d'atteindre (toujours avec les boules tirées de l'urne) l'objectif en question, qui était la moyenne de la loi de Gauss (25). Dans les années 1950, lorsque les statistiques furent enseignées aux cadres de l'industrie (26), les dispositifs expérimentaux revêtirent une importance considérable dans la pédagogie.

Or Dumas ne s'est pas contenté, dans son article, de donner des résultats théoriques : il a cherché à valider ses raisonnements au moyen de ce que nous appellerions actuellement une simulation (lui-même parle de « *vérification expérimentale* », terme impropre qui lui vaudra d'ailleurs les remontrances d'un lecteur). Le travail qu'il a effectué est impressionnant. Il construit une sorte d'urne mécanisée avec laquelle il effectue six mille tirages dont la liste figure à la fin de son article. Les jetons de l'urne sont au nombre de cinq cents, marqués chacun d'une lettre entre A et E et d'un chiffre de 1 à 100. Il se sert ensuite de ces tirages pour établir des statistiques sur les résultats de différentes conditions de recette et vérifier qu'ils sont raisonnablement en accord avec ses calculs. Il conclut par l'affirmative.

Le dispositif mécanique par lui-même mérite une description : une caisse contenant des jetons tourne en permanence autour d'un axe horizontal, réalisant un brassage constant ; l'axe est creux et permet d'extraire des jetons un par un. Le numéro du jeton prélevé est noté, il est remis dans l'urne. L'expérience est répétée six mille fois. Pour ce qui concerne la réalisation matérielle, il est probable que Maurice Dumas, qui était alors en poste à l'arsenal de Toulon, a recouru aux services d'ouvriers de l'établissement.

En publiant ces séries de tirages, Dumas estime fournir un moyen qui servirait « *à tous ceux qui chercheraient à vérifier expérimentalement des lois de probabilité* ». Il rappelle en effet que, si le calcul des probabilités « *est exposé constamment avec l'aide d'urnes et de*

(21) Selon le témoignage de M. DUMAS dans un entretien en 1990.

(22) En 1925, DUMAS soumet son article du *Mémorial* à la revue *La Technique moderne*, dont le public est plus large et plus industriel mais il est refusé.

(23) Lettre de H.G. ROMIG à DUMAS (21 oct. 1968), suite à une lettre de ce dernier s'enquérant de sa priorité concernant la publication de ces courbes d'efficacité : « *Your 1925 Probability of Acceptance Curves evidently are among the earliest* ». Tout en mentionnant plusieurs travaux internes, Romig ne peut citer une véritable publication. Pourtant, les ingénieurs de AT&T disposaient d'un support privilégié, le *Bell System Technical Journal*, dans lequel Shewhart fit paraître plusieurs articles dès 1924.

(24) STIGLER S.M. : *The History of Statistics*, Cambridge (Mass.) : Harvard University Press, 1986, p. 277. D'autres exemples de simulations au XIX<sup>e</sup> siècle sont présentés dans : STIGLER S.M. : *Statistics on the Table*, Cambridge (Mass.) : Harvard University Press, 1999

(25) Communication de Bernard BRU.

(26) Par exemple au CERESTA, organisme de formation continue fondé par Georges DARMOIS, qui était parallèlement directeur de l'Institut de statistique de l'Université de Paris (ISUP).

*boules de différentes couleurs, on ne cite guère comme expérience faite que les pièces de monnaie lancées en l'air par Descartes et par Buffon* ». Les « séries de Dumas » semblent toutefois être restées ignorées (alors que bien d'autres séries expérimentales telles que les dés de Weldon ont marqué l'histoire des probabilités).

Concernant la « *vérification expérimentale* », il considère que les jetons portant la lettre C représentent les défectueux, qui sont donc dans une proportion de 20 %. Il compte ensuite, sur les séries de nombres obtenues expérimentalement, le nombre de séries qui auraient entraîné l'acceptation du lot si cette série avait représenté les résultats d'un échantillon. Par exemple, une suite de quatre tirages ne comportant pas la lettre C est considérée comme l'essai d'un échantillon de quatre fusées sans échec (condition du genre A : lot accepté s'il ne se produit aucun échec). Le nombre de telles séries est comparé au nombre théorique, et cela pour chaque genre de condition de recette. Dumas, ne maîtrisant pas les outils statistiques qui lui auraient permis de déterminer si les écarts sont significatifs ou non, ne peut guère aller plus loin. Il se borne donc à constater un accord apparent entre la théorie et l'expérience.

