Généralité sur les tests

1. Définition d'une hypothèse de base notée H_0 et dite l'hypothèse nulle.
2. L'hypothèse alternative est notée H_1.
3. On postule que la vérité se trouve dans H_0 ou H_1 et qu'une seule des deux est vraie.
4. Un test statistique est une procédure permettant de trancher entre H_0 et H_1 au vu des observations.
5. On distingue deux types d'erreurs possibles pour une décision

	« Vérité »	
Décision	H_0	H_1
H_0	OK $(1-\alpha)$	**erreur de type II β**
H_1	**erreur de type I α**	OK $(1-\beta)$

6. (6a) α est la probabilité de rejeter de H_0 à tort (erreur de type I). Elle est fixée a priori et est appelée le **seuil du test**.
   $1-\alpha$ est souvent appelé **le niveau du test** et

$$1-\alpha = \mathbb{P}\{\text{conservation de } H_0 \mid H_0 \text{ est vraie}\}$$

   C'est une mesure de la capacité d'un test à ne pas remettre en cause une hypothèse initiale vraie

   (6b) β est la probabilité de rejeter de rejeter de H_0 à tort, ou de conserver H_0 à tort (erreur de type II).
   La probabilité $1-\beta$ est appelée **la puissance du test** mais est souvent difficile à calculer.
   Notons que $\beta = \mathbb{P}\{\text{conserver } H_0 \mid H_1 \text{ est vraie}\}$ et

$$1-\beta = \mathbb{P}\{\text{rejet de } H_0 \mid H_1 \text{ est vraie}\}$$

   autrement dit, la puissance du test est une mesure de la capacité d'un test à remettre en cause une hypothèse initiale fausse
7. En général, les deux risques α et β et donc le niveau et la puissance d'un test évoluent en sens contraire ! En général, pour un risque fixé, on recherche un test de puissance maximale.

8. Il y a donc généralement une asymétrie dans le traitement des deux hypothèses H_0 et H_1 par un test. Le choix de H_0 est souvent guidé par les principes suivants :
   (a) H_0 est une hypothèse solidement établie et qui n'a pas été réfutée par des expériences antérieures
   (b) H_0 correspond à une hypothèse de prudence dans le cas où choisir H_1 à tort aurait des conséquences dramatiques pour l'utilisateur
   (c) H_0 est la seule hypothèse facile à formuler !
   (d) H_0 est une hypothèse à laquelle on tient particulièrement et pas nécessairement pour des raisons avouables !

Structure générale d'un test

1. Formulation des deux hypothèses H_0 et H_1
2. Fixer le niveau du seuil α du test
3. Choix d'une statistique T (en général un estimateur pour un test sur la valeur d'un paramètre)
4. Déterminer une région de conservation $A_{1-\alpha}(H_0)$ (ou son complémentaire, une région de rejet ou critique $\mathrm{R}_{1-\alpha}(H_0)$) de l'hypothèse H_0, c'est à dire de valeurs observables de T qui soient compatibles (incompatibles) avec H_0, telle que

$$\mathbb{P}\{T \in A_{1-\alpha}(H_0)\} \geq 1-\alpha \quad \text{ou encore} \quad \mathbb{P}\{T \in \mathrm{R}_{1-\alpha}(H_0)\} \leq \alpha$$

5. Calculer T à partir des données et utiliser la règle de décision suivante
   (C) **Conservation de** H_0 si la valeur observée de T est dans $A_{1-\alpha}(H_0)$
   (R) ou encore **Rejet de** H_0 si la valeur observée de T n'est pas dans $A_{1-\alpha}(H_0)$ (i.e. est dans la région critique $\mathrm{R}_{1-\alpha}(H_0)$)

- La même démarche que pour le croisement des rats nous permet de construire un test du type

$$H_0 : g(\theta) = g(\theta_0)$$

contre

$$H_1 : g(\theta) \neq g(\theta_0) \text{ ou } H_1 : g(\theta) < g(\theta_0) \text{ ou } H_1 : g(\theta) > g(\theta_0)$$

associé à une région de confiance $\mathrm{R}_{1-\alpha}(g(\theta))$ pour le paramètre $g(\theta)$ par la règle

$$\widehat{g(\theta)}_n \in A_{1-\alpha}(H_0) \Longleftrightarrow g(\theta_0) \in \mathrm{R}_{1-\alpha}(g(\theta))$$

en utilisant une région de confiance $\mathrm{R}_{1-\alpha}(g(\theta))$ respectivement bilatérale, unilatérale inférieure, unilatérale supérieure

- Nous nous contentons de formuler la règle de décision pour certains d'entre eux.

