Les tests statistiques servent à vérifier si les données obtenues sont compatibles avec une propriété (par exemple, une moyenne théorique).

Différence observée entre une distribution est dû au hasard de l'échantillon.

* H0 la différence vient de l'échantillonnage.
* H1 différence entre l'échantillon et la population.

|  |  |  |
| --- | --- | --- |
| Réalité  Choix | H0 | H1 |
| H0 |  |  |
| H1 |  | (puissance du test) |

Interprétation : Au risque est accepté H1 alors que c’est faux. Cela correspond au rejet à tort de l’hypothèse nulle.

Rmq : le risque est généralement fixé à 5%.

Règle de décision :

* L’appartenance à intervalle de confiance.
* Par des statistique du test soit car le seuil soit par la p-value (par comparaison du risque).

## Intervalle de confiance

Exemple de comparaison d’une moyenne observée avec une moyenne de référence

avec la moyenne théorique.

Intervalle de confiance

### Comparer deux paramètres observés

Pour comparer la valeur de deux échantillons, il faut

* H0 la moyenne observée des deux échantillons est la même.
* H1 la moyenne observées des deux échantillons est différente.

Cela revient à faire deux tests, regarder si la moyenne observée est dans l’intervalle pour chaque moyenne observée.

## Les tests

|  |  |  |
| --- | --- | --- |
| Hypothèse | Seuil | P-value |
| On ne peut pas rejeter H0 au risque  inconnu |  |  |
| On accepte H1 au risque |  |  |

Par exemple, au risque alpha de 5%,

### Comparaison la moyenne d’un échantillon observée et une moyenne de référence

### Comparaison entre de la distribution de deux échantillons (Student)

p-valeur est la probabilité d’obtenir une valeur aussi extrême sous l’hypothèse h0. Énoncer de façon différentes, elle indique dans quelle mesure les données sont conformes à l’h0.

Rejeté l’hypothèse nul mais pas l’hypothèse nulle ne peut pas être acceptés. Il est possible au mieux de ne pas la rejeter.

Rmq : une valeur est significativement différente si l’hypothèse h0 (les valeurs sont égales) est rejetée.

Calcul de la probabilité sous h0 que l’événement tiré est lieu. S’il est inférieur au seuil, h0 n’est pas retenu.

Exemple : on effectue des tirages de pile ou face. H0 la piece n’est pas biaisée

4 tirages et 4 piles : % on ne rejette pas h0.

5 tirages et 5 piles : % on rejette h0

Test Du khi 2

H0 les données suivent une loi de probabilité/

Calcul de la différence (ou distance) entre les données et la loi.

Test

# ANOVA

ANalysis Of Variance

La moyenne des groupes est issue d’une même population.

Il existe des différences signaficative entre les moyennes.

Utilité

Mesurer l'influence de variables qualitatives sur une variable quantitative.

MANOVA : dans le cas ou il existe plusieurs variables quantitatives à expliquer on utilise MANOVA (Multivariate analysis of variance).

Avec une constante

l’erreur qui suit une loi normale