Dumas n'est pas le seul à expérimenter avec les urnes. W.A. Shewhart, le physicien américain qui a développé la théorie du contrôle statistique de processus, a lui aussi beaucoup pratiqué ce mode de simulation et l'a même utilisé pour la validation. C'est essentiellement par ce moyen qu'il a éprouvé la robustesse de la méthode des cartes de contrôle dans les cas où les distributions statistiques ne sont pas gaussiennes (distributions rectangulaire et triangulaire). Lui aussi a publié, dans son manuel classique (27), les résultats de quatre mille tirages expérimentaux. Cette convergence des pratiques expérimentales est un phénomène de même nature que les doublons en histoire des sciences, mais le fait qu'elle se produise au même moment est particulièrement intrigant. En outre, Dumas et Shewhart ne sont pas les seuls : en Angleterre, LHC Tippett publie une longue série de tirages aléatoires qui, selon E.S. Pearson (28), ouvre un nouveau champ d'investigation : l'étude empirique des distributions d'échantillonnage. Dumas, bien qu'étant isolé des milieux savants de la statistique mathématique, participe sans le savoir à ce renouveau ; c'est ainsi que l'on peut interpréter sa naïve prise de conscience qu'aucune expérimentation réellement nouvelle depuis Descartes et Buffon n'aurait marqué avant lui le calcul des probabilités.

(27) SHEWHART W.A. : *op. cit.*, Annexe II, Table A, B, C, pp. 442-471

(28) PEARSON, E.S. : *op. cit.*

L'expérimentation a cependant un coût. En milieu industriel, l'expérimentation en « vraie grandeur », sur les équipements de production eux-mêmes, comporte des coûts indirects liés aux perturbations qu'elle introduit dans les programmes de production. C'est d'ailleurs pour des raisons de ce type que Gosset-Student développa la théorie des petits échantillons, afin d'exploiter des données obtenues en un temps limité sur des cuves de brassage (il travaillait comme ingénieur aux brasseries Guinness). Le contrôle statistique apporte avec lui une sensibilité au coût de l'information, au coût de son acquisition. Les urnes de Shewhart et Dumas ont l'avantage de permettre une simulation à peu de frais de dispositifs industriels dont la mobilisation est très coûteuse et difficilement justifiable pour des fins de recherche. La rapidité des manipulations expérimentales est aussi un élément important. À cet égard, W.E. Deming notait en 1939 que les applications industrielles des statistiques pouvaient être validées bien plus rapidement que les applications agronomiques, ces dernières étant conditionnées par des rythmes saisonniers (29).

---

Que valent pour l'action les mesures de qualité proposées ?

Voyons maintenant la question de l'insertion dans les réseaux métrologiques existants : comment l'information acquise lors de la prise en recette acquiert une valeur pratique, par raccordement à d'autres évaluations ou mesures.

Estienne écrit sans détour que sa méthode du « *calcul des cotes* » est inapplicable dans le cas des marchés d'armement. En effet, elle conduit à des conditions de recette différentes, selon les fournisseurs, en fonction de la confiance qu'a l'acheteur dans la qualité de leurs fabrications respectives, confiance qui est mesurée par la « *cote* » du fabricant : « *Il est impossible de songer à imposer par un cahier des charges général des conditions différentes aux concurrents ; on peut le regretter mais il faut s'y résigner* ». La méthode pour mesurer les performances des fabricants, que propose Estienne, est donc en complète contradiction avec le principe d'égalité qui prévaut dans le soumissionnement des marchés publics. Il propose également de faire évoluer la « *marque de fabrique* » d'un fabricant (son indice de qualité) en fonction des résultats observés, et donc d'en garder une mémoire, ce qui est tout aussi

---

29) Préface à : SHEWHART W.A. : *Statistical Method from the Viewpoint of Quality Control*, The Graduate School, U.S. Department of Agriculture, Washington, 1939. Traduction française : *Les fondements de la maîtrise de la qualité*, Economica, Paris, 1989.

contraire aux principes des marchés publics. Remarquons que de tels systèmes sont maintenant devenus courants dans l'industrie privée, dans le cadre de l'assurance qualité (« *démérites* »).

