

L'ÉTHIQUE DE LA MODÉLISATION DANS UN MONDE OÙ LA NORMALITÉ N'EXISTE PLUS

Arthur Charpentier

Professeur, Université de Rennes 1

Le mécanisme de couverture des catastrophes naturelles a été créé pour indemniser « les dommages matériels directs non assurables ayant eu pour cause déterminante l'intensité anormale d'un agent naturel » (article L. 125-1 alinéa 3 du Code des assurances). Toujours sur le plan légal, la Cour de cassation a formulé, en novembre 1986, un principe selon lequel « nul ne doit causer à autrui un trouble anormal de voisinage ». Et pour pouvoir prétendre à une indemnisation suite à une détention provisoire, il est nécessaire que le demandeur apporte la preuve que la détention lui a causé un « préjudice manifestement anormal et d'une particulière gravité » (article 149 du Code de procédure civile). Mais que signifie cette « anormalité » présente dans tous ces articles ? Selon le dictionnaire, l'anormalité est définie comme « contraire à l'ordre habituel des choses » (on pourrait y voir une notion empirique, statistique), « contraire à l'ordre juste des choses » (cette notion de « juste » appelle probablement à une définition normative) mais aussi « non conforme au modèle ». Définir une norme n'est déjà pas simple si on ne s'intéresse qu'à l'aspect descriptif, empirique, comme peuvent le faire les actuaires lorsqu'ils construisent un modèle (en particulier en grande dimension, où comme nous le verrons la normalité n'existe plus), mais si on intègre en plus une dimension de justice et d'éthique, on se demande si la tâche n'est pas impossible...

L'homme moyen de Quetelet et de Galton

Au XIX^e siècle, si plusieurs astronomes mesuraient la vitesse d'un même objet céleste, ils obtenaient (souvent) plusieurs mesures différentes. Pour savoir laquelle utiliser dans leurs calculs, l'idée d'utiliser « la méthode

des moyennes » s'est rapidement imposée – comme le rappelle Stahl [2006], et surtout Sheynin [1973] – cette moyenne ayant une précision plus grande que n'importe quelle autre grandeur (ou dirait aujourd'hui statistique). À partir d'un ensemble d'observations, on pose

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

On peut noter que cette grandeur est aussi la solution du problème d'optimisation

$$\bar{x} = \underset{m}{\operatorname{argmin}} \left\{ \sum_{i=1}^n [x_i - m]^2 \right\}$$

qui montre l'importance des « moindres carrés ». Adolphe Quételet fut, semble-t-il, le premier à appliquer ce calcul de moyennes à des mesures humaines, introduisant son fameux concept d'« homme moyen ». Si on définit la moyenne à l'aide d'une minimisation d'erreur quadratique, on a une interprétation en termes de prévision : la taille moyenne est la taille que devrait mesurer une personne tirée au hasard (à une variation aléatoire – et imprévisible – près). En 1846, dans une [lettre](#) (1) Adolphe Quételet utilise l'image de la statue du gladiateur pour expliquer ce que peut être l'homme moyen : « Supposons qu'on ait employé un millier de statuaires pour copier le gladiateur avec tout le soin imaginable. Votre Altesse ne pense certainement pas que les mille copies qui auront été faites, reproduiront chacune exactement le modèle, et qu'en les mesurant successivement, les mille mesures que j'obtiendrais seraient aussi concordantes que si je les avais prises toutes sur la statue du gladiateur même. Aux premières chances d'erreur viendraient se joindre les inexactitudes des copistes ; en sorte que l'erreur probable serait peut-être très grande. Malgré cela, si les copistes n'ont pas travaillé avec des idées préconçues, en exagérant ou en diminuant certaines proportions d'après des préjugés d'école, et si leurs inexactitudes ne sont qu'accidentelles, les mille mesures, groupées par ordre de grandeur, présenteront encore une régularité remarquable et se succéderont dans l'ordre que leur assigne la loi de possibilité. Je vois sourire Votre Altesse ; elle me dira sans doute que de pareilles assertions ne me compromettront pas, attendu qu'on ne sera pas disposé à tenter l'expérience. Et pourquoi pas ? Je vais peut-être bien l'étonner, en disant que l'expérience est toute faite. Oui vraiment, on a mesuré plus d'un millier de copies d'une statue que je n'assurerai pas être celle du gladiateur, mais qui, en tout cas, s'en éloigne peu : ces copies étaient même vivantes, en sorte que les mesures ont été prises avec toutes les chances d'erreur possible : j'ajouterai,

de plus, que les copies ont pu se déformer par une foule de causes accidentelles. On doit donc s'attendre, ici, à trouver une erreur probable très sensible. »

