J. BONITZER

Réflexions sur les modèles statistiques de décision

Revue de statistique appliquée, tome 32, n° 1 (1984), p. 9-37
<http://www.numdam.org/item?id=RSA_1984__32_1_9_0>

REFLEXIONS SUR LES MODELES STATISTIQUES DE DECISION

J. BONITZER

TABLE DES MATIERES

I. INTRODUCTION

Problèmes-carrefours où confluent des intérêts multiples (intellectuels, mais aussi matériels : financiers, économiques, technologiques, sociaux, politiques), les problèmes de décision sont explicitement posés en tant que tels depuis un peu plus de trois siècles. Mon ambition n'est pas d'en faire ici une théorie exhaustive, mais plutôt une analyse en profondeur sur quelques points, dont la compréhension peut influer sur l'attitude à prendre dans des cas d'applications pratiques, en mettant spécialement l'accent sur le rôle de la statistique – étant entendu, et c'est même un point sur lequel j'insisterai fortement, que la statistique n'est qu'un des moyens de réponse à la disposition des praticiens en mal de décision (surévaluer son rôle est un écueil à éviter).

Mots-Clés : **Décision, Statistique, Modèles.**

La technique mathématique des statistiques à mettre en œuvre pour leurs applications ne présente pas de difficulté (je ne parle pas, faut-il le souligner? de la *théorie* en statistique mathématique). Au contraire, la *signification pratique* des opérations de caractère statistique est *très difficile* à comprendre. Il en résulte que la mise en œuvre des statistiques revêt beaucoup trop souvent un caractère formel et mécanique, qu'elle consiste à "faire un test", "faire une estimation", "faire une analyse de régression", "faire une analyse de données" sans se demander pourquoi, ni à quoi elles servent.

Les réflexions qui vont être exposées dans cet article ne seront donc pas destinées à proposer sur des points précis un progrès des techniques statistiques, mais à attirer l'attention des praticiens sur des questions qu'ils ont intérêt à se poser, et sur les conséquences concrètes et pratiques qu'elles peuvent avoir.

Je toucherai inévitablement au débat bien connu qui oppose "subjectivistes" et "fréquentistes". L'objet de ce débat est intéressant en soi (bien qu'il ait un peu tendance à tourner en rond). Mais d'une part, je me refuse à considérer qu'il faille nécessairement adhérer à l'une ou l'autre de ces deux thèses(1) (je les récuse l'une et l'autre pour des raisons différentes) : d'autre part je soulignerai, aussi souvent que ce sera nécessaire, que le débat a avant tout à mes yeux une portée pratique ; et que, sur des points importants, on ne se comporte pas pratiquement de la même façon selon les réponses qu'on fait aux questions d'"interprétation" du raisonnement statistique.

Afin que cet exposé ne se perde pas dans l'abstraction, je discuterai d'ailleurs, non seulement des problèmes généraux de la décision, mais aussi de situations particulières, tirées de ma propre expérience de praticien, à propos desquelles des questions délicates se posent.

Je commencerai par un passage en revue des étapes du raisonnement statistique classique, décrites et analysées dans un ordre *logique*, caractérisé par une description de plus en plus détaillée de la "situation décisionnelle" concernée (cet ordre n'est pas identique à l'ordre *historique* de l'introduction des différentes techniques, qui s'est faite d'ailleurs par morceaux avec des changements de perspective au cours de ses étapes successives).

A l'exception de la Théorie statistique des décisions, il s'agira de techniques suffisamment connues de tous les praticiens pour que leur description n'entraîne pas de difficultés de lecture. En ce qui concerne la Théorie statistique des décisions, des explications un peu plus techniques seront inévitables. Elles sont indispensables, en outre, pour faire comprendre la complication fondamentale, "incontournable" (pour employer un terme à la mode), des problèmes de décision, quand on veut en approfondir la compréhension.

II. LES ETAPES LOGIQUES DU RAISONNEMENT STATISTIQUE

L'orsqu'un ingénieur affronte un problème pratique, il ne peut jamais le faire en tenant compte de *tous* ses détails. Il y en a qu'il considère comme des données sur lesquelles il est sans action ; souvent comme des données globales, dans les-

(1) Pas plus qu'à la bizarre position "logiciste" qu'on y ajoute parfois (cf. [1] p. 526 ouvr.).

quelles on intègre des quantités de caractéristiques secondaires qu'il n'y a pas lieu de détailler. Il en est d'autre au contraire que l'on considère comme significatifs ; ceux en fonction desquels on prendra des décisions. Nous admettrons que le terme de "hasard" se réfère à un certain type de phénomènes (que nous essayerons de caractériser). Les phénomènes du hasard, ou aléatoires, ne font pas exception à la règle : on doit choisir les caractéristiques qu'on intègrera dans leur représentation. Le vrai problème est toujours de savoir ce qu'on considère comme des données "ultimes", globales, non analysables, et ce qu'au contraire on doit considérer comme significatif, et soumettre à l'analyse. Le progrés de la connaissance revient toujours à intégrer dans l'analyse une partie de ce qui en était autrefois exclu. Nous allons voir que les étapes (logiquement) successives du raisonnement statistiques, à savoir :

- l'estimation ponctuelle ;
- l'estimation par intervalle de confiance ;
- le test statistique "classique", au sens de NEYMAN et PEARSON ;
- le choix des stratégies de décision, au sens de la Théorie Statistiques des Décisions de WALD

ont en effet pour caractéristiques d'intégrer successivement dans le champ du raisonnement des éléments que les étapes précédentes traitent comme des données "ultimes".

Mais nous verrons aussi que ces étapes successives correspondent à chaque fois à un éclairage nouveau du problème général de la décision, et que la dernière implique un véritable bouleversement dans la compréhension d'ensemble des problèmes décisionnels, dont il ne me semble pas qu'il ait été toujours perçu dans sa vraie dimension.

III. L'ESTIMATION PONCTUELLE

III.1. L'échantillonnage aléatoire

Toute opération statistique commence par le choix des objets sur lesquels on va chiffrer les caractéristiques. Il en est ainsi, pour commencer, de l'estimation ponctuelle, mais ce qui va en être dit concerne aussi bien toutes les formes plus élaborées du raisonnement statistique.

L'estimation ponctuelle se présente comme une forme de la mesure (au sens métrologique du terme (2)). Mais c'en est une forme très particulière, du fait que la caractéristique estimée n'est jamais *directement* celle qui intéresse le spécialiste concerné, mais une caractéristique *transformée* par l'opération *d'échantillonnage aléatoire*. A la caractéristique directement intéressante est ainsi substituée une autre caractéristique, qui lui ressemble fort – et qui même, espère-t-on, lui est égale – à savoir un paramètre issu d'une distribution de probabilité.

(2) Le terme de *mesure* a aussi une signification mathématique abstraite, qu'il importe d'autant plus de distinguer de son sens métrologique que la Théorie *mathématique* de la Mesure est la base de toute la Théorie contemporaine des Probabilités. Lorsque j'emploierai (très rarement) le terme "mesure" dans son sens mathématique, ce sera toujours dans l'expression "Théorie mathématique de la mesure".

Premier exemple

Nous supposons qu'un ingénieur des Mines veut savoir quelle est la richesse d'un filon en un certain métal, (pour éclairer une décision d'exploitation). Nous supposons en outre que ce filon est entièrement délimité, et que son volume est V (connu). Ce qui intéresse l'ingénieur, c'est la teneur totale du filon, soit $V\theta$; ou encore sa teneur moyenne θ. Il est patent qu'*elle n'a rien d'aléatoire.*

Pour la mesurer par une estimation ponctuelle, cet ingénieur peut procéder ainsi(3): il divise idéalement le filon en N petits morceaux disjoints de même volume, et il tire au sort n de ces petits morceaux (ou "prises"), qui seront analysés. Soit $X_1\ X_2 \ldots X_n$ les teneurs respectives en métal des prises. L'ingénieur prendra pour estimateur de la teneur moyenne θ la moyenne arithmétique $\overline{X} = \dfrac{\sum_i X_i}{n}$.

On peut faire à ce sujet deux remarques :

1) Dans l'échantillonnagre aléatoire, le hasard est *complètement exogène* par rapport au problème pratique posé. Il peut être introduit par l'usage de tables de nombres au hasard (dont l'origine se perd dans les méandres de leur génération), par un programme informatique de génération de suites "pseudo-aléatoires"(4) par un dispositif mécanique *ad hoc* (tel que le quartage, ou les boules de la loterie nationale), ou encore par des simulations des types les plus divers, comme celles qui sont mises en œuvre par les organismes de sondages.

2) Quoi qu'il en soit, la mesure se fait en deux temps :

– d'abord on passe de θ à θ', espérance mathématique de $\overline{X}$, par l'échantillonnage aléatoire ; il faut espérer que celui-ci est bon (mais l'est-il jamais absolument ?) ; sinon, il introduit une erreur systématique, ou *biais* ;

– ensuite on passe de θ' à $\overline{X}$; on peut dire, en remontant la chaîne, que $\overline{X}$ est pris pour mesure de θ' ; et donc, indirectement, de θ ; quand on passe de θ' à $\overline{X}$, il s'intordruit une erreur aléatoire ; dans les cas (possibles) où l'espérance mathématique de l'estimateur n'est pas exactement égale au paramètre qu'on veut estimer, il peut s'introduire un second biais.

Deuxième exemple

Un ingénieur chargé de construire un pont en béton précontraint reçoit un lot de fils d'acier de précontrainte. Il veut en caractériser la qualité.

La différence avec le premier exemple tient à ce que les caractéristiques qui intéressent pratiquement l'ingénieur ne se résument pas à une simple moyenne. En fait, il a besoin de connaître plus ou moins une *distribution* des caractéristiques du fil dans le lot, et cette distribution ressemble plus ou moins à une distribution de probabilités. On peut donc avoir l'illusion, faisant une estimation, de travailler *directement* sur un objet aléatoire. On peut d'ailleurs discuter si cet objet

(3) Naturellement, la procédure réelle est beaucoup plus complexe et raffinée, et tient compte des connaissances qu'on a sur la formation et la structure des filons ; elle est ici outrancièrement simplifiée pour les besoins de l'exposé.

(4) La distinction de ce qui est "essentiellement" aléatoire et de ce qui est "pseudo-aléatoire" renvoie à des problèmes mathématiques de très haut niveau. Pour la compréhension des situations décisionnelles, elle n'a aucune importance.

est vraiment aléatoire, ou seulement apparemment. C'est une question de point de vue(5). Quoi qu'il en soit, l'opération d'échantillonnage aléatoire à laquelle l'ingénieur va se livrer sur le lot va engendrer une autre distribution de probabilité, dont on souhaitera naturellement qu'elle soit dans des relations précises avec la distribution des caractéristiques dans le lot, de sorte que puisse être interprété clairement la passage de celle-ci à celle-là(6). Un second passage sera effectué ensuite, de la distribution de probabilité issue de l'échantillonnage aux estimateurs qui auront pour fonction de la représenter pour les besoins pratiques : moyennes, variances, fractiles et valeurs extrêmes, étendues, etc.

Ce qu'il faut retenir, c'est que *l'échantillonnage aléatoire implique une modélisation*, (qui consiste en la représentation d'un objet par une distribution de probabilité) ; c'est-à-dire, en fait, les premiers linéaments de l'expression d'un point de vue pratique dans toute opération relevant de la statistique (point essentiel sur lequel je reviendrai longuement dans la suite). Il y a des usages relativement élaborés de la statistique où le recours à des modèles est évident. C'est le cas des analyses de plans d'expériences (modèles linéaires), des statistiques de processus (modèles de BOX et JENKINS, ou autres), etc. Dans l'estimation ponctuelle d'un paramètre numérique unique, il se réduit à si peu de chose (la représentation d'une caractéristique originelle, quelle que soit sa nature, par un paramètre issu d'une loi de probabilité) qu'il en devient presque invisible.

