

son choix, mais l'éventualité de cas tels que ceux décrits ci-dessus, et encore plus tels que celui décrit à la figure 4, force cette conclusion attentiste.

Une attitude prudente

Point de vue 3. Dans une situation où seules les données des tableaux 1, 2 et 3 sont disponibles, aucune prescription ne s'impose absolument, du fait de l'existence possible d'une partition contradictoire comme celle donnée par les tableaux 4 et 5. Celui qui dispose des données des tableaux 6, 7, 8 et 9, qui indiquent pour chaque catégorie précise (sexe connu et couleur des yeux connue) si le médicament est meilleur que le placebo pourra en revanche prescrire un traitement à un patient dont on connaît le

sexe et la couleur des yeux : il se référera au tableau le concernant. Ce jugement prudent est confirmé par deux arguments.

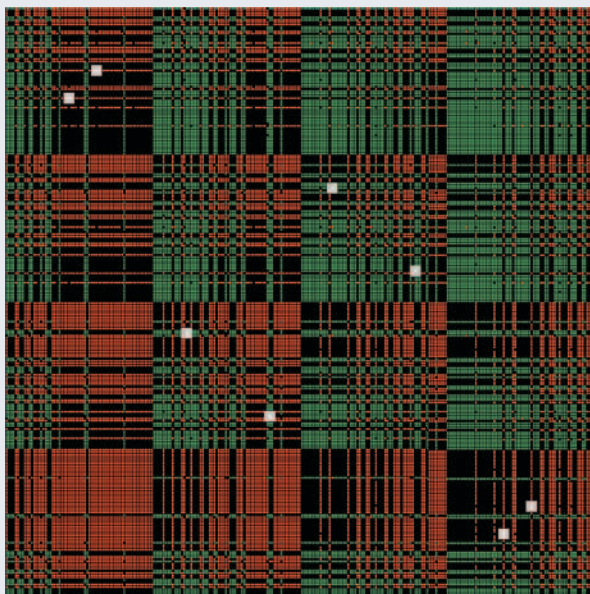
Argument 1. Quand on dispose de données analogues à celles du tableau 1 (indiquant que « le taux de guérison est meilleur avec le médicament qu'avec le placebo ») et que les effectifs ne sont pas trop petits (supérieurs à 3), il est toujours possible de faire une partition des patients en deux catégories complémentaires de façon à avoir des tableaux 2 et 3 produisant des conclusions opposées à celle du tableau 1 (« le placebo est meilleur »).

En effet, composons la catégorie 1 de patients guéris par le placebo (on en garde au moins un pour la catégorie 2) et aucun de la catégorie que le placebo ne guérit pas, auxquels on ajoute tous les patients

que le médicament guérit et quelques patients que le médicament ne guérit pas. Ce choix assure que dans la catégorie 1, le taux de réussite du placebo est 100 %, et qu'il ne l'est pas pour le médicament. Dans la catégorie 2, il n'y a aucun patient que le médicament guérit (ils ont tous été mis dans la catégorie 1) et donc le taux de guérison avec le médicament est nul, ce qu'il n'est pas pour le placebo puisqu'on a réservé au moins un patient guéri par le placebo pour la catégorie 2. Aussi bien dans la catégorie 1 que 2, le placebo sera donc strictement meilleur que le médicament. Si les coefficients du tableau 1 sont $\{2, 2, 2, 3\}$, la méthode décrite donne $\{2, 1, 1, 0\}$ pour le tableau 1 et $\{0, 1, 1, 3\}$ pour le tableau 2.