Mais Estienne est un homme d'action. Concédant que sa théorie est inapplicable dans sa pureté originelle, il propose une troisième solution : « *l'appel à la concurrence réduite, dont l'effet est précisément de permettre de prime abord, sans discussion, l'élimination des fabricants auxquels on attribuait a priori une cote par trop faible* » (p. 447).

À l'appui de cette troisième voie qui va dans le même sens que ses calculs, il invoque également le bon sens fondé sur l'expérience : « *Nous avons été chargés pendant plusieurs années de la réception des tubes de canon de provenances diverses, suivant les chances d'adjudications où la concurrence au rabais (30) avait introduit des fournisseurs de marques notablement différentes ; nous n'étions pas en possession, à cette époque, des résultats numériques précis que nous publions aujourd'hui, mais parfois le bon sens nous a donné la conviction absolue qu'il eût été préférable de recevoir, sans épreuve, un tube sorti d'une bonne usine, plutôt que de recevoir, comme nous y obligeait le cahier des charges, un tube ayant strictement satisfait aux épreuves réglementaires, mais venu d'une usine médiocre (bon et médiocre étant définis par rapport à ce que l'on sait des fournitures précédentes)* » (p. 448)

Dans ces conditions, l'essai sur échantillon devient surtout un moyen de « *vérifier si la bonne opinion qu'on avait a priori sur une fabrication n'est pas accidentellement injustifiée pour le lot considéré* ».

La théorie d'Estienne est tout à fait remarquable car elle réussit, d'une façon très rigoureuse, à allier l'usage du calcul à une notion de confiance fondée sur des relations sociales. S'il démystifie l'idée que la prise en recette garantit la bonne qualité, c'est pour se rabattre sur un autre critère de qualité, présupposant des relations à long terme avec les fournisseurs et une capitalisation de l'expérience acquise. Il affiche un grand scepticisme vis-à-vis des jugements basés uniquement sur des mesures. L'article de Detchessahar & Lemarchand [2001] donne la même impression : Estienne s'oppose aux démarches purement techniques et fait prévaloir les rapports avec des acteurs représentatifs ou pertinents. Par exemple, il estime que la qualité d'un char d'assaut, du point de vue de sa conception, doit être jugée par de simples soldats possédant l'expérience du front, porteurs de critères d'usage au combat, et non seulement par les ingénieurs de l'Administration, porteurs de critères tech-

(30) En termes modernes : au moins disant

niques. Il apparaît ainsi comme un opposant à la « convention de qualité » égalitariste, celle qui met tous les concurrents sur un même plan et leur applique les mêmes épreuves numériques standardisées. Il ne semble cependant pas qu'il ait eu l'occasion de mettre en œuvre ses idées en matière de marchés publics (31).

Dumas, quant à lui, se range totalement du côté du calcul. Il établit la vérité toute nue en montrant que les conditions de recette en vigueur n'offrent qu'une faible garantie contre les lots de qualité insuffisante (selon sa définition, qui peut être contestée). Il ne donne guère d'indications sur la manière d'insérer la méthode d'échantillonnage dans les pratiques. En fait, nous l'avons vu, la taille des échantillons à essayer (et donc à détruire) le rend très sceptique. Ceci l'amène à prendre en considération la méthode d'Estienne, qui conduit en effet à des échantillons de taille plus réduite, mais qu'il rejette en définitive comme impossible pour les raisons déjà évoquées d'inégalité de traitement des fournisseurs. Il conclut que, si l'on ne veut pas sacrifier le nombre de pièces nécessaires pour obtenir une « *certitude morale* », on aura au moins l'explication du fait que les lots reçus seront mauvais. Décevante conclusion. Son article lui vaut quelques éloges et témoignages de sympathie, mais il ne lui est pas proposé de passer à la mise en œuvre. Sa critique des conditions de recette est-elle trop radicale pour être prise en compte sérieusement ?