La modélisation est un moyen de connaissance d'une extrême généralité ; elle n'est pas particulière à la statistique. La première caractéristique de celle-ci, bien au-delà de l'estimation ponctuelle, est de recourir à des modèles formés par des distributions de probabilité. C'est ainsi qu'en Théorie Statistique des Décisions, on sera amené à considérer un espace dont les éléments sont des distributions de probabilité, et qu'on appelle conventionnellement "espace des états de la nature". Il doit être évident qu'il faut considérer qu'il s'agit là d'une appellation contractée ; cet espace est en fait celui des "modèles des états de la nature".

La statistique présente en outre une seconde caractéristique.

Dans les usages *directs* du Calcul des Probabilités, on a affaire à un phénomène aléatoire qu'on représente par une distribution de probabilité (c'est-à-dire un objet mathématique non aléatoire). Dans les usages *statistiques* du Calcul des Probabilités, on a affaire à une distribution de probabilité, qu'on représente par une variable (ou un vecteur, ou un processus) aléatoire. Les problèmes de statistiques se rattachent ainsi à la vaste classe des *problèmes inverses.* Mais, comme les distributions de probabilité sont de simples objets mathématiques, qu'on ne rencontre pas à l'état natif dans les phénomènes naturels, on est bien obligé de les fabriquer *ad hoc*. C'est la fonction de l'échantillonnage aléatoire. Dans toute application de la Statistique, il y a nécessairement au moins *trois* types d'objets : des "objets de la nature", des modèles probabilistes de ces objets, et des valeurs observées de variables aléatoires.

---

(5) Sur l'importance du "point de vue", voir plus loin (§ VI.2). Si on se place du point de vue de la technique de fabrication du fil, il se peut que l'ensemble des caractéristiques de celui-ci soit à considérer comme une occurrence d'un processus ergodique, c'est-à-dire comme une distribution non aléatoire reproduisant en un sens précis une distribution aléatoire.

(6) Une situation comparable existe dans toutes les applications démographiques de la Statistique, où l'on étudie la distribution de caractéristiques (âge, taille, poids, etc.) dans une population. Cette distribution ne doit pas être confondue avec celle de l'échantillon, extrait de la population par un acte aléatoire ou pseudo-aléatoire, sur lequel on fait des mesures ou des observations.

III.2. Ce qu'apporte, et ce que n'apporte pas, l'estimation ponctuelle.

Nous avons vu que l'estimation ponctuelle fournit une mesure approchée d'un paramètre (ou d'un ensemble de paramètres) d'un modèle. Elle peut en outre fournir un *indicateur* de l'erreur commise – mais non pas une borne supérieure – sous la forme d'une variance, d'une variance estimée, ou d'une étendue de l'échantillon. L'indicateur variance, notamment, permet de comparer divers estimateurs d'un même paramètre, et de choisir le plus précis (le plus "efficace").

D'autre part, l'estimation ponctuelle ne se prête pas à sa traduction en consignes d'action (Elle favorise même l'illusion qu'on peut faire des Statistiques "à toutes fins", sans aucune idée des applications qu'on en tirera).

L'estimation par intervalle de confiance va avoir pour objet de remédier au premier défaut, le test statistique, jusqu'à un certain point, au second.

IV. L'ESTIMATION PAR INTERVALLE DE CONFIANCE

L'estimation par intervalle de confiance est directement inspirée de la conception élémentaire de l'erreur physique de mesure. l'exemple le plus simple, celui de l'estimation de la moyenne d'une variable normale de variance connue, est parfaitement représentatif.

Soit $X = (X_1, X_2 \ldots X_n)$ un échantillon de n variables normales indépendantes, de même distribution, de moyenne m (inconnue) et de variance σ^2 (connue). Soit $\overline{X} = \dfrac{\sum_i X_i}{n}$. Pour toute probabilité ϵ, on peut définir un nombre positif t_ϵ tel que l'intervalle $(\overline{X} - t_\epsilon\sigma, \overline{X} + t_\epsilon\sigma)$ contienne la moyenne m avec une probabilité $1 - \epsilon$. Autrement dit, $\overline{X}$ est une mesure approchée de m, avec une erreur dont $t_\epsilon\sigma$ est une borne supérieure, avec la probabilité $1 - \epsilon$.

On voit bien ce qu'apporte l'estimation par intervalle de confiance par rapport à l'estimation ponctuelle : des précisions sur l'ordre de grandeur de l'erreur commise, qui permettent des raisonnements informels, véritables embryons de raisonnements décisionnels : par exemple, on peut tester la validité d'un modèle qui affecte à la moyenne m une valeur m_0. Si m_0 est hors de l'intervalle de confiance calculé, on sera enclin à rejeter l'hypothèse.

Mais on voit en même temps les questions qu'elle soulève.

D'abord, celle de *l'interprétation* du concept de probabilité, qui a fait couler des flots d'encre dans des batailles épiques entre "subjectivistes" et "fréquentistes".

Pour un "subjectiviste", comme SAVAGE, la probabilité ϵ a la signification d'une évaluation *personnelle* (subjective) du "degré de croyance rationnelle" (selon une formule due à Keynes) dans la réalisation de l'évènement auquel elle s'applique (en l'occurrence, cet évènement est : "l'intervalle de confiance ne contient pas la moyenne réelle")(7). Pour un "fréquentiste", comme von MISES, elle est la

(7) La thèse de SAVAGE est que l'évaluation "personnelle" des probabilités combine une évaluation initiale, purement originelle et ultime (sans fondement rationnel, pur produit de la "croyance") et l'information fournie par l'observation. Elle revient en fait à supprimer le concept de probabilité, pour le remplacer par celui d'estimateur bayésien de la probabilité, jugé seul opérationnel.

limite de la fréquence de réalisation de l'évènement dans une suite aléatoire infini d'estimations répétées.

J'indique dès maintenant (et je préciserai plus loin pourquoi) que le point de vue subjectiviste me paraît devoir être fermement rejeté – bien que son attrait soit explicable, et que certaines idées qu'on lui rattache soient validables dans une autre approche – ; mais que les réponses "fréquentistes", bien qu'inspirées d'idées plus saines, et finalement plus intelligibles, restent entachées de faiblesses qui ne leur permettent pas de convaincre. La faiblesse du "fréquentisme" tient d'abord à ce qu'il a de commun avec le "subjectivisme", à savoir une excessive focalisation sur le seul concept de probabilité d'un évènement particulier. La Théorie moderne des Probabilités ne raisonne plus ainsi. Ses objets son des *espaces* ou des *distributions de probabilité*, où l'on associe de façon indissoluble les probabilités à un ensemble d'évènements (une tribu d'évènements), dont on ne peut sans dommage méconnaître la signification, comme on le verra plus loin ([8]).

Quoi qu'il en soit, l'estimation par intervalle de confiance pose une autre question, apparemment plus mineure, mais dont les conséquences vont se montrer à très longue portée : Pourquoi, dans le cas illustratif choisi, prendre un intervalle *symétrique* par rapport à l'estimateur $\overline{X}$? Pourquoi pas une demi-droite, à droite ou à gauche, – ce qu'on fait, d'ailleurs, dans d'autres cas – ? Ou même, n'importe quel intervalle $(\overline{X} - t_1 \sigma, \overline{X} + t_2 \sigma)$, pour peu seulement qu'il y ait une probabilité $1 - \epsilon$ qu'il contienne la moyenne m ?

C'est d'abord une réponse à cette question que la Théorie classique des tests va nous proposer. Mais sa réponse va se trouver être en même temps l'amorce (seulement l'amorce) d'une très importante évolution.

V. LA THEORIE CLASSIQUE DES TESTS DE NEYMAN ET PEARSON

V.1. Le test d'une hypothèse simple contre une contre-hypothèse simple.

NEYMAN et PEARSON, vers 1930, démontrent un théorème qui apporte une réponse à la question du choix de l'intervalle de confiance, parmi tous ceux de même risque.

On a vu que l'estimation par intervalle de confiance se prêtait plus ou moins à des applications décisionnelles en permettant le cas échéant de rejeter *une* hypothèse, ou, dans le cas contraire, de lui maintenir (ou peut être de consolider) son statut d'hypothèse envisageable.

Le théorème de NEYMAN et PEARSON va plus loin, en opposant l'une à l'autre *deux* hypothèses respectivement appelées *l'hypothèse* et la *contre-*

(8) Historiquement, le fréquentisme a buté sur un obstacle instructif : la difficulté très grande qu'il y a à définir formellement la notion de suite aléatoire (la définition n'est pas unique, il y a plusieurs façons possibles, pour une suite, d'être aléatoire ; et il y faut des mathématiques d'un niveau très avancé). De sorte que l'ambition de von MISES de déduire de cette notion une axiomatique des Probabilités aisément intelligible s'est trouvée ruinée. Cela dit, la tentative de von MISES n'a pas peu contribué à lancer un courant de recherches mathématiques fécondes. Moralité : il est illusoire de vouloir fonder une axiomatique sur des présupposés philosophiques ; une axiomatique, c'est une synthèse de travaux scientifiques. Mais les idées philosophiques peuvent suggérer des champs de recherches.

hypothèse. Nous allons indiquer son contenu et voir sa mise en œuvre d'après un exemple illustratif.

Nous imaginons un ingénieur chargé de réceptionner un lot de ciment pour construire un ouvrage délicat. Il a besoin d'un très bon ciment, bien défini. Or, c'est un ingénieur très expérimenté et, d'après l'aspect du produit livré, il a l'impression que c'est un autre ciment qu'on lui a livré, lui aussi bien défini, mais impropre à l'usage qu'il veut en faire. Il sait qu'on peut les distinguer par une analyse chimique facile et assez précise portant, disons, sur la silice soluble (dans l'acide chlorhydrique dilué). Le premier doit avoir une teneur θ_0, et le second θ_1. Pour faire le test, on va échantillonner le ciment (correctement, nous le supposons), et analyser n petites prises, et on obtiendra un ensemble de n résultats d'essai, $X = (X_1, X_2 \ldots X_n)$ qu'on peut considérer comme des variables aléatoires indépendantes en probabilité, de même distribution, dont la moyenne est θ_0 si le ciment est le bon ciment, et θ_1 si c'est le mauvais. Nous admettons en plus que la distribution des X_i est normale, ce qui est plausible s'il s'agit d'erreurs de mesures expérimentales.

L'ingénieur doit ici prendre une décision : accepter la livraison, ou la refuser ; ou, en d'autres termes, il doit choisir entre deux hypothèses : l'hypothèse H_0 que $\theta = \theta_0$ et l'hypothèse H_1 que $\theta = \theta_1$. La théorie nous dit qu'il peut le faire en fonction de la valeur de la variable aléatoire $\overline{X} = \dfrac{\sum_i X_i}{n}$, et selon que cette valeur sera ou non dans un certain sous-ensemble A des valeurs possibles de $\overline{X}$. Par exemple, si $\overline{X}$ est dans A, on acceptera la livraison, et on la refusera dans le cas contraire. A sera appelé ensemble d'acceptation de l'hypothèse H_0.

Plus précisément, le théorème de NEYMAN et PEARSON nous permet de déterminer une classe d'ensembles d'acceptations possibles, tels que A, qui jouissent d'une propriété optimale que nous allons énoncer maintenant :

De toutes façons, du fait du caractère aléatoire de la mesure, il subsiste toujours une possibilité de se tromper ; nous appelons α (ou *risque de première espèce*) la probabilité de refuser la fourniture alors qu'elle est bonne ; et β (ou *risque de seconde espèce*) la probabilité de l'accepter alors qu'elle est mauvaise.

