Tests d'hypothèses

N. Jégou

Université Rennes 2

M1 2SEP

Big Picture

- Idée (très) générale :
  Examiner une question au regard de données
- Formalisation du problème :
  - Modèle statistique
  - Des données (sensées suivre le modèle)
  - Deux hypothèses
    - Hypothèse nulle : $\mathcal{H}_0$
    - Hypothèse alternative : $\mathcal{H}_1$
- Conclusion pratique :
  - Calcul d'une statistique de test
  - Décision
    - On conserve $\mathcal{H}_0$
    - On rejete $\mathcal{H}_0$

Examen d'une question

- Objectif : répondre à une question
  - Le traitement est-il efficace sur les rats de laboratoire ? (Exemple 1)
  - Le médicament est-il plus efficace qu'un placebo à faire baisser la pression artérielle ? (Exemple 2)
  - ...
- Les données doivent aider à répondre
  - Les données donnent-elle plus de crédit à la réponse oui ou à la réponse non ?
  - Problème : les données sont échantillonnées donc aléatoires
    - Elles peuvent conforter l'une des réponses sur l'échantillon dont on dispose
    - Mais un nouvel échantillon pourrait conforter l'autre réponse

Modèle statistique

Modèle statistique : hypothèse(s) sur la loi dans la population

- Exemple 1 La guérison d'un rat est une variable aléatoire suivant une loi de Bernoulli :

$$X \sim \mathcal{B}(p)$$

- Exemple 2 La pression artérielle X suit une loi normale :

$$X_{\text{med}} \sim \mathcal{N}(\mu_{\text{med}}, \sigma^2) \qquad X_{\text{pla}} \sim \mathcal{N}(\mu_{\text{pla}}, \sigma^2)$$

Les données

Les données sont supposées être i.i.d. selon la loi spécifiée dans le modèle :

- Exemple 1 Des rats sont soumis traitement :

$$X_i \sim \mathcal{B}(p) \qquad i = 1, \dots n$$

- Exemple 2 Des patients reçoivent le médicament, d'autres[1] le placébo :

$$X_{i,\text{med}} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_{\text{med}}, \sigma^2) \qquad X_{i,\text{pla}} \underset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_{\text{pla}}, \sigma^2)$$

[1] Les échantillons sont alors indépendants

Hypothèse Nulle

L'hypothèse nulle $\mathcal{H}_0$

- doit amener des éléments de réponse à la question
- porte sur le(s) paramètre(s) des lois du modèles
- est simple[2] :

  Exemple 1 $\mathcal{H}_0 : p = 0.8$

  Exemple 2 $\mathcal{H}_0 : \mu_{\text{med}} = \mu_{\text{pla}}$ (pas d'effet du médicament)

[2] La raison est qu'on sera amené à considérer une loi sous $\mathcal{H}_0$

Hypothèse alternative

L'hypothèse nulle $\mathcal{H}_1$

- ne recouvre pas l'hypothèse nulle
- présente un intérêt par rapport à la question
- peut être composite

Exemple 1 $\mathcal{H}_1 : p < 0.8$

Exemple 2 $\mathcal{H}_1 : \mu_{\text{med}} < \mu_{\text{pla}}$ (effet du médicament)

Statistique de test

- C'est une variable aléatoire fonction des données :

$$Z = f(X_1, X_2, \dots, X_n)$$

- On dispose de sa (seule) réalisation sur les données de l'échantillon :

$$z = f(x_1, x_2, \dots, x_n)$$

- Elle est construite pour pouvoir utiliser sa loi sous $\mathcal{H}_0$
  - loi qui donne la distribution des valeurs attendues de T sous $\mathcal{H}_0$
  - situer la réalisation z dans cette distribution permet de confronter les données à $\mathcal{H}_0$

