

Les paradoxes de la segmentation et de la discrimination en assurance



**ARTHUR
CHARPENTIER**
Professeur, Université du Québec à Montréal (UQAM)

Arthur Charpentier présente, dans cet article documenté, les tensions entre segmentation actuarielle et discrimination en assurance. Les biais nombreux, et notamment les biais algorithmiques – difficilement détectables sans accès à des données sensibles –, complexifient la lutte contre la discrimination. L'auteur analyse également pour *Risques* les paradoxes juridiques existants entre protection des données et lutte contre les inégalités.



La décision ne peut pas être raciste puisqu'elle a été prise sans aucune information sur l'origine ethnique de la personne.» On a tous entendu ce genre d'affirmation un jour ou l'autre. Que ce soit sur le racisme, l'âgisme ou le sexism. Que cela concerne des décisions humaines, des modèles ou des algorithmes. Pourtant, la « loi de Kranzberg » nous rappelle que si la technologie n'est ni bonne ni mauvaise, elle n'en est pas neutre pour autant¹. La neutralité risque de ne venir qu'à un certain prix. Et il est peut-être temps de revenir sur les grands principes autour de la segmentation et de l'équité, en assurance, pour mieux comprendre de quoi on parle quand on évoque la question de la discrimination.

L'équité actuarielle

Avant de définir l'équité actuarielle, ou « actuarial fairness » comme on dit en anglais, il est important de revenir sur ce qu'est l'assurance. Nous définissons généralement l'assurance comme la contribution financière du plus grand nombre au malheur de quelques-uns. L'assurance est donc, dans son essence, un mécanisme de partage des risques, qui ne peut avoir de sens que collectivement, par mutualisation des risques. Mais, au niveau plus concret, ou légal, de droit des contrats, l'assurance est avant tout un transfert de risque, d'un assuré vers une compagnie d'assurance, puisqu'en échange d'une contribution financière (la prime ou la cotisation), l'assuré achète la promesse d'une indemnisation future, en cas d'événement malheureux. Et ces deux perspectives sont à l'origine d'une certaine confusion.

L'équité actuarielle est liée à la seconde dimension, c'est-à-dire le transfert de risque. Elle remonte au XVII^e siècle, lorsque les mathématiciens ont introduit la valeur moyenne dans le contexte d'un partage équitable des jeux. En fait, elle remonte à Aristote et au concept de justice dans l'échange, comme le rappellent Antonio José Heras Martínez, David Teira et Pierre-Charles Pradier. C'est pourquoi les annuités, en assurance-vie, depuis Christiaan Huygens ou Jean de Witt, sont calculées sur la base des valeurs espérées des pertes futures, où les probabilités de versement de capital sont

calculées sur la base de tables de mortalité. Et il faudra attendre les années 1960 pour que Kenneth Arrow introduise explicitement le terme d'« équité actuarielle », dans le domaine de l'assurance maladie, pour dire que la contribution des assurés doit être égale à la valeur attendue des pertes. Il n'y a aucune connotation morale mais simplement un principe économique, dans cette perspective de transfert de risque, entre les assurés et la compagnie d'assurance. Faire payer plus serait injuste pour les assurés, faire payer moins serait injuste pour la compagnie d'assurance. Dans l'ensemble, la prime perçue doit donc être égale à la perte payée par les assureurs : il s'agit d'un simple principe d'équilibre (en moyenne).

Mais on peut aussi s'interroger, en revenant sur la première dimension, et considérer les problèmes d'équité dans le partage des risques. Une fois que les assureurs savent qu'ils doivent collecter une certaine somme pour, en moyenne, indemniser tout le monde, comment a priori décider d'un partage équitable du gâteau entre les assurés ? À quelle hauteur chacun doit-il contribuer ? Où placer le curseur entre une prime moyenne identique, parfaitement solidaire, et une prétendue individualisation des contributions, pour tenir compte des risques individuels ?

Discrimination directe, ou indirecte ?