Remarquons que l'ensemble d'acceptation A, dont le choix fait partie de la procédure de décision, est exactement un intervalle de confiance au niveau de confiance $1 - \alpha$ (9). Pour α donné, nous avons une infinité de choix possible. Le théorème de NEYMAN et PEARSON nous dit alors comment choisir celui qui rend minimum le risque de seconde espèce. En l'occurrence, ce sera une demi-droite, de la forme $(\overline{X} < a)$ (10) si $\theta_0 < \theta_1$; ou $(\overline{X} > a)$ dans le cas contraire.

Ce que je viens de décrire, c'est *l'embryon* de la théorie classique des tests ; ce qu'on appelle le *test d'une hypothèse simple contre une hypothèse simple.* Il appelle déjà des commentaires très instructifs.

---

(9) A ceci près que, au départ, ce n'est même pas forcément un intervalle ; ce peut être n'importe quel sous-ensemble (plus précisément, n'importe quel sous-ensemble *mesurable*, c'est-à-dire auquel on peut attribuer une probabilité). *Pratiquement* ce sera en général un intervalle, ou un ensemble fini d'intervalles.

(10) Nous ne raffinons pas : en particulier nous ne demandons pas si cette demi-droite est ouverte ou fermée.

1) *C'est la présence d'une contre-hypothèse déterminés* H_1 $(\theta = \theta_1)$ *en face de l'hypothèse testée* H_0 $(\theta = \theta_0)$ *qui fixe le choix de l'ensemble d'acceptation* (qui est en même temps intervalle de confiance de l'hypothèse H_0) *parmi tous ceux qui réalisent le risque de première espèce supposé donné.* Une des lacunes de la théorie de l'estimation par intervalle de confiance est donc comblée.

2) Au niveau embryonnaire qui vient d'être décrit, la théorie des tests possède une caractéristique qu'elle va tendre à perdre dans ses développements ultérieurs, et qui sera récupérée dans la théorie statistique des décision : les deux hypothèses sont traitées de façon globalement symétrique. Même si le théorème de NEYMAN et PEARSON privilégie le risque de première espèce α, pour fixer ensuite la valeur du risque de seconde espèce β, il est parfaitement applicable à l'envers, et donne le même résultat. En fait, il existe *trois paramètres*, les risques α et β, et *l'effectif d'échantillon* n, parmi lesquels deux peuvent être choisis librement ; le troisième en découle.

Cette symétrie est un caractère intéressant. Elle signifie que, si l'on rejette l'hypothèse H_0, on peut donner à la décision de rejet un *contenu positif*: celui de l'*adoption* de l'hypothèse H_1. (et non pas le contenu purement négatif du *rejet* de H_0). Autrement dit, la porte est déjà entr'ouverte à la considération d'une *multiplicité* de choix possibles ; il ne s'agira plus d'accepter ou de refuser une hypothèse privilégiée, mais de *choisir* parmi plusieurs hypothèses possibles.

3) L'élaboration des relations qui définissent un test statistique peut conduire à les interpréter comme des relations entre un "signal" et un "bruit de fond".

Par exemple, si les variables aléatoires observées, qui sont les coordonnées de l'échantillon, sont des variables normales, de variance σ^2 et de moyenne θ_0 ou θ_1, selon que l'hypothèse réalisée est H_0 et H_1, on aura

$$\theta_1 - \theta_0 = h(\alpha, \beta) \frac{\sigma}{\sqrt{n}} \tag{1}$$

(où $h(\alpha, \beta)$ est une fonction décroissante de α et β).

Si nous nous donnons les *trois* paramètres α, β et n, nous voyons que nous ne pourrons obtenir que les risques soient inférieurs ou au plus égaux à α et β que si $\theta_1 - \theta_0$ est supérieur ou au plus égal au second membre de (1). Nous ne pouvons donc *discriminer* deux hypothèses que si leur distance $|\theta_1 - \theta_0|$ est supérieure à un nombre fixé par les risques, la variance σ^2 et l'effectif d'échantillon. On peut dire encore que ces deux derniers paramètres caractérisent un "bruit de fond", et que la distance entre les deux hypothèses est un "signal" qui ne peut être détecté, *compte tenu des risques* α et β, que s'il est assez grand.

N.B. : Les notions de "signal" et de "bruit de fond" sont donc *relatives à des risques donnés.*

4) Nous n'avons encore aucun moyen de déterminer "rationnellement" ces risques.

Lorsqu'il s'agit d'applications à des contrôles de qualité, dans le cadre de procédures normalisées(ou réglementées autrement), les risques α et β sont parfois spécifiés, et l'ingénieur n'a pas à se préoccuper spécialement de la façon dont ils ont été choisis. En fait, ils résultent d'un usage empirique, qui a montré qu'ils étaient d'un bon ordre de grandeur.

Mais, dans des applications moins standardisées, (et d'ailleurs, aussi bien, pour spécifier les normes et règlements), il faut bien s'interroger sur la façon de fixer les risques. Les praticiens savent intuitivement que ceux-ci doivent tenir compte :

– du coût de décisions erronées ;
– de l'information préalable disponible sur les vraisemblances des hypothèses du test.

Ils doivent être d'autant plus petits que le coût des erreurs est plus élevé et que les hypothèses sont plus vraisemblables.

Sont ainsi soulevées deux questions : celle du coût des décision erronées, et celle de la vraisemblance des hypothèses ; mais il est évident qu'elles ne sont pas explicitement posées par la théorie classique d'une manière qui permette de progresser vers leur solution.

V.2. Extensions de la théorie

La théorie du test d'une hypothèse simple contre une hypothèse simple ne s'applique qu'à des cas tellement particuliers qu'il est impossible d'en rester là. On a donc cherché à l'étendre, mais sans arriver à une théorie véritablement synthétique (ce qui n'empêche qu'au passage, on arrivait, dans le cadre de cette théorie quelque peu difforme, à des résultats fondamentaux importants(11)). (D'une part, on a cherché à traiter le concept d'hypothèses composite, c'est-à-dire posant l'appartenance du paramètre testé à un ensemble, plutôt que simplement son égalité à une valeur.On a introduit pour cela le concept, techniquement utile, de *fonction de puissance* (ou celui, équivalent, de courbe d'efficacité).

D'autre part, on a cherché dans certains cas (par exemple pour les techniques de plans d'expériences) à construire des *tests d'hypothèses multiples,* c'est-à-dire permettant de considérer plus de deux hypothèses. Cependant, en règle générale, il ne s'agit pas, pour l'utilisateur du test, de choisir entre différentes hypothèses *exclusives* les unes des autres, mais de démolir une hypothèse composite *par morceaux, inclus les uns dans les autres*. Par exemple, une analyse de régression portera sur, disons, 4 paramètres, dont on voudra tester l'égalité à 0 (ce qui signifiera qu'ils sont sans effet sur la valeur de la variable dépendante). Un test global montre que l'hypothèse doit être globalement rejetée ; mais cela n'empêche pas que peut-être certains de ces paramètres sont égaux à zéro, tandis que d'autres en diffèrent. Un test d'hypothèse multiple (tel que les tests S de SCHEFFE ou T de TUKEY) permet de décider (avec un risque d'erreur) lesquels peuvent être considérés comme nuls, et lesquels non.

Remarque

Dans ses développements, la Théorie des Tests a systématiquement tendu à estomper la symétrie de l'hypothèse et de la contre-hypothèse, qui caractérisait le test d'une hypothèse simple contre une hypothèse simple, pour privilégier l'une des hypothèses.

Il faut bien voir que les deux schémas, le symétrique et le dissymétrique, correspondent au moins en première analyse à des visées pratiques différentes.

Même si les deux objectifs se combinent le plus souvent, il est important de distinguer celui qui vise *l'acquisition de connaissances* de celui qui vise *le choix d'un acte de portée pratique* parmi un ensemble d'actes possibles connues d'avance.

---

(11) Tels que ceux concernant, par exemple, le concept de "résumé exhaustif".

La théorie des tests met en avant principalement le premier objectif, lorsqu'elle se donne comme problème d'accepter ou de rejeter *une* hypothèse, privilégiée par rapport à la contre-hypothèse. Cette hypothèse est toujours énoncée du point de vue de la connaissance, jamais directement du point de vue de l'action, ou de la décision. Si l'hypothèse est acceptée, on considère que son énoncé est consolidé, mais non définitivement validé. Si elle est rejetée, on considère qu'il est refuté [12]. Mais, dans un cas comme dans l'autre, on nous dit ce qu'il faut *penser* ; on ne nous dit pas ce qu'il faut *faire*. Un schéma symétrique peut certes être compris dans le même sens. Mais il se prête de façon beaucoup plus directe à l'application pratique. A chaque hypothèse peut en effet être associé une décision d'agir parmi toutes celles, données à l'avance, qui sont autorisées par la situation pratique dans laquelle on se trouve.

Du point de vue d'un chercheur scientifique, on peut discuter des mérites respectifs des deux approches (et nous y reviendrons). Du point de vue d'un praticien, il est assez évident que la traduction des hypothèses en décisions d'actes pratiques est toujours préférable. Cependant, comme je l'ai signalé, on ne peut faire qu'un praticien ne se trouve parfois dans une situation de chercheur scientifique, à court d'explication si son hypothèse de travail est infirmée, et par conséquent, semble-t-il, ne sachant *que faire*. Nous verrons plus clairement, à propos de la Théorie Statistique des Décisions, que cette situation est plus apparente que réelle.

V.3. Ce qu'apporte, et ce que n'apporte pas la théorie des tests

En résumé, la Théorie des Tests apporte deux perfectionnements par rapport à l'estimation par intervalle de confiance :

1) A seuil de risque donné, elle précise le choix de l'intervalle de confiance, *alias* ensemble d'acceptation.

2) Elle permet de caractériser un pouvoir discriminant de l'opération statistique.

Elle présente trois insuffisances ou limitations sérieuses :

1) Elle ne fournit aucune règle pour le choix des risques et des effectifs d'échantillon.

2) Elle ne s'étend pas aux cas d'hypothèses multiples exclusives les unes des autres, et par conséquent :

3) Elle ne se prête pas aisément à des choix d'actes pratiques, sauf dans des situations très simples.

Telle était la situation vers la fin des années trente, lorsque WALD proposa, sous une forme d'une généralité ambitieuse la Théorie Statistique des Décisions (TSD), dont nous allons voir que, même si sa mise en œuvre formelle est difficile, et si elle ne répond pas à toutes les questions pratiques, elle permet d'énormes progrès dans la compréhension du raisonnement statistique.

---

(12) La théorie des tests est ainsi directement inspirée de ce qu'on appelle le "schéma hypothético-déductif" selon lequel la voie du progrès scientifique passe par l'invention des hypothèses scientifiques, suivie d'expérimentations tendant à les confirmer ou à les infirmer. Ce schéma a été longuement développé par des théoriciens de la science tels que DUHEM ou POINCARE en France, RUSSELL, CARNAP ou POPPER à l'étranger. Je n'y insisterai pas, sauf pour signaler que sa portée, sinon sa validité, peut être discutée.

VI. LA THEORIE STATISTIQUE DES DECISIONS DE WALD

Il faut d'emblée décrire le cadre formel de cette Théorie, qui est d'une certaine complication. Cette complication, semble-t-il, est "dans la nature des choses", et on ne voit pas bien comment l'éviter.

VI.1. Les concepts de base

1. L'espace des "états de la nature" Θ

L'espace des "états de la nature" est un ensemble dont les éléments sont des distributions de probabilité, dont chacune est censée correspondre à un "état de la nature" possible. D'après ce qui a été dit plus haut, on comprendra que la terminologie est impropre, et qu'il s'agit en réalité d'un espace des modèles probabilistes des états de la nature. On le désignera par Θ, et ses éléments par θ.

