Université de Rennes

ENSAI École nationale de la statistique et de l'analyse de l'information

Chapitre 5. Tests statistiques - Introduction raisonnée

Claude Petit, université de Rennes - claude.petit@univ-rennes.fr

Nov. 2026

Pourquoi faire un test statistique ?

Une observation peut s'écarter d'une valeur annoncée à cause du hasard d'échantillonnage.

Question fondamentale

L'écart observé est-il dû aux fluctuations aléatoires ?

Un test propose une règle de décision, fixée à l'avance, avec un risque d'erreur contrôlé.

Exemple : remplissage

Une machine annonce une quantité moyenne de 1 l. On mesure 25 bouteilles et on trouve : $\overline{x} = 1{,}04$ l. Faut-il parler de sur-remplissage ?

Tout dépend de la variabilité des remplissages.

Deux hypothèses à confronter aux données

θ paramètre inconnu θ.

Hypothèses statistiques

$$\mathscr{H}_0 : \theta \in \Theta_0 \quad \text{(hypothèse nulle)}, \quad \mathscr{H}_1 : \theta \in \Theta_1 \quad \text{(alternative)}.$$

En général, $\Theta_0 \cup \Theta_1 = \Theta$.

Interprétation

- $\mathscr{H}_0$ est l'affirmation que l'on met à l'épreuve.
- $\mathscr{H}_1$ décrit l'écart que l'on cherche à mettre en évidence.

Machine de remplissage

$\mathscr{H}_0 : \mu \leq 1$ l

$\mathscr{H}_1 : \mu > 1$ l

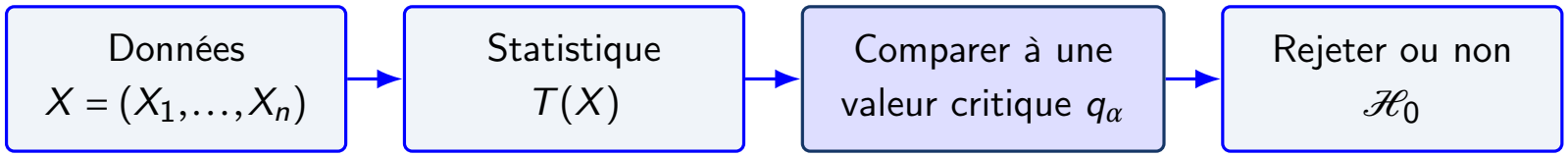
On rejète $\mathscr{H}_0$ ou pas : ne pas rejeter n'est pas prouver qu'elle est vraie.

Test unilatéral ou bilatéral ?

Unilatéral droit	Unilatéral gauche	Bilatéral
$\mathscr{H}_0 : \theta \leq \theta_0$	$\mathscr{H}_0 : \theta \geq \theta_0$	$\mathscr{H}_0 : \theta = \theta_0$
$\mathscr{H}_1 : \theta > \theta_0$	$\mathscr{H}_1 : \theta < \theta_0$	$\mathscr{H}_1 : \theta \neq \theta_0$
Recherche d'un excès.	Recherche d'un déficit.	Recherche d'un écart, dans les deux sens.

Une hypothèse est simple si elle fixe une seule valeur de θ ; elle est composite si elle en autorise plusieurs.

Comment les observations conduisent à une décision



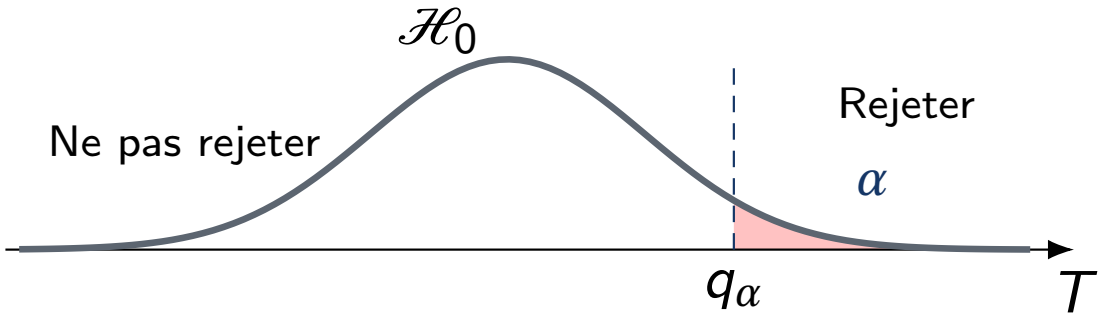
Statistique de test et région critique

On choisit $T(X)$ dont la loi sous $\mathscr{H}_0$ est connue (ou approchable). La région de rejet est :

$$R = \{x : T(x) \in R_T\} \text{ et } \phi(X) = \mathbb{1}_{\{X \in R\}} \in \{0, 1\}.$$

$\phi = 1$ signifie : rejeter $\mathscr{H}_0$; $\phi = 0$: ne pas rejeter $\mathscr{H}_0$.

