

TD 2 : Causalité et démarche scientifique

Exercice 1 : Magnétiseurs, arnaques ou vrai pouvoir ?

Le **toucher thérapeutique sans contact (TTSC)** est une méthode consistant à sentir, sans contact physique, un supposé champs énergétique humain, et à l'influencer à des fins thérapeutiques. Les pratiquants de cette méthode sont appelés rebouteux, magnétiseurs ou encore "praticien en soin énergétique" (PSE).

En 1996 dans le cadre d'un projet scolaire, **Emily Rosa**, âgée de seulement 9 ans, mène une étude mettant en cause la validité du TTSC. Deux ans plus tard, elle publie son travail dans la revue de médecine *Journal of the American Medical Association*, revue à fort facteur d'impact. Elle devient alors la plus jeune publiante de cette prestigieuse revue.

Voici l'**expérience d'Emily** : elle se positionnait derrière un miroir sans tain, et levait aléatoirement (à l'aide d'un pile ou face) la main gauche ou la main droite. Le PSE de l'autre côté du miroir devait "détecter" quelle main Emily avait levée, ce test était reproduit 5 fois pour chaque PSE. Elle a récolté les réponses de 10 PSE et a observé un taux de réponses correctes de 53%¹.

1. Quelle est l'hypothèse nulle H_0 d'Emily ? Quel autre modèle tente-elle de discréditer ?
2. En accord avec H_0 , quel est le taux le plus probable de bonne réponse ? On le qualifiera par la suite de taux de bonne réponse attendu.
3. Nous avons bien sûr besoin d'une statistique pour réaliser un test d'hypothèse. Parmi les statistiques suivantes, laquelle utiliseriez vous ?
 - la différence entre le taux de réponse attendu et le taux de réponse observé.
 - la somme entre le taux de réponse attendu et le taux de réponse observé.
 - la valeur absolue de la différence entre le taux de réponse attendu et le taux de réponse observé.

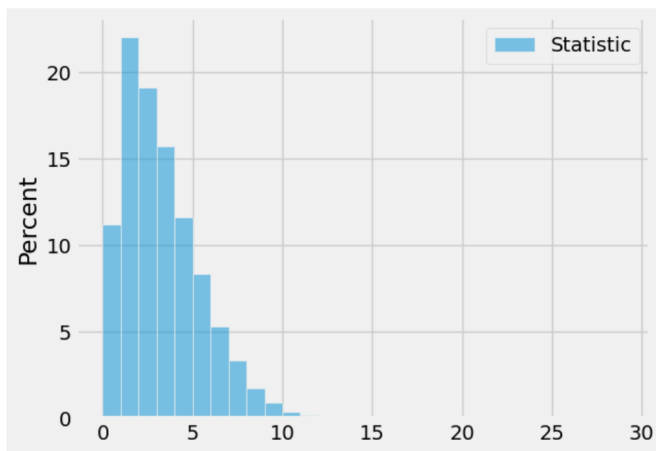
Justifiez et précisez comment s'appelle cette statistique. Calculez la valeur de la statistique correspondant à votre réponse à la question 2.

4. Il nous faut maintenant une distribution de cette statistique correspondant à l'hypothèse H_0 . A cette fin, vous allez reproduire à l'aide de votre voisin l'expérience d'Emily. Placez un objet dans votre main et demandez à votre voisin de deviner la main dans laquelle le papier est placé. Répétez l'opération 50 fois et comptabilisez ses bonnes réponses. En déduire la valeur de la statistique observée et communiquez votre résultat à l'enseignant. Si votre voisin n'est pas rebouteux, l'hypothèse H_0 sera respectée ! Si vous préférez travailler seul, vous pouvez tirer à pile ou face 50 fois, ou utiliser une autre méthode aléatoire de votre choix. L'essentiel est d'obtenir une séquence aléatoire de 0 ou de 1 contenant cinquante éléments.

⚠ En attendant que l'enseignant collecte tous les résultats, vous devez commencer l'exercice 2.

¹A des fins pédagogiques, certains nombres ainsi que le protocole expérimental ont été changés, l'article original est disponible ici <https://jamanetwork.com/journals/jama/fullarticle/187390>

5. L'enseignant a maintenant collecté les résultats de toute la classe dans un tableau et a tracé l'histogramme empirique de la statistique sous H_0 . Pouvez-vous conclure à partir de cette histogramme ?
6. La distribution avec un nombre beaucoup plus importants de répétition (20000) est donnée en figure 1. Le tableau à côté donne le contenu en pourcentage de chaque colonne. Pouvez-vous conclure à présent ?
7. Déterminez au-dessus de quelle valeur de la statistique, le test serait statistiquement significatif en faveur de l'hypothèse alternative (valeur $p < 5\%$). Même question pour un test "hautement significatif sur le plan statistique" (valeur $p < 1\%$). .
8. Quelles critiques peut-on formuler à l'égard de l'étude d'Emily ?



Numéro de bac	Hauteur du bac (%)
0	11.4
1	21.9
2	18.4
3	15.4
4	12.2
5	8.6
6	5.6
7	3.1
8	1.7
9	0.8
10	0.4
11	0.2
12	0
...	...

Figure 1: Distribution empirique de notre statistique selon H_0 obtenues avec 20000 répétitions.

Table 1: Hauteur des colonnes de l'histogramme donnée en Figure 1.

Exercice 2 : Distribution empirique et probabilités

Dans des cas simples, la distribution de la statistique selon H_0 peut aussi être retrouvée de manière analytique à l'aide des probabilités. C'est le cas de la distribution de la figure 1.

1. Quelle est la probabilité suivant H_0 d'avoir n bonnes réponses si on vous teste n fois ? On prendra naturellement $n = 50$ dans nos applications numériques.
2. Quelle est la probabilité d'avoir $n - 1$ bonnes réponses si on vous teste n fois ?
3. Montrez que la probabilité d'avoir p bonnes réponses si on vous teste n fois est de

$$P(p \text{ bonnes réponses}) = \frac{n!}{p!(n-p)!2^n}$$

4. Calculez la hauteur de la colonne du bac 0 et du bac 1 à l'aide de votre réponse précédente. Comparez avec les valeurs contenues dans le tableau 1. Commentez.
5. Expliquez pourquoi le bac 0 est environ deux fois plus petit que le bac 1.
6. Vérifiez à présent le résultat de la question 7 de l'exercice 1.