Une autre hypothèse est le caractère abstrait de la représentation de la situation, présentée comme une unité en soi et coupée du reste du monde. La situation est définie localement, elle est étroitement circonscrite à la remise d'une livraison par un fabricant à un client, sans tenir compte de la durée de la relation ou de ses à-côtés tels que les connaissances sur les besoins réciproques. Cette coupure est probablement une condition nécessaire à la mise en forme probabiliste de la situation, à sa modélisation (pour employer un terme moderne), à sa représentation abstraite et au déroulement du raisonnement mathématique. Le travail de Dumas est d'ailleurs réduit au statut d'exercice intellectuel par un correspondant à qui il a adressé son article, professeur à l'École des Ponts et Chaussées qui « *remercie son jeune camarade (...) de son très intéressant et très utile article du Mémorial et le félicite d'avoir*

(31) Il en parle encore au conditionnel en 1925, dans une lettre à DUMAS, qui lui a envoyé son article : « *Mon exemple des épreuves de recette de fusées avait pour objet de mettre en évidence les avantages de l'appel à la concurrence réduite opposée à la concurrence au rabais pour les fabrications de guerre. Je n'ai pas mon texte sous les yeux mais j'ai dû dire cela. Je crois que les industriels qui n'ont encore fait aucune fourniture à la Guerre admettraient fort bien que leur entrée dans les appels à une adjudication fût payée par des épreuves plus dures que celles des industriels déjà éprouvés.* » (lettre à M. DUMAS du 8 avril 1925).

*traité avec succès un sujet très difficile*». Il aurait donc manqué un travail inverse de l'abstraction, que l'on pourrait appeler de mise en situation (plutôt qu'« application », terme qui passe sous silence toutes ces difficultés), consistant à « convoquer » des acteurs et actants réels afin qu'ils se saisissent de la représentation proposée et fassent l'expérience de l'insérer dans leurs pratiques.

Mais pourquoi les acteurs se tourneraient-ils vers de nouvelles représentations ? Pourquoi se produirait-il un changement dans des conventions de mesure qui, en usage depuis longtemps, ne semblent pas soulever de problème bien sérieux ? De ce point de vue, l'exemple américain de AT&T est instructif en nous montrant le rôle joué par la direction de l'entreprise (32). La mauvaise qualité des fabrications soulevait depuis longtemps des problèmes entre clients et fournisseurs, notamment pour l'imputation des frais d'inspection des produits (qui étaient considérables). Dès 1906, AT&T a mis en place une politique de contrôle de qualité, sans cependant recourir aux concepts probabilistes. Les problèmes ont persisté et, en 1923, le département d'ingénierie a été chargé de redéfinir les méthodes de travail en couvrant l'ensemble des activités de l'entreprise, depuis les bureaux d'études de conception jusqu'aux chantiers d'installation chez les clients. Du personnel de haut niveau a été recruté et mis à plein temps sur le sujet (Shewhart, PhD de physique, et Dodge, ingénieur MIT). Il y avait donc une volonté politique des instances dirigeantes de l'entreprise, ce qui a conduit à l'implication de nombreux acteurs occupant des positions très différentes dans l'organisation. Le rôle de la direction, ici, est de désigner le problème comme important (parmi l'infinité de problèmes que rencontre une entreprise), d'affecter des moyens, de soutenir le projet en le légitimant.

Un deuxième volet des problèmes d'insertion d'une nouveauté technique est la cohérence et la continuité avec ce qui existe déjà ou se met en place en parallèle. Chez AT&T, aux Laboratoires Bell, c'est un ensemble cohérent de méthodes qui a été construit. Le contrôle de réception de Dodge et Romig était complété par le contrôle de processus de Shewhart et s'appuyait sur des résultats obtenus à partir de ce dernier. C'est un point extrêmement important mais qui passe souvent inaperçu : Dodge et Romig font l'hypothèse que l'on connaît le taux moyen de défectueux dans les fournitures soumises au contrôle et ils imposent ce taux au client, dans