Pratiquement, on admettra très souvent que la forme analytique des éléments de Θ est connue, et que chaque θ est entièrement déterminé par les valeurs d'un ensemble fini de paramètres (par exemple, les moyenne et variance, si les distributions sont normales). Si ces paramètres sont au nombre de p, l'espace des états de la nature pourra prendre la forme d'un espace euclidien à p dimensions.

Dans cet espace Θ, on peut définir des évènements, par des sous-ensembles de Θ. Dire que θ appartient à l'un de ces sous-ensembles, soit Θ_A c'est une façon d'exprimer une *hypothèse* sur les états de la nature.

Enfin, on attribue à ces hypothèses des probabilités, dites *à priori*. Ou, en termes abstraits, on définit une distribution de probabilité P sur l'ensemble, noté $\mathcal{A}$, des hypothèses définies dans Θ. L'espace des états de la nature est ainsi un triplet $(\Theta, \mathcal{A}, P)$, c'est-à-dire un *espace de probabilité* (au sens de la théorie moderne des Probabilités). Il faut évidemment bien distinguer la distribution de probabilité P des éléments θ et Θ, qui sont aussi des distributions de probabilité (on a une distribution de probabilités sur un ensemble de distributions de probabilités ; pourquoi pas ?).

2. Espace des observations Ω

Le statisticien dispose d'un ensemble d'observations, prenant la forme d'un ensemble de n variables aléatoires (ou éventuellement d'une infinité), dont les distributions de probabilité dépendent de θ (sinon , leur observation ne fournirait aucune information sur θ).

Nous avons donc à tenir compte ;

– de l'ensemble des valeurs possibles du multiplet des n variables aléatoires observées (c'est un espace à n dimensions) ; soit Ω cet ensemble ;

– des évènements possibles à l'occasion de leur observation (représentés par des sous-ensembles de Ω) ; soit $\mathcal{B}$ l'ensemble de ces évènements ;

– d'une distribution de probabilité P_θ, attribuant une probabilité (dépendant de θ) à chaque évènement.

Nous considérons donc en fait une *famille* d'espace de probabilité $(\Omega, \mathcal{B}, P_\theta)$, indexé par l'"état de la nature" θ. L'indexation porte uniquement sur la probabilité P, non sur l'ensemble d'évènements $\mathcal{B}$, qui est toujours le même.

3. Espace des décisions Δ

C'est un ensemble de décisions possibles. On peut représenter des "évènements décisionnels" par des sous-ensembles de Δ. L'ensemble de ces "évènements décisionnels" sera désigné par $\mathcal{D}$. Le couple $(\Delta, \mathcal{D})$ définit un *espace probabilisable* (quant on y aura ajouté une probabilité, définie pour tout "évènement décisionnel", ce sera un espace de probabilité).

4. Fonction de décision

C'est une *variable aléatoire* $d(X)$, faisant correspondre une décision à chaque valeur X prise par le multiplet d'observations. C'est donc une application de $(\Omega, \mathcal{B}, P_\theta)$ dans $(\Delta, \mathcal{D})$. (cf. Fig. 1).

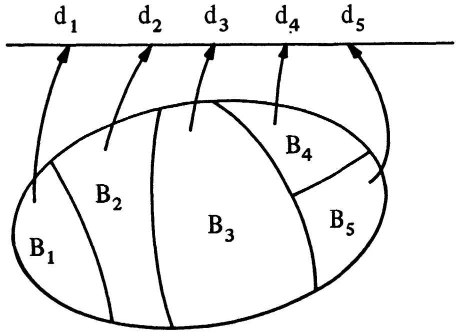

Figure 1

Supposons, comme dans le cas de la figure 1, qu'il y ait un nombre fini de décisions possibles. A chacune de ces décisions, nous pouvons faire correspondre un évènement dans $(\Omega, \mathcal{B}, P_\theta)$: celui qui contient toutes les valeurs de X qui conduisent à la décision correspondante. Comme nous connaissons pour chacun la probabilité $P_\theta(B_i)$ pour chaque évènement B_i et pour chaque valeur de θ, nous pouvons *transporter* une distribution de probabilité sur $(\Delta, \mathcal{D})$, en posant $P_\theta^d(B_i) = P_\theta(B_i)$.

A ce point de l'exposé, on comprend que le problème de la décision est ramené à celui du choix de la fonction de décision – ce qu'on appellera en d'autres termes une *stratégie*.

5. Fonction de coût

Il reste un dernier élément à définir pour que le problème de la décision puisse recevoir une solution à savoir une fonction de coût, dépendant de l'observation et de l'état de la nature θ.

En fait, cette fonction de coût se décompose en une somme de deux termes :

$C_1[d(\omega), \theta]$, coût direct de la décision

$C_2(\omega, \theta)$, coût propre de l'expérimentation ([13]).

Si l'on retient que les décisions d sont prises en fonction de l'observation aléatoire X, on voit que le coût total $C = C_1 + C_2$ est une variable aléatoire, dont

(13) C_2 dépend de ω, notamment, dans le cas des tests séquentiels.

on peut calculer l'espérance mathématique, sous la forme d'un *nombre* entièrement déterminé, d'abord par la distribution de X (fonction de θ), qui détermine celle de la fonction de décision d(X) ; puis par celle des états de la nature θ.

Solution optimale

Une solution est optimale si elle minimise l'espérance mathématique de C.

Illustration des concepts de coût

Reprenons le cas de l'ingénieur réceptionnant un ciment (§ V.1) dans l'esprit de la T.S.D. La décision d_0, c'est l'acceptation de l'hypothèse H_0 ($\theta = \theta_0$; le ciment est bon) ; concrètement, l'ingénieur accepte le ciment et le met en œuvre. Le coût C_1 (d_0 , θ_0) peut être pris égal à 0 (on a pris la "bonne" décision). Le coût $C_1(d_0 , \theta_1)$ correspond au cas où l'on met en œuvre le "mauvais" ciment. Il peut en résulter des défauts de l'ouvrage, nécessitant un surcoût d'entretient, peut-être des réparations plus ou moins grosses, peut-être même une démolition et une reconstruction prématurées. C'est le coût de tous ces inconvénients qui forme C_1 (d_0 , θ_1).

La décision d_1, concrètement, c'est le refus de la fourniture. Si celle-ci est effectivement mauvaise ($\theta = \theta_1$), on peut admettre que le coût est nul (C_1 (d_1 , θ_1) = 0). Si elle est bonne, il en résulte pour le fournisseur un dommage, en dépit de ses soins. Il doit se garantir contre ce dommage, par exemple en majorant ses prix pour en tenir compte, ou en s'assurant. C'est ce mécanisme qui déterminera $C_1(d_1 , \theta_0)$; on sent bien que l'évaluation chiffrée peut en être délicate, mais elle n'est pas inconcevable.

Le coût C_2 est simplement celui des essais du laboratoire de chantier.

Remarques

a) L'intégration des coûts est un pas décisif dans la compréhension de la décision. Elle ne procède pas d'une idée *a priori*, mais de la pratique, mettant en évidence les insuffisances de la théorie des tests, et suggérant leur origine.

b) L'existence d'un optimum, c'est-à-dire d'une décision rationnelle dans les cas relevant de la T.S.D. tient au fait que les espèces mathématiques des deux composantes du coût, C_1 et C_2, tendent à varier en sens inverse. *Il n'y aurait pas d'optimum, et partant, pas de "décision rationnelle", s'il n'existait pas un coût d'acquisition des connaissances* (il est clair qu'on aurait alors intérêt à développer à l'infini l'effort de connaissance, dans tous les sens possibles).

Sur la base de cette description, on peut faire un constat des différences entre les conceptions mises en œuvre dans la Théorie Classique des Tests et la Théorie Statistique des Décisions. C'est d'autant plus utile qu'on a souvent tendance à les estomper. Or, le constat est impressionnant, et chacun de ses points va appeler de longs commentaires. Enumérons-les d'abord :

– Dans la Théorie Statistique des Décisions, les hypothèses – alias décisions (le changement de terminologie est significatif) – sont traités de façon symétrique. Aucune n'est privilégiée ;

– La Théorie Statistique des Décisions introduit explicitement les coûts des décisions (variable en outre selon l'"état de la nature").

– Elle introduit, sous forme embryonnaire mais explicite, le coût de l'expérimentation ;

– Elle prend en compte explicitement l'information préalable, sous forme d'une distribution de probabilité a priori des "états de la nature".

Nous allons discuter les significations de ces caractéristiques.

VI.2. L'espace des décisions

Nous avons déjà signalé (§ V.2) la signification générale d'un traitement symétrique de l'ensemble des décisions d'un problème décisionnel.

Il faut encore ajouter ceci :

La définition d'un ensemble de décisions possibles, qui permet d'induire, sur l'espace des résultats d'expériences, des ensembles d'acceptation correspondants, traduit abstraitement l'existence d'un *point de vue pratique* sur tout phénomène aléatoire impliqué dans un problème décisionnel. C'est ce point de vue pratique qui détermine l'ensemble des décisions possibles de l'acteur d'une activité pratique concerné par l'expérience scientifique en question (auquel s'ajoute la donnée des coûts qu'il est appelé à subir et de l'information préalable qu'il détient).

LA NOTATION DE *POINT DE VUE PRATIQUE* EST **ABSOLUMENT ESSENTIELLE.** ELLE NE DOIT JAMAIS ETRE OUBLIEE, CAR, AU MOINS DANS LES APPLICATIONS DE LA STATISTIQUE, *LE POINT DE VUE PRATIQUE FAIT PARTIE DE LA DEFINITION MEME DU HASARD.* UN UTILISATEUR DE LA STATISTIQUE DOIT DONC TOUJOURS SE DEMANDER DE QUEL POINT DE VUE, C'EST-A-DIRE AU NOM DE QUOI IL OPERE.

Si on ne le fait pas, ce n'est d'ailleurs pas tellement parce qu'on l'oublie, mais parce que ce n'est pas toujours facile à déterminer. Nous aurons à nous demander *qui* ou *qu'est-ce qui* prend un point de vue sur une opération statistique ; quel est ce *quoi* "au nom de quoi" on opère.

Quoi qu'il en soit, pour conforter les affirmations catégoriques de l'encadré ci-dessus, on peut remarquer que la notion de "point de vue pratique" est présente sous des formes diverses, plus ou moins évidentes ou masquées, dans *toutes* les interprétations du raisonnement statistique. Elle est identifiable dans l'interprétation subjectiviste, sous la forme d'un "point de vue subjectif" du statisticien dans l'évaluation des coûts et de l'information préalable (voir ci-après § VI.5). Elle est implicite aussi dans le formalisme objectiviste de R. von MISES, selon lequel la probabilité d'un évènement est par définition la limite de la suite des fréquences de réalisation de cet évènement dans une suite de référence (ou "référentiel") formée d'une infinité d'épreuves aléatoires identiques. Il est facile de voir que le point de vue est alors représenté précisément par le référentiel.

Enfin, elle est représentée dans la Théorie moderne des Probabilités avant tout à partir du rôle qu'y joue le concept de variable aléatoire (et ses extensions : vecteurs et processus aléatoires). Ce concept, qui pouvait paraître peu fondamental dans la conception du Calcul des Probabilités du XIX^e siècle[14], prend une tout autre importance au XX^e siècle, dès lors que la Théorie des Probabilités vient se

(14) Bien qu'il remonte au milieu du XIX^e siècle (TCHEBYCHEV), il est encore si peu enraciné au début du XX^e siècle que le terme même de variable aléatoire est absent du Cours de Probabilité de J. BERTRAND, qui faisait autorité en 1900.

greffer sur la Théorie mathématique de la mesure et de l'intégration ; puisqu'il n'est autre, dorénavant, que la transposition probabiliste du concept de fonction intégrable. Or, définir une variable, un vecteur ou un processus aléatoire, c'est choisir une ou un ensemble de caractéristiques, qui vont exprimer ce qui dans le phénomène auquel on a affaire est effectivement pratiquement important (en Statistique : pratiquement important pour la décision). Une variable aléatoire détermine en somme le résultat pratique d'une épreuve aléatoire. La Théorie moderne des Probabilités n'en reste d'ailleurs pas là. Elle fait correspondre au concept de variable aléatoire le concept plus abstrait de "tribu" d'évènements (ensemble d'évènements satisfaisant à certaines règles de stabilité combinatoire) qui est encore une autre représentation (plus abstraite) du "point de vue pratique" sur un phénomène aléatoire.

