

Comment démêler les « causes » des inégalités dans les statistiques ?

16 avril 2021



La mesure des inégalités repose généralement sur des comparaisons statistiques entre différentes catégories de populations (riches ou pauvres, femmes ou hommes, immigrés ou non, etc.) dans un domaine donné (par exemple, l'espérance de vie). Comment savoir quel est le facteur qui explique l'écart parmi un grand nombre ?

Pour le comprendre, on peut par exemple [comparer la réussite à l'école des enfants de parents étrangers et celle dont les parents sont nés français](#). Les élèves qui ont au moins un parent d'origine étrangère réussissent en général moins bien que ceux dont les deux parents sont nés français. Les chiffres sont suffisamment stables pour que cette réalité ne soit pas mise en doute : c'est statistiquement vrai, mais ce résultat appelle plusieurs remarques.

La première remarque est celle du degré de « déterminisme » des statistiques. Ces chiffres ne veulent pas dire que tous les enfants de parents étrangers échouent à l'école, loin de là ! Beaucoup d'entre eux réussissent aussi bien, voire mieux que des enfants de parents français. On ne fait que comparer des moyennes.

Il faut toujours garder à l'esprit la différence entre une liaison statistique entre deux variables (les statisticiens parlent de « corrélation ») et l'existence d'un lien de causalité entre elles. Dans notre exemple, les difficultés scolaires sont bien corrélées à l'origine des parents : l'origine étrangère fait légèrement baisser la probabilité de réussite. Mais il n'y a pas de « causalité ». On ne peut pas dire que l'échec soit la conséquence directe de l'origine des parents. Pour pouvoir affirmer cela, il faudrait que tous les enfants de parents étrangers échouent (et que la nationalité française assure la réussite scolaire !). Malgré des tendances perceptibles dans les statistiques, la relation entre origine et réussite n'est pas aussi mécanique.

La seconde remarque est un peu plus compliquée, mais plus fondamentale. Il faut se demander comment, en pratique, la nationalité étrangère des parents influence la réussite scolaire. D'où viennent les différences que l'on observe dans la réussite ? Qu'est-ce qui joue en défaveur des enfants de parents étrangers ? Est-ce une question de niveau de maîtrise de la langue française ? De connaissance des rouages de l'institution scolaire ? Difficile de répondre...

Les inégalités sont toujours le produit de plusieurs facteurs qui se mélangent et peuvent se renforcer mutuellement. Dans la France contemporaine, où l'immigration est massivement liée au travail peu qualifié, le fait d'avoir des parents étrangers est *corrélé* à toutes sortes de choses : parents moins diplômés, familles plus nombreuses, concentration dans des quartiers périphériques, etc. Chacun de ces facteurs peut avoir un effet négatif sur la réussite scolaire. Comment être sûr que, derrière la nationalité des parents, ce ne sont pas ces autres facteurs qui agissent ? Par exemple, une fille ou un fils d'ambassadeur étranger en France, vivant dans les beaux quartiers et fréquentant les meilleurs établissements scolaires, a toutes les chances de réussir bien mieux à l'école que des enfants de parents français issus d'un milieu populaire. La nationalité n'est donc qu'un attribut parmi d'autres, et il se peut qu'elle ne joue en fait aucun rôle par elle-même.

Il existe une approche mathématique qui vise à isoler ces différents facteurs pour savoir quel est l'« effet propre » de chacun : le raisonnement « toutes choses égales par ailleurs ». Découvrons-en la logique.

Raisonner toutes choses égales par ailleurs

Le raisonnement « toutes choses égales par ailleurs » s'inspire de la méthode expérimentale venue des sciences naturelles. Prenons un exemple simple. Si l'on veut tester l'efficacité d'un engrais, on va prendre deux graines issues d'une même plante, donc parfaitement identiques, et les faire pousser en contrôlant tous les paramètres environnementaux (même terre, même température, même pression atmosphérique, même hygrométrie ambiante et même arrosage). À l'une des deux graines, on donnera l'engrais testé (sujet « test »), à l'autre non (sujet « témoin »). Au bout d'un temps donné, on pourra mesurer leurs tailles respectives. Si la plante qui a reçu l'engrais est plus développée, on pourra mesurer la différence et dire que « toutes choses égales par ailleurs », l'engrais augmente la croissance de X %. Bien sûr, pour attester réellement l'effet de l'engrais, il faudra renouveler l'expérience un certain nombre de fois et calculer l'effet moyen. Mais on aura bien prouvé une relation de cause à effet entre l'engrais et la croissance de la plante.

