UNIVERSITÉ PARIS OUEST NANTERRE LA DÉFENSE
U.F.R. SEGMI — Année universitaire 2014 – 2015
L2 Économie — Cours de B. Desgraupes

**Méthodes Statistiques**

**Séance 05: Introduction aux tests d'hypothèses**

---

# Table des matières



---

# 1 Généralités sur les tests d'hypothèse

La plupart des problèmes de statistique sont des problèmes de décision ou de choix. On dispose d'un modèle statistique dans lequel les variables suivent des lois de probabilité. Ce modèle est en général un modèle d'échantillonnage car on travaille sur un échantillon plutôt que sur la population entière.

Le choix (ou la décision), dans le cas des tests d'hypothèses, consiste à déterminer, parmi deux éventualités, laquelle est la bonne. On formule une éventualité et on cherche à décider s'il faut l'*accepter* ou la *rejeter*.

C'est un problème de prise de décision et il faut donc disposer d'une *règle de décision*.

Les lois de probabilités du modèle statistique ne sont pas forcément connues. On distingue donc deux situations :

1. si on (suppose qu'on) connaît les lois suivies par l'échantillon extrait de la population, on dit qu'on a affaire à des *tests paramétriques*. Cela provient du fait que les lois sont souvent définies en fonctions de paramètres (comme par exemple le $m$ et le $\sigma$ d'une loi normale $\mathcal{N}(m, \sigma)$ ou le $\lambda$ d'une loi exponentielle) et que la formulation du test porte sur un de ces paramètres.

2. si on ne connaît pas les lois de l'échantillon, on parle de *test non-paramétrique*. La procédure du test ne repose alors pas sur des propriétés de lois de probabilité.

## 1.1 Hypothèses d'un test

Afin de pouvoir décider entre plusieurs hypothèses possibles, on met en avant une hypothèse particulière que l'on appelle l'*hypothèse nulle* (notée $H_0$).

On formule aussi une *hypothèse alternative* qui est notée $H_1$.

Souvent l'hypothèse $H_1$ est le contraire de l'hypothèse $H_0$ mais ce n'est pas nécessairement le cas. Il arrive que l'hypothèse $H_1$ soit plus restrictive.

Par exemple, si l'hypothèse $H_0$ est

$$\boxed{H_0 \,:\, a = b}$$

l'hypothèse $H_1$ pourrait être

$$\boxed{\begin{array}{c} H_1 \,:\, a \neq b \\ H_1 \,:\, a < b \\ H_1 \,:\, a > b \end{array}}$$

Dans le premier cas, on parle de *test bilatéral* et dans les deux autres cas de *test unilatéral*.

Le fait que les hypothèses $H_0$ et $H_1$ ne sont pas l'opposé l'une de l'autre peut sembler surprenant mais il faut bien comprendre deux choses :

1. le choix fait par le test est basé sur l'hypothèse $H_0$ : il consiste à décider s'il faut rejeter ou accepter $H_0$ et non pas à choisir entre $H_0$ et $H_1$.
2. l'hypothèse $H_1$ sert à formuler la *règle de décision* du test.

Le résultat d'un test est "*rejeter $H_0$*" ou bien "*ne pas rejeter $H_0$*". On ne conclut jamais par "*rejeter $H_1$*" et encore moins par "*accepter $H_1$*".

**Exemple**

Le montant moyen des achats par client dans toutes les succursales d'une enseigne commerciale sont de 50 € avec un écart-type de 5 €. Afin de vérifier les performances d'un point de vente, le directeur prélève un échantillon des achats effectués par 20 consommateurs. Les résultats observés sont les suivants :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 44.49 | 43.66 | 48.32 | 48.95 | 49.30 | 53.05 | 44.52 | 51.50 | 46.94 | 50.28 |
| 47.73 | 53.38 | 43.72 | 49.95 | 45.78 | 48.56 | 38.14 | 46.82 | 50.78 | 49.38 |

Les résultats fournis par cet échantillon sont-ils conformes aux valeurs globales ?

On suppose que les achats suivent une loi normale $\mathcal{N}(50, 5)$.

On calcule facilement que $\bar{X}_{20} = 47.7625$. Cette moyenne étant inférieure à la moyenne nationale, le directeur du point de vente s'inquiète de savoir si il y a effectivement une différence.