Tout ceci renforce puissamment l'idée qu'*il n'y a pas de hasard en soi*, mais que *la présence d'un point de vue est inhérente à toute manifestation du hasard* ([15]).

VI.3. Le coût des décisions et le "principe de RAMSAY"

Le coût des décisions intervient dans la Théorie Statistique des Décisions sous la forme de son espérance mathématique.

La tradition probabiliste sur ce point est immuable depuis PASCAL, et la solution qu'il a apportée au "problème des partis". Sa réponse, on le sait, était la suivante : pour qu'un jeu de hasard soit équitable, la mise de chaque joueur doit être égale à l'espérance mathématique de son gain.

Or, l'expérience la plus courante, dans le nombreux cas, semble infirmer la règle de Pascal, et appeler des correctifs.

L'un des exemples les plus couramment invoqués est celui de la loterie ([16]). On sait que, dans la plupart des loteries, l'espérance mathématique du gain est inférieure au prix du billet. Puisque cependant les billets trouvent acheteurs, il faut bien, semble-t-il, que de deux choses l'une : ou bien le comportement de ceux-ci est irrationnel, ou bien il repose sur une rationalité impliquant d'autres règles que celle de Pascal. On suggère alors que le gain d'un gros lot a en réalité une "utilité subjective" distincte de son évaluation monétaire : l'"utilité subjective"

---

(15) On peut se demender s'il en est encore ainsi même si l'on est en présence d'un phénomène "naturellement" aléatoire, et non pas d'un problème décisionnel. La réponse me paraît devoir être positive. Tout phénomène aléatoire naturel est un phénomène complexe présentant à la fois un aspect macroscopique et un aspect microscopique. Ou, en d'autres termes, associant un processus macroscopique et un processus microscopique, ayant chacun ses lois propres. Le "point de vue" est alors celui du processus macroscopique sur le processus microscopique. De même que, lorsqu'un intérêt pratique est en jeu, le "point de vue" est représenté par l'ensemble des caractéristiques d'un phénomène qui importent pour le déroulement de l'activité pratique correspondante ; de même, lorsqu'il s'agit d'un phénomène naturel, le "point de vue" du phénomène macroscopique sur le phénomène microscopique est représenté par l'ensemble des caractéristiques de ce dernier qui concourent à déterminer le premier.

(16) Je ne discuterai pas ici le paradoxe de Saint-Petersbourg (très souvent invoqué) parce qu'il repose sur une espérance mathématique *infinie*. Or, on sait bien que l'application de la loi forte des grands nombres à des espérances mathématiques infinies repose sur une extension, à manier avec précaution. Même s'il est vrai que le gain moyen, dans le "jeu de Saint-Petersbourg", tend vers l'infini avec la répétition indéfinie du jeu, sa distance à l'infini reste toujours infinie. "Tendre vers l'infini" n'a pas la même signification pratique que "tendre vers une valeurs finie".

de devenir multimillionnaire, et ainsi d'échapper à une vie de gène, ou même de misère, serait supérieure à l'évaluation pure et simple du nombre de millions qui permet de la réaliser (17).

Des auteurs éminents, tels que l'économiste anglais RAMSAY (dans les années vingt), puis von NEUMANN et MORGENSTERN, et enfin SAVAGE, ont proposé de conserver la règle de Pascal, mais en l'appliquant non plus à des sommes monétaires, mais à des "utilités" subjectives. C'est ce que j'appellerai le "principe de RAMSAY".

Cette idée a été fortement discutée, notamment à l'occasion d'un Séminaire du CNRS sur la Théorie du Risque en Econometrie. On lui a fait l'objection (18) qu'elle était incompatible avec l'existence supposée d'une préférence ou d'une répulsion subjective pour le risque. Nous verrons que cette objection, bien que sa formulation soit tout-à-fait contestable, n'est pas sans un certain fondement.

En fait, deux objections bien plus graves peuvent lui être faites :

– la première, c'est que cette idée rend incompréhensible la signification de l'espérance mathématique ;

– la seconde, c'est qu'elle est inutile, et qu'il est tout-à-fait possible, de se ramener à une formulation en termes de coûts de production, ou de valeurs d'échange.

En ce qui concerne la première objection, il faut rappeler d'abord ce qui fonde la signification de l'espérance mathématique, quand elle est appliquée à des coûts monétaires.

1. L'application de l'espérance mathématique à des coûts monétaires.

La justification de l'application de l'espérance mathématique à des coûts monétaires est aisément intelligible, comme reposant sur la *loi forte des grands nombres* et de façon générale sur le concept de *convergence presque sûre.* Elle suppose *l'additivité* des coûts monétaires, – c'est-à-dire que les sommes de coûts monétaires ont une signification pratique.

Rappelons d'abord l'énoncé de la loi forte des grands nombres de KOLMOGOROV-KHINTCHINE (19) :

Soit des variables aléatoires $X_1, X_2 \ldots\ldots X_n \ldots\ldots$, indépendantes en probabilité, toutes de même distribution, ayant une moyenne (commune) m ; soit $\overline{X}_{(n)}$ la moyenne arithmétique

$$\overline{X}_{(n)} = \frac{\sum_i X_i}{n}$$

Alors $\overline{X}_{(n)}$ converge vers m quand n tend vers l'infini, avec une probabilité égale à 1.

(N.B. : $\overline{X}_{(n)}$ peut *logiquement* ne pas converger vers m ; par exemple, rien n'em-

---

(17) On notera que cette "interprétation" est contradictoire de l'idée, souvent avancée par les subjectivistes, d'un rendement décroissant, en utilité, des sommes monétaires. Mais en fait, le joueur d'une loterie n'a aucune idée de ce que c'est d'être multimilionnaire, parce qu'il n'en a aucune expérience. Et comment pourrait-il évaluer l'"utilité" d'une situation dont il n'a aucune expérience ?

(18) Spécialement M. ALLAIS.

(19) La première apparition du concept remonte à un énoncé de E. BOREL, 1909, applicable à des variables aléatoires ayant une variance.

pêche logiquement que tous les X_i soit égaux entre eux, et différents de m ; mais la probabilité de cet évènement est nulle, et on peut admettre *pratiquement* qu'il ne se produira pas). Certes, ce qui intéresse le praticien, c'est la *somme* des X_i, et non pas leur moyenne. Mais l'espérance mathématique reste toujours la meilleure évaluation possible, au long terme, de cette variable aléatoire, puisqu'elle ne s'en écarte qu'avec l'ordre de grandeur de $\sqrt{n}$, alors que toute autre évaluation s'en écarterait de l'ordre de grandeur de n.

La loi forte des grands nombres repose sur le concept de convergence presque sûre (c'est-à-dire de convergence avec une probabilité égale à 1), qui peut être étendu à bien d'autres situations que celle, très particulière, de la loi forte des grands nombres, qui peuvent être caractérisées par des additions de variables aléatoires non indépendantes, ou non identiquement distribuées.

Dans tous les cas, cependant, la signification du recours à l'espérance mathématique, qui est par définition une fonctionnelle *linéaire*, découle de son applicable à des grandeurs *additives*.

On peut dire que tout acte aléatoire d'un "décideur" est un élément de la chaîne de tous les actes aléatoires qu'il fait dans son existence, et que *le hasard se manifeste comme un phénomène de compensation au long terme entre les résultats de tous ces actes, chiffrés en termes monétaires*, pourvu seulement que les coûts et les bénéfices affectent la même caisse.

Il est évident que cette signification ne peut être maintenue, dès lors que l'espérance mathématique s'appliquerait à des grandeurs non additives, comme le sont essentiellement des utilités subjectives. C'est pourquoi RAMSAY confère à cette application la valeur d'un principe premier, vérifiable selon lui par des tests psychologiques, mais impossible à fonder sur des bases plus profondes.

Or, nous allons voir maintenant que cette idée, bloquante vis-à-vis de tout progrès ultérieur de la connaissance, est en outre inutile.

2. La rationalité du joueur de loterie peut-elle s'exprimer en termes de coûts monétaires ?

Un subjectiviste ne nous en voudra pas, je suppose, de partir du cas d'un héros de roman, puisque les romanciers sont des maîtres de l'analyse psychologique.

Le docteur HASSELBACHER [20], vieux médecin "paumé" mais lucide, émigré à Cuba dans les années cinquante, a coutume de jouer chaque semaine à la loterie nationale de l'île. Il explique que, ce faisant, il goûte successivement l'ivresse d'être millionnaire (entre l'achat du billet et le tirage) et le vertige d'être ruiné (au moment du tirage).

C'est, semble-t-il, assez bien vu. En plus de l'espérance du gain, le joueur de loterie achète un autre bien précis : un lot de fantasmes, une part de rêve.

Et il faut ajouter ceci : cet autre bien, cette part de rêve, il l'achète, 1) sur le marché du rêve, 2) au prix courant du rêve. Il n'y a pas qu'une seule façon de rêver fortune et ruine, qui serait celle de la loterie. En France, par exemple, on peut jouer à la roulette, au poker, au tiercé, sans parler d'autres substituts plus éloignés. Et il y a nécessairement une certaine concurrence entre ces divers "produits", dont chacun a son coût de production (frais de gestion, impôts), et par conséquent une certaine péréquation de leurs prix sur le marché.

---

(20) Héros secondaire de "Our man in Havana", de Graham GREENE.

En résumé, on peut énoncer les conclusions suivantes :

a) L'écart entre le prix de vente d'un billet de loterie et l'espérance mathématique du gain des joueurs n'est pas réglé par l'"utilité subjective" du gain, mais par le coût de production de la loterie. Lorsqu'il achète un billet de loterie, le joueur achète en somme deux biens économiques, à la fois distincts et conjoints. (Il y a dans le domaine de la consommation finale des situations analogues à celles de productions conjointes dans le domaine de la production). Rien dans sa situation ne requiert d'appliquer l'espérance mathématique à des "utilités subjectives".

Il reste vrai que les joueurs décident d'acheter ou de ne pas acheter un billet de loterie en fonction de la double considération de leur budget, et de la place de l'utilité qu'ils attribuent, à la fois au gain espéré et à la possession du billet pour d'autres usages que l'espérance du gain. C'est cela qui détermine le volume du marché (dont on sait bien qu'il contribue aussi à déterminer les coûts de production).

b) Il est donc légitime de vouloir aller plus loin. Il faut alors prendre conscience de l'existence d'un vrai problème non résolu. La notion d'"utilité subjective" a de fortes affinités avec le concept de "fonction de satisfaction" des théories économiques néo-marginalistes. Or, non seulement ces théories sont fragiles, mais, dans les limites de leur validité (que je ne discuterai pas ici), leur objet concerne uniquement les choix économiques *globaux* des individus, et leur agrégation dans un équilibre *global*. Autrement dit, la consommation des individus est décrite par un ensemble de biens sans autre structure que celle d'un ordre partiel. Le "principe de RAMSAY", lui, concerne des choix *locaux*. Il suppose donc une structuration plus fine, dont sa discussion donne à penser que ce n'est pas la bonne.

Celle du "cas du docteur HASSELBACHER" nous suggère que la "bonne" structuration pourrait être à chercher dans une élaboration formelle de la notion d'*usage* des biens économiques, permettant de distinguer jusqu'au niveau approprié des usages tels que ceux de nourriture, de logement, de santé publique, de culture, etc. (l'"usage" que le docteur fait du billet de loterie, c'est de répondre à son besoin de rêve). Cette notion a certainement un fondement réel, comme l'atteste sa mise en œuvre empirique dans les études de marché.