On adopte le même raisonnement en médecine pour tester l'efficacité d'un médicament. On prend deux groupes de « cobayes » qu'on va essayer de rendre aussi similaires que possibles. À l'un sera donné le médicament, à l'autre un « placebo ». Au bout d'un certain temps, on regarde si les membres du groupe qui a reçu le médicament vont significativement mieux que les autres.

Les sciences humaines s'inspirent de cette logique même si... nous ne sommes pas des légumes ! Comparer les destins des humains est d'une complexité infiniment supérieure à l'analyse de la croissance des plantes. Il n'est pas possible, ni souhaitable de réaliser des « expériences » sur la vie des gens. On se contente donc de s'inspirer de la logique du raisonnement expérimental, en utilisant certains outils statistiques qui permettent de s'en approcher. Les statisticiens appellent cela des « régressions multiples » ou « modélisations ».

Intuitivement, la logique est assez simple, et on peut la présenter grossièrement sans rentrer dans les détails mathématiques. Au départ, il faut avoir un grand nombre d'individus (plusieurs centaines au minimum) à comparer et le plus possible d'informations sur chacun. Imaginons que nous avons quelques milliers de personnes dont nous connaissons par exemple l'âge, le sexe, le niveau de diplôme, la profession des parents, la nationalité, la taille de la commune de résidence, etc.

On va d'abord définir ce qu'on cherche à comparer : le fait de faire quelque chose par exemple, ou bien une valeur (revenus, nombre d'enfants). On peut par exemple se pencher sur le fait d'avoir son baccalauréat.

Il faut ensuite choisir une situation de départ, dite « de référence », qui servira de point de comparaison. On va par exemple dire que le profil de référence est une femme, de milieu aisé, dont les deux parents sont français, vivant dans une grande agglomération, etc. C'est en quelque sorte l'équivalent de la plante « témoin » (celle qui n'a pas d'engrais) dans l'approche expérimentale évoquée plus haut. On prend en compte, pour ce groupe de femmes, la valeur du phénomène testé, par exemple, dans cette population, le pourcentage de celles qui ont le baccalauréat.

Le raisonnement « toutes choses égales par ailleurs » consiste à mesurer comment évolue la réussite au baccalauréat par rapport à la situation de référence. Si, par exemple, au lieu d'être une femme, on est un homme, mais que toutes les autres caractéristiques sont conservées. Dans notre exemple : un homme – plutôt qu'une femme donc - de milieu aisé, dont les deux parents sont français, vivant dans une grande agglomération, etc. On fait ainsi défiler une à une les caractéristiques en comparant à la situation de référence, mais en conservant toujours toutes les autres caractéristiques inchangées.

Lorsque les résultats sont significatifs, on peut dire que, « toutes choses égales par ailleurs », le fait d'être ceci plutôt que cela augmente ou diminue la probabilité de réalisation du phénomène étudié, c'est-à-dire indépendamment des autres facteurs. Dans le cas d'inégalités ou de discriminations, cela permet par exemple de montrer l'effet propre de chaque caractéristique des individus. Dans notre exemple, on pourrait ainsi savoir si le fait d'être un homme joue sur le fait d'avoir ou non le bac, après avoir mis de côté les autres facteurs.