Cet homme moyen a beaucoup plu à l'époque, en particulier au sein de l'école anglaise eugéniste, dirigée par Francis Galton, même si ce dernier s'intéresse surtout aux déviations par rapport à cette norme (déviation vers le haut et déviation vers le bas). Comme le rappelle Bulmer [2004], « *the deviations from that average – upwards towards genius, and downwards towards stupidity – must follow the law that governs deviations from all true averages* ». Les travaux de Galton ont visé à comprendre ces déviations. Si Florence Nightingale affirmait que « *the average man is God's will* » (2), Galton de son côté s'intéressait davantage au caractère héréditaire de la déviation qu'à la moyenne. Mais cet homme moyen a-t-il pour autant du sens ?

■ À la recherche de la personne « moyenne »

Rose [2016] présente deux exemples dans son ouvrage *The End of Average*. Le premier est tiré de problèmes rencontrés par l'armée américaine dans les années 1950. Lors de la conception des postes de pilotage des avions de chasse, des ingénieurs avaient utilisé les dimensions de plus de 4 000 pilotes pour positionner de manière optimale le siège par rapport aux pédales, le manche à balai, la hauteur du pare-brise, mais aussi la forme du siège, du casque, etc. Ces mesures ont permis de calculer les mensurations du pilote « médian » dans une dizaine de dimensions. Par exemple la taille moyenne des pilotes était de 179 cm, ce qui a permis de définir la taille d'un pilote moyen entre 175 et 185 cm. Si une majorité des pilotes est de taille moyenne, parmi les 4 000 pilotes, aucun n'était « moyen » dans toutes les dimensions. Comme l'affirmait Daniels [1952], « concevoir un poste de pilotage pour le pilote moyen revenait en fait à n'en concevoir pour aucun ».

Le second exemple est lié à deux statues, celles de *Norma* et *Normann* (exposées historiquement à

Cleveland, aujourd'hui conservées à la bibliothèque de Harvard). L'artiste Abram Belskie et l'obstétricien Robert Latou Dickinson ont réalisé ensemble ces statues, en 1943. Leur particularité est qu'aucun modèle n'a été représenté. En fait, il s'agissait de représenter une femme et un homme qui avaient les mensurations moyennes de l'époque (à partir de mesures effectuées sur des milliers de sujets). Une fois ces statues réalisées, un concours a été organisé pour trouver qui ces statues pouvaient bien représenter. Plusieurs milliers de personnes de l'Ohio ont envoyé leurs mensurations, mais aucune ne correspondait à celles des statues. Certes, plusieurs centaines avaient la même taille. Plusieurs centaines avaient le même tour de poitrine. Mais aucune n'avait toutes les bonnes mesures. Car comme l'explique Todd Rose, l'homme n'est pas unidimensionnel : on le mesure en plusieurs dimensions. Et chercher à le résumer en une grandeur unidimensionnelle est beaucoup trop réducteur. C'est ce qu'il montre dans son livre sur les tests d'intelligence, par exemple, où un même QI peut être associé à deux personnes très différentes. Pareil pour décider de recruter quelqu'un, se focaliser sur un seul indicateur n'a pas de sens. Le souci quand on travaille dans un contexte multivarié, c'est que la moyenne perd de son sens. En fait, d'un point de vue probabiliste, être moyen peut être extraordinaire.

■ La malédiction de la dimension

En fait, ce problème est bien connu des statisticiens, sous le nom de « fléau de la dimension ». Prenons un exemple simple : supposons qu'une grandeur d'intérêt suive une loi normale (μ, σ^2), par exemple le poids, la taille, le tour de poitrine, etc. On pourrait dire qu'être dans la norme, c'est se trouver dans un intervalle $[\mu \pm 1,5\sigma]$. Si on a une loi normale, cette situation survient dans 85 % des cas. Et les 15 % qui ne sont pas dans cet intervalle seront vus comme « anormaux ». Les grandeurs peuvent être anormalement petites, ou anormalement grandes. On peut également regarder deux dimensions, le poids et la taille, par exemple. La norme serait que dans les deux dimensions, on soit dans l'intervalle

