

LA REPRODUCTIBILITÉ

LA CRISE DE L'INSTITUTION SCIENTIFIQUE

MARC BOURDEAU

« Personally, the writer prefers to set a low standard of significance at the 5 percent point. [...] A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level of significance. »

R.A. Fisher (1935)

La reproductibilité et les conflits d'intérêt

La reproductibilité est la règle d'or du progrès scientifique (Popper, 2002, [orig. 1959], Kenett & Schmucl, 2017) et surtout Ioannidis (2001, 2005, 2007) qui a attaché le grelot depuis plus de dix ans.

À cet égard, la première chose à remarquer est que toute probabilité de dépassement, $p_{\text{dép}}$, est une variable aléatoire qu'on réalise une seule fois dans toute expérience statistique. Elle a donc une loi, une moyenne, un écart type, etc. Ces propriétés ne sont pas simples à déterminer (Boos & Stefansky, 2011 ; Sackrowitz & Samuel-Cahn, 1997). Cependant, sous des hypothèses assez générales, Boos & Stefansky (p.221) déduisent un taux de non-reproductibilité aux environs de 33%, cela en admettant l'honnêteté des chercheurs...

Ce taux contraste fortement avec les résultats de la gigantesque étude du *Reproducibility Project*, rapportée en septembre 2015 par le *NY Times* ainsi que par *Le Monde*, à la suite de la revue *Science* (Nosek & al., 2015) qui estime cette proportion aux deux-tiers ! Cette étude jette tout un pavé dans la mare. On apporte ici un sérieux discrédit sur les Sciences humaines et sociales & les Sciences de la santé (SHS&Santé). Ioannidis l'estime à environ 60%. Young (2008) en annonce au moins 80%...

Les sociétés, souvent multinationales, et autres organismes subventionnaires ont bien du pouvoir sur les chercheurs (Marsan & Laliberté, 2015 ; Steneck & al., 2015; Foucart, 2015; Angell, [divers articles](#) du *NYReview of Books*). L'honnêteté est une vertu qui se perd.

Limiter les dégâts

C'est ainsi que la statistique est plus que jamais en état de crise. On a contesté sur une base philosophique l'inférence statistique qu'il ne fallait pas confondre avec l'inférence scientifique. Soit, mais c'est surtout l'organisation scientifique qui est en crise, et non le fondement théorique.

On a surtout retenu l'importance de publier les intervalles de confiance associés aux paramètres testés, et pas seulement les probabilités de dépassement. Ces intervalles de confiance bilatéraux, équivalents aux tests sur hypothèse nulle (THN), donnent une meilleure idée de la valeur d'un effet obtenu (Krantz, Berkson, 2003). On trouvera d'autres conseils chez Gibbons & Pratt (1975), et chez Ellis (2010).

Cependant, les enquêtes sur la reproductibilité initiées par Ioannidis exigent impérativement d'autres importants ajustements.

On recommande de plus en plus de publier même les études non significatives, ce qui serait un grand progrès car c'est seulement par des études répétées qu'on peut établir les propriétés en Sciences humaines et sociales et en Sciences de la santé (SHS&Santé). L'utilisation de la théorie des méta-analyses est sans cesse handicapée par ce *biais de publication*... Sans compter que pareil biais de publication est un incitatif à la tricherie. Les définitions et propriétés des protocoles expérimentaux devraient bien sûr être déclarées dans les publications, mais surtout au préalable i.e. *ante hoc*, et les données devraient être accessibles (Bailar the III & Mosteller, 1988).

On pourrait penser aussi développer un site internet, à déclaration obligatoire, non modifiable une fois les informations enregistrées, pour regrouper les protocoles *ante hoc* des projets d'expérimentation statistique, i.e. les descriptions complètes des expériences statistiques, ainsi que les données recueillies et les analyses complètes *post hoc*. Il faudrait de plus, une fois les protocoles évalués et agréés par des experts, garantir aux auteurs des articles la publication de leurs résultats, même pour ceux qui ne « montrent pas d'effet », i.e. même si $p_{dép}$ est trop grand pour être 'significatif' au seuil prescrit.

En aucun cas toutefois, les fraudes ne deviennent-elles impossibles... C'est l'honnêteté intellectuelle qui est en jeu ici. Et l'honnêteté n'étant pas la chose du monde la plus répandue... Rien n'empêche les chercheurs de publier des données trafiquées au préalable pour obtenir des résultats significatifs mais pas très... question d'assurer une publication, et de ne pas passer pour un fraudeur advenant une étude à l'effet contraire... qu'on ne risque pas beaucoup de voir tant qu'on ne publie que des études significatives. Publier ou périr !

On remarque que la non reproductibilité ne sévit pas beaucoup en dehors des SHS&Santé. Dans l'industrie de fabrication par exemple, on utilise énormément des plans d'expériences parfois très complexes pour garantir les connaissances hypothétiques et pour progresser. L'inférence statistique classique y est indispensable. Et pas question de tricher. Attention toutefois aux impératifs commerciaux en amont dénoncés par Marcia Angell ainsi que dans le Oreskes & Conway, pour l'industrie pharmaceutiques (les *Big Pharma*), l'industrie du tabac, l'industrie des énergies fossiles, de la chimie de l'agriculture... Les vertus se perdent. L'institution scientifique est bien en crise.

Références

On peut télécharger le fichier des références citées et quelques autres [ici](#).