

# Signification statistique



Jonathan Jarry

Des résultats dits « significatifs » ne veulent sans doute pas dire ce que vous pensez qu'ils disent. La méprise est commune.

Une nouvelle étude a examiné la question suivante : le supplément « Gwynex » est-il associé à une perte de poids ? Un groupe de recherche l'a testé sur des volontaires et il a obtenu des résultats statistiquement significatifs. Eh oui, le Gwynex fait perdre du poids ; ils ont obtenu une valeur-p juste sous le seuil important de 5 %. Cela signifie qu'ils ont 95 % de chances que leur hypothèse soit vraie... n'est-ce pas ? Pas du tout !

Cela veut-il dire :

- Que les résultats de leur étude doivent être considérés sérieusement ?
- Que si vous répétez l'expérience, vous obtiendrez des résultats significatifs 95 % du temps ?

Non plus !

## Une expérience scientifique typique

Une expérience scientifique est presque toujours construite autour de deux thèses opposées : l'hypothèse « nulle », soit que votre expérience n'aura pas d'effet, et l'hypothèse « alternative », soit qu'elle aura un effet. Ou, si vous aimez l'hyperbole, l'hypothèse ennuyeuse et celle qui change la donne !



Voici une expérience au cours de laquelle 20 personnes ont pris du Gwynex et 20 personnes ont pris un placebo. Nous les pesons six semaines plus tard. L'hypothèse nulle : il n'y a pas de différence entre les deux groupes, ils ont en moyenne perdu le même poids. L'hypothèse alternative : le groupe ayant pris du Gwynex a perdu plus de poids que le groupe placebo.

Lorsque l'expérience est terminée, nous devons procéder à une analyse particulière pour déterminer la signification statistique des résultats. Il n'est pas toujours évident de décider si nos résultats favorisent l'hypothèse nulle – pas de différence suffisante entre les deux groupes – ou l'hypothèse alternative – le Gwynex, ça marche !

## Une notion très mal comprise

De nombreux scientifiques portent un culte sacrosaint à la valeur-p. Elle vous dit à quel point vos résultats sont significatifs. Elle détermine souvent dans quelle revue sera publié votre article ou même s'il sera publié tout court. Le seuil significatif est souvent fixé à 0,05 – c'est-à-dire à 5 % (ou inversement à 95 %). Mais quelle est la définition exacte de cette fameuse valeur-p ?

Vous serez troublé d'apprendre que la plupart des étudiants ne saisissent pas vraiment ce concept de base. Même des chercheurs en psychologie et ceux qui enseignent la signification statistique comprennent souvent très mal cette statistique si commune.

Permettez-moi d'insister sur l'incompréhension commune au sujet de cette mesure statistique. C'est hilarant si l'on veut montrer quelqu'un du doigt. Et vraiment terrifiant si l'on travaille dans un domaine de recherche comme moi.

Un bref questionnaire a été envoyé aux départements de psychologie de six universités allemandes. Qui a répondu ? Voici : 30 instructeurs en méthodologie, ce sont eux qui enseignent la valeur-p aux étudiants ; 39 chercheurs en psychologie, qui n'enseignent pas les statistiques, mais s'en servent dans leur travail ; et 44 étudiants en psychologie.

Ils devaient répondre à six questions sur le sens de la valeur-p. En voici une : « Vous avez trouvé la probabilité que l'hypothèse nulle soit vraie. »

L'astuce ? Les six énoncés étaient faux.

Combien de personnes ont répondu qu'au moins une réponse était vraie, alors qu'elles étaient toutes fausses ? Les étudiants en psychologie, 100 % ; les chercheurs en psychologie (qui n'enseignent pas les statistiques), presque 90 % ; les instructeurs de statistiques, 80 %. Il importe de remarquer que ces derniers n'étaient toutefois pas des professeurs, mais des étudiants avancés.

