

Mathématiques et Epidémiologie

(article proposé à la Jaune et la Rouge, numéro de novembre 2011)

par Bernard Beauzamy
PDG, Société de Calcul Mathématique SA

Tout lecteur, parcourant ce numéro, en ressortira avec la conviction que l'épidémiologie repose sur des mathématiques. Les conclusions dépendent de tests statistiques, et, pour alimenter ces tests, il faut des données, qui obéissent normalement aux lois de la statistique. L'épidémiologie utilise aussi des mathématiques pour modéliser la propagation d'une épidémie, par exemple sous la forme d'équations différentielles.

Tout ceci paraît de très haut niveau. Mais ce n'est qu'apparence : l'épidémiologie fait bien appel à ces outils, mais tant leur conception que leur utilisation feront hurler de rage n'importe quel mathématicien professionnel, s'il a le malheur de les découvrir comme cela a été mon cas.

Les mathématiques sont la discipline la plus ancienne (environ 6 000 ans), mais en même temps elle compte très peu de professionnels, et ils sont très récents (60 ans). Comme discipline, elle a engendré des métiers, dont le plus vieux métier du monde ; c'est en comptant ses moutons, et en constatant qu'il lui en manquait, que le berger a dit à sa femme : "essaie de nous rapporter un peu d'argent". Mais un professionnel est celui qui vit de sa prestation ; ce n'est pas le cas des enseignants-chercheurs, qui sont fonctionnaires. Après la seconde guerre mondiale, sont apparues des sociétés commerciales proposant des prestations mathématiques (souvent d'optimisation), et la SCM s'inscrit dans cette lignée.

Un professionnel, bien entendu, est attaché à la qualité du résultat : celui-ci doit être prédictif, explicatif, à l'égard des lois de la Nature. Nous avons une obligation de résultat. C'est la Nature qui nous juge ; pour nous, l'appartenance à des sociétés savantes, les honneurs, le jugement par les pairs, sont des mots vains. Les mathématiques sont non-brevetables : nous exposons nos démonstrations et nos résultats à la vue de tous. C'est de là que doit provenir la qualité : il faut montrer que tous les cas ont été considérés, que toutes les déductions sont correctes, etc. Il faut tout mettre au grand jour, sans jamais se réfugier derrière un argument d'autorité.

Ma découverte de l'épidémiologie s'est faite progressivement à partir de 2004 : j'ai été invité par le CEA, puis par l'IRSN, puis par RTE, à donner mon avis de professionnel à propos de certaines études épidémiologiques ; il s'agissait de l'action des faibles doses de radioactivité sur la santé, ou encore des maladies susceptibles d'apparaître au voisinage des lignes à haute tension.

Et, immédiatement, cela a été l'horreur que j'évoquais plus haut. Il ne m'a pas fallu dix minutes pour m'apercevoir que l'ensemble des études que j'avais à expertiser reposait sur des outils mathématiques totalement factices, fabriqués par des non-professionnels, pour la circonstance, à partir de connaissances de statistiques mal digérées. Comme disait Boileau "et d'un tronc fort illustre une branche pourrie".

A quoi donc cela tient-il ? C'est fort simple : les experts du domaine se sont érigés en caste, et ont déclaré qu'eux seuls avaient le droit de porter cette dignité. Si vous vous approchez, on vous déclarera que vous n'êtes pas épidémiologiste, pas même médecin.

Bien des professions ont érigé de pareilles barrières : les astrologues, les prêtres, etc. Il faut, pour y appartenir, une consécration interne.

Dans les exemples que nous avons traités, l'absurdité était flagrante dès la première ligne. Une étude affirme, par exemple, que la probabilité de développer la maladie d'Alzheimer est plus grande au voisinages de lignes à haute tension qu'au loin. Tout ceci s'appuie sur des tests statistiques en apparence fort savants. Mais prenons les chiffres bruts, tels que l'article les publie en première page : il y a moins d'Alzheimer, en proportion de la population, à proximité des lignes qu'au loin ! Il faut, dans ces conditions, des outils bien complexes pour parvenir à la conclusion inverse.

Une mode d'investigation souvent utilisé par les épidémiologistes est l'étude "cas témoin". Un groupe de personnes atteintes d'une maladie (par exemple Alzheimer) est comparé à un "groupe témoin". Voici une sorte de définition :

"Les études cas-témoins sont basées sur la comparaison de deux groupes, le premier, constitué des "cas" : individus atteints de la maladie étudiée ; l'autre, les "témoins", est choisi pour qu'il soit comparable au premier pour toutes les caractéristiques (par exemple âge, sexe, poids), à l'exception de celle qui est étudiée."