Dans les deux cas, on se place dans une vision économique du problème. Gary Becker, prix Nobel d'économie et pionnier de l'analyse économique des comportements sociaux, s'est penché dans *The Economics of Discrimination* sur ces questions de discrimination, dans un contexte plus général que celui de l'assurance (principalement le marché du travail). Pour Gary Becker, il existe une forme de discrimination correspondant à une préférence, ou un goût (souvent un dégoût). Les agents peuvent éprouver une désutilité à interagir avec des membres d'un groupe donné (par exemple, minorités ethniques, femmes, etc.). Bien souvent, elle n'a rien de rationnel, et elle est inefficace, économiquement. Un employeur qui refuse d'embaucher une personne compétente parce qu'elle est noire, âgée, ou homosexuelle se retrouvera pénalisé face à des compétiteurs qui ne discriminent pas, par simple antisélection.

Mais il y a une autre forme de discrimination, étudiée plus en détail par Kenneth Arrow et Edmund Phelps, appelée « discrimination statistique », et qui repose sur une logique d'information imparfaite et de rationalité. Il s'agit d'un comportement calculé, pas d'un préjugé. Ici, les agents utilisent alors les caractéristiques moyennes d'un groupe (sexe, origine ethniques, âge, etc.) comme proxy pour prédire certains traits qui les intéressent : productivité pour un employeur, risque pour un actuaire. Cette approche permet de légitimer les contrôles au faciès par la police, appelés aussi « profilage ethnique ». Mathias Risse et Richard Zeckhauser rappellent que dans un cadre utilitariste basé sur une approche coût-bénéfice, cette approche est pertinente. Dans une analyse décisionnelle, où les coûts attendus des erreurs (de type I) d'indication erronée sont mis en balance avec les coûts attendus des erreurs (de type II) de non-indication erronée, il est possible d'y voir une rationalité, à court terme tout du moins. En effet, bien souvent, la discrimination statistique n'est pas efficiente² à long terme. Elle reproduit et renforce les inégalités, créant une prophétie autoréalisatrice inefficace pour la société dans son ensemble. Kenneth Arrow et Edmund Phelps ont justifié cette discrimination instrumentale, statistique, que nous

En assurance, les biais de sélection fonctionnent de la même manière, et sont peut-être la forme de biais la plus connue.

pourrions appeler « actuarielle », qui part d'un objectif souvent péculiaire, supposé légitime, supposé indépendant de la discrimination en cause. Elle est justifiée dans un contexte d'incertitude pour Kenneth Arrow, et dans un modèle de rationalité limitée pour Edmund Phelps, mais, en tant que frein à la mobilité sociale, elle est souvent inefficace à long terme.

Les biais sont partout

Un biais est généralement caractérisé comme étant un écart systématique par rapport à l'objectivité, qui conduit à des résultats, des décisions ou des perceptions faussés. Les biais peuvent exister dans les données, les algorithmes et le jugement humain. Et ils influencent souvent les résultats de manière involontaire. Les biais de survie montrent qu'ils peuvent induire des lectures de données et des décisions à l'opposé de ce qu'il convient de faire³. En assurance, les biais de sélection fonctionnent de la même manière, et sont peut-être la forme de biais la plus connue. Ils surviennent lorsqu'un échantillon n'est pas représentatif de la population cible, généralement à cause d'un processus de sélection qui favorise certains groupes d'individus par rapport à d'autres.

Or, dans le contexte des données massives et de l'algorithme croissante des décisions, ces biais peuvent prendre une forme plus subtile, et plus redoutable encore. Un algorithme de recrutement, par exemple, peut apprendre à éviter les CV de candidates portant des prénoms féminins, simplement parce que les données passées (historiquement biaisées) montrent un taux plus faible de femmes dans certains postes. L'algorithme ne discrimine pas consciemment : il réplique, amplifie et légitime un biais historique, tout en se réclamant de la neutralité mathématique. Ce biais de sélection dans les données se traduit automatiquement par un biais d'apprentissage dans les modèles. Ces biais sont d'autant plus difficiles à identifier qu'ils se logent dans des couches techniques (pondérations, variables latentes, matrices de covariance) souvent inaccessibles aux non-spécialistes. De ce point de vue, la discrimination indirecte devient invisible, systémique, et auto-renforcée.