Statistique de test : Exemple 1

- Le nombre de guérison Z suit une loi binomiale :

$$Z = \sum_{i=1}^{n} X_i \sim \mathcal{B}(n, p)$$

- Loi inconnue car p inconnu
- Mais sous $\mathcal{H}_0$: $p = p_0$,

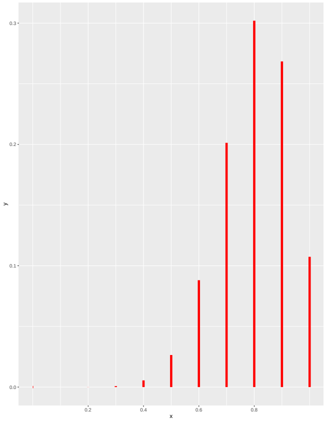
$$Z = \sum_{i=1}^{n} X_i \sim \mathcal{B}(n, p_0)$$

- La loi de $\hat{p}$, fréquence empirique de guérison, dérive de celle de Z :

$$\hat{p} = Z/n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Distribution attendue sous $\mathcal{H}_0$: Exemple 1

- Cas $n = 10$ et $\mathcal{H}_0 : p = 0.8$



- Sous $\mathcal{H}_0$ on s'attend à $\hat{p} \approx 0.8$

Statistique de test : Exemple 2

- La statistique de test est

$$Z = \frac{(\bar{X}_{\text{med}} - \bar{X}_{\text{plac}}) - (\mu_{\text{med}} - \mu_{\text{pla}})}{\hat{\sigma} \times \sqrt{\frac{1}{n_{\text{med}}} + \frac{1}{n_{\text{pla}}}}}$$

- Avec

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n_{\text{med}}} (X_i - \bar{X}_{\text{med}})^2 + \sum_{i=1}^{n_{\text{pla}}} (X_i - \bar{X}_{\text{pla}})^2}{n_{\text{med}} + n_{\text{pla}} - 2}$$

- Elle fondée sur
  - l'écart entre moyennes de groupes $\bar{X}_{\text{med}} - \bar{X}_{\text{pla}}$
  - les effectifs de groupes n_{med} et n_{pla}
  - l'estimateur de la variance $\hat{\sigma}^2$

Statistique de test : Exemple 2

- La loi de Z est inconnue car μ_{med} et μ_{pla} sont inconnus
- Sous $\mathcal{H}_0$: $\mu_{\mathrm{med}} = \mu_{\mathrm{pla}}$,

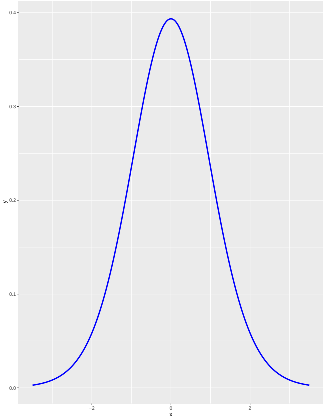
$$Z = \frac{(\bar{X}_{\mathrm{med}} - \bar{X}_{\mathrm{pla}})}{\hat{\sigma} \times \sqrt{\frac{1}{n_{\mathrm{med}}} + \frac{1}{n_{\mathrm{pla}}}}}$$

- Z suit alors une loi de Student :

$$Z \sim \mathcal{T}_{(n_{\mathrm{med}} + n_{\mathrm{pla}-2})}$$

Distribution attendue de Z sous $\mathcal{H}_0$

- Cas $n_{\mathrm{med}} = n_{\mathrm{pla}} = 10$, distribution de Student à 18 d.d.l.



- Sous $\mathcal{H}_0$, on s'attend à une valeur de la stat. de test ≈ 0

Statistique observée vs Distribution sous $\mathcal{H}_0$

Exemple 1

- Si 5 rats guérissent : $\hat{p} = 0.5$

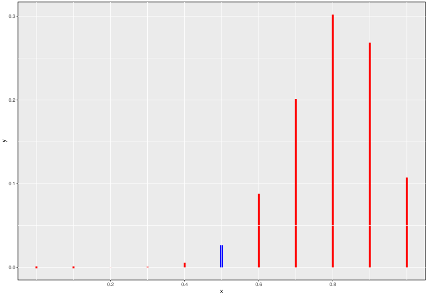
- $\mathcal{H}_0$ reste acceptable ?
  $\hat{p}$ faible car l'échantillon est "anormal" ?