Il s'agit de tester si la moyenne – estimée par la moyenne empirique $\bar{X}_{20}$ fournie par l'échantillon – est significativement différente de 50. Les hypothèses d'un tel test sont :

$$\begin{cases} H_0 : \quad m = 50 \\ \\ H_1 : \quad m < 50 \end{cases}$$

On a choisi ici une hypothèse $H_1$ unilatérale car on soupçonne que les résultats sont inférieurs à la valeur attendue.

On reprendra cet exemple plus loin. □

## 1.2 Risques d'erreur

On dispose en général d'une information insuffisante puisqu'on a des informations sur un échantillon seulement et non pas sur toute la population.

La prise de décision encourt donc un double risque d'erreur:

- on peut décider que $H_0$ est fausse alors qu'elle est vraie. C'est le *risque de première espèce*, noté $\alpha$ (lu *alpha*).

- on peut décider que $H_0$ est vraie alors qu'elle est fausse. C'est le *risque de deuxième espèce*, noté $\beta$ (lu *beta*).

Le risque $\alpha$ intéresse l'*utilisateur* du test : pour lui, $H_0$ est rejetée ou pas au risque $\alpha$. Le risque $\alpha$ est le risque de rejeter à tort.

Il est courant de fixer $\alpha = 0.05$ ou $\alpha = 0.01$.

Le risque $\beta$ intéresse le *concepteur* du test : c'est le risque d'accepter à tort. Par conséquent, pour lui, $1 - \beta$ représente la *puissance du test* puisque c'est le risque d'accepter à juste titre.

Le tableau suivant résume la situation :

| Décision\\$H_0$ | *Vraie* | *Fausse* |
|---|---|---|
| *Accepter* | Décision correcte | Erreur de 2ème espèce |
| *Rejeter* | Erreur de 1ère espèce | Décision correcte |

En termes de probabilité, cela donne :

| Décision\\$H_0$ | *Vraie* | *Fausse* |
|---|---|---|
| *Accepter* | $1 - \alpha$ | $\beta$ |
| *Rejeter* | $\alpha$ | $1 - \beta$ |

• **Remarques :**

1. les procédures de tests fixent une limite supérieure au risque de première espèce. On prend souvent la valeur 5% (significatif) ou 1% (très significatif). Cette valeur (aussi appelée *seuil*) représente le *niveau de signification* du test.

2. on souhaite minimiser à la fois les risques $\alpha$ et $\beta$ mais, pour un échantillon donné, une diminution du risque $\alpha$ conduit à une augmentation du risque $\beta$.

3. l'erreur de seconde espèce diminue si la taille de l'échantillon augmente.

4. un test unilatéral est plus puissant qu'un test bilatéral.

## 1.3 Région de rejet

Les tests procèdent tous schématiquement de la même manière : on dispose d'une *variable de décision* $X$ qui suit, *lorsque l'hypothèse* $H_0$ *est vraie*, une loi théorique $P$.

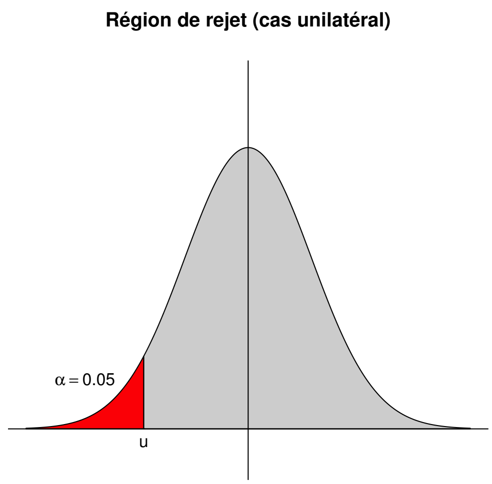
La *région de rejet* est l'ensemble des valeurs de la variable de décision qui sont "improbables" lorsque $H_0$ est vraie. Ce sont des valeurs tellement extrêmes qu'elles conduiraient à rejeter cette hypothèse. Mais, si $H_0$ était vraie, ce serait un rejet à tort. Or justement c'est la probabilité que mesure le risque $\alpha$.

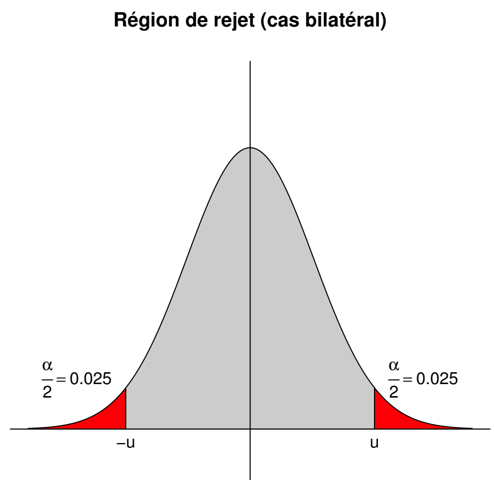
Par conséquent, le seuil $\alpha$ définit la taille de la région de rejet. C'est une région située sous la courbe de la distribution d'échantillonnage et dont la surface vaut $\alpha$.