Test exact de la moyenne de $\mathcal{N}(m, \sigma^2)$ avec σ^2 inconnue

1. $H_0 : m = m_0$ (où m_0 est une valeur fixée) contre $H_1 : m \neq m_0$
2. Le seuil α est fixé
3. Région de conservation de H_0 :

$$A_{1-\alpha}(m_0) = \left[\, m_0 \pm t_{n-1,1-\alpha/2} \sqrt{\frac{\tilde{s}^2_X}{n}}\, \right] \quad (45)$$

où $t_{n-1,1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ d'une loi Student à $(n-1)$ degré de liberté.

4. **Règle de décision :**
   (D) **Conservation de** H_0 si la valeur observée de $\overline{X}_n \in A_{1-\alpha}(H_0)$:

$$\overline{X}_n \in \left[\, m_0 \pm t_{n-1,1-\alpha/2} \sqrt{\frac{\tilde{s}^2_X}{n}}\, \right]$$

et rejet dans le cas contraire.

Formulation classique

La condition d'appartenance à $A_{1-\alpha}(H_0)$ est reformulée de manière équivalente en termes de la fonction $D_n(m_0)$ dite de décision : on conserve H_0 avec un risque de type I au plus de α ssi

$$D_n(m_0) = \frac{\overline{X}_n - m_0}{\sqrt{\frac{\tilde{s}^2_X}{n}}} \in [-t_{n-1,1-\alpha/2}, t_{n-1,1-\alpha/2}]$$

Test asymptotique de la moyenne m d'une loi pour de grands échantillons

1. $H_0 : m = m_0$ (où m_0 est une valeur fixée) contre $H_1 : m \neq m_0$
2. Le seuil α est fixé
3. **Règle de décision :**
   (D) (σ^2 connue) **Conservation de** H_0 si la valeur observée de $\overline{X}_n \in A_{1-\alpha}(H_0)$:

$$\overline{X}_n \in \left[\, m_0 \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}\, \right]$$

et rejet dans le cas contraire.

   (D) (σ^2 est inconnue) **Conservation de** H_0 si la valeur observée de $\overline{X}_n \in A_{1-\alpha}(H_0)$:

$$\overline{X}_n \in \left[\, m_0 \pm z_{1-\alpha/2} \sqrt{\frac{\tilde{s}^2_X}{n}}\, \right]$$

et rejet dans le cas contraire.

où $z_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ d'une loi normale centrée-réduite

Formulation classique

Même statistique de décision et même forme que précédemment en remplaçant les quantiles de Student par ceux de la loi normale centré-réduite

Test exact de la variance d'une $\mathcal{N}(m, \sigma^2)$ avec m inconnue

1. $H_0 : \sigma^2 = \sigma_0^2$ (où σ_0^2 est une valeur fixée) contre $H_1 : \sigma^2 \neq \sigma_0^2$
2. Le seuil α est fixé
3. Région de conservation de H_0 :

$$A_{1-\alpha}(\sigma_0^2) = \left[\, \sigma_0^2 \frac{x_{n-1,\alpha/2}}{n-1}, \sigma_0^2 \frac{x_{n-1,1-\alpha/2}}{n-1}\, \right] \quad (46)$$

si $x_{n-1,\alpha/2}, x_{n-1,1-\alpha/2}$ sont les quantiles d'ordre $\alpha/2, 1-\alpha/2$ d'une loi du chi-deux à $(n-1)$ degrés de liberté.

4. **Règle de décision :**
   (D) **Conservation de** H_0 si la valeur observée de $\tilde{s}^2_X \in A_{1-\alpha}(H_0)$:

$$\tilde{s}^2_X = \frac{1}{n-1} \sum_i (x_i - \overline{X}_n)^2 \in \left[\, \sigma_0^2 \frac{x_{n-1,\alpha/2}}{n-1}, \sigma_0^2 \frac{x_{n-1,1-\alpha/2}}{n-1}\, \right]$$

et rejet dans le cas contraire.

Formulation classique

La condition d'appartenance à $A_{1-\alpha}(H_0)$ est reformulée de manière équivalente en termes de la fonction de décision $D_n(\sigma_0^2)$ dite de décision : on conserve H_0 avec un risque de type I au plus de α ssi

$$D_n(\sigma_0^2) = \frac{(n-1)\tilde{s}^2_X}{\sigma_0^2} \in [x_{n-1,\alpha/2}, x_{n-1,1-\alpha/2}]$$

Test d'égalité de λ à une valeur donnée pour $\mathrm{Exp}(\lambda)$

1. $H_0 : \lambda = \lambda_0$ (où λ_0 est une valeur fixée) contre $H_1 : \lambda \neq \lambda_0$
2. Le seuil α est fixé
3. Région de conservation de H_0 :

$$A_{1-\alpha}(\lambda_0) = \left[\ \lambda_0 \frac{2n}{x_{2n,1-\alpha/2}}, \lambda_0 \frac{2n}{x_{2n,\alpha/2}}\ \right]$$

si $x_{2n,\alpha/2}, x_{2n,1-\alpha/2}$ sont les quantiles d'ordre $\alpha/2, 1-\alpha/2$ d'une loi du chi-deux à $2n$ degrés de liberté.