Région critique : visualisation



Test unilatéral droit

T grand contredit $\mathcal{H}_0$:

$$R_T =]q_\alpha, +\infty[.$$

$$\mathbb{P}_{\theta_0}[T > q_\alpha] \leq \alpha.$$

Le seuil critique q_α est construit à partir de la loi de T sous l'hypothèse nulle. La figure représente ici une loi normale centrée réduite.

Deux erreurs possibles

	$\mathscr{H}_0$ vraie	$\mathscr{H}_1$ vraie
$\neg$ rejeter $\mathscr{H}_0$	Bonne décision	Erreur de type II
Rejeter $\mathscr{H}_0$	Erreur de type I	Bonne décision

Erreur de première espèce

Fausse alerte : rejeter $\mathscr{H}_0$ alors qu'elle est vraie.

$$\alpha(\theta) = \mathbb{P}_\theta[\phi = 1],\ \theta \in \Theta_0.$$

Erreur de seconde espèce

Non-détection : ne pas rejeter $\mathscr{H}_0$ alors que $\mathscr{H}_1$ est vraie.

$$\beta(\theta) = \mathbb{P}_\theta[\phi = 0],\ \theta \in \Theta_1.$$

Le niveau : contrôler les fausses alertes

Définition

La taille du test est le risque maximal de première espèce :

$$\alpha^* = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta[\phi(X) = 1].$$

Un test est de niveau α si $\alpha^* \leq \alpha$. Typiquement : $\alpha = 5\%$.

Fixer $\alpha = 5\%$ signifie que, si $\mathscr{H}_0$ est vraie, on accepte un risque de fausse alerte d'au plus 5%.



Cela ne signifie pas que $\mathscr{H}_0$ a 95% de chances d'être vraie !

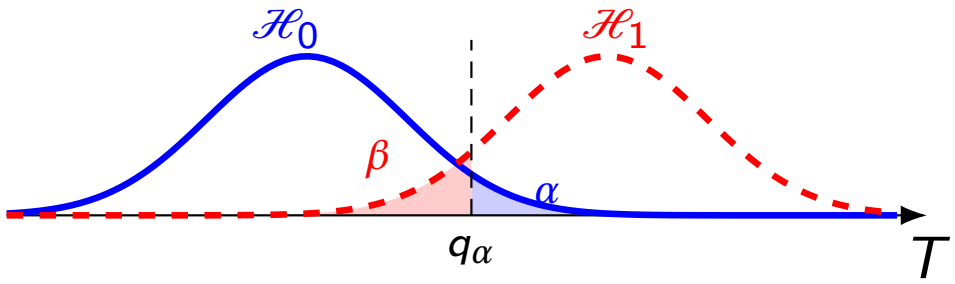
La puissance : détecter une fausse hypothèse nulle

Puissance de $\theta \in \Theta_1$

$$\pi(\theta) = \mathbb{P}_\theta[\phi = 1] = 1 - \beta(\theta).$$

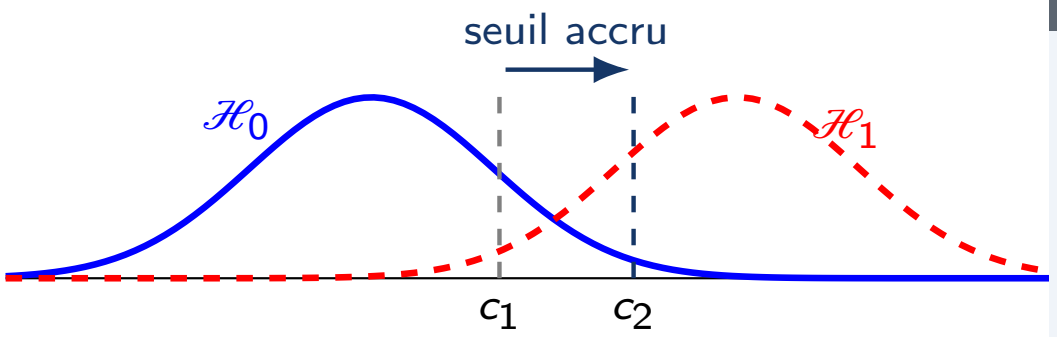
Mesure la probabilité de détecter un écart réel.

- Puissance élevée ⇒ moins de non-détections.
- Dépend de la taille de l'écart, de n et de la variabilité.



Sous $\mathscr{H}_0$: $\mathcal{N}(0,1)$; sous une alternative : $\mathcal{N}(3,1)$.

Le compromis entre les deux risques



Seuil plus sévère

On rejète moins :

- α diminue ;
- β augmente ;
- la puissance diminue.

