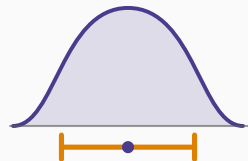


Statistique inférentielle

Chapitre 15 : La p-value



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Ce que c'est

La forme du test

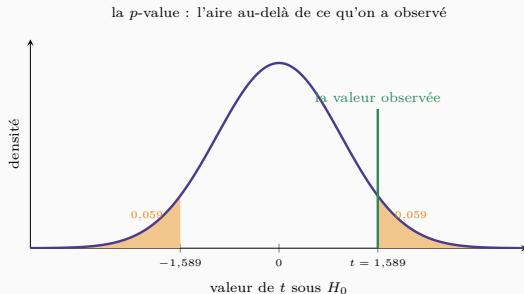
Les trois contresens

Les mauvaises pratiques

Ce qu'il faut retenir

Ce que c'est

L'aire au-delà de ce qu'on a observé



La définition, mot pour mot

La probabilité d'obtenir un résultat *au moins aussi extrême* que celui observé, si H_0 est vraie.

Ce que la p -value n'est pas

- la probabilité que H_0 soit vraie ;
- la probabilité de se tromper en rejetant ;
- une mesure de l'importance de l'effet.

Une seule chose est vraie : plus elle est petite, moins les données sont compatibles avec H_0 .

Ici 0,119 en bilatéral, 0,059 en unilatéral : la conclusion dépend de la question posée avant de voir les données.

Valeur-p

La ***p*-value** est la probabilité d'obtenir un résultat au moins aussi extrême que celui observé, **si H_0 est vraie**.

$$p < \alpha \iff \text{on rejette } H_0$$

Les quatre derniers mots sont l'essentiel

Tout se calcule **sous l'hypothèse que H_0 est vraie**. La *p*-value est une probabilité *conditionnelle* à H_0 .

Elle ne dit donc rien sur la probabilité que H_0 soit vraie — c'est le conditionnement inverse, et le chapitre 8 des probabilités a montré à quel point les deux diffèrent.

Le calcul complet

$t = 1,589$ avec 49 degrés de liberté.

$$p_{\text{unilatéral}} = P(T > 1,589) = 0,059 \quad p_{\text{bilatéral}} = 2 \times 0,059 = 0,119$$

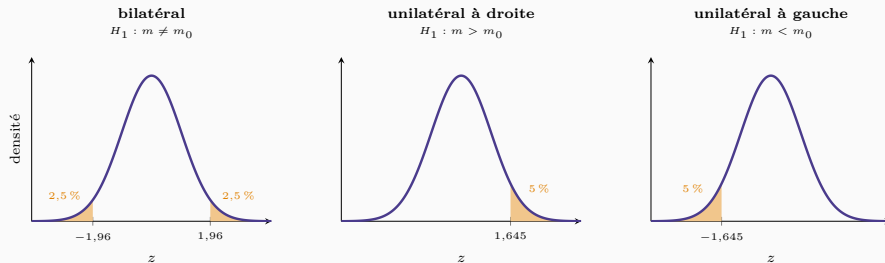
Le cas le plus instructif du cours

À $\alpha = 5\%$, on ne rejette pas. À $\alpha = 10\%$, on rejetterait.

0,059 et 0,05 ne décrivent pas deux mondes différents. **Le seuil est une convention, pas une frontière naturelle** — et présenter 0,051 et 0,049 comme deux conclusions opposées est un abus.

La forme du test

Trois façons de répartir le risque



Le même risque $\alpha = 5\%$, réparti de trois façons.

Le seuil unilatéral 1,645 est *plus facile à franchir* que le seuil bilatéral 1,96 : à données identiques, un test unilatéral conclut plus souvent au rejet.

D'où la règle déontologique : la forme du test se choisit *avant* de regarder les données, et se justifie par la question posée. Passer au unilatéral après avoir vu le signe de l'écart est une faute — c'est en pratique se donner 10 % de risque en en affichant 5 %.