La catégorie et les stéréotypes

En 1983, la Cour suprême des États-Unis⁴ avait souligné que « même une véritable généralisation concernant une catégorie ne peut justifier un traitement fondé sur la catégorie », mettant ainsi en lumière la tension profonde entre savoir statistique et justice individuelle, entre la normalité statistique et la norme éthique, entre ce qui est et ce qui devrait être. Car catégoriser, c'est forcément simplifier. C'est attribuer à un individu les caractéristiques supposées moyennes d'un groupe auquel il est rattaché, parfois arbitrairement, et souvent indépendamment de sa volonté. Cette opération est au cœur de la construction des stéréotypes, qui, bien que parfois basés sur des régularités empiriques, produisent une forme d'injustice épistémique : l'individu est réduit à une variable, nié dans sa singularité. Cette logique s'exacerbe dans les contextes de classification des données, où l'on ne juge plus une personne en tant que telle, mais en tant que membre d'une catégorie prédefinie, souvent corrélée à des facteurs sensibles (origine, genre, lieu de résidence, etc.).

Ce mécanisme est particulièrement visible dans le domaine de l'assurance et de l'actuarial, où l'analyse du risque repose classiquement sur la segmentation des assurés en groupes homogènes. Le rôle des actuaires est bien souvent de proposer une tarification différenciée selon des critères comme l'âge, le sexe ou le code postal – autant de catégories qui, bien qu'efficaces d'un point de vue actuariel, posent des problèmes éthiques et juridiques dès lors qu'elles servent à justifier des écarts de traitement sans lien direct avec le comportement réel de l'individu. Et bien souvent, d'autres considérations entrent en jeu. Si on peut justifier qu'un jeune conducteur devrait payer trois fois plus cher qu'un autre sa prime d'assurance, uniquement en vertu de son âge (et de l'accidentologie élevée de cette classe d'âge, simplement à cause de son manque d'expérience), doit-on pour autant

Les questions éthiques, et surtout légales, sont importantes quand on commence à s'interroger sur la segmentation et la construction de modèles prédictifs.

appliquer une telle surprime? Au sein de l'Union européenne, la Cour de justice a déjà jugé illégal (arrêt Test-Achats, 2011) le recours au sexe comme facteur de tarification, au nom du principe d'égalité. Cela montre que le droit, contrairement à la logique actuarielle, ne tolère pas que des généralisations statistiques, fussent-elles fondées, deviennent des critères de traitement différencié, sans démonstration de nécessité et de proportionnalité.

L'équité, ou l'égalité?

Ce principe d'égalité rigide ne cesse d'interroger, et finit par faire écho à la légende de Procuste (Προκρύστης). Dans la mythologie grecque, Procuste était un brigand qui vivait sur la route entre Athènes et Éleusis. Il invitait les voyageurs à passer la nuit chez lui, leur promettant un lit parfaitement adapté à leur taille. Mais ce lit avait une particularité sinistre : si l'hôte était trop grand, Procuste lui tranchait les jambes; s'il était trop petit, il l'étirait violemment pour qu'il y rentre. Autrement dit, le

voyageur devait s'adapter au lit, et non l'inverse. Cette égalité réduite à une uniformité rigide, où l'on traite tout le monde de la même manière, au mépris des différences individuelles est peut-être un peu extrême. Or, en matière de justice, le principe d'égalité ne signifie pas que tous doivent être traités identiquement, mais qu'ils doivent être traités équitablement, c'est-à-dire selon leurs situations, leurs besoins, et parfois leurs vulnérabilités. Comme Frederick Schauer, on aurait presque envie de dire que l'égalité ne consiste pas à traiter de la même manière ceux qui se ressemblent, mais qu'au contraire, ce principe exige peut-être qu'on traite de la même manière ceux qui ne se ressemblent pas. En France, l'égalité d'accès aux installations publiques exige généralement que les personnes soient traitées de la même manière, même si leurs qualifications pour utiliser ces installations varient; et l'égalité de citoyenneté, qui sous-tend la démocratie, donne à chacun la possibilité de s'exprimer, une personne, une voix.

Si l'égalité correspond à une forme d'uniformité, l'équité serait au contraire basée sur une forte individualisation (certains auteurs parlent de « particularisation »), pour tenir compte des caractéristiques propres de chacune et chacun. Au lieu d'interdire aux pilotes de ligne de piloter au-delà d'un certain âge, une approche individualisante consisterait à tester la vue, l'ouïe et les réflexes de chaque pilote, plutôt que de se fier à l'âge comme indicateur d'une diminution des facultés.

Quels objectifs pour les actuaires ?