le sens où celui-ci n'est pas supposé demander une proportion de défectueux dans les lots qui soit inférieure au taux moyen du process (33) ; en clair, le client doit s'accommoder de la qualité produite par le fournisseur. Ce principe se comprend mieux si on considère l'ensemble de la démarche : le contrôle de processus de Shewhart conduit à une stabilisation du processus de fabrication et on peut alors considérer le taux de défectueux comme une grandeur physique significative et le mesurer ; s'il est jugé trop élevé, il faut engager des actions systématiques pour l'améliorer, mais cela ne relève pas du contrôle de réception. Nous avons vu qu'Estienne employait la métaphore du fleuve, qui est une représentation similaire, mais sans l'idée qu'il fût possible d'agir sur le processus pour amener le taux de défectueux à rester stationnaire. Du côté américain, la mise en pratique de méthodes prototypes, en situation de chantier réel, la comparaison avec les méthodes anciennes, ont conduit à définir des grandeurs acceptables pour les différentes caractéristiques des plans d'échantillonnage : risque du client fixé à 10 % (c'est la « certitude morale » d'éliminer un lot mauvais que Dumas fixait arbitrairement à 5 %, valeur qui augmente beaucoup la taille de l'échantillon), taux de défectueux tolérés nettement supérieur au taux moyen de défectueux du process.

Mais le facteur probablement le plus lourd de conséquences dissuasives pour Dumas est le choix malheureux du contrôle destructif comme support du raisonnement : c'est un problème trop difficile, qui n'admet pas de solution miraculeuse. Dodge et Romig, au contraire, se placent dans une configuration où l'on peut pratiquer l'inspection en totalité des lots rejetés en remplaçant les éléments défectueux par des bons. Il s'agit de composants pour centraux téléphoniques ; le contrôle de qualité contribue à la constitution d'un équipement durable et il n'est pas le seul mode de contrôle mis en œuvre. Le contrôle sur échantillon aléatoire se présente alors, par rapport à l'inspection en totalité, comme un substitut plus économique et plus rapide. Le fait qu'il soit toujours possible de revenir à l'inspection à 100 % a contribué à faire accepter la méthode par les acteurs méfiants. En outre, on a pu comparer les performances de ces deux méthodes, et le contrôle sur échantillon a été jugé satisfaisant, compte tenu de la situation.

La comparaison France-États-Unis nous permet de conclure au caractère crucial des situations où les

(32) Bayart D. : « Savoir organisationnel... », *op. cit.* Voir aussi le chapitre 9, « Quality Assurance », dans : Fagen M.D. (éditeur) : *A History of Engineering and Science in the Bell System, The Early Years (1875-1925)*, vol. 1, Bell Telephone Laboratories, 1975.

(33) Et même plus : le taux de défectueux dans l'échantillon, tel qu'il est prévu dans les tables, ne peut prendre des valeurs inférieures à 1,66 fois le taux moyen du process. Les moyens de calcul publiés à l'époque sont construits ainsi : on ne pouvait y échapper à moins de refaire tous les calculs, ce qui aurait demandé un investissement considérable.

méthodes sont conçues : alors que les méthodes de Dumas et de Dodge-Romig semblent presque identiques sur le plan formel, elles ont pris des significations radicalement différentes en fonction de l'objet auquel elles étaient appliquées et du cadrage qui leur était assigné. En France, la signification est restée comme bloquée sur l'exemple des munitions et l'idée de l'essai destructif. Comme si l'induction statistique était réservée à des grandeurs que l'on ne saurait pas mesurer directement. On peut voir une confirmation de cette orientation dans l'autre exemple favori des théoriciens français du calcul des probabilités (34), celui de l'estimation de la durée de vie des ampoules : là aussi, le contrôle est destructif et on estime une variable qui n'est pas mesurable autrement que par la destruction des objets.

Les ingénieurs des Laboratoires Bell adoptent une approche qui n'est pas centrée d'emblée sur l'estimation de grandeurs telles que le taux de défectueux dans le lot. Shewhart la présente plutôt comme un dispositif organisationnel qui converge, par itérations successives, vers une production conforme à des spécifications (35). Le but visé par le contrôle n'est pas en premier lieu de produire une mesure de grandeur mais d'agir sur un système, en le régulant à travers son organisation. Cette approche pragmatique est bien caractérisée par la remarque d'un expert du domaine : « *Dans le contrôle de réception lot par lot, la question pratique immédiate est ce qu'on fait avec chacun des lots. On doit prendre une décision, soit l'accepter, soit le rejeter, peut-être l'inspecter en totalité. Des critères précis sont établis pour régler ces questions. Habituellement, l'estimation de la qualité du lot à partir de la composition de l'échantillon est vue comme une question académique* » (36).