La règle déontologique

La forme du test — unilatérale ou bilatérale — se décide **avant** de regarder les données, et se justifie par la question posée.

Pourquoi c'est une faute de la changer

Passer au unilatéral après avoir vu le signe de l'écart revient à se donner 10 % de risque tout en en affichant 5 %.

Sur notre exemple, la manœuvre ferait passer p de 0,119 à 0,059 — et changerait une conclusion prudente en conclusion favorable. **Le risque réel a doublé sans que personne ne le voie.**

Les trois contresens

Ce que la p-value n'est pas

Les trois erreurs, dans l'ordre de fréquence

1. Ce **n'est pas** la probabilité que H_0 soit vraie.
2. Ce **n'est pas** la probabilité de se tromper en rejetant.
3. Ce **n'est pas** une mesure de l'importance de l'effet.

La seule chose qui soit vraie

Plus la p -value est petite, moins les données observées sont **compatibles** avec H_0 .

Rien de plus. Elle mesure une compatibilité, pas une croyance, et certainement pas une ampleur.

Le troisième contresens est le plus coûteux

Deux études, deux conclusions opposées à la lecture

Étude	n	Écart estimé	p
A	50	+800 dh	0,08
B	50 000	+12 dh	0,001

Laquelle intéresse le gestionnaire

L'étude B a la plus petite p -value et l'effet le plus dérisoire. L'étude A a l'effet le plus important et une p -value au-dessus du seuil.

Une p -value ne se lit jamais seule. Il faut l'écart estimé, son intervalle de confiance, et la taille de l'échantillon.

Les mauvaises pratiques

Le mécanisme

Tester vingt hypothèses indépendantes au seuil de 5 % garantit presque une découverte :

$$P(\text{au moins un rejet à tort}) = 1 - (0,95)^{20} = 0,64$$

Deux chances sur trois de trouver quelque chose qui n'existe pas.

Ce que cela produit en entreprise

Un service marketing croise trente segments de clientèle et publie les deux qui « ressortent ». Statistiquement, il en aurait trouvé autant sur des données purement aléatoires.

La correction existe — on divise le seuil par le nombre de tests — mais la vraie règle est ailleurs : **annoncer les hypothèses avant de regarder les données.**

Ce qu'un rapport statistique doit indiquer

1. Le nombre de tests **réellement effectués**, pas seulement ceux qui ont abouti.
2. L'**ampleur** de l'effet et son intervalle de confiance, à côté de la p -value.
3. Le seuil α et la forme du test, **fixés à l'avance**.

La conclusion du chapitre

La p -value est un outil correct, employé de travers. Elle a été conçue comme un indicateur parmi d'autres, et on en a fait un verdict automatique.

Un test répond à une question posée; il ne fouille pas des données à la recherche d'une question.
C'est la différence entre confirmer et prospecter — et les deux démarches n'ont pas le même statut.

Ce qu'il faut retenir

Six points

1. La p -value est la probabilité d'un résultat aussi extrême, **si H_0 est vraie**.
2. Règle : $p < \alpha$ conduit au rejet. Sur nos données, $p = 0,059$ en unilatéral.
3. Le seuil est une **convention** : 0,051 et 0,049 ne décrivent pas deux mondes différents.
4. La forme du test se fixe **avant** de voir les données.
5. Trois contresens : ce n'est ni $P(H_0)$, ni le risque d'erreur, ni une mesure d'importance.
6. Vingt tests à 5 % donnent 64 % de chances d'une fausse découverte.

La suite

La théorie est complète. La sixième partie l'applique aux quatre situations que rencontre effectivement un gestionnaire : une moyenne, deux groupes, plusieurs groupes, un tableau d'effectifs.

Le chapitre 16 commence par le plus simple — et boucle au passage avec l'intervalle de confiance.