4. **Règle de décision :**
   (D) **Conservation de** H_0 si la valeur observée de $\widehat{\theta}_n = n/T_n \in A_{1-\alpha}(H_0)$:

$$\frac{n}{T_n} \in \left[\ \lambda_0 \frac{2n}{x_{2n,1-\alpha/2}}, \lambda_0 \frac{2n}{x_{2n,\alpha/2}}\ \right]$$

et rejet dans le cas contraire.

À la lumière de ces quelques tests on peut faire les quelques remarques suivantes (analogues à celles faites sur les $\mathbb{R}$) :

1. **Si n est fixé :** quand le risque α diminue, ou encore le niveau $1-\alpha$ du test augmente, alors le domaine des valeurs des observations compatibles avec H_0 s'élargit, i.e. on augmente la largeur de sorte à déclarer « normales » de plus plus de valeurs des observations : on adopte une attitude de plus en plus conservatrice via à vis de H_0.
   Il faut alors faire attention à la puissance du test, i.e. à sa capacité à détecter que l'hypothèse de base est fausse.
2. **Pour un niveau $1-\alpha$ fixé :** si n devient de plus en plus grand alors la largeur du domaine de conservation de H_0 a tendance à diminuer. L'incertitude diminue avec la taille des données de sorte que si H_0 est vraie, alors des échantillons de taille de plus en plus importantes doivent donner des estimations des paramètres de plus en plus proches de celles proposées par H_0.

Une spécifié des tests : la p-valeur

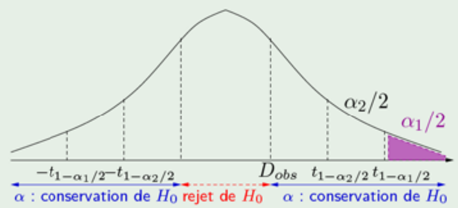
- Pour chaque test, on fixe le risque α de première espèce à savoir, α est la probabilité de rejeter à tort l'hypothèse nulle H_0.
- La **p-valeur** α_c **(ou probabilité critique) d'un test** peut être définie comme la plus petite valeur du seuil α conduisant au rejet de H_0. Ou encore, c'est la valeur minimale du risque de type I que l'on prend en rejetant H_0. Autrement dit,
  - pour tout $\alpha \geq \alpha_c$, on rejette H_0 au profit de H_1
  - et pour tout $\alpha < \alpha_c$, on conserve H_0.
- Un intérêt direct est que cette p-valeur permet de connaître la conclusion du test pour toute valeur du risque α.
  La p-valeur n'est pas toujours simple à calculer mais elle est fournie par tous les logiciels donnant accès à la méthodologie des tests.
- En pratique, on interprète la p-valeur comme suit :
  - une petite valeur de la p-valeur indique que l'hypothèse H_0 a peu de chances d'être vraie au vu des données.
  - Au contraire, une assez grande valeur de la p-valeur indique qu'au vu des données, l'hypothèse H_0 peut difficilement être remise en cause.

Exemple de calcul de la p-valeur pour le test sur la moyenne d'une $\mathcal{N}(m, \sigma^2)$

On part de la formulation classique : on conserve H_0 si la valeur observée de $D_n(m_0)$

$$D_{n,obs}(m_0) = \frac{\overline{X}_n - m_0}{\sqrt{\frac{s^2_X}{n}}} \in [\,-t_{n-1,1-\alpha/2}, t_{n-1,1-\alpha/2}]$$

Un représentation graphique de la situation :



Une fois avoir placé la valeur $D_{n,obs}$ sur l'axe, tant que les quantiles de la loi de Student encadrent $D_{n,obs}$ on conserve H_0 mais dès que l'un des deux croisent la valeur $D_{n,obs}$, on rejette. D'où la valeur de la probabilité critique α_c correspond à

$$\alpha_c = \mathbb{P}\{|st_{n-1}| > D_{n,obs}\} = 2[1 - F_{st_{n-1}}(D_{n,obs})]$$

où $F_{st_{n-1}}$ est la fonction de répartition d'une Student st_{n-1}.

Quid du cas non-gaussien et/ou petits échantillons ?

- Faire des statistiques inférentielles avec un modèle statistique qui ne soit pas constitué d'une famille de lois spécifiques :
  - statistiques pour des modèles semi-paramétriques
  - statistiques pour des modèles non-paramétriques
- Test d'adéquation à une loi, test d'indépendance, estimation non-paramétrique de densité, ...
- Mais apporter des réponses pertinentes avec des outils de « statistiques traditionnelles » lorsqu'on travaille avec des petits échantillons est un pb délicat, voire impossible dans certain cas car il y a souvent une incompatibilité méthodologique.
  - ↪ Adopter le point de vue Baysesien des statistiques [Rob06]