Partie IV : Introduction aux tests d'hypothèse

Introduction

Jeu de Pile ou Face et triche

Karl et Ronald jouent à Pile ou Face.
Karl parie systématiquement sur Pile et Ronald sur Face.

- Au bout de 6 lancers :
  - Karl obtient 1 fois Pile
  - Ronald obtient 5 fois Face

  $\Rightarrow$ Cela vous semble-t-il suspect ?
- Ils continuent. Au bout de 18 lancers :
  - Karl obtient 4 fois Pile
  - Ronald obtient 14 fois Face

  $\Rightarrow$ Cela vous semble-t-il suspect ?

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Introduction

Jeu de Pile ou Face et triche

- Au bout de 6 lancers :
  - Karl obtient 1 fois Pile
  - Ronald obtient 5 fois Face

$\Rightarrow$ Cela vous semble-t-il suspect ? Si $X \sim \mathcal{B}(6, \frac{1}{2})$ alors

$$\mathbb{P}(\mathbf{X} \geq \mathbf{5}) = \mathbf{1} - \mathbf{pbinom}(\mathbf{4}, \mathbf{6}, \mathbf{0.5}) \approx \mathbf{0.109}$$

- Ils continuent. Au bout de 18 lancers :
  - Karl obtient 4 fois Pile
  - Ronald obtient 14 fois Face

$\Rightarrow$ Cela vous semble-t-il suspect ? Si $X \sim \mathcal{B}(18, \frac{1}{2})$ alors

$$\mathbb{P}(\mathbf{X} \geq \mathbf{14}) = \mathbf{1} - \mathbf{pbinom}(\mathbf{13}, \mathbf{18}, \mathbf{0.5}) \approx \mathbf{0.015}$$

Karl a 985 chances sur 1000 de ne pas se tromper en refusant d'attribuer au hasard seul sa perte au jeu.

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Formulation des hypothèses

Jeu de Pile ou Face et triche

On souhaite déterminer si Ronald est un tricheur ou est honnête. On confronte alors ces deux hypothèses :

- H_0 : Ronald est honnête, chaque lancer a une chance sur deux de faire Face. (Hypothèse nulle)
- H_1 : Ronald est un tricheur, il utilise une pièce qui a plus de chances de faire Face. (Hypothèse alternative)

Ces deux hypothèses ne jouent pas des rôles symétriques :

- la première suppose que Ronald n'a pas d'effet sur le jeu, que seul le hasard intervient ;
- tandis que la second considère qu'un processus supplémentaire (par exemple la triche, utilisation d'une pièce truquée) modifie les résultats par rapport au premier cas de figure.

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Formulation mathématiques des hypothèses

Jeu de Pile ou Face et triche

Le modèle probabiliste doit permettre de voir si l'échantillon observé est une « exception » ou s'il ne diffère pas significativement de la **majorité** des autres échantillons choisis au hasard.

Soit X la v.a comptant le nombre de Face obtenues après n lancers (ici $n = 18$), les hypothèses H_0 et H_1 peuvent se réécrire comme des hypothèses sur la loi de X, ce qui se traduit par un **test paramétrique** :

- H_0 : $X \sim \mathcal{B}(n, p)$ où $p = \frac{1}{2}$. (Hypothèse nulle)
- H_1 : $X \sim \mathcal{B}(n, p)$ où $p > \frac{1}{2}$. (Hypothèse alternative)

$\Rightarrow$ p n'est connue que dans le cas H_0 où Ronald est honnête, on ne peut donc calculer explicitement de probabilité que dans le cas de l'hypothèse H_0. On dit que H_0 est **testable** et que H_1 ne l'est pas directement.

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Principe du test d'hypothèse

Analogie avec un procès en justice

- On se place sous l'hypothèse H_0 (présomption d'innocence) pour voir s'il est raisonnable de maintenir cette hypothèse au vu des données observées (éléments de l'enquête).
- À l'issue du test statistique (après enquête), on pourra prendre la décision de **rejeter l'hypothèse** H_0 (de condamner Ronald) si l'on considère les résultats de l'expérience comme incompatibles avec cette l'hypothèse, jugée fortement improbable au vu des données.
- Si au contraire les résultats sont compatibles avec l'hypothèse, on dira que l'on ne rejette pas H_0 (Ronald est acquitté).