Les questions éthiques, et surtout légales, sont importantes quand on commence à s'interroger sur la segmentation et la construction de modèles prédictifs. Pourtant, il n'est pas rare d'entendre des actuaires dire que leur unique objectif est l'exactitude ou la précision des prédictions, « accuracy », dirait-on en anglais. Comme les informaticiens qui créent des algorithmes identifiant des chiens et des chats sur des images. Sauf que sur la photo, il y a réellement un chien, et on peut savoir si l'algorithme a mal identifié l'animal. L'actuaire en tarification n'a pas pour but de deviner qui sera

Le flou et les incohérences juridiques

En plus de la difficulté de choisir le bon modèle, la bonne prédition, il est difficile de comprendre ce que veut la législation. Par exemple, la plupart des États américains ont adopté une forme d'interdiction de base contre la discrimination injuste entre individus de même catégorie et exposés essentiellement au même risque, à moins que la limitation ou l'écart repose sur des principes actuariels solides (*« where the refusal, limitation, or rate differential is based on sound actuarial principles or is related to actual or reasonably anticipated experience »*, comme le dit le Code de l'État de New York, section 4224). Et personne ne sait vraiment ce que sont ces principes actuariels solides.

En Europe, l'article 5-1 de la directive 2004/113/CE du Conseil du 13 décembre 2004 a mis en œuvre le principe d'égalité de traitement entre les hommes et les femmes dans l'accès aux biens et services et la fourniture de biens et services. Il dispose que « l'utilisation du sexe comme facteur actuariel dans le calcul des primes et des

Un des paradoxes les plus notables du droit contemporain réside dans la tension entre l'exigence de non discrimination et les interdictions de traitement des données sensibles.



prestations ne doit pas entraîner de différences dans les primes et les prestations des personnes». L'article 5-2, cependant, autorisait une dérogation à l'interdiction si l'utilisation du sexe était « fondée sur des données actuarielles et statistiques pertinentes et exactes », un peu comme dans la législation que l'on retrouve dans plusieurs États en Amérique du Nord. Dans l'affaire historique de l'arrêt Test-Achats, de 2011, la Cour de justice de l'Union européenne a déclaré l'article 5-2 invalide, ce qui signifie que le sexe ne peut plus être pris en compte, même s'il est pertinent d'un point de vue actuariel. Comme la Commission européenne l'a précisé par la suite⁵, il reste possible pour les assureurs de proposer des produits d'assurance spécifiques au sexe

pour couvrir des risques spécifiques associés au sexe de l'assuré, comme le cancer de la prostate ou le cancer du sein. En revanche, et pour compliquer encore les choses, cette option est interdite en matière de grossesse et de maternité, compte tenu du mécanisme de solidarité spécifique créé par l'article 5-3.

En fait, un des paradoxes les plus notables du droit contemporain réside dans la tension entre l'exigence de non-discrimination et les interdictions de traitement des données sensibles. En Europe, le Règlement général sur la protection des données (RGPD) interdit en principe le traitement des données révélant l'origine raciale, les opinions politiques, les convictions

religieuses ou l'orientation sexuelle (article 9). Cette interdiction vise légitimement à protéger les individus contre les usages abusifs de ces informations. Toutefois, dans le contexte des décisions algorithmiques, cette règle peut devenir contre-productive. En effet, sans accès à ces variables sensibles, il devient impossible de tester si un algorithme prend des décisions discriminatoires fondées – directement ou indirectement – sur ces mêmes catégories. En d'autres termes, l'impossibilité de collecter certaines données rend invisible la discrimination, et empêche les audits algorithmiques efficaces. Ce paradoxe juridique crée un espace de non-droit technique, où l'opacité algorithmique se trouve renforcée par les contraintes de protection des données elles-mêmes.

Cette incohérence juridique est d'autant plus préoccupante qu'elle finit par opposer deux impératifs également fondamentaux : la protection de la vie privée et la lutte contre les discriminations. L'un empêche de connaître, l'autre exige de savoir. La jurisprudence peine encore à articuler ces deux logiques. Ces incohérences finissent par construire des formes d'irresponsabilité technologique, où les discriminations ne sont pas commises de manière intentionnelle mais apparaissent comme les sous-produits mécaniques d'un système juridiquement aveugle. Ce flou normatif fragilise la capacité du droit à jouer pleinement son rôle de rempart contre l'injustice systémique dans un contexte de modèles prédictifs de plus en plus complexes.