La régulation d'un système suppose l'établissement d'un réseau métrologique dont le périmètre couvre au moins ce système. Au contraire, l'estimation d'une grandeur est une opération qui, en apparence du moins, a un caractère local : le résultat est une mesure avec une indication de l'approximation, c'est un nombre qui est généralement vu comme autosuffisant. Le système métrologique qui rend en réalité possible l'établissement de cette mesure est relégué au second plan, voire même complètement occulté au profit de la seule opération statistique d'estimation. Plus on essaie d'éléments, plus l'estimation de la

variable inconnue est précise, voilà ce qui obnubile le regard, tout au moins dans ces premiers travaux. Or une mesure isolée n'intéresse personne, dans le monde industriel ; une mesure ne prend de signification que mise en rapport avec d'autres grandeurs, en particulier économiques. Les pionniers français du contrôle statistique, tout en ayant tenté de suivre cette piste, ne pouvaient à eux seuls effectuer un travail qui, dans les pays voisins, a mobilisé des équipes entières placées en situation dans des contextes industriels authentiques, en vraie grandeur, plutôt que simulés.

---

## DES AMATEURS À LA DÉCOUVERTE D'UNE NOUVELLE CULTURE DU CONTRÔLE

Même si Estienne et Dumas n'ont pas rencontré un grand succès à leur époque, les méthodes de contrôle statistique ont pris une importance considérable par la suite : elles se sont largement répandues dans l'industrie européenne après la deuxième Guerre mondiale, sous le couvert du Plan Marshall, des Missions de productivité, de la modernité industrielle. En 1946, par exemple, un cours est rédigé par Jean Mothes, administrateur de l'INSEE, à l'intention de l'École nationale de la statistique et Dumas est invité par le professeur Fréchet à faire une présentation au séminaire de calcul des probabilités de l'Institut Henri Poincaré. C'est également, au Japon, le début du mouvement qui aboutira à la fameuse « *qualité japonaise* » (37). Le contrôle statistique est un élément fondateur de la « *culture du contrôle* » (38) dont le développement s'est poursuivi ensuite avec la cybernétique, la recherche opérationnelle, la gestion des systèmes. En ce sens, son apparition serait un phénomène inéluctable, point de passage obligé pour la manipulation, la connaissance, la gestion des longues séries d'objets manufacturés que la société industrielle produit de plus en plus nombreux.

On s'attendrait à ce que les innovateurs, Estienne et Dumas, soient « *connectés* » aux milieux industriels, à ce qu'ils répondent à une demande au moins tacite, reflétant un besoin encore mal cerné mais néanmoins présent. Or tel n'est pas le cas, et il faut bien constater que les aptitudes de nos artilleurs au raisonnement probabiliste ne se sont pas « *transférées* » à l'industrie, malgré les bénéfices qui auraient certainement pu en

---

(34) Voir par exemple : MARCH L. : *Les Principes de la méthode statistique, avec quelques applications aux sciences naturelles et à la science des affaires*, Alcan, Paris, 1930.

(35) SHEWHART, 1939, op. cit.

(36) GRANT E.L. & LEAVENWORTH R.S. : *Statistical Quality Control*, International Student Edition, McGraw-Hill Kogakusha, Tokyo, 1972, p. 385

---

(37) voir GOGUE J.M. : « Chronique de la qualité au Japon », *Annales des Mines, série Gérer et Comprendre*, n° 21, décembre 1990, pp. 59-65

(38) LEVIN M.R. (ed) : *Cultures of Control*, Harwood Academic Publishers, Amsterdam, 2000.