$\Rightarrow$ Cela ne signifie pas que l'on ait la certitude que H_0 soit vrai (être acquitté est différent que d'être innocent), mais que l'**on ne dispose pas d'assez de preuves pour la rejeter** (c'est-à-dire ici pour accuser Ronald).

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Risques d'erreur

Quand on prend des décisions en se basant sur des tests statistiques, on n'est pas à l'abri de commettre des erreurs. Elles sont de deux types :

- **Erreur de type I** : Rejeter à tord H_0, cela revient à accuser un innocent (erreur judiciaire).
- **Erreur de type II** : Accepter à tord H_0, cela revient à innocenter un coupable.

Etat \ Décision	Accepter H_0	Rejeter H_0
H_0 vraie	Pas d'erreur	Erreur de type I
H_1 vraie	Erreur de type II	Pas d'erreur

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Risques d'erreur

- **Erreur de type I** : Probabilité de rejeter H_0 alors que H_0 est vraie :

$$\alpha = \mathbb{P}(\text{rejeter } H_0 \mid H_0 \text{ est vraie})$$

  Proposition contraire : accepter H_0 alors que H_0 est vraie (vraisemblance) :$1 - \alpha = \mathbb{P}(\text{accepter } H_0 \mid H_0 \text{ est vraie})$

- **Erreur de type II** : Probabilité d'accepter H_0 alors que H_1 est vraie :

$$\beta = \mathbb{P}(\text{accepter } H_0 \mid H_1 \text{ est vraie})$$

  Proposition contraire : rejeter H_0 lorsque que H_1 est vraie (puissance) :$1 - \beta = \mathbb{P}(\text{rejeter } H_0 \mid H_1 \text{ est vraie})$

Etat \ Décision	Accepter H_0	Rejeter H_0
H_0 vraie	Pas d'erreur (proba $1 - \alpha$)	Erreur (proba α)
H_1 vraie	Erreur (proba β)	Pas d'erreur (proba $1 - \beta$)

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Région de rejet

La prise de décision se fera en fonction de l'appartenance des données observées à une certaine région de valeurs. Ici, on a envie :

- D'accuser Ronald de tricherie si le nombre de Faces est très élevé.
- De ne pas l'accuser si le nombre de Faces est raisonnable.

On cherche donc une région, que l'on notera W_α appelée **région critique**, composée de valeurs élevées, dans laquelle on a peu de chances de tomber si jamais H_0 est vraie :

$$\mathbb{P}(X \in W_\alpha \mid H_0) \leq \alpha$$

On choisit de **rejeter** H_0 dans cette région.

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Région de rejet

Dans notre exemple, on prend par exemple $\alpha = 0.05$ et on rejette H_0 si le nombre de Faces observé est trop grand au niveau α, c'est-à-dire s'il est plus grand qu'une valeur seuil k_α qui dépend du risque d'erreur que l'on est prêt à accepter.

Pour trouver cette région W_α la plus grande possible, on doit chercher tous les k tels que

$$\mathbb{P}(X \geq k) \leq \alpha$$

et prendre la plus petite parmi elles. Par exemple pour $X \sim \mathcal{B}(6, \frac{1}{2})$ on a

$$\mathbb{P}(X \geq 6) = 0.0156, \quad \mathbb{P}(X \geq 5) = 0.109$$

donc $W_\alpha = \{6\}$, tandis que pour $X \sim \mathcal{B}(18, \frac{1}{2})$ on a

$$W_\alpha = \{13, 14, 15, 16, 17, 18\}$$

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Notion de p-valeur

Si on prend un risque $\alpha = 0.01$ alors la région critique sera

$$W_\alpha = \{15, 16, 17, 18\}$$

qui ne contient pas la valeur observée 14. Donc **entre les deux niveaux de risques il y a une valeur α où on change de décision**, **cette valeur s'appelle la p-valeur.**