Conclusion

Le marin Shadok affirmait que « s'il n'y a pas de solution, c'est qu'il n'y a pas de problème ». On peut aussi se dire que s'il n'y a pas de solution, c'est peut-être avant tout parce que le problème est mal posé. Les actuaires doivent aujourd'hui naviguer à l'aveugle entre des injonctions techniques, où on leur demande de segmenter davantage tout en cherchant du sens qui ne peut exister que par mutualisation, et des injonctions légales, où on leur demande de respecter la vie privée tout en luttant contre les discriminations.



Notes

¹ La loi de Kranzberg, énoncée par l'historien des technologies Melvin Kranzberg en 1986, a été formulée dans un article intitulé « Technology and History: 'Kranzberg's Laws' », publié dans la revue *Technology and Culture*. Six lois sont proposées, mais la première est la plus connue, rappelant que les effets de la technologie dépendent du contexte social, politique, économique et culturel dans lequel elle est utilisée : « Technology is neither good nor bad ; nor is it neutral. »

² Anglicisme qui traduit probablement mieux le terme « efficient » que la traduction « efficace », au sens utilisé par Elizabeth Popp Berman dans son livre *Thinking like an economist: How efficiency replaced equality in US public policy*.

³ Abraham Wald (1902-1950) a montré que l'analyse des zones endommagées des bombardiers revenant de mission était biaisée, car elle ignorait ceux qui avaient été abattus, donc ne prenait pas en compte les parties vulnérables des avions. Il a recommandé de renforcer les zones non endommagées, comme les moteurs et la queue, car elles étaient probablement les plus critiques pour la survie des avions.

⁴ 463 U.S. 1073 (1983), « *The use of sex-segregated actuarial tables to calculate retirement benefits violates Title VII whether or not the tables reflect an accurate prediction of the longevity of women as a class, for under the statute, «[e]ven a true generalization about [a] class» cannot justify class-based treatment. »*

⁵ Lignes directrices sur l'application de la directive 2004/113/CE du Conseil dans le secteur des assurances, à la lumière de l'arrêt de la Cour de justice de l'Union européenne dans l'affaire C-236/09 (arrêt Test-Achats)



Références

- Aristote. (1990). *Éthique à Nicomaque* – Livre V (trad. J. Tricot), Vrin (écrit v. 350 av. J.-C.)
- Arrow, K. (1971). «The Theory of Discrimination « *Labor Economics*, Vol. 4, Princeton University, Department of Economics, Industrial Relations Section.
- Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.
- Berman, E. P. (2022). *Thinking like an economist: How efficiency replaced equality in US public policy*. Princeton University Press.
- Charpentier, A. (2018). « Les modèles prédictifs peuvent-ils être loyaux et justes ? ». *Risques*, n° 113, 91-96.
- Charpentier, A., Barry, L. & Gallic, E. (2020). « Quel avenir pour les probabilités prédictives en assurance ? ». *Annales des Mines*, 2020 (1), 74-77.
- Charpentier, A. (2024). *Insurance, biases, discrimination and fairness*. Springer.
- Frezel, S., & Barry, L. (2020). *Fairness in uncertainty: Some limits and misinterpretations of actuarial fairness*. *Journal of Business Ethics*, vol. 167, 127-136.
- Kranzberg, M. (1986). « Technology and history : « Kranzberg's laws » ». *Technology and Culture*, vol. 27(3), 544-560.
- Harcourt, B. E. (2006). *Against Prediction : Profiling, Policing, and Punishing in an Actuarial Age*. University of Chicago Press.
- Schauer, F. (2006). *Profiles, probabilities, and stereotypes*. Harvard University Press.
- von Mises, R. (1928). *Wahrscheinlichkeit Statistik und Wahrheit*. Springer-Verlag.
- Hellman, D. (1998). « Two Types of Discrimination : The Familiar and the Forgotten ». *California Law Review*, vol. 86 : 315-361.
- Winston, K. (1974). *On Treating Like Cases Alike*. *California Law Review*, vol. 62: 1-39.
- Hellman, D. (2008). *When is Discrimination Wrong?*