S'il est vrai que les usages ne sont pas indépendants les uns des autres, et qu'ils prennent leur signification dans le cadre de *modes de vie* dont la structuration peut être l'objet d'une recherche socio-économique qui reste largement à faire, il n'est pas invraisemblable cependant que la théorisation de celle-ci puisse mettre en œuvre un nouveau concept d'utilité, jouissant de propriétés d'additivité partielle. On ne voit pas pourquoi des utilités d'usages distincts ne seraient pas additives, dès lors que la disposition ou la non-disposition de l'un d'eux n'affecterait pas de façon sensible l'utilité de l'autre.

Nous pouvons enfin remarquer que la notion d'"usage" présente un côté *objectif*. Sa mise en œuvre empirique dans les études de marché suffit à l'attester. Lorsque nous aurons avancé dans l'identification du "sujet de vue", nous aurons peut-être du même coup quelques lueurs sur le sujet de l'évaluation des utilités, et sur la nature des cas où on peut les additionner.

3. La portée du concept d'espérance mathématique dans la Théorie moderne des Probabilités.

Le passage de la Théorie classique des tests à la Théorie Statistique des Décisions est caractérisé *avant tout* (bien plus que par la prise en compte de probabilité *a priori*, sur laquelle la plupart des commentateurs insistent préfèrentiellement) par un déplacement dans la hiérarchie des concepts, des probabilités de risques vers l'espérance mathématique du gain.

Ce déplacement est entièrement conforme à l'esprit général de la Théorie moderne des Probabilités, telle qu'elle a été formalisée en 1933 par KOLMOGOROV dans les *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Ainsi que KOLMOGOROV y insiste lui-même, l'axiomatique de la Théorie des Probabilités est dorénavant directement dérivée de celle de la Théorie de l'Intégration de LEBESGUE, au point de presque s'y identifier. Cela ne va pas sans que l'espérance mathématique, qui n'est autre que la traduction probabiliste de l'intégrale de LEBESGUE, passe du même coup au premier rang des concepts de la nouvelle théorie.

Cette prépondérance nouvelle est patente dans les développements de la Théorie eux-mêmes. D'une part, il n'est pas seulement commode de considérer la probabilité d'un évènement donné A comme l'espérance mathématique de sa fonction indicatrice I_A (21) ; cette commodité a une racine fondamentale, que révèle la théorie de l'espérance et de la probabilité *conditionnelles* : on peut dériver la seconde de la première, non le contraire (au moins dans le cas général). D'autre part, alors que l'axiomatique de KOLMOGOROV reposait sur le concept de probabilité définie comme une "mesure sur un espace abstrait", le développement de la Théorie des processus aléatoires conduit à lui substituer dans bien des cas une définition comme "mesure de Radon" (dont le théorème de représentation de RIESZ garantit l'équivalence dans un large domaine), – c'est-à-dire comme une abstraction cachée derrière un ensemble d'espérances mathématiques.

Enfin, on peut aussi noter le glissement opéré par la Théorie moderne à propos de la définition de la convergence en loi, plutôt comme convergence d'un ensemble d'espérances mathématiques que comme celle d'une suite de fonctions de répartition.

Pratiquement, ce déplacement d'accent, inséparable de celui qui pousse vers les premiers rôles le concept de variable aléatoire (puisque l'espérance mathématique est celle d'une variable aléatoire – cf. § VI.1), traduit, ou trahit, la présence latente d'un point de vue dans la définition même du hasard. Loin d'avoir un simple sens formel dans la structure de la théorie, il revêt une signification fondamentale pour la compréhension du hasard.

Historiquement, il est intéressant de rappeler que les fondateurs du Calcul des Probabilités, à commencer par PASCAL, *ignoraient le concept de probabilité*, qui ne fut introduit par J. BERNOULLI qu'au début du XVIII^e siècle, et raisonnaient uniquement en termes d'espérance mathématique. Avec l'axiomatique de KOLMOGOROV, on ne fait en somme qu'assister à l'un de ces développements en spirale, fréquents dans l'histoire des sciences.

Ces remarques éclairent d'une lumière instructive les impasses de réflexions philosophiques, "subjectivistes" ou "objectivistes", qui continuent à spéculer imperturbablement sur un état dépassé des connaissances théoriques, à base de probabilités d'évènements isolés.

---

(21) Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité, et un évènement A. Si l'on désigne par ω l'élément courant de Ω, $I_A(\omega)$ est la fonction qui prend la valeur 1 si $\omega \in A$, et 0 si $\omega \notin A$.

VI.4. Le coût de l'expérimentation

Contrairement à la Théorie classique des Tests, la Théorie Statistique des Décisions ne considère pas l'effectif d'échantillon comme une donnée. Il est très remarquable que, dans un cadre ainsi élargi, il ne puisse exister d'optimum que si l'on y intègre, fût-ce encore très partiellement, et sous une forme élémentaire, un *coût de production des connaissances*. Il en est ainsi parce que le coût *global* de la décision est alors la somme d'un coût croissant (généralement linéairement) en fonction de l'effectif d'échantillon – le coût de l'exécution des mesures matérielles effectuées sur les éléments de l'échantillon – et d'un coût décroissant (selon une fonction convexe) – le coût direct des conséquences des décisions selon l'état de la nature (l'un et l'autre exprimés par leur espérance mathématique).

La portée de cette caractéristique peut être considérablement élargie. Tout acte tendant à la connaissance a un coût, et pas seulement l'exécution d'essais dans une opération statistique. Nous y reviendrons.

Lorsqu'elle a affaire à un problème ne comportant que deux décisions (ou hypothèses) possibles, la Théorie Statistique des Décisions nous fournit une règle de décision formellement identique à celle de la Théorie classique des Tests : c'est-à-dire un ensemble d'acceptation de l'un des hypothèses (22) (dont le complémentaire est l'ensemble d'acceptation de l'autre hypothèse). De la donnée de cet ensemble d'acceptation, et de l'effectif d'échantillon, on peut déduire les risques du Test classique correspondant (ce n'est d'ailleurs pas simple du tout). La Théorie Statistique des Décisions, par l'élargissement de son horizon, attire alors notre attention sur une particularité des Tests Statistiques assez souvent méconnue : à savoir que ce qui fixe pratiquement (et, en règle générale, par le canal d'une pratique purement empirique) le niveau des risques d'un test statistique, c'est la conjonction de *trois* éléments : l'information préalable, les coûts directs des décisions selon l'état de la nature, *et le coût des mesures.*

VI.5. L'information préalable

C'est très souvent sur la seule façon de prendre en compte l'information préalable qu'on oppose la Théorie Statistique des Décisions à la Théorie Classique des Tests, en mettant l'accent de façon unilatérale sur la filiation bayésienne. On l'annexe ainsi, à peu de frais, à la conception subjectiviste du Calcul des Probabilités.

Il est vrai que la Théorie Statistique des Décisions inclut un espace des états de la Nature, revêtant la forme d'un ensemble de distributions de probabilités *a priori*, et y figurant à l'état de donnée apparemment ultime. Il est assez tentant alors de conférer à cette donnée "ultime" la signification du produit d'une évaluation par un sujet psychologique.

Or, cette conception n'est pas seulement profondément choquante au regard de la conception habituelle (au point d'être quasi-universelle) du travail scientifique, considéré comme travail de connaissance de phénomènes objectifs, elle est dans bien des cas *pratiquement* inacceptable. Elle l'est en particulier lorsque l'on veut établir des *normes* formelles d'emploi des statistiques, comme il en figure dans

(22) De même que la théorie des tests fournissait une règle de décision formellement identique à la construction d'un intervalle de confiance.

tous les corpus de normalisation : il est impossible en effet d'établir des contrats, des cahiers des charges de prestations, incluant la spécification de contrôles statistiques sur des bases subjectives.

Avant de préciser la critique du subjectivisme, et d'indiquer quel est, à mes yeux, son défaut central, il faut cependant constater que, dans bien des situations, ce défaut peut être sans portée pratique, et le subjectivisme, comme toute, localement défendable. Il en est ainsi, en particulier, lorsque la décision d'un acteur quelconque d'une activité sociale repose entièrement sur sa responsabilité personnelle, et n'a pas à recevoir une sanction objective. Même lorsque cette décision est soumise à un contrôle contractuel, et que l'on doit justifier les raisons pour lesquelles on a choisi tel niveau de risque, on peut s'en tirer par des procédures d'évaluation "intersubjectives" des distributions de probabilité, par des groupes d'experts, dont la qualité d'expert est objectivement reconnue par des titres ou des références, et dont on peut objectivement constater l'accord. On peut sans aucun doute argumenter plus subtilement, et s'interroger sur le fondement des évaluations "subjectives" (se demander en particulier si une évaluation "intuitive" n'a pas en réalité un contenu d'analyse logique informelle qui pourrait être explicité) ; on pourra surtout, nous allons le voir, s'interroger sur la nature des points de vue représentés dans l'évaluation ([23]) (et donc se demander quelles doivent être les règles objectives présidant au choix des experts). C'est sur ce dernier point que je vais surtout insister.

1. Un contre-exemple instructif

Les subjectivistes se réfèrent volontiers à l'autorité de von NEUMANN et MORGENSTERN pour soutenir leurs thèses relatives aux utilités subjectives. C'est pourtant la partie la plus banale, la plus contestable et la moins intéressante de *Theory of Games* qu'ils invoquent ainsi, négligeant le contenu des chapitres relatifs à la Théorie des Jeux stratégiques proprement dite dans ce qu'il a de plus profond ; et notamment à la Théorie des jeux à deux joueurs, de somme nulle (théorie tout à fait indépendante de la thèse de ces mêmes auteurs relative à l'utilité).

On connaît le jeu "de la pierre, des ciseaux et du papier" – disons, en bref, le "jeu PCP". Deux joueurs choisissent chacun, dans l'ignorance du choix de leur adversaire, l'un de ces trois objets.

La pierre bat les ciseaux (elle les ébrèche); les ciseaux battent le papier (ils le coupent) ; le papier bat la pierre (elle l'enveloppe). Le perdant paye 1 franc au gagnant. La théorie des jeux stratégiques montre que les joueurs, mus par le désir de gagner, *doivent* en arriver à jouer, ou à faire comme s'ils jouaient au hasard, en choisissant à chaque fois l'objet sur lequel ils jouent avec la probabilité 1/3 (tous les coups étant en outre indépendants en probabilité)([24]). Et non seulement il *doivent* le faire, s'ils ne veulent pas perdre, mais dans ce cas très simple, *c'est ce qu'ils font pratiquement*, même si leur comportement résulte d'un ajustement empirique, et non d'un raisonnement formel. Il est facile de comprendre, en effet,

(23) "Point de vue" a un sens beaucoup plus large qu'"intérêt". Ce terme caractérise ici tout ce qui détermine l'expérience personnelle de l'expert, en fonction de son rôle social, c'est-à-dire la connaissance qu'il a, en fonction de ce rôle, des règles qui s'imposent à la pratique sociale dont il est un agent. Naturellement, ce qui concerne directement les intérêts matériels dont il est le représentant joue un rôle de premier plan dans son "point de vue" (tout n'y est pas sur le même plan). Mais ce serait une grave erreur d'interprétation de le limiter à cela.