$[\mu \pm 1,5\sigma]$. Si les grandeurs sont indépendantes, la probabilité que les deux grandeurs soient « normales » est de 75 % $P[X_1 \in [\mu_1 \pm 1,5\sigma_1^2], X_2 \in [\mu_2 \pm 1,5\sigma_2^2]] = P[X \in [\mu \pm 1,5\sigma^2]]^2 = 0,85^2 \sim 0,7225$. Autrement dit, en dimension deux, 75 % des observations sont globalement normales, et 25 % sont alors anormales. En dimension 3, on passe à 65 %, soit plus du tiers d'observations anormales. En dimension cinq, on passe en dessous de 50 %, autrement dit, être dans la norme dans les cinq dimensions n'est plus le cas de la majorité. Et en dimension vingt, ceux qui sont normaux sont plutôt atypiques, avec une proportion de l'ordre de 5 %. Ainsi, en grande dimension, la normalité n'est plus aucunement associée à l'idée d'une majorité. Ce problème est celui que rencontrent aujourd'hui les actuaires lorsqu'ils utilisent des données en très grande dimension, en tarification par exemple : il devient très difficile de caractériser une classe tarifaire (en disant à quoi ressemble l'assuré moyen dans cette classe).

Normalité, statistiques et normes

■ De la normalité statistique à la norme, de la régularité à la règle

D'un point de vue empirique, descriptif, être dans la norme, ne signifie rien d'autre qu'être dans la moyenne, ne pas se détacher trop de cette moyenne. On aura alors tendance à définir la norme comme la fréquence de ce qui se produit le plus souvent, comme l'attitude la plus fréquemment rencontrée ou la préférence la plus régulièrement manifestée. Mais cette normalité n'est pas la normativité, et « être dans la norme », être exemplaire, relève alors d'une dimension différente, qui se rattache cette fois non plus à une description du réel mais à une identification de ce vers quoi il doit tendre. On passe donc du registre de l'être à celui du devoir-être, du « *is* » au « *ought* » pour reprendre la terminologie de Hume [1739]. Il est en effet difficile d'envisager le modèle (ou la

normalité) sans glisser vers ce second sens qu'on peut trouver au concept de norme, et qui déploie quant à lui une dimension à proprement parler normative. Cette vision mène à une confusion entre normes et lois, même si toute la normativité n'est pas épuisée par les lois. Hume constate ainsi que, dans tous les systèmes de morale, les auteurs passent de constats de faits, c'est-à-dire énonciatifs de type « il y a », à des propositions qui comportent une expression normative, comme « il faut », « on doit ». Ce que Hume conteste, c'est le passage d'un type d'affirmation à un autre : pour lui, ce sont là deux types d'énoncés qui n'ont rien à voir les uns avec les autres, et qu'on ne peut donc pas enchaîner logiquement les uns avec les autres, en particulier d'une norme empirique vers une règle normative. Pour Hume, une affirmation qui ne serait pas normative ne peut pas donner lieu à une conclusion de type normatif. Cette affirmation de Hume a suscité nombre de commentaires et d'interprétations, en particulier parce qu'énoncée telle quelle, elle semble être un obstacle à toute tentative de naturalisation de la morale – comme le détaillent McIntyre [1959] ou Rescher [1990]. En ce sens, on trouve la distinction forte entre la norme dans la régularité (normalité) et la règle (normativité).

■ Les lois statistiques, du micro au macro

La loi statistique relève de ce qui « est » parce qu'on l'a observé ainsi (par exemple, « Les hommes sont plus grands que les chiens. »). La loi humaine (divine ou judiciaire) relève de ce qui « est » parce qu'on l'a décrété, et donc « doit être » (« Les hommes sont libres et égaux. » ou « L'homme est bon. »). Enfin, la loi physique relève de ce qui « est » parce qu'on peut le montrer (« Les planètes sont attirées entre elles. »), souvent dans le cadre d'hypothèses. On voit que les trois concepts peuvent être liés. Par exemple, la loi de Kepler a été établie historiquement à l'aide d'observations (et entrainé historiquement dans la première catégorie), avant d'être démontrée dans le cadre du modèle copernicien (et de passer alors dans la troisième). On peut d'ailleurs associer un concept d'équilibre à cette loi, cette « norme ».

Toutefois, comme le souligne Hilpinen [1971], les lois probabilistes posent toutefois bon nombre de questions, il suffit de penser aux lancers de dés ou aux attentes : que signifie « il est normal d'attendre cinq minutes le bus à l'arrêt », ou plus dérangeant d'un point de vue éthique, « il est normal qu'une personne placée en détention provisoire soit emprisonnée dix-huit mois » ?