Dans l'introduction d'une étude sur les leucémies au voisinage des lignes à Haute Tension, on lit :

"29 081 enfants atteints de cancer, incluant 9700 atteints de leucémie. Enfants de moins de 14 ans et nés en Angleterre et au Pays de Galles entre 1962 et 1995. Les sujets témoins étaient individuellement appariés sur le sexe, la date de naissance approximative et la zone d'enregistrement de la naissance."

Autrement dit, les cas-témoin sont choisis au hasard. Mais un probabiliste demande immédiatement : "selon quelle loi ?". C'est là que le bât blesse, parce qu'un épidémiologiste ne sait pas ce qu'est une loi de probabilité, tandis que pour un probabiliste professionnel, le mot a un sens très clair. Pour l'épidémiologiste, "au hasard" signifie : n'importe comment dans l'annuaire du téléphone. C'est là qu'on voit l'absurdité du raisonnement, car enfin ce qu'on cherche, c'est précisément une loi de probabilité, celle de l'apparition de la leucémie sachant qu'on est proche des lignes (loi conditionnelle) à comparer avec la même loi, sachant qu'on est loin. Or ces lois, par définition, sont inconnues ! On va donc comparer une loi inconnue avec une loi faite "au pif", à partir de l'annuaire du téléphone. Tout ceci, bien entendu, est dissimulé sous un abondant jargon scientifique. On note aussi, dans l'étude en question, que la zone de résidence est assimilée à la zone d'enregistrement de la naissance, mais là on sort de l'outrage mathématique.

Que peut-on y faire ? A peu près rien ; quoi que l'on puisse dire, de toute façon les épidémiologistes n'écouteront pas un mathématicien ; ils sont complètement convaincus de connaître suffisamment de mathématiques. Pour eux, le cœur du problème n'est pas dans le traitement statistique qui est fait ; ils appuient sur un quelconque bouton d'un quelconque logiciel, et un résultat sort. Si le résultat n'est pas conforme à leurs attentes, ils recommencent avec un autre bouton, puis avec un autre logiciel, jusqu'à avoir le résultat attendu, celui qui montre un vrai péril. Les journalistes le publieront à l'envi, et les politiques légifèreront à qui mieux mieux. C'est ainsi que s'est propagée la grippe H1N1, et toutes sortes de maladies avant elles.

Faire émerger les bonnes pratiques

Après avoir réalisé le travail que l'on me demandait, c'est-à-dire montré que les études épidémiologiques que l'on m'avait soumises étaient dépourvues de sens, j'ai tenté d'alerter les responsables des différents organismes, en leur disant ceci : vous auriez intérêt à faire "front commun", en faisant émerger de bonnes pratiques.

Des pratiques communes, partagées d'abord entre organismes compétents, puis étendues internationalement, parviendraient à créer un socle commun, opposable aux absurdités qui fleurissent un peu partout. A l'inverse, tant que ces bonnes pratiques n'existent pas, chacun a beau jeu de présenter les siennes comme excellentes.

J'ai complètement échoué dans ma démarche. A l'exception de Alain Bugat, à l'époque administrateur général du CEA, qui s'est montré intéressé, les autres responsables n'ont rien fait ; bien au contraire, ils se sont opposés à toute démarche commune. A RTE, par exemple, le responsable de ces questions a déclaré qu'il ne fallait pas mélanger, dans l'esprit du public, les dangers causés par les champs électromagnétiques avec ceux causés par les radiations, comme si la définition correcte de lois de probabilité pouvait mettre la population en péril.

Lorsque j'ai proposé de mettre sur notre site web, sous ma responsabilité, les conclusions auxquelles j'étais parvenu, ce même responsable s'y est opposé, de crainte de "créer un buzz sur Internet". Comme si ce "buzz" n'existait pas déjà, comme si Internet ne regorgeait pas déjà de commentaires désobligeants, aussi bien sur les ondes que sur les radiations et sur tous les organismes qui produisent de l'électricité ou la transportent. Le simple fait de se défendre est une faute aujourd'hui. C'est la politique de l'autruche, chaque petite autruche ayant sa petite tête dans son petit trou.

Comme disait Von Neumann, le mathématicien n'est pas responsable du monde tel qu'il est. Ma propre contribution au débat se limitera donc à ceci : avoir proposé ce numéro à thème pour "La Jaune et la Rouge", en espérant qu'il se trouvera encore parmi les lecteurs un petit nombre que l'obscurantisme et la pusillanimité ambiants aura épargnés.