(24) C'est ce qu'enseigne du *théorème du minimax*. La stratégie qu'il définit est celle qui *mini*mise, en espérance mathématique, le *maxi*mum de la perte du joueur qui l'adopte, face à l'ensemble des stratégies possibles de son adversaire.

que si l'un des joueurs choisit un des objets avec une fréquence plus grande que les autres, il suffit à son adversaire de choisir à tous les coups l'objet qui le domine pour être gagnant au long terme ; mais aussi que si ce joueur adopte une règle logique quelconque pour choisir le déroulement de ses coups, (par exemple, celle de la permutation circulaire), et si son adversaire la découvre, celui-ci pourra gagner à tous les coups, en appliquant la même règle logique au même ensemble d'objets transformé par permutation circulaire. Si les joueurs jouent assez longtemps l'un contre l'autre, ils finiront par comprendre cette situation, et par en tirer empiriquement la conclusion. Il est absolument vain de se demander s'ils joueront "réellement" au hasard ou s'ils "simuleront" le hasard : il faut en tout cas que l'effet pratique soit le même.

Le jeu perd ainsi son caractère de combat psychologique (décrit par E. POE dans un passage célèbre de "la Lettre volée", à propos d'un jeu encore peu plus simple), pour prendre celui d'un jeu de hasard, ont les probabilités sont déterminées par les règles *purement objectives* du jeu. La stratégie ainsi définie prend le double caractère d'une stratégie *stable*, et d'une stratégie *effective*.

Cela va nous permettre de mieux comprendre, maintenant, une situation de plus grande portée pratique, et tout-à-fait instructive.

2. Un cas de contrôle de qualité

Nous considérons le cas suivant : un ingénieur, qui doit construire un pont en béton armé, reçoit d'un fournisseur un lot de fils d'acier destiné à fabriquer ses armatures. Ce fil d'acier est normalisé, et l'ingénieur le soumet à un essai de réception, lui-même normalisé ; cet essai comporte le prélèvement d'un échantillon, suivi de diverses mesures (diamétre du fil, résistance à la traction, allongement à la rupture). Il se conclut par la réception ou le rejet du lot.

Au moins deux personnes sont intéressées au premier chef par cet essai : le fournisseur et le client. Indirectement, il y a aussi les responsables de la normalisation et de la rédaction des règlements techniques (il y a intérêt pour eux à connaître, dans sa réalité pratique, la mise en œuvre des normes et règlements), et sans doute d'autres. Mais, pour la clarté de l'exposé, nous nous bornerons aux deux premiers acteurs, directement en présence.

Dans l'esprit de la TSD, nous devons tenir compte de distributions *a priori* des caractéristiques mesurées. Ce ne sont pas les mêmes pour le fournisseur et pour le client.

Commençons par le fournisseur. La qualité d'un produit dépend des conditions de sa fabrication ; plus précisément, des réglages de cette fabrication, portant sur un ensemble fini de paramètres. Ces réglages n'excluent pas une certaine variabilité aléatoire des caractéristiques, du fait de l'intervention d'autres paramètres, sur lesquels le fabricant est sans action (soit que leur effet soit mineur, soit qu'il ne sache pas les maîtriser, soit qu'il ne se soit pas aperçu qu'ils sont à la fois importants et maîtrisables).

Cette variabilité aléatoire est de nature entièrement objective. Il se trouve que, dans ce cas, on a affaire à un processus objectif qu'on peut décrire naturellement par une distribution de probabilité (on a vu plus haut, § III.1, que ce n'était pas le cas général ; mais ici, indiscutablement, ce l'est). Et il va de soi que, dans l'établissement de sa stratégie industrielle, pour régler sa fabrication au mieux de ses intérêts, le fournisseur doit chercher à la connaître le mieux possible. Il tâchera de s'en approcher par tous les moyens disponibles, c'est-à-dire par l'observation Statistique, *mais aussi par l'analyse technique des conditions de la fabrication.*

Le fournisseur connaît plus ou moins les spécifications des contrôles de qualité de l'ensemble de ses clients. Il sait que ses coûts de fabrication seront d'autant plus élevés que les réglages de la fabrication seront plus précis ; mais, d'un autre côté, il ne peut se permettre de subir un taux de rebuts trop élevés. Il fixera donc le niveau de son contrôle interne de fabrication de manière à optimiser le coût global, combinant le coût direct de fabrication (incluant les frais de contrôle technique) et le coût indirect des rebuts par la clientèle (qui dépend de la structure de celle-ci, de l'existence éventuelle de clients moins exigeants à qui on peut refiler des productions de seconde qualité, et aussi des frais de transport).

Du côté du client, la situation diffère sur plusieurs points. D'une part, s'imposent à lui les points de vue de la normalisation et des règlements techniques. Même s'il connaissait les conditions spécifiques de la production de son fournisseur, le contrôle qu'il exerce est déterminé (contractuellement) avant le choix de celui-ci. Il doit donc tenir compte de la distribution pondérée des caractéristiques des fournitures de l'ensemble des fournisseurs possibles. Il y a donc bien, en un sens, une "personnalisation" des distributions de probabilité du fournisseur et du client. Mais cette personnalisation n'est pas d'origine psychologique. Elle tient à des raisons entièrement objectives. De plus, les distributions de probabilité, vues respectivement des points de vue du fournisseur et du client, ne sont pas indépendantes ; la première est, par rapport à la seconde, une distribution conditionnelle, liée par la condition que le fournisseur est un fournisseur déterminé.

Il reste au client (ou aux responsables de la normalisation et de la réglementation technique) à mettre en œuvre des moyens de connaissance de la distribution objective à laquelle ils ont affaire. Il opérera pour cela comme le fournisseur : il mettra en œuvre à la fois une observation statistique de longue durée, et tous les autres moyens à sa disposition lui permettant d'analyser ce qui contribue à déterminer cette distribution objective ; c'est-à-dire, en particulier, une connaissance d'une part des relations techniques entre le coût et la qualité des productions ; d'autre part, de la structure du marché déterminant entre fournisseurs et clients une véritable situation relevant de la théorie des jeux.

Pour conclure, on peut dire ceci :

D'une part, dans tous les cas de ce genre, il s'agit d'adapter le processus de décision à une réalité objective. La "personnalisation" des distributions de probabilité n'est pas elle-même le produit de dispositions psychologiques différentes des acteurs, mais de positions objectives différentes dans un processus social d'ensemble.

D'autre part, pour approcher la connaissance de la réalité objective à l'œuvre dans une situation décisionnelle, tous les moyens sont bons ; non seulement l'observation statistique, mais aussi l'analyse technique, économique, sociologique du processus dans lequel la décision doit venir prendre place.

3. Le cas du jeu de roulette

Le défaut des conceptions subjectivistes est donc double : d'une part, de méconnaître le poids réel des réalités objectives ; d'autre part, de survaloriser le rôle de l'observation statistique par rapport à celui de l'analyse de ces réalités. Sur le premier point, le subjectivisme est indéfendable ; sur le second, il peut faire valoir qu'il ne méconnaît pas l'intérêt d'une analyse de la réalité relevant d'autres moyens que l'observation statistique ; mais que, l'analyse étant faite, il faut bien retourner à l'observation statistique pour la meubler de chiffres précis attribuables aux éléments nouveaux qu'elle dégage.

Aussi bien ne peut-on l'incriminer sur ce point d'une sous-estimation *de principe*, mais *de fait.* Il n'est pas seul à s'en rendre coupable. L'exemple suivant va montrer de façon précise comment les conceptions usuelles du hasard peuvent détourner l'attention de ce qui est essentiel au profit de ce qui est secondaire.

L'exemple choisi est celui du jeu de roulette. La doctrine, classique jusqu'au début du XX^e siècle, attribuait l'équiprobabilité des résultats à la symétrie mécanique de la roulette. L'idée de rattacher l'équiprobabilité à la symétrie a été depuis lors très généralement critiquée, et quasiment abandonnée. Peut-être à tort, d'ailleurs, car les développements de la Physique du XX^e siècle ont rendu aux propriétés de symétrie un rôle de premier plan, dans un cadre élargi par la Théorie des Groupes ; et, dans les situations de Théorie des Jeux, il est évident que les probabilités qui définissent les stratégies optimales découlent des symétries de la règle du jeu ([25]).

Cependant, dans le cas du jeu de roulette, allèguer la symétrie mécanique de la roulette pour expliquer l'équiprobabilité masque un élément autrement important. En fait elle est elle-même garantie par l'intervention d'une multiplicité de personnages, dont chacun a son point de vue sur la question : le P.D.G. du casino (il a intérêt à déjouer les manœuvres d'éventuels "joueurs scientifiques", qui profiteraient d'un écart trop grand par rapport à l'équiprobabilité pour s'enrichir à son détriment) ; le technicien qui fabrique la roulette (il est dans la situation du fournisseur de l'exemple précédent, et doit soumettre son produit à des essais de réception) ; l'inspecteur de la brigade des jeux (il doit prévenir les tricheries). Le joueur peut *croire* à l'équiprobabilité des numéros, non seulement parce que la roulette "a l'air" symétrique, mais bien davantage parce qu'il *sait* qu'il vit dans une société policée. Jacques MONOD, dont la compréhension du hasard est loin d'être à la hauteur de son génie de biologiste, écrit quelque part qu'"à la roulette, le hasard est inessentiel", sous le prétexte que, si l'on connaissait avec une précision suffisante les conditions initiales du mouvement mécanique de la roulette et de la bille, on pourrait prédire le résultat avec certitude. Non seulement cet argument est peut-être purement et simplement faux (car il n'est pas sûr que les solutions des équations de la Mécanique soient dotées des propriétés de continuité requises pour sa validité) ; mais, bien davantage, il suffit de remarquer que le jeu de roulette n'existe que parce qu'il est un jeu de hasard. Si par aventure on y supprimait le hasard, on supprimerait le jeu du même coup (cf. les menaces que les ordinateurs font peser sur l'avenir du jeu d'échecs). Si le mot essentiel a un sens, il est difficile de faire plus essentiel.

S'il faut tirer une conclusion de ces remarques, c'est avant tout celle-ci :

Chaque personne appelée à prendre une décision dans une situation de hasard, le fait comme représentant d'un intérêt particulier, défini objectivement, dans le cadre d'une activité sociale, elle-même réglée par des contraintes objectives. Il est important que cet intérêt soit clairement identifié, et qu'il n'y ait pas de méprise à son sujet.

---

(25) On notera à ce sujet une manifestation de plus de la prégnance du point de vue. Il n'y a pas de symétrie *en soi.* En Physique quantique, les symétries sont relatives à des interactions données. Dans un domaine d'expérience plus quotidienne, un dé à jouer est cubique du point de vue du joueur ; il ne l'est pas du point de vue du peintre cubiste.

VII. MODELES STATISTIQUES DE DECISION ET RECHERCHE SCIENTIFIQUE

VII.1. Hypothèses scientifiques et actes de recherche

J'ai appelé l'attention, plus haut, sur la distinction entre démarches "heuristique" et "optimisante". Je voudrais maintenant aller un peu plus loin, et montrer que toute démarche heuristique s'accompagne de manière indissociable d'une démarche optimisante, que la recherche scientifique est une activité à deux faces, une face "cognitive" et une face "pratique", et que toute application que l'on y fait de modèles statistiques de décision relève, en dernière analyse, de sa face pratique, et d'une démarche optimisante. Et que c'est une illusion d'espérer l'"interpréter" en simples termes de progrès des connaissances.

Pour m'en expliquer clairement, je vais d'abord exposer un argument classique des subjectivistes (et des logicistes) contre les fréquentistes ; il pourra sembler qu'il est poussé jusqu'à la caricature, mais le trait est à peine forcé.

Imaginons une modélisation de l'expérience de MICHELSON en terme de théorie des tests. Nous avons face-à-face une hypothèse : "l'éther existe" et une contre-hypothèse : "l'éther n'existe pas, et la vitesse de la lumière est invariante" ; et notre problème est de décider quelle hypothèse est vraie. Si nous voulons maintenant passer à l'interprétation de l'expérience en termes de théorie statistique des décisions, nous nous trouvons devant un double problème.