On peut voir la norme comme une régularité de cas, observée à l'aide de fréquences (ou de moyennes), par exemple, sur la taille des individus, la durée du sommeil, autrement dit des données qui constituent la description d'individus. Les données anthropométriques ont ainsi permis de définir une taille moyenne des individus dans une population donnée, en fonction de leur âge ; par rapport à cette taille moyenne, un écart de 20 % en plus ou en moins détermine le gigantisme ou le nanisme. Si l'on pense aux accidents de la route, il peut être considéré comme anormal d'avoir un accident de la route une année donnée, à un niveau individuel (micro), car la majorité des conducteurs n'ont pas d'accident. Néanmoins, du point de vue de l'assureur (macro), la norme est que 10 % des conducteurs aient un accident. Il serait donc anormal que personne n'ait d'accident. C'est l'argument que l'on retrouve dans Durkheim [1897]. De l'acte singulier qu'est le suicide, s'il est considéré du point de vue de l'individu qui le commet, Durkheim tente de le voir comme un acte social, relevant alors d'une réelle régularité, au sein d'une société donnée. Dès lors, le suicide devient, selon Durkheim, un phénomène normal. Les statistiques permettent alors de quantifier la tendance au suicide dans une société donnée, dès lors que l'on observe non plus l'irrégularité qui apparaît dans la singularité d'une histoire individuelle, mais une normalité sociale du suicide.

■ Norme, convention et aspects éthiques

Si on prend un point de vue évolutionniste, ce qui est normal est ce qui est le plus capable de s'adapter, de répondre aux besoins, de fournir un modèle pour

la résolution des situations (la nature faisant disparaître l'anormalité), et la normalité tend vers la normativité, et il devient difficile de faire la distinction entre les deux aspects. En fait, David Hume aborde ce point dans l'exemple bien connu des rameurs, qui montent dans le même bateau pour traverser un fleuve et qui rament en cadence (cet exemple est longuement discuté dans Mackie [1980]). Les deux rameurs ajustent peu à peu leurs coups de rame, l'un par rapport à l'autre, et il n'est pas nécessaire d'obtenir un accord explicite (qui formulerait la norme) qu'ils respecteraient. La loi, qui consiste à imposer une norme peut être utile en cas de conflit (si un des rameurs refuse de ramer, ou deux rameurs de capacités physiques très différentes), mais bien souvent, il n'est pas nécessaire de formuler explicitement cette norme inhérente à leur conduite. L'observateur extérieur observera une régularité (lorsque le rythme de croisière sera atteinte) qu'il pourra modéliser, mais ce rythme normal observé n'est pas forcément imposé par une loi. Dans le cas de rameurs, on retrouve la notion d'équilibre évoquée précédemment. Construire un modèle, c'est extraire le signal du bruit (pour reprendre la distinction de Silver [2015]), c'est chercher une norme, au sens statistique. Mais cela va plus loin s'il l'on construit un modèle prédictif, la réalité devant alors se conformer au modèle, comme l'espèrent souvent les actuaires.

Notes

1. *Publiée dans l'ouvrage* Lettres sur la théorie des probabilités, appliquée aux sciences morales, 1846.
2. *La valeur moyenne relèverait d'une volonté divine.*

Bibliographie

- BULMER M., *Francis Galton: Pioneer of Heredity and Biometry*. Johns Hopkins University Press, 2004.
- DANIELS G., "The Average Man", *Air Force Aerospace Medical Research Lab*, vol. 53, n° 7, 1952.
- DURKHEIM E., *Le suicide*, 1897.
- HILPINEN R., *Deontic Logic: Introductory and Systematic Readings*, Dordrecht, D. Reidel Publishing Company, 1971.
- HUME D., *Traité de la nature humaine. Tome III : de la morale*, 1739.
- MCINTYRE D.C., "Hume on 'Is' and 'Ought'", *The Philosophical Review*, vol. 68, n° 4, Duke University Press, 1959, pp. 451-468.
- MACKIE J.L., *Hume's Moral Theory*, Routledge & Kegan Paul Books, 1980.
- RESCHER N., "How Wide Is the Gap Between Facts and Values?", *Philosophy and Phenomenological Research*, vol. 50, 1990, pp. 297-319.
- ROSE T., *The End of Average: How We Succeed in a World That Values Sameness*, HarperOne, 2016.
- SHEYNIN O., "Mathematical Treatment of Astronomical Observations (A Historical Essay)". *Archive for History of Exact Sciences*, vol. 11, 1973, pp. 97-126.
- SILVER N., *The Signal and the Noise: Why So Many Predictions Fail – But Some Don't*, Penguin Press, 2015.
- STAHL S., "The Evolution of the Normal Distribution", *Mathematics Magazine*, vol. 79, 2006, pp. 96-113.