Version revue par Hubert Jacquet, rédacteur en chef de la Jaune et la Rouge,
le "Patouillet" du 21^{ème} siècle :

Parce que vous bâclez un journal de bigotes
Pensé par Escobar, écrit par Patouillet ;

(Victor Hugo, Les Châtiments)

Mieux **utiliser** les outils **mathématiques**

par Bernard Beauzamy (68) PDG de la Société de calcul mathématique (SCM)

L'épidémiologie fait largement appel aux outils mathématiques, aussi bien pour les statistiques que la modélisation. En pratique, l'emploi de ces outils conduit souvent à des résultats faux ou contestables, faute d'une discipline et d'une méthodologie de mise en œuvre rigoureuse. Le développement de bonnes pratiques, partagées par des organismes compétent, est indispensable pour sortir de ces errements.

Repères

Les mathématiques sont une discipline très ancienne (environ 6 000 ans), mais elle compte très peu de professionnels, et ils sont très récents (60 ans). Comme discipline, elle a engendré de nombreux métiers. Mais les professionnels sont des spécialistes de l'application des mathématiques dans des domaines les plus divers. Ils vivent de leurs prestations ce qui n'est pas le cas des enseignants-chercheurs, qui sont fonctionnaires. Après la seconde guerre mondiale, sont apparues des sociétés commerciales proposant des prestations mathématiques.

Il est communément admis que l'épidémiologie repose sur des mathématiques. Les conclusions dépendent de tests statistiques, et, pour alimenter ces tests, il faut des données, qui obéissent normalement aux lois de la statistique. L'épidémiologie utilise aussi des mathématiques pour modéliser la propagation d'une épidémie, par exemple sous la forme d'équations différentielles. Tout ceci paraît de très haut niveau. Mais ce n'est qu'apparence : l'épidémiologie fait bien appel à ces outils, mais tant leur conception que leur utilisation feront sursauter n'importe quel mathématicien professionnel.

Obligation de résultat

Un professionnel, bien entendu, est attaché à la qualité du résultat : celui-ci doit être prédictif, explicatif, à l'égard des lois de la Nature. Nous avons une obligation de résultat. C'est la Nature qui nous juge ; pour nous, l'appartenance à des sociétés savantes, les honneurs, le jugement par les pairs, sont des mots vains. Les mathématiques sont non-brevetables : nous exposons nos démonstrations et nos résultats à la vue de tous. C'est de là que doit provenir la qualité : il faut montrer que tous les cas ont été considérés, que toutes les déductions sont correctes, etc. Il faut tout mettre au grand jour, sans jamais se réfugier derrière un argument d'autorité.

Outils factices

Ma découverte de l'épidémiologie s'est faite progressivement à partir de 2004 : j'ai été invité par le CEA, puis par l'IRSN, puis par RTE, à donner mon avis de professionnel à propos de certaines études épidémiologiques ; il s'agissait de l'action des faibles doses de radioactivité sur la santé, ou encore des maladies susceptibles d'apparaître au voisinage des lignes à haute tension. Il ne m'a pas fallu dix minutes pour m'apercevoir que l'ensemble des études que j'avais à expertiser reposait sur des outils mathématiques totalement factices, fabriqués par des non-professionnels, pour la circonstance, à partir de connaissances de statistiques mal digérées.

Esprit de caste

A quoi donc cela tient-il ? C'est fort simple : les experts du domaine se sont érigés en caste, et ont déclaré qu'eux seuls avaient le droit de porter cette dignité. Bien des professions ont érigé de pareilles barrières : les astrologues, les prêtres, etc. Il faut, pour y appartenir, une consécration interne.

Absurdité flagrante

Une étude affirme que la probabilité de développer la maladie d'Alzheimer est plus grande au voisinage de lignes à haute tension qu'au loin. Tout ceci s'appuie sur des tests statistiques en apparence fort savants. Mais prenons les chiffres bruts, tels que publiés par cette étude elle-même: il y a moins d'Alzheimer, en proportion de la population, à proximité des lignes qu'au loin ! Il faut, dans ces conditions, des outils bien complexes pour parvenir à la conclusion inverse.

Cas témoins

Une mode d'investigation souvent utilisé par les épidémiologistes est l'étude "cas témoin". Un groupe de personnes atteintes d'une maladie (par exemple Alzheimer) est comparé à un "groupe

témoin". Voici une sorte de définition : "Les études cas-témoins sont basées sur la comparaison de deux groupes, le premier, constitué des "cas" : individus atteints de la maladie étudiée ; l'autre, les "témoins", est choisi pour qu'il soit comparable au premier pour toutes les caractéristiques (par exemple âge, sexe, poids), à l'exception de celle qui est étudiée."