- $\mathcal{H}_0$ n'est pas crédible ?
  $\Rightarrow$ Rejet de $\mathcal{H}_0$?

Statistique observée vs Distribution sous $\mathcal{H}_0$

Exemple 1

- $\mathbb{P}(\hat{p} = 0.5|\mathcal{H}_0) \approx 0.026$
- $\mathbb{P}(\hat{p} \leq 0.5|\mathcal{H}_0) \approx 0.033$



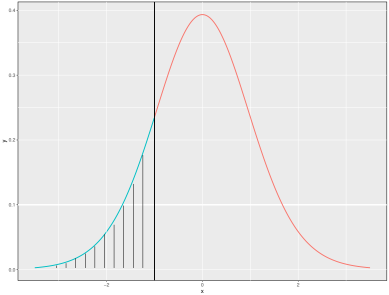
Statistique observée vs Distribution sous $\mathcal{H}_0$

Exemple 2

- Supposons que l'on observe $z = -1$ sur l'échantillon
- Est-ce une valeur suffisament faible pour rejeter $\mathcal{H}_0$ et considérer le médicament efficace ?
- Au contraire, est-elle trop peu éloignée de 0 pour remettre en cause $\mathcal{H}_0$?

Statistique observée vs Distribution sous $\mathcal{H}_0$

Exemple 2 $\mathbb{P}(Z < -1|\mathcal{H}_0) \approx 0.17$



Probabilité critique : p-value

- Le logiciel R donne une probabilité pour aider à la décision
- La probabilité, sous $\mathcal{H}_0$, d'observer une réalisation au moins aussi défavorable à $\mathcal{H}_0$ que celle dont on dispose
- Elle permet de juger de l'accord des données avec $\mathcal{H}_0$

- C'est la probabilité critique ou `p-value`
- Exemple 1
  $\mathrm{p.value} = \mathbb{P}\left(\hat{p} \leq 0.5 | \mathcal{H}_0\right) \approx 0.033$
- Exemple 2
  $\mathrm{p.value} = \mathbb{P}(Z < -1 | \mathcal{H}_0) \approx 0.17$

Prise de décision et erreur

- α : risque de première espèce
- β : risque de deuxième espèce

Décision / Réalité

	$\mathcal{H}_0$ conservée	$\mathcal{H}_0$ rejetée
$\mathcal{H}_0$ vraie		$\boldsymbol{\alpha}$
$\mathcal{H}_0$ fausse	$\boldsymbol{\beta}$	

Définition : Erreur de première espèce

Definition (Erreur de première espèce α)

- Probabilité de rejeter à tort $\mathcal{H}_0$
- Probabilité de rejeter $\mathcal{H}_0$ alors que $\mathcal{H}_0$ est vraie

Choix de α

- Fixée par l'utilisateur : en général $\alpha = 5\%$ (ou 1%)
- C'est le risque maximum que l'on accorde à la possibilité de se tromper en rejetant $\mathcal{H}_0$

Définition : Erreur de seconde espèce

Definition (Erreur de seconde espèce β)

- Probabilité de conserver $\mathcal{H}_0$ alors que c'est $\mathcal{H}_1$ qui est vraie
- Probabilité de conserver $\mathcal{H}_0$ à tort
- Probabilité de rejeter $\mathcal{H}_1$ alors que $\mathcal{H}_1$ est vraie
- Probabilité de rejeter $\mathcal{H}_1$ à tort

Remarque : Cette erreur est non contrôlée car on ne connaît pas (en général) la loi sous $\mathcal{H}_1$

Prise de décision

- On se fixe α (souvent $\alpha = 5\%$ ou $\alpha = 1\%$)

- On compare la probabilité critique p_{value} à α
  - Si la probabilité critique $p_{\text{value}} \leq \alpha$
    Les données sont jugées trop peu en accord avec $\mathcal{H}_0$
    $\rightarrow$ Rejet de $\mathcal{H}_0$

  - Si la probabilité critique $p_{\text{value}} > \alpha$
    Les données ne sont pas jugées trop en désaccord avec $\mathcal{H}_0$
    $\rightarrow$ Conservation (non rejet) de $\mathcal{H}_0$