Chapitre 7.
Tests d'hypothèse nulle

Yann Barsamian

École Européenne de Bruxelles 1

Année scolaire 2023–2024

Plan du chapitre

- Est-ce le fruit du hasard ?
- Une nouvelle notion : tester une hypothèse
- Test d'hypothèse nulle à l'aide de la loi binomiale
- Test d'hypothèse nulle à l'aide de la loi normale

Introduction

Lundi 18 mars, il y avait 8 absents sur 17 élèves. Est-ce le fruit du hasard ? Est-ce justifier de supposer que des élèves ont séché le test ?

Depuis SMS, si je fais la moyenne sur les 17 élèves, depuis début septembre vous avez un pourcentage d'assiduité de $82,16\%$.

$X =$ nombre d'élèves absents à une période donnée. On modélise X par la loi binomiale $\mathcal{B}(17, 0.1784)$.

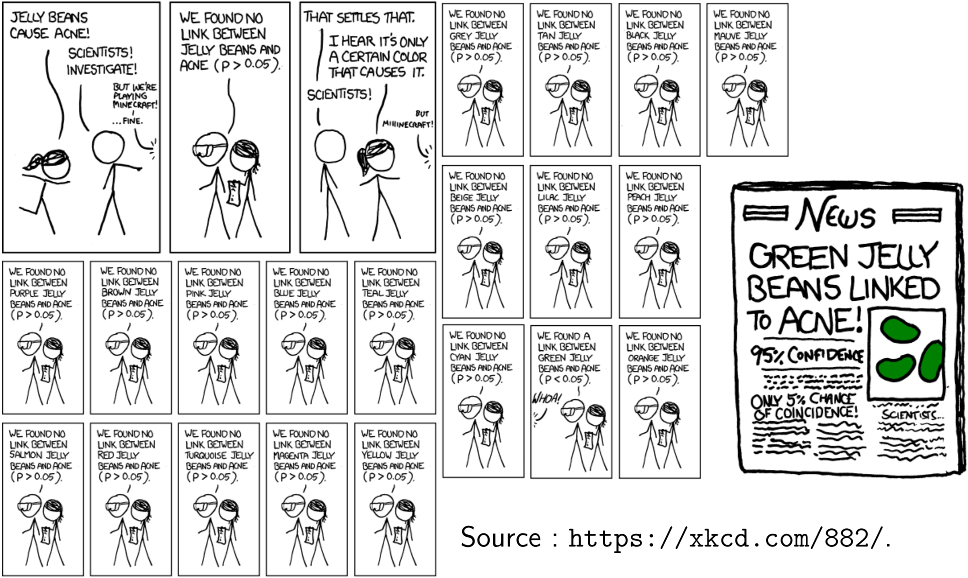
1. Pourquoi l'utilisation de la loi binomiale est-elle justifiée ?
2. Calculez $P(X \geq 8)$
3. Pourquoi ai-je demandé $P(X \geq 8)$ et pas $P(X = 8)$?
4. Considérez-vous que l'événement « $X \geq 8$ » est suffisamment rare pour penser que ce n'est pas le hasard ?
5. Sur les 8 absents, 2 ont été envoyés à l'infirmerie, il restait donc 6 absences injustifiées (selon SMS, le lundi soir à 23h22). Refaire les calculs.

Introduction

$X =$ nombre d'élèves absents à une période donnée. On modélise X par la loi binomiale $\mathcal{B}(17, 0.1784)$.

1. Répétition (17 fois) à l'identique d'événements indépendants (ou quasiment).
2. $P(X \geq 8) \approx 0,0054 = 0,54\%$
3. On veut ici savoir si l'événement devrait être considéré comme rare ou pas ; or avoir encore plus que 8 absents, c'est encore plus rare, et ça devrait encore plus encourager le doute sur le caractère aléatoire des absences de lundi. On va donc considérer l'événement « $X \geq 8$ » et pas « $X = 8$ ».
4. Un événement est décrété « rare » lorsque sa probabilité est $\boxed{\text{plus petite que } 5\%}$ (seuil arbitraire dû à Fisher, 1890–1962). C'est donc rare, et il est raisonnable que j'ai des doutes.
5. Ici on prend Y qui suit $\mathcal{B}(15, 0.1784)$ et on regarde $P(Y \geq 6) \approx 0,0372 = 3,72\%$. C'est encore assez rare pour avoir des doutes.