Dans l'introduction d'une étude sur les leucémies au voisinage des lignes à Haute Tension, on lit : "29 081 enfants atteints de cancer, incluant 9700 atteints de leucémie. Enfants de moins de 14 ans et nés en Angleterre et au Pays de Galles entre 1962 et 1995. Les sujets témoins étaient individuellement appariés sur le sexe, la date de naissance approximative et la zone d'enregistrement de la naissance."

Lois de probabilité

Autrement dit, les cas-témoin sont choisis au hasard. Mais un probabiliste demande immédiatement: "selon quelle loi ?". C'est là que le bât blesse, parce qu'un profane ne sait pas ce qu'est une loi de probabilité, tandis que pour un probabiliste professionnel, le mot a un sens très clair. Pour un non professionnel, "au hasard" signifie : n'importe comment dans l'annuaire du téléphone. C'est là qu'on voit l'absurdité du raisonnement, car enfin ce qu'on cherche, c'est précisément une loi de probabilité, celle de l'apparition de la leucémie sachant qu'on est proche des lignes (loi conditionnelle) à comparer avec la même loi, sachant qu'on est loin. Or ces lois, par définition, sont inconnues ! On va donc comparer une loi inconnue avec une loi faite "au pif", à partir de l'annuaire du téléphone. Tout ceci, bien entendu, est dissimulé sous un abondant jargon scientifique. On note aussi, dans l'étude en question, que la zone de résidence est assimilée à la zone d'enregistrement de la naissance, mais là on sort de l'outrage mathématique.

Dérives scientifiques

Les épidémiologistes ont parfois du mal à écouter les mathématiciens : ils sont complètement convaincus de connaître suffisamment de mathématiques. Pour eux, le cœur du problème n'est pas dans le traitement statistique qui est fait; ils appuient sur un quelconque bouton d'un quelconque logiciel, et un résultat sort. Si le résultat n'est pas conforme à leurs attentes, ils recommencent avec un autre bouton, puis avec un autre logiciel, jusqu'à avoir le résultat attendu, celui qui montre un vrai péril. Les journalistes le publieront à l'envi, et les politiques légifèreront à qui mieux mieux. C'est ainsi que s'est propagée la grippe H1N1, et toutes sortes de maladies avant elles.

Faire émerger les bonnes pratiques

Après avoir réalisé le travail que l'on me demandait, c'est-à-dire montré que les études épidémiologiques que l'on m'avait soumises étaient dépourvues de sens, j'ai tenté d'alerter les responsables des différents organismes, en leur disant ceci : vous auriez intérêt à faire "front commun", en faisant émerger de bonnes pratiques. Des pratiques communes, partagées d'abord entre organismes compétents, puis étendues internationalement, parviendraient à créer un socle commun, opposable aux absurdités qui fleurissent un peu partout. A l'inverse, tant que ces bonnes pratiques n'existent pas, chacun a beau jeu de présenter les siennes comme excellentes.

J'ai malheureusement échoué dans ma démarche. Un Administrateur général du CEA s'est montré intéressé . D' autres responsables n'ont rien fait ; bien au contraire, ils se sont opposés à toute démarche commune, déclarant parfois qu'il ne fallait pas mélanger, dans l'esprit du public, les dangers causés par les champs électromagnétiques avec ceux causés par les radiations, comme si la définition correcte de lois de probabilité pouvait mettre la population en péril.

Craintes du buzz

Lorsque j'ai proposé de mettre sur notre site web, sous ma responsabilité, les conclusions auxquelles j'étais parvenu sur l'usage des outils mathématiques en épidémiologie, un des responsables consulté s'y est opposé, de crainte de "créer un buzz sur Internet". Comme si ce "buzz" n'existait pas déjà, comme si Internet ne regorgeait pas déjà de commentaires désobligeants, aussi bien sur les ondes que sur les radiations et sur tous les organismes qui produisent de l'électricité ou la transportent. Le simple fait de se défendre est une faute aujourd'hui. C'est la politique de l'autruche, chaque petite autruche ayant sa petite tête dans son petit trou.

Comme disait Von Neumann, le mathématicien n'est pas responsable du monde tel qu'il est. J'espère qu'il se trouvera -parmi les lecteurs un petit nombre que l'obscurantisme ambiant aura épargnés.