Dans notre exemple, pour quel niveau α a-t-on

$$W_\alpha = \{14, 15, 16, 17, 18\} \text{ ?}$$

On obtient

$$\boxed{\alpha = \mathbb{P}(X \geq 14) = 0.015}$$

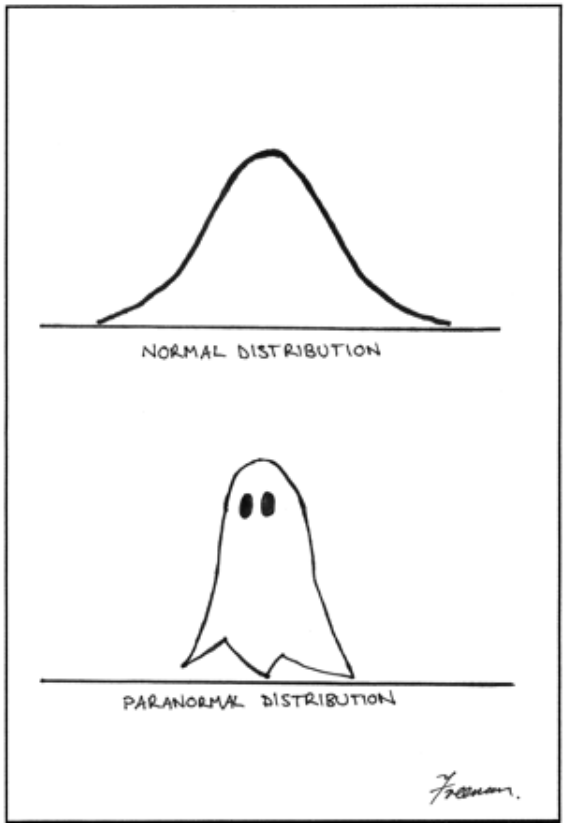
La p-valeur est donc par définition la probabilité sous l'hypothèse nulle d'observer des données au moins aussi grandes que la donnée observée.

Crédits : A. Jebrane (cours statistiques L2 psychologie)

Notion de *p*-valeur

Un domaine d'application : l'étude du paranormal

Afin de tester un potentiel don d'un individu, on peut soumettre ce dernier à une épreuve aléatoire (par exemple : le test des cartes de Zener vu en TP), et quantifier *via* la *p*-valeur si le résultat obtenu peut être considéré comme extraordinaire sous l'hypothèse nulle.