Revenons au cas des inégalités scolaires liées à l'origine étrangère des parents. Si l'on observe les données brutes, on constate que les enfants de parents d'origine étrangère réussissent moins bien que les enfants de parents français. Mais plusieurs études [1] ont montré que [toutes les nationalités d'origine n'ont pas, « toutes choses égales par ailleurs »](#) les mêmes effets sur la probabilité de réussite scolaire. Mieux : pour un certain nombre de nationalités, le fait d'avoir des parents étrangers a même un effet positif – « toutes choses égales par ailleurs » - sur la réussite [2]. Autrement dit, si l'on prend des enfants de même âge et de même sexe, qui viennent des mêmes milieux sociaux, qui grandissent dans des conditions semblables, ceux qui ont des parents étrangers peuvent avoir de plus grandes chances de réussite à l'école que les enfants de parents français. Ce résultat - qui semble paradoxal par rapport aux données brutes de départ - ne s'explique pas par une « discrimination positive » réservée aux étrangers à l'école mais par le fait que la migration s'accompagne souvent de fortes ambitions placées sur les enfants par leurs parents, de qui l'on attend qu'ils retrouvent, grâce à l'école, un niveau social au moins équivalent à celui que les parents ont abandonné dans leur pays d'origine. Leur moins bonne réussite dans les données brutes n'est pas liée au fait d'être issus de l'immigration mais d'un milieu social défavorisé.

Les limites

Les « modélisations » permettent de démêler les différents facteurs qui s'intriquent pour produire les inégalités. Elles permettent aussi d'attester l'existence d'une inégalité spécifique, comme par exemple lorsqu'on mesure [« toutes choses égales par ailleurs » les différences de salaire entre les femmes et les hommes](#). Attention, il ne faut pas croire pour autant que le raisonnement « toutes choses égales par ailleurs » soit une méthode magique permettant de répondre à toutes les questions. Plusieurs limites doivent être prises en compte.

La méthode, d'abord, suppose, par le jeu des combinaisons entre facteurs, d'évaluer des situations parfois très improbables, comme des personnes très qualifiées occupant des postes qui le sont très peu, etc. Plus fondamentalement, l'affirmation « toutes choses égales par ailleurs » suppose que l'ensemble des facteurs pesant sur le phénomène sont pris en compte et évalués. Or, non seulement il est rare qu'on possède toutes les informations pertinentes pour évaluer une discrimination, mais certaines ne sont pas quantifiables simplement. C'est par exemple le cas de la couleur de peau, qui joue un rôle dans les discriminations, mais qui échappe généralement aux approches statistiques pour des raisons pratiques et légales. Mais prendre en compte la couleur de la peau entraîne des débats complexes.

Nous travaillons bien avec des humains et non des plantes. Certains comportements sont bien difficiles à classer. Prenons le cas de la mesure des écarts de salaire entre femmes et hommes. Pour mesurer la discrimination, on ne compare que des temps complets. Si une personne travaille moins qu'une autre, il est normal qu'elle gagne moins. Dans certains cas, le temps partiel est imposé par l'employeur ou la personne est contrainte de l'accepter (par exemple faute de modes de garde pour les enfants). En recherchant la mesure des inégalités pour des temps complets, le statisticien laisse de côté une partie du problème.

Au bout du compte, les modèles nous permettent d'avancer dans la connaissance, de mieux connaître l'influence des différents facteurs. Mais, devant un tableau de chiffres, il est important de toujours se poser la question : « au fond, quel est le facteur qui joue ? » et de toujours garder en tête qu'une corrélation n'est pas forcément une cause. Ceci dit, il faut aussi être conscient que ces modèles, appliqués aux êtres humains, laissent toujours une part de la réalité inobservée. On ne devrait pas dire « toutes choses égales par ailleurs » mais bien « toutes choses prises en compte égales par ailleurs »... Les sciences humaines ont encore de belles découvertes à faire.

Alex Alber
Maître de conférences en sociologie à l'université de Tours

Photo / © Ridofranz

[1] Voir par exemple « [Les enfants d'immigrés ont des parcours scolaires différenciés selon leur origine migratoire](#) », Yaël Brinbaum, Laure Mogueurou et Jean-Luc Primon, in *Immigrés et descendants d'immigrés en France*, coll. Insee Références, Insee, 2012.

[2] Voir le travail devenu classique de Louis-André Vallet : « [L'assimilation scolaire des enfants issus de l'immigration et son interprétation : un examen sur données françaises](#) », *Revue française de pédagogie* n° 117, 1996

Emplacement :

[Accueil](#) > [Outils](#) > [Mesurer les inégalités](#) >

Adresse de cet article : <https://www.inegalites.fr/Comment-demeler-les-causes-des-inegalites-dans-les-statistiques>