Cette région peut prendre deux formes différentes :

- si on fait un test unilatéral, elle est entièrement à une extrêmité de la distribution de probabilité ;
- si on fait un test bilatéral, elle est en deux morceaux de surface $\alpha/2$ à chaque extrêmité de la distribution.

Voici deux illustrations graphiques dans le cas où la variable de décision $X$ suit une loi normale centrée réduite $\mathcal{N}(0,1)$.






• Dans le cas d'un test unilatéral, la limite de la région de rejet est le quantile $u$ tel que

$$P(X \leq u) = 0,05$$

Par exemple, pour une distribution normale, on trouve $\boxed{u = -1,64}$. Si le test était unilatéral de l'autre côté, on aurait $u = +1,64$.

• Dans le cas d'un test bilatéral (et d'une distribution normale), les limites de la région de rejet sont les quantiles $-u$ et $u$ tels que

$$P(-u \leq X \leq u) = 0,05$$

On trouve, pour une distribution normale, $\boxed{u = 1,96}$.

## 1.4 Région d'acceptation

La *région d'acceptation* est le complémentaire de la région de rejet. C'est l'ensemble des valeurs observées de la variable de décision pour lesquelles l'hypothèse $H_0$ est acceptable.

La règle de décision des tests consiste à regarder si la valeur de la variable de décision se trouve dans la région d'acceptation ou dans la région de rejet.

**Exemple**

Reprenons l'exemple des montants des achats par client.

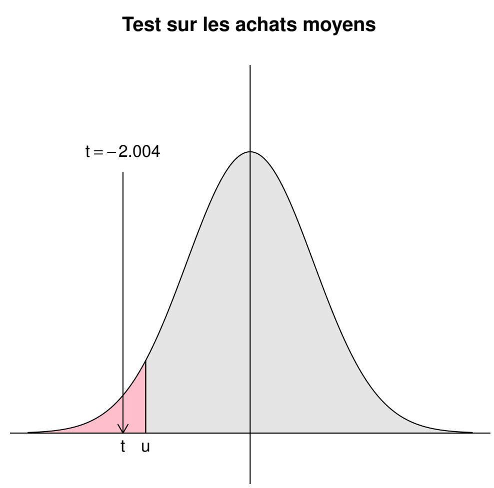
On a vu dans la séance sur l'échantillonnage, que la variable $\bar{X}$, lorsqu'elle est issue d'une population normale avec moyenne $m$ et variance $\sigma^2$ connues, suit une loi normale $\mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right)$.

Posons $Z = \frac{\sqrt{n}\,(\bar{X} - m)}{\sigma}$.

Sous l'hypothèse $H_0$, la moyenne $m$ vaut 50. On calcule

$$Z = \sqrt{20}\,\frac{47.7625 - 50}{5} = -2.0013$$

Cette valeur est en dessous du quantile $u = -1.64$ et se trouve donc dans la région de rejet. On doit donc rejeter l'hypothèse $H_0$ au risque 5% de se tromper. L'échantillon indique des résultats significativement inférieurs à la valeur moyenne $m = 50$.




## 1.5 Notion de p-valeur

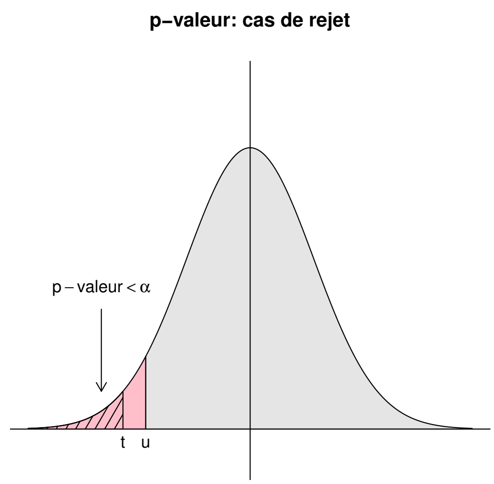
Les logiciels de calcul statistique expriment souvent le résultat d'un test en fournissant une grandeur appelée *p-valeur* (en anglais *p-value*).