Comment améliorer la détection ?

À niveau α fixé, un échantillon + grand n augmente la puissance.

La p-valeur : mesurer l'extrême sous $\mathscr{H}_0$

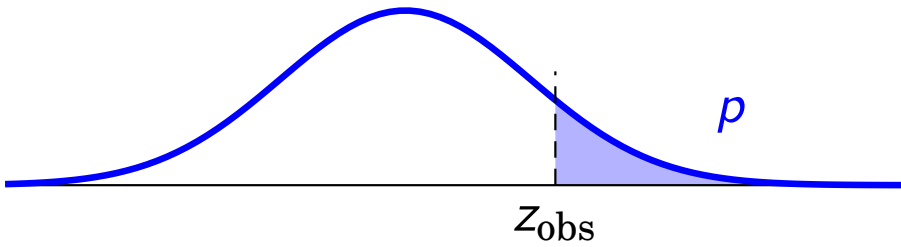
Définition intuitive

La p-valeur est la probabilité, sous $\mathscr{H}_0$, d'obtenir une statistique au moins aussi extrême que celle observée.

Pour un test unilatéral droit avec hypothèse simple :

$$p(x) = \mathbb{P}_{\theta_0}[T(X) \geq T(x)].$$

On rejette au niveau α si $\boxed{p(x) \leq \alpha}$ (tests correctement calibrés).



Unilatéral droit : aire dans la queue droite à partir de la valeur observée.

Bien interpréter la p-valeur

Une autre caractérisation

La p-valeur est le plus petit niveau auquel l'observation conduit au rejet, pour une famille de tests emboîtés.

Si $p = 0.023$

- au niveau 5% : rejet ;
- au niveau 1% : non-rejet.

Deux contre-sens à éviter

La p-valeur n'est ni :

- $\mathbb{P}(\mathscr{H}_0 \text{ vraie} \mid \text{données})$;
- la probabilité que la décision de rejet soit erronée.

Pour une hypothèse nulle composite, le calcul doit garantir le niveau pour tous les paramètres de Θ_0.

Exemple : modèle et hypothèses

Remplissage des flacons

On suppose $X_1, \ldots, X_{25}$ indépendantes, de loi

$$X_i \sim \mathcal{N}(\mu, 10^{-2}),\ \overline{x} = 1{,}04 \text{ l}.$$

L'écart-type $\sigma = 0{,}1$ l est supposé connu.

Hypothèses

$$\mathcal{H}_0 : \mu \leq 1, \qquad \mathcal{H}_1 : \mu > 1.$$

Test unilatéral droit au niveau $\alpha = 5\%$.

Statistique

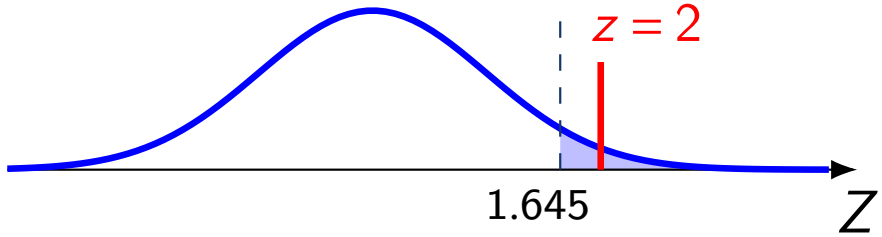
$$Z = \frac{\overline{X} - 1}{0{,}1/\sqrt{25}} = \frac{\overline{X} - 1}{0{,}02}.$$

Sous la frontière $\mu = 1$:
$Z \sim \mathcal{N}(0,1)$.

Exemple fil rouge : la décision à 5%

Seuil : $z_{0.95} \simeq 1.645$.

$$Z > 1.645 \iff \overline{X} > 103.29.$$



Statistique observée :

$$z_{\text{obs}} = \frac{1{,}04 - 1}{0{,}02} = 2.$$

p-valeur :

$$p = 1 - \Phi(2) \simeq \boxed{0.0228}.$$

Conclusion

$0.0228 < 0.05$: on rejette $\mathscr{H}_0$. Les données fournissent un argument statistique en faveur d'un sur-remplissage.

Exemple : quelle est la puissance ?

Si la vraie moyenne était $\mu = 1{,}06$ l, la règle de test serait la même : rejeter pour $\overline{X} > 1{,}03$.

Calcul

Sous $\mu = 1{,}06$, on a $\overline{X} \sim \mathcal{N}(1{,}06; 0{,}02^2)$, donc

$$\begin{aligned} \pi(1{,}06) &= \mathbb{P}_{1{,}06}(\overline{X} > 1{,}03) \\ &= 1 - \Phi\left(\frac{1{,}03 - 1{,}06}{2}\right) \\ &\simeq \boxed{0.912}. \end{aligned}$$

Interprétation

Si $\mu = 1{,}06$ ml, le test détecte le sur-remplissage dans environ 91 % des échantillons de 25 flacons.