D'une part, nous devons nous demander ce que va nous coûter l'adoption d'une hypothèse lorsque l'autre est vraie. Et, lorsque nous posons le problème en termes purement cognitifs, nous n'avons aucun moyen de répondre. D'autre part, nous devons nous demander quelles sont les probabilités *a priori* des hypothèses. Le subjectiviste croit alors triompher ; il serait effectivement idiot d'interprêter ces probabilités *a priori* en termes de fréquences : tant de fois sur cent l'éther existerait, et tant de fois sur cent la vitesse de la lumière serait invariante. Conclusion : la signification de la probabilité est purement subjective (26).

Mais les choses ne se passent pas, en recherche scientifique, aussi simplement en termes d'acceptation ou de rejet des hypothèses. Deux remarques au moins peuvent être faits à son sujet.

La première est que beaucoup de difficultés, insolubles lorsqu'on raisonne en termes de contenu de connaissance des hypothèses scientifiques, cessent de l'être lorsqu'on raisonne en termes d'*actes de recherche* (27). La recherche scientifique est elle-même une activité pratique ayant ses règles propres, où le chercheur, en même temps qu'il navigue entre la vérité et l'erreur, voit offrir à son choix des actes classables en *types d'actes* : publier un résultat, faire une autre expérience, renforcer sa documentation, arrêter une recherche etc. Il a lui-même ses propres spécialités : il est plus ou moins théoricien, expérimentateur, constructeur de matériel scientifique ; il est physicien, chimiste, mathématicien, biologiste, etc. Un type d'actes très important est le transfert du problème à d'autres spécia-

(26) J'y insiste encore, le trait de caricature est à peine forcé. En témoigne par exemple SAVAGE, lorsqu'il "met au défi" un fréquentiste de donner un sens à l'énoncé : "il est probable que le whisky fait plus de mal que de bien contre les morsures de serpent" ([3] p. 176).

(27) Poser le problème de la connaissance en termes d'actes de recherches a été une caractéristique intéressante d'un courant philosophique un peu dédaigné aujourd'hui, le pragmatisme (en particulier dans l'œuvre de Ch. PEIRCE). Il ne s'agit pas de "revenir au pragmatisme", mais de reconnaître l'intérêt de certaines de ses conceptions.

listes : un autre encore, une analyse critique en commun. Enfin, le résultat d'une expérience, ou plutôt d'une suite d'expériences scientifiques, peut être transféré à la sphère de la production industrielle ou agricole, pour y être appliqué à des besoins sociaux extra-scientifiques.

Le contenu d'une décision scientifique est tout autant, sinon plus, le choix d'un acte donné de pratique scientifique, que la décision d'adopter ou de rejeter une hypothèse scientifique. Il est clair que, si l'on considère les chose ainsi, le problème du coût des décisions commence à se poser dans des termes moins choquants.

Du même coup, disparaît la singularité de l'hypothèse scientifique. Un acte de recherche, en tant qu'il appartient à un type d'actes, est le produit d'une expérience accumulée ; on connaît plus ou moins bien, empiriquement, le genre de succès que l'on peut en attendre (et le genre d'échecs) et sa probabilité ([28]).

D'un autre côté, les notions même de vérité et d'erreur, dans la connaissance scientifique, ne peuvent se traiter en termes de logique pure, indépendamment de tout point de vue pratique. La connaissance scientifique est à tout jamais une connaissance *approchée*, jamais entièrement vraie. Tout statisticien sait parfaitement bien que toute hypothèse d'un test sera démontrée fausse, quels que soient les risques choisis, pour peu que l'effectif de l'échantillon soit assez grand.

Les façons qu'ont les hypothèses scientifiques d'être approchées sont infiniment variées ; l'approximation peut affecter aussi bien le modèle théorique que la mesure expérimentale ou le calcul mathématique. Ce qu'on demande à un ensemble de connaissances scientifiques, c'est de fournir des informations convenables pour un ensemble d'actes pratiques – dont les actes pratiques de la recherche scientifique elle-même, qui sont, au moins pour certains d'entre eux, les plus exigeants en précision.

S'introduit ici, pour une activité pratique donnée (de son point de vue), la distinction de ce qui est *bruit de fond* et de ce qui est *signal*. Cette activité pratique dépend d'une infinité de paramètres, allant des plus déterminants aux plus négligeables. Seuls, un nombre fini d'entre eux seront retenus dans les modèles de décision auxquels elle donnera lieu. Les autres seront confondus dans un bruit de fond dont on ne cherche pas à les extraire.

La question à régler est alors celle de la position de la frontière entre bruit de fond et signal. On sait que, dans les écoles d'ingénieurs, les professeurs enseignent aux élèves à ne pas formuler les résultats de leurs calculs avec trop de décimales, en général sans justifier cet enseignement autrement que dogmatiquement : "ce serait absurde". Mais il ne suffit pas d'allèguer les caractéristiques du traitement statistique du signal pour fournir la justification. La raison décisive est ailleurs, et on peut l'énoncer ainsi : c'est que, pour l'application pratique envisagée, le bénéfice de la précision supplémentaire ne payerait même pas l'encre, le temps et le papier qu'il faut pour écrire une décimale de plus, et la peine qu'on infligerait aux lecteurs pour en prendre connaissance ([29]).

Cet exemple appelle notre attention sur les actes de coût infime, apparemment négligeables, qui parsèment la vie quotidienne du chercheur, en quête de

---

(28) H. REICHENBACH, qui a tenté de systématiser l'approche "fréquentiste" de von MISES, a assez bien senti l'ambiguité de toute hypothèse scientifique (il existe non seulement des types d'actes, mais des types d'hypothèses). Mais il est resté prisonnier d'une approche purement cognitive, méconnaissant le point de vue pratique.

(29) La réduction (minime) du bruit de fond de la mesure a pour contre-partie son augmentation à la lecture.

confirmation ou de refutation d'une hypothèse scientifique. Il ne passera aux actes coûteux que lorsque les actes de coût négligeables ne lui apporteront plus rien qui justifie même leur coût infime. Il faudra bien alors qu'il évalue *en termes de coût monétaire* l'intérêt des actes coûteux : engager une longue recherche théorique, peut-être stérile ; une expérience scientifique lourde, aux résultats incertains ; la construction d'un appareil scientifique nouveau. Le chercheur, ou l'autorité qui prendra la décision sur sa proposition, doit alors évaluer au moins le type d'avantage qu'on peut en retirer, et l'ordre de grandeur de sa probabilité ([30]). Avantage toujours lié au passage d'une ou d'un groupe de caractéristiques du phénomène étudié du bruit de fond dans le signal.

La décision scientifique peut alors être posée en termes de rapport coût/efficacité.

Lorsqu'on raisonne en termes de vérité et d'erreur, il est difficile de comprendre pourquoi la décision en matière de recherche scientifique peut être empirique, et la frontière du bruit de fond et du signal relativement floue.

Mais si l'on raisonne en termes pratiques, si l'on comprend que l'un des problèmes du scientifique est en somme de hiérarchiser des actes de recherche dont les rapports coût-efficacité varient dans des proportions fantastiques ; que seuls un petit nombre d'entre eux peuvent appartenir au groupe de tête ; que l'on se trouve devant un problème d'optimisation et que par conséquent il n'y a rien de mortel, en général, si l'on arrive pas tout-à-fait à l'optimum – si l'on comprend cela, on comprend du même coup que la pratique scientifique puisse s'autoriser pour sa propre conduite un empirisme qu'elle cherche à éliminer pour les autres.

Il faut enfin, pour répondre à une dernière objection, faire une dernière remarque.

VII.2. Connaissance non formalisée et connaissance rationnelle

Beaucoup des traits gênants des conceptions classiques du hasard et de la Statistique me paraissent tenir à la séparation excessive, trop radicale, que l'on y instaure entre connaissance empirique et connaissance formalisée.

La connaissance n'est jamais ni totalement sans fondement rationnel, ni totalement rationnelle. Mais elle a toujours une fonction pratique.

Nous n'avons donc pas à viser une rationalité totale du raisonnement, mais une *rationalisation* toujours inachevée, et toujours susceptible de progrès, parce qu'elle renforce notre efficacité pratique.

Notre question devient alors : comment, sur quels points faut-il faire porter nos efforts de connaissance pour agir mieux ?

Le raisonnement statistique, considéré comme un processus mis en œuvre par un sujet donné, pour les fins de ce sujet (quelles qu'elles soient, et de quelque façon qu'elles soient déterminées) dans des circonstances données, peut réussir, ou ne pas réussir. C'est une question qui ne peut avoir de réponse qu'au long terme,

(30) Lorsqu'on lit des informations sur une entreprise scientifique gigantesque, comme par exemple la construction du LEP du CERN, on ne peut qu'être frappé par la *certitude* qu'ont les scientifiques d'obtenir des informations d'intérêt primordial, quelles qu'elles soient. D'autre part, il n'y a pas besoin de beaucoup réfléchir pour comprendre que l'intérêt social des progrès de la connaissance fondamentale en Physique est évalué sur la base de tout l'acquis du XX^{e} siècle. Il n'était pas évalué de la même façon il y a 20 ans, 40 ans, 60 ans.

comme une réponse à un problème de convergence. Peut-être devrions-nous nous intéresser plus intensément aux conditions objectives qui gouvernent la convergence des processus décisionnels.

BIBLIOGRAPHIE

La bibliographie concernant les problèmes abordés dans cette étude est extrêmement abondante. Je me bornerai à renvoyer à quelques ouvrages, où l'on peut trouver l'essentiel :

[1] *Encyclopédie de la Pléïade – Logique – Epistémologie.*

L'article de B. MATALON "Epistémologie des Probabilités" fournit un tableau bien équilibré des idées dans leur diversité.

Le point de vue subjectiviste est exprimé avant tout dans

[2] L. SAVAGE. – *The Foundations of Statistics* (Dover – Publications, 2^e éd. 1972).

D'autres développements intéressants peuvent être trouvés dans des articles de E. BOREL, F.P. RAMSAY, B. de FINETTI, SAVAGE et autres, rassemblés dans

[3] KYBURG et SMOKLER. – *Studies in subjective probability*, Willey. éd.

La question de l'utilité subjective, dans ses liaisons avec la théorie économétrique, est discutée dans le

[4] *Séminaire du CNRS sur la théorie du risque en économétrie* (1952).

Les contributions de la Théorie des Jeux sont exposées dans l'ouvrage fondamental.

[5] J. von NEUMANN et O. MORGENSTERN. – *Theory of Games and Economic Behavior,* Princeton University Press.

Le point de vue objectiviste est développé notamment dans

[6] R. von MISES. – *Wahrscheinlichkeit, Statistik und Wahrheit*, Springer – Verlag, 4^e éd. 1971.

[7] H. REICHENBACH. – *The theory of Probability*, University of California Press – Berkeley 1949.

Si l'on veut absolument prendre connaissance du point de vue logiciste, on peut le faire en lisant.

[8] K. POPPER. – *Logique de la découverte scientifique*, Payot, éd.
(Un bon tiers de ce gros ouvrage à succès concerne les Probabilités. Les idées de CARNAP, que POPPER se donne pour tâche de combattre, procèdent en fait de la même inspiration, à des détails près).

Enfin pour aborder la Théorie Statistique des Décisions, plutôt qu'aux ouvrages de vulgarisation, qui en accentuent excessivement le côté bayésien, le mieux est de se reporter à la source théorique :

A. WALD. – *Statistical décision functions,* WILEY éd. (pas très facile à lire) ou à l'excellent exposé, plus condensé, et très clair, dans

D.A.S. FRASER. – *Non parametric methods in Statistics*, Wiley éd.

N.B. : (Ces deux ouvrages demandent un minimum de connaissances de la Théorie mathématique de la Mesure, exposée dans tous les bons ouvrages modernes de Théorie des Probabilités).