Ce type de partages est certes artificiel, mais il montre que le risque de paradoxe de

4. Visualisation de la fréquence du paradoxe



27 = 0×64 + 1×16 + 2×4 + 3×1 = 0123 en base 4				216 = 3×64 + 1×16 + 2×4 + 0×1 = 3120 en base 4			
27	Guéri	Non guéri	Taux de guérison	216	Guéri	Non guéri	Taux de guérison
Médicament	1	3	25 %	Médicament	4	3	57 %
Placebo	2	4	33 %	Placebo	2	1	66 %
Le placebo est meilleur				Le placebo est meilleur			
MATRICE SOMME							
243	Guéri	Non guéri	Taux de guérison				
Médicament	5	6	45,45 %				
Placebo	4	5	44,44 %				
Le médicament est meilleur							

Cette figure représente l'ensemble des cas possibles lorsqu'on considère le nombre de guérisons pour des patients hommes prenant le médicament ou le placebo et pour des patientes femmes prenant le médicament ou le placebo (cela correspond aux tableaux 2 et 3 du texte). On a envisagé tous les cas lorsque les huit éléments des deux tableaux parcourent toutes les valeurs entières entre 1 et 4. Cela fait $4^8 = 65\,536$ cas différents.

En regroupant les tableaux hommes et femmes d'un cas, on obtient un tableau de synthèse (hommes et femmes réunis) qui donne parfois une situation de type « paradoxe de Simpson » : les tableaux hommes et femmes vont dans le même sens (le médicament est meilleur, ou le placebo est meilleur), mais le tableau synthétique donne une conclusion opposée.

Chacun des 65 536 cas est représenté par un pixel. Les quatre coefficients du tableau pour les hommes (première colonne : a, b ; deuxième colonne : c, d) fournissent l'abscisse du pixel en considérant le nombre entier dont l'écriture en base 4 a pour chiffres $(a-1), (b-1), (c-1), (d-1)$. Les quatre coefficients

du tableau pour les femmes donnent, selon le même codage, l'ordonnée du pixel. La couleur du pixel est choisie selon les règles suivantes.

– Si on a un cas « normal » (les tableaux hommes et femmes donnent une même conclusion, qui est aussi celle du tableau de synthèse), le pixel est vert ou marron selon que le placebo l'emporte dans le cas général ou non.

– Si on a un cas sans grand intérêt (les tableaux hommes et femmes se contredisent, ou il y a égalité entre les succès du médicament et du placebo), le pixel est noir.

– Si on a un cas de paradoxe de Simpson (le tableau de synthèse s'oppose aux deux tableaux hommes et femmes qui vont dans le même sens), le pixel est clair, entouré d'un halo pour aider à le repérer. Il n'y a que huit cas, correspondant aux huit pixels (27, 216), (39, 228), (78, 114), (114, 78), (141, 177), (177, 141), (216, 27) et (228, 39), le premier étant détaillé ci-contre.

Cette représentation, imaginée et réalisée par Jean-François Colonna, montre qu'heureusement le paradoxe de Simpson se produit assez rarement.

5. Le paradoxe de Simpson et l'émergence de la coopération

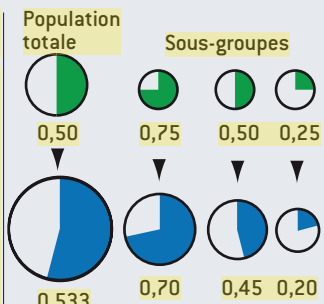
Dans la compétition entre organismes biologiques, le maintien des « biens communs » est une question importante. Lorsqu'il y a un producteur de ce bien commun et un consommateur (un organisme qui en bénéficie sans le produire), les consommateurs ont tendance à proliférer aux dépens des producteurs. L'oxygène de l'air est un exemple de bien commun produit seulement par certains organismes et consommé par d'autres qui ne le produisent pas. En 2009, des chercheurs de l'Université Rockefeller à New York (John Chuang, Olivier Rivoire, Stanislas Leibler) ont créé et étudié une telle situation avec deux souches de la bactérie *Escherichia coli*. L'une des souches produisait un antibiotique utile aux deux souches et l'autre en profitait sans

le produire. Les chercheurs ont composé plusieurs groupes de bactéries, chaque groupe comportant les deux types de bactéries. Une situation apparemment paradoxale est apparue : les non-producteurs croissaient plus vite dans chaque groupe (ce qui n'est pas une surprise) bien qu'au total, les non-producteurs voyaient leur effectif global décroître en proportion. Il s'agissait d'un phénomène de type paradoxe de Simpson.