Cependant, ces instructeurs enseignent cette méthode et ils commettent souvent des erreurs ! Si 80 % des professeurs de droit ne pouvaient enseigner correctement le type de preuves admissibles en cour de justice, notre système de justice aurait de graves problèmes.

### Ce que la valeur-p ne signifie pas

Avant de vous expliquer ce que la valeur-p veut vraiment dire, ce que ce test de signification statistique décrit réellement, voici une courte liste de ce qu'il ne veut pas dire :

- Il ne nous donne pas la probabilité que ces résultats seraient reproductibles dans les mêmes conditions.
- Il ne nous dit pas que les manipulations expérimentales réalisées n'ont eu aucun effet.
- Il ne prouve ni ne réfute une hypothèse particulière.
- Il ne nous donne pas la probabilité qu'une hypothèse soit vraie.

De nombreux scientifiques pensent que ce dernier énoncé est exact, soit que la valeur-p nous donne la probabilité que prendre du Gwynex va résulter en une perte de poids, compte tenu des résultats de l'étude. Pas du tout !

### Ce que la valeur-p signifie

Voici, finalement, ce que veut dire la valeur-p, si souvent citée en recherche, si souvent relayée par les médias lorsqu'il est question de résultats significatifs : la valeur-p nous donne la probabilité des résultats

obtenus, ou de résultats encore plus extrêmes, **en supposant que l'hypothèse nulle soit vraie.**

Une valeur-p de 0,05 (ou de 5 %) représente la probabilité d'avoir les résultats obtenus, lorsque nous évaluons la perte de poids entre le groupe Gwynex et le groupe placebo, en supposant qu'il n'y a pas de différence entre les deux groupes.

La probabilité de nos données, compte tenu de l'hypothèse nulle, est 5 %. Le degré de vérité de l'hypothèse nulle n'est pas mis en question. Nous supposons qu'elle est vraie.

Si nous inversons l'ordre des suppositions, nous obtenons une situation vraiment différente, qui ne peut être évaluée en utilisant la valeur-p. Ce serait la probabilité que notre hypothèse (nulle) soit vraie, étant donné les résultats que nous avons obtenus. C'est ce que les scientifiques veulent savoir ! Mais cette probabilité dépend d'une inférence bayésienne requérant une évaluation de plausibilité préalable.

Encore une fois. La valeur-p est la probabilité des résultats obtenus en supposant qu'il n'y a pas de différence entre le Gwynex et le placebo. D'autre part, les statistiques bayésiennes nous donnent la probabilité qu'il n'y ait pas de différence entre le Gwynex et le placebo, étant donné les résultats obtenus durant l'expérience.

### Se servir du bon outil statistique

Vous ne voyez pas la différence ? Croyez-moi, il y en a une. Ces deux méthodes statistiques sont importantes ; elles sont toutes deux des outils devant être soigneusement choisis. Elles sont parfois considérées comme des ustensiles : un couteau n'est pas très utile pour manger de la soupe, une cuillère ne coupe pas bien.

Il s'avère que de nombreux scientifiques, de très nombreux scientifiques, mangent leur soupe avec un couteau. Lorsque vous entendrez dans les médias qu'une nouvelle étude a atteint un haut niveau de signification statistique, il est probable que l'on se réfère à la valeur-p. Cela ne veut peut-être pas dire ce qu'on laisse entendre... 🤔

#### Source :

The Body of Evidence :  
<http://www.bodyofevidence.ca>

Significance :  
<https://www.youtube.com/watch?v=wIKWwCoA3PU&t=197s>

Traduction de Louis Dubé

**Jonathan Jarry** a une maîtrise en biologie moléculaire. Il a, notamment, passé plusieurs années en laboratoire à mettre au point des tests diagnostiques pour le cancer. En tant que rationaliste et communicateur scientifique, il estime que toute personne devrait posséder des connaissances suffisantes pour faire des choix éclairés au sujet de sa santé et du monde qui l'entoure.