Attention quand on fait des expériences et/ou quand on interprète des résultats d'expérience...



Source : `https://xkcd.com/882/`.

L'hypothèse nulle, notée H_0, est une hypothèse à propos d'une caractéristique de la population, qui est initialement supposée vraie.

Exemple : En 2010, le salaire moyen annuel d'un étudiant sortant de bachelier aux U.S.A. était de 48 351 \$[1]. Une hypothèse pourrait être H_0 : « le salaire moyen annuel d'un étudiant sortant de bachelier aux U.S.A. était le même en 2023 qu'en 2010 ».

L'hypothèse alternative, notée H_a, est une hypothèse qui vient contredire cette hypothèse nulle.

Exemple : dans l'exemple précédent, H_a serait : « en 2023, le salaire moyen était plus grand que 48 351 \$ »

Quand on teste H_0 contre H_a, on rejettera H_0 en faveur de H_a si l'échantillon permet d'avoir suffisamment de doutes contre H_0. Si l'échantillon n'est pas suffisamment suggestif, H_0 ne sera pas rejeté. Ainsi un test statistique ne peut que :

- rejeter H_0 ou ;
- ne pas rejeter H_0.

1. Source : « Winter 2010 Salary Survey », `www.naceweb.org`

Exercice 1. Un fabricant de balles de tennis a calibré sa machine pour produire des balles de diamètres 3 pouces. Malheureusement, après plusieurs semaines de travail, il a l'impression que les balles ne sont plus conformes. Cela dit, recalibrer la machine est long et coûteux, donc il fait d'abord un test.

Formuler H_0 ainsi que H_a.

Exercice 2. Les lampes fluorescentes compactes utilisent moins d'énergie que les lampes à incandescence, et sont plus chères. L'emballage mentionne « Durée moyenne : 8 000 heures ». Un groupe d'utilisateurs a peur que la publicité soit mensongère. Ils font un test pour en avoir le cœur net.

Formuler H_0 ainsi que H_a.

H_0 est toujours une égalité. Par exemple en appelant μ la durée moyenne d'une lampe en heures : $H_0 =$ « $\mu = 8\ 000$ »

H_a est soit :

- une hypothèse à gauche : $H_a =$ « $\mu < 8\ 000$ » ;
- une hypothèse à droite : $H_a =$ « $\mu > 8\ 000$ » ;
- une hypothèse bilatérale : $H_a =$ « $\mu \neq 8\ 000$ ».

Cela dépend de ce qu'on cherche, et du point de vue.

Exercice 3. Pour chaque paire, indiquer si elle respecte ou pas les règles de formulation d'hypothèses, et pourquoi.

1. $H_0 : \mu = 15,\ H_a : \mu = 15$
2. $H_0 : p = 0,4,\ H_a : p > 0,6$
3. $H_0 : \mu = 123,\ H_a : \mu < 123$
4. $H_0 : \mu = 123,\ H_a : \mu = 125$

Exercice 4. Aux U.S.A., un chercheur se demande si l'autorisation de porter des armes dissimulées permettait de réduire les crimes. Il écrit « The strongest thing I could say is that I don't see any strong evidence that they are reducing crime »(San Luis Obispo Tribune, 23/01/2003). A-t-il testé :

1. H_0 : les armes dissimulées réduisent les crimes contre H_a : les armes dissimulées ne réduisent pas les crimes, ou ;
2. H_0 : les armes dissimulées ne réduisent pas les crimes contre H_a : les armes dissimulées réduisent les crimes ?

Est-ce que l'hypothèse nulle a été rejetée ou pas ? Expliquez.

Exercice 5. Monsieur Tran veut changer le système de notes en S1–S3, et rajouter des « + » et des « - » (afin de pouvoir noter A^+, C^- etc.) Il ne le fera que si plus de 60% des professeurs sont d'accord. Si p représente la proportion de professeurs d'accord, quel test est-ce que monsieur Tran devrait effectuer ?

1. $H_0 : p = 0,6$ contre $H_a : p < 0,6$
2. $H_0 : p = 0,6$ contre $H_a : p > 0,6$