Et pour un test bilatéral ?

$\neq$ dans les deux sens

$$\mathscr{H}_0 : \mu = \mu_0, \ \mathscr{H}_1 : \mu \neq \mu_0.$$

Sous $\mathscr{H}_0$, si $Z \sim \mathcal{N}(0,1)$, on rejette si

$$|Z| > z_{1-\alpha/2}.$$

La région critique comporte deux queues de probabilité $\alpha/2$.

p-valeur : $p = 2\left[1 - \Phi(|z_{\text{obs}}|)\right]$.

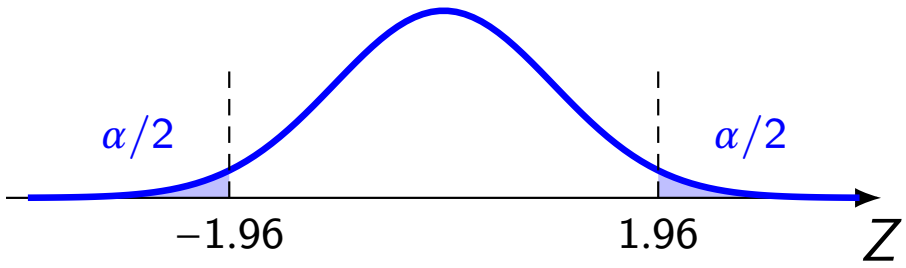


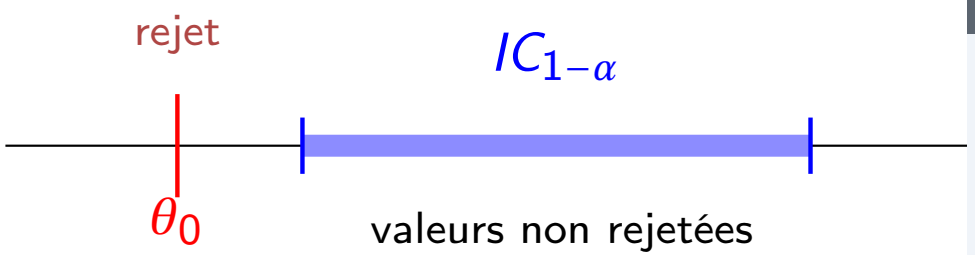
Illustration pour $\alpha = 5\,\%$.

Test et intervalle de confiance : une dualité

Propriété

Pour une famille de tests bilatéraux de niveau α, on obtient une région de confiance en réunissant les valeurs θ_0 non rejetées :

$$\boxed{\theta_0 \notin IC_{1-\alpha}(X) \iff \mathscr{H}_0 : \theta = \theta_0 \text{ rejetée.}}$$



Ex : variance connue

$$IC_{1-\alpha}(\mu) = [\overline{X} - q,\ \overline{X} + q]$$

$$q = z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Test unilatéral : inverser les bornes de confiance unilatérales.

La méthode pratique en six étapes

1. Définir le modèle.
2. Formuler $\mathscr{H}_0$ et $\mathscr{H}_1$ (unilatéral / bilatéral).
3. Fixer le niveau α à l'avance.
4. Choisir $T(X)$ et sa loi sous $\mathscr{H}_0$.
4. Calculer la région critique et la p-valeur.
5. Conclure : rejeter / ne pas rejeter ; discuter l'effet et la puissance.

À retenir

Un test ne mesure pas la probabilité que $\mathscr{H}_0$ soit vraie : il mesure la compatibilité des données avec $\mathscr{H}_0$ dans un cadre probabiliste défini, en contrôlant les fausses alertes.

À retenir

 Test statistique : réponse péremptoire à une question mal posée.

- hypothèse nulle : $\theta \in \Theta_0$ vs alternative $\theta \in \Theta_1$.
- $\phi(x) = \mathbb{1}_{[X \in R]}$.
- $\alpha = \mathbb{P}_\theta[\phi(X) = 1],\ \theta \in \Theta_0$.
- $\beta = \mathbb{P}_\theta[\phi(X) = 0],\ \theta \in \Theta_1$.
- $\pi = \mathbb{P}_\theta[\phi(X) = 1] = 1 - \beta,\ \theta \in \Theta_1$.
- $p(x) = \mathbb{P}_{\theta_0}[T(X) \geq T(x)]$.
- $p(x) \leq \alpha \Rightarrow$ on rejète $\mathscr{H}_0$ au niveau α.
- $\theta_0 \notin IC_{1-\alpha}(X) \iff \mathscr{H}_0 : \theta = \theta_0$ rejetée.