TABLEAU COMPARATIF			
Comparaison des méthodes Estienne, Dumas, Dodge & Romig			
Caractère	Estienne	Dumas	Dodge & Romig
Date de publication	1903	1925	1929
Définition de la qualité de l'élément	bon/mauvais	bon/mauvais	bon/mauvais
Nature de l'essai	destructif	destructif	non destructif
Possibilité de l'inspection à 100 %	non	non	oui
Connaissance <i>a priori</i>	cote du	néant fournisseur (confiance)	taux moyen de défectueux du process
Résultat du contrôle sur échantillon	rejet/acceptation du lot, nouvelle cote du fournisseur	rejet/acceptation du lot	si rejet, tri du lot et remplacement par de bons éléments
Critère optimisé, critère de choix du test	non explicité, pas d'optimisation : démarche exploratoire	1) tout en fixant proba de refuser un lot mauvais, maximiser proba d'accepter un lot bon 2) minimiser le nombre de pièces examinées (suggéré)	minimiser le coût du contrôle (nombre d'éléments examinés)
Degré de pénétration dans la pratique	uniquement théorique(calculs et raisonnements)	théorie et simulation expérimentaleréels	théorie, méthode développée en situation (chantier) avec acteurs
Proportion de défectueux définissant un lot bon	1% (dans l'exemple numérique du tableau)	1% (exemple du calcul d'une taille d'échantillon)	supérieure à 1,66 fois le taux moyen de déf. du process
Risque du client (probabilité tolérée d'accepter un lot mauvais)	non précisé	5%	10%
Outils cognitifs facilitant le raisonnement	formules algébriques	formules algébriques, graphiques (courbes d'efficacité)	formules algébriques, abaques donnant les paramètres opérationnels

**Tableau 3 : Principales caractéristiques des méthodes de contrôle de réception**

être tirés. Cet insuccès des méthodes probabilistes dans le domaine des pratiques industrielles tient à un ensemble de facteurs trop complexe pour que nous puissions en proposer ici une analyse globale. L'histoire de la réception scientifique et sociale de ces méthodes est elle-même passablement tourmentée. Nous soulignerons seulement les éléments qui ont

trait à la formation des ingénieurs français de l'époque. Bien peu d'entre eux avaient une connaissance, même élémentaire, du calcul des probabilités. Pour ce qui concerne les Polytechniciens (ce que sont Estienne, Dumas, Lhoste), le calcul des probabilités est enseigné à Polytechnique de façon très irrégulière, et pratiquement pas entre 1894 et 1919 (39). De plus,



il s'agit le plus souvent de la théorie des erreurs et de la méthode des moindres carrés (enseignées en astronomie), non de la « doctrine des chances ». Estienne et Dumas font exception. Estienne, pour sa part, a certainement bénéficié de l'excellent cours de probabilités dispensé par le colonel Henry à l'École d'artillerie de Fontainebleau. Et la promotion 1919S (celle de Dumas) est la première à recevoir un enseignement de probabilités plus moderne, par le grand mathématicien Paul Lévy. Cela n'expliquerait-il pas un peu la hardiesse que montre Dumas, « inventant » le problème qu'il traite et l'abordant selon plusieurs voies théoriques différentes ? Dumas ne semble pas se confronter à un problème proposé par d'autres, pour lequel il manquerait une solution satisfaisante, mais plutôt prendre prétexte d'une question assez abstraite (le titre de son article, « *Sur une interprétation des conditions de recette* », ne semble pas destiné aux praticiens) pour développer des raisonnements et calculs probabilistes. Au fond, Estienne et Dumas nous apparaissent comme des amateurs – au sens noble du terme – de formalisation, qui se livrent par plaisir au calcul des probabilités : Estienne intitule « *Loisirs d'artilleur* » un livre (40) rassemblant plusieurs essais (dont l'*Essai sur l'art de conjecturer*) relevant de ce qu'on appellerait actuellement « modélisation ». Dumas a également consacré aux probabilités et statistiques beaucoup de ses loisirs tout au long de sa vie. Ces artilleurs paraissent tout autant sensibles à l'esthétique des raisonnements qu'à leurs applications pratiques. Les revues professionnelles (*Revue d'artillerie*, *Mémorial de l'artillerie française*) ont publié depuis le début du XIX<sup>e</sup> siècle bon nombre d'articles sur les probabilités, non seulement dans leurs applications au tir mais aussi sous des angles plus théoriques, reflétant les débats intenses entre « *artilleurs-savants* » (41), théoriciens ou expérimentalistes, dans le domaine de la balistique. En dépit de ces difficultés à s'insérer dans les pratiques industrielles, l'approche statistique développe les possibilités de contrôle, et surtout elle en change les modalités. Ainsi, il est difficile de dire en général qu'elle permet « plus » de contrôle. Certes, elle permet de contrôler des objets nouveaux, tels les fabrications