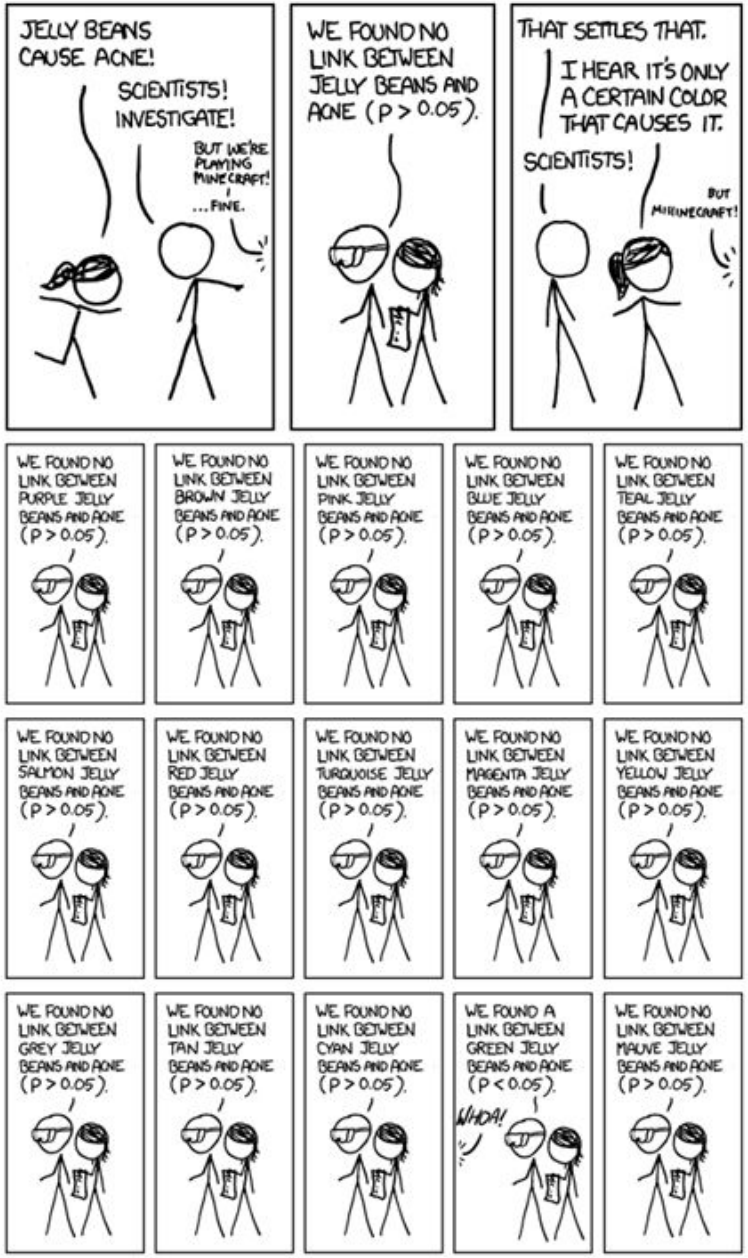


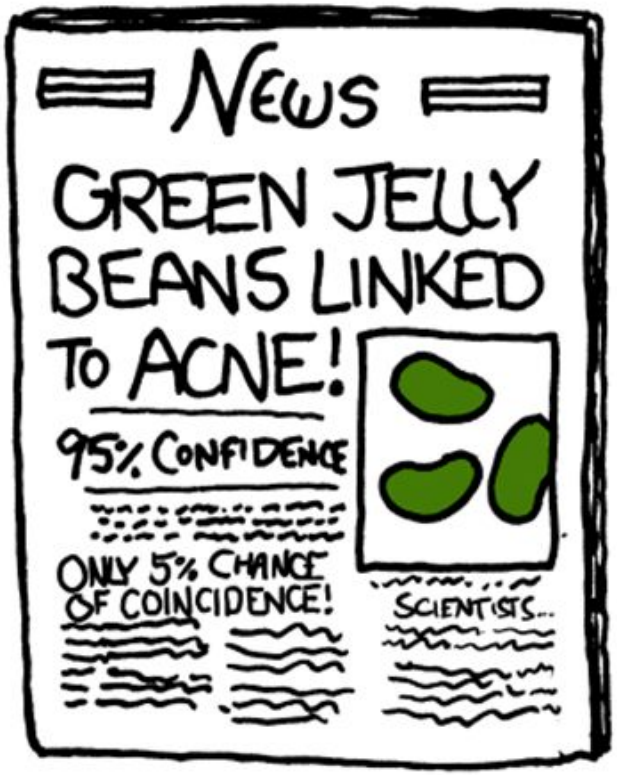
Notion de p-valeur en vidéo

https://www.youtube.com/watch?v=xVIt51ybvu0

https://www.youtube.com/watch?v=PRtwo1j0y2I

Attention aux tests multiples





http://xkcd.com/882/

Quelques mots de conclusion pour ce cours

Ce qu'il faut retenir de tout cela

- Savoir résumer et présenter des données.
- Qu'est qu'une variable aléatoire ? une densité de probabilité ? une fonction de répartition ?
- Savoir manipuler la règle de la somme, du produit et de Bayes.
- Connaitre et reconnaitre la loi normale et autres lois usuelles .
- Étudier les propriétés des estimateurs (biais, variance, etc.) ainsi que leur loi.
- Comprendre et expliciter les intervalles de confiance, tests d'hypothèse et calcul de p-valeur.
- Interpréter des résultats statistiques et éviter les écueils mentionnés dans ce cours.