Pour un biologiste, ce système est un exemple frappant de conflit entre niveaux de sélection. Les bactéries productrices du bien commun sont les bénéficiaires de l'ensemble du système quand on les considère comme un tout, alors qu'à un niveau individuel (celui auquel opère la sélection), elles sont désavantagées,

puisque dans chaque groupe leur proportion diminue.

Grâce aux effets du paradoxe de Simpson, un trait qui bénéficie à la population considérée comme un tout peut ainsi se trouver sélectionné, bien qu'à un niveau individuel le trait soit désavantageux. La chose est étonnante : ce qui est mauvais au niveau individuel se trouve au total favorisé par l'effet mécanique d'un paradoxe de Simpson. La réalisation concrète de l'expérience par les chercheurs démontre que cet effet sélectif paradoxal n'est pas seulement théorique, mais doit être pris en compte par les spécialistes de l'évolution. Cette dynamique doit être envisagée comme un mécanisme de sélection de traits individuels favorables à la coopération et à l'altruisme. Utiliser le paradoxe



En vert, la proportion initiale de bactéries de type 1.
En bleu, la proportion finale de bactéries de type 1.

de Simpson est l'une des ruses que la sélection naturelle mettrait en œuvre pour favoriser les traits coopératifs et faire ainsi émerger et prospérer des entités collectives d'individus coopérateurs et solidaires.

Simpson est toujours présent : tout tableau composé de données pas trop petites est susceptible d'être séparé en deux tableaux créant un paradoxe de Simpson.

Argument 2. Il existe des situations que l'on nommera « Double-Simpson », dont la première a été découverte par Jean-François Colonna, du Centre de mathématiques appliquées de l'École polytechnique, où :

- Le tableau 1 suggère que le médicament est meilleur que le placebo.
- Les tableaux 2 et 3 suggèrent que le placebo est meilleur que le médicament.
- Les tableaux 4 et 5 suggèrent que le médicament est meilleur que le placebo.
- Il existe deux jeux de données (tableaux 6A-9A et tableaux 6B-9B) compatibles avec les tableaux 1-5 et tels que pour chaque catégorie précise (il y en a quatre), ce que laissent supposer les données A est l'inverse de ce que suggèrent les données B (voir l'encadré 3).

Cette possibilité de jeux A et B conduisant à des prescriptions opposées pour chacune des quatre catégories n'était pas connue avant les expériences numériques de J.-F. Colonna. Elle établit définitivement

que dans certaines situations, celui qui veut des certitudes absolues ne peut agir à cause du paradoxe de Simpson.

Heureusement, dans les cas que nous avons envisagés, celui qui est prêt à prendre le risque de se tromper (en médecine le prescripteur est souvent dans une telle situation) réussira à prendre des décisions à peu près sûres. On montre en effet que celui qui connaît un tableau comme notre tableau 1 et le sépare en deux catégories complémentaires ne tombera sur des résultats opposés (tels les tableaux 2 et 3) qu'avec une probabilité de 1,92 %. De même, en fusionnant deux tableaux donnant des indications dans le même sens, la probabilité que le tableau résultant contredise les deux premiers est faible. De même encore, la probabilité que trois tableaux, dont l'un est la somme des deux autres, présentent un paradoxe de Simpson est faible.

La conclusion est que même si les cas où le paradoxe de Simpson se produit sont vraiment gênants (c'est ce qu'établit le « Double-Simpson »), ils sont heureusement assez rares. On peut donc prendre le risque... de ne pas y penser !

L'AUTEUR



J.-P. DELAHAYE est professeur à l'Université de Lille et chercheur au Laboratoire d'informatique fondamentale de Lille (LIFL).

BIBLIOGRAPHIE

J. Pearl, *Simpson's paradox : An anatomy*, 2011 : <http://escholarship.org/uc/item/3s62r0d6>

H. Goltz et M. Smith, *Yule-Simpson paradox in research, Practical Assessment, Research & Evaluation*, vol. 15, 2010 : <http://pareonline.net/pdf/v15n15.pdf>

J. Chuang et al., *Simpson's paradox in a synthetic microbial system*, *Science*, vol. 323, pp. 272-275, 2009.

M. Pavlides et M. Perlman, *How likely is Simpson's paradox ? The American Statistician*, 2009 : <http://bit.ly/10xToug>