Cours N°7 : Tests statistiques

[Dans ce cours, je ne traite pas les nombreux exemples utilisés par la prof pour vous faire comprendre les buts des tests en vous donnant des cas concrets. ALLEZ EN COURS, c'est important pour vous !]

I. Théorie des tests : approche décisionnelle

A. Introduction

Lorsqu'on effectue une **comparaison** entre deux ou plusieurs séries de données, on observe toujours une **différence plus ou moins grande** entre les paramètres mesurés.
On sait de plus que **si l'on répétait cette comparaison** avec de nouvelles séries de données, o**n n'observerait pas la même chose**.
Le but d'un test est de **déterminer si la différence observée est simplement due au hasard**, par les fluctuations d'échantillonnage, **ou si au contraire la différence observée est bien réelle**.

Avant d'analyser un résultat de test, il faut s'assurer que l'**expérience a été bien faite**, c'est-à-dire que d'une part **les groupes expérimentaux sont comparables** (pour un paramètre donné) et d'autre part que **le nombre de sujets inclus dans chaque groupe est suffisant**.

B. Théorie générale des tests

La théorie des tests correspond au fait que **pour une loi de probabilité X**, **deux hypothèses H_A et H_B, entre lesquelles on doit trancher**, **existent**.
L'hypothèse H_A affirme qu'X est distribué selon $f_A(x)$.
L'hypothèse H_B affirme qu'X est distribué selon $f_B(x)$.

Étant donné que ce sont les **deux seules hypothèses**, elles doivent être considérées comme **contraire**. On a ainsi : $\mathbf{P(H_A) + P(H_B) = 1}$

Pour réaliser le test et trancher entre nos deux hypothèses, on va **tirer au sort un échantillon et observer n valeurs de X**.

C. Règle de décision et évaluation de la règle de décision

La **règle de décision** consiste à établir une **partition à priori de l'ensemble des valeurs observables possibles en deux « régions »** :

- Celle où l'on rejette H_A.
- Celle où l'on rejette H_B.

Des **critères différents** peuvent être choisis pour **évaluer cette règle de décision**, notamment la **probabilité de se tromper en l'utilisant** (on va donc chercher à avoir une probabilité la plus faible possible). Cependant, calculer cette probabilité nécessite de **connaître les probabilités « à priori » de H_A et H_B**.

D. Limites de l'approche décisionnelle

Il n'est en général **pas facile de déterminer les probabilités des hypothèses en concurrence** (même si en connaître une revient à connaître les deux).
De plus, **ces probabilités sont difficiles à évaluer**.

II. Tests de NEYMAN-PEARSON

En science, **on n'accepte les hypothèses que provisoirement. On progresse à chaque fois qu'on rejette une hypothèse**.
On n'accepte jamais l'hypothèse de référence, **on dit qu'on ne la rejette pas**.

A. Hypothèse nulle H_0

L'**hypothèse nulle H_0** est l'**hypothèse de référence** (celle dont on part et qu'**on n'accepte jamais**).
Elle consiste à supposer que les paramètres ou les distributions comparées sont identiques.
C'est l'hypothèse qu'**on veut avoir « peu de chance » de rejeter si elle est vraie** (mais qu'on veut rejeter).

On définit alors le **risque de première espèce α**, encore appelée **erreur de type I**, qui est la **probabilité de rejeter H_0 alors qu'H_0 est vraie** :

$$\boxed{\alpha = P(\text{rejet } H_0 / H_0 \text{ vraie})}$$

Par convention, α = 5%.

B. Hypothèse alternative H_1

L'**hypothèse alternative H_1** est l'**autre hypothèse**, **celle qu'on accepte quand on rejette H_0**.

Elle permet de définir :

- Le **risque de deuxième espèce β** ou **erreur de type II**, qui est la **probabilité de ne pas rejeter H_0 alors qu'H_1 est vraie** :

$$\boxed{\beta = P(\text{non rejet } H_0 / H_1 \text{ vraie})}$$

- La « **puissance du test** », qui est la **probabilité de rejeter l'hypothèse nulle si l'hypothèse alternative est vraie**.

$$\boxed{\text{Puissance} = 1 - \beta = P(\text{rejet } H_0 / H_1 \text{ vraie})}$$

La théorie de Neyman-Pearson vise à fournir les tests les plus puissants possibles pour un risque α fixé par défaut à 5%.

Il existe plusieurs sortes d'hypothèses alternatives :

- L'**hypothèse alternative bilatérale** (**toujours pour des tests**) : exemple $\pi \neq 0,75$.
- L'**hypothèse alternative unilatérale** (pas au programme) : exemple $\pi < 0,75$.
- L'**hypothèse alternative simple** (**pour des calculs de puissance**) : exemple $\pi = 0,5$.
- L'**hypothèse alternative composite**, **composée de plusieurs hypothèses simples**.

C. Test $\pi = \pi_0$ contre $\pi \neq \pi_0$

Soit π une proportion (inconnue) **de personnes concernées par un certain paramètre**. On pose :

- **$H_0 : \pi = \pi_0$** (ce qu'on veut rejeter)
- **$H_1 : \pi \neq \pi_0$**

On observe **p**, **proportion de personnes concernées par le paramètre étudié sur un échantillon de taille n**.

On calcule ensuite le **paramètre z du test** :

$$\boxed{z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}}$$

- **Si $|z| > 1,96$: On rejette H_0. La fréquence π est différente de π_0 !**
- **Si $|z| \leqslant 1,96$: On ne rejette pas H_0.** La **zone de non rejet se trouve à l'intérieur de l'intervalle de pari à 95%**. L'expérience ne permet pas de rejeter H_0 mais **on ne l'accepte pas**. On dit généralement qu'on n'a pas pu trouver de différence significative.
- Il faut **vérifier les conditions de validité !**
  Il suffit de montrer que $n\pi_0 \geqslant 5$ **ET** $n(1-\pi_0) \geqslant 5$

D. Degré de signification

Le **degré de signification** est la **plus petite valeur du risque de première espèce α qu'on aurait pu choisir avec lequel on aurait continué de rejeter H_0.**

Attention : le **risque de première espèce est défini à priori**. En pratique, il est toujours de **5%**. Le **degré de signification est défini à postériori**. En théorie, il est toujours **inférieur à 5%**.

[La valeur de z trouvée correspond à notre ε_α. On cherche ensuite la valeur α qui lui correspond. Si on a une valeur comprise dans un intervalle, on prend la valeur la plus grande des deux bornes de l'intervalle, et pas la plus petite, pour être sur que notre degré de signification est juste (ou du moins possible) !]

E. Puissance d'un test

La **puissance** est la **probabilité pour que l'expérience conduise à rejeter l'hypothèse nulle lorsqu'elle est fausse**. Elle est définie par la relation :

$$\boxed{Puissance = 1 - \beta = P\left(rejet\ H_0 / H_1 vraie\right)}$$

On calcule la puissance pour une **hypothèse alternative simple** jugée **« intéressante »**. **Il faut connaître H_1 pour calculer une puissance**.

On cherche à utiliser des **expériences « puissantes »**, effectuées sur des **nombres assez grands de sujets**, avant de se lancer dans un test statistique.

Ce document, ainsi que tous les cours P1, sont disponibles gratuitement sur http://coursp1bichat-larib.weebly.com