(39) cf. CRÉPEL P. : « Calcul des probabilités : de l'arithmétique sociale à l'art militaire », p. 197-215, in : BELHOSTE B., DAHAN DALMEDICO A., PICON A. (dir.) : *La formation polytechnicienne : 1794-1994*, Paris : Dunod, 1994. (42) WALD est l'auteur de la théorie du « contrôle séquentiel », qui est la modalité la moins coûteuse en inspection, d'un point de vue théorique au moins (car sur le terrain industriel, elle a souvent été considérée comme trop complexe à mettre en œuvre). WALD A. : *Sequential Analysis*, New York : Wiley, 1947

(40) ESTIENNE J.E. : *Loisirs d'artilleur*, Berger-Levrault, Paris, 1906

(41) BRU B. : « L'École d'application de l'artillerie et du génie de Metz », *Actes de la journée d'étude du 2 nov. 1995*, Paris : Musée des Plan-Reliefs, 1996, pp. 61-69.

en grande série, mais c'est surtout le « grain » du tamis qui change : pour contrôler une série prise comme un tout, il n'est plus nécessaire de contrôler chacun des éléments (contrôle exhaustif) car on sait calculer une taille d'échantillon de façon à obtenir une protection aussi bonne que souhaitée. On sait donc ajuster le grain du contrôle à la protection que l'on veut obtenir – dans la limite, bien sûr, où le comportement de l'objet contrôlé reste compatible avec les hypothèses du modèle. Le contrôle au niveau du collectif pris comme tel représente un saut conceptuel majeur par rapport au contrôle exhaustif des éléments car on admet la possibilité de « *laisser passer des erreurs* », sous condition que le phénomène soit calculable.

Un autre aspect important, et complémentaire du précédent, est la prise en compte du coût de l'acquisition d'information, c'est-à-dire de la dimension économique du contrôle. L'acquisition d'information a un coût, à mettre en balance avec les avantages que l'on retire de l'information. Cette « *théorie statistique de la décision* » sera formalisée entre 1940 et 1950 par le mathématicien et économiste Abraham Wald, après qu'il eut lui aussi théorisé le contrôle statistique (42). Shewhart a constamment insisté sur la dimension économique du contrôle statistique (son ouvrage majeur s'intitule « *Economic Control of Quality of Manufactured Products* ») mais sans arriver à l'articuler très clairement avec le raisonnement statistique. Estienne et Dumas, chacun à leur façon, perçoivent cette dimension économique et tentent de la traduire dans leurs conceptualisations respectives, avec un succès mitigé.

Enfin, le « regard » de la recherche en gestion nous montre que les phénomènes d'ordre technique, scientifique et administratif demandent à être traités conjointement si l'on veut comprendre la genèse de telle ou telle technique d'organisation. Le travail d'Estienne en témoigne tout particulièrement : ses propositions en matière de concurrence restreinte relèvent d'une théorie de l'action administrative qui conditionne le statut et la valeur de la connaissance qu'apporte l'échantillonnage. La connaissance tirée de l'induction statistique ne possède pas en elle-même de valeur mais elle en acquiert une par sa mise en relation avec des actions, ou avec d'autres connaissances (ou mesures) qui sont liées à – ou représentent – des actions. Pour en saisir toute la signification et la valeur, c'est l'ensemble du réseau des relations à travers lesquelles elle circule qu'il faut reconstituer.

(42) WALD est l'auteur de la théorie du « contrôle séquentiel », qui est la modalité la moins coûteuse en inspection, d'un point de vue théorique au moins (car sur le terrain industriel, elle a souvent été considérée comme trop complexe à mettre en œuvre). WALD A. : *Sequential Analysis*, New York : Wiley, 1947