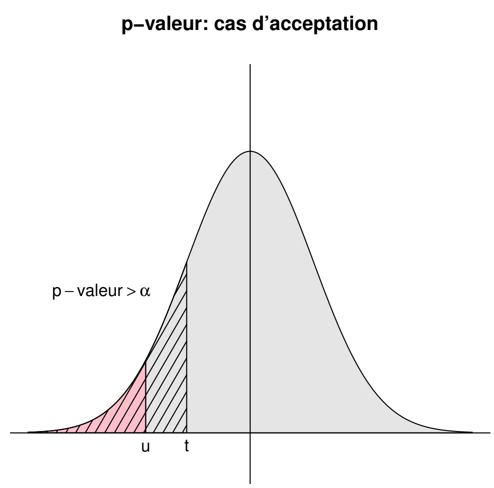
La p-valeur est le niveau de signification réel du test pour l'échantillon donné.

Prenons l'exemple d'un test unilatéral à gauche. Si $t$ est la valeur calculée de la variable de décision, la p-valeur est la probabilité que $X$ soit inférieure à la valeur $t$ calculée :

$$\boxed{\text{p-value} = P(X \leq t)}$$

Cela correspond à la zone hachurée sur le graphique suivant.




Dans le cas d'un test unilatéral à droite, la p-valeur est la probabilité que $X$ soit supérieure à la valeur $t$ calculée :

$$\boxed{\text{p-value} = P(X \geq t)}$$

Dans le cas d'un test bilatéral à droite, la p-valeur est :

$$\boxed{\text{p-value} = P(|X| \geq |t|)}$$

Exprimée en termes de surfaces, la règle de décision consiste à rejeter l'hypothèse $H_0$ dès que la p-valeur est inférieure au risque $\alpha$ choisi, c'est-à-dire dès que le niveau de signification réel est inférieur au seuil (qui est le niveau de signification nominal).

**Exemple**

Dans le cas de l'exemple étudié, on calcule la p-valeur facilement au moyen d'une table de la loi normale ou d'un logiciel statistique. On trouve :

$$P(Z < -2.0013) = 0.02268$$

Puisque $0.02268 < 0.05$, on rejette l'hypothèse $H_0$.

Voici, à titre d'exemple, la sortie du logiciel statistique R (`http://www.r-project.org`) correspondant au test précédent (couramment appelé un "*test z*") :

```
    One-sample z-Test

data:  achats
z = -2.0013, p-value = 0.02268
alternative hypothesis: true mean is less than 50
95 percent confidence interval:
      NA 49.6015
sample estimates:
mean of x
  47.7625
```

## 1.6 Types de tests

Dans les séances qui suivent, nous étudierons trois catégories de tests d'hypothèse :

1. les *tests de conformité* sont des tests à 1 échantillon dans lesquels on compare la valeur d'un paramètre avec une valeur théorique.
2. les *tests de comparaison* sont des tests qui comparent la valeur d'un paramètre entre 2 échantillons.
3. les *tests d'adéquation* (appelés aussi *tests d'ajustement*) cherchent à vérifier si la distribution d'un échantillon est conforme à une loi de probabilité donnée. On les appelle *tests d'homogénéïté* lorsqu'on compare la distribution de deux échantillons entre eux.

## 1.7 Choix des tests

Tous les tests possèdent des conditions d'application. Il est indispensable de vérifier que les conditions sont bien remplies pour que le test soit valide. Celles-ci dépendent du contexte, de connaissances préalables concernant les données statistiques étudiées, de la taille des échantillons, etc.

Pour un même problème, il peut exister plusieurs tests et il faut choisir lequel ou lesquels sont appropriés.

Les tests paramétriques sont en général plus puissants que les tests non paramétriques mais ils comportent des contraintes très précises qu'il n'est pas toujours facile de vérifier (comme la normalité d'une population), surtout si les échantillons sont de taille réduite.

Si la distribution de la population n'est pas connue, il faut avoir recours à un test non paramétrique.

Pour résumer :

Si la p-valeur, c'est-à-dire la probabilité associée à la valeur observée de $X$ (ou à une valeur encore plus extrême) est inférieure ou égale au seuil $\alpha$, on rejette $H_0$.

**Remarque :**

En cas de rejet, on conclut le test en disant "*on rejette l'hypothèse $H_0$ au risque $\alpha$ de se tromper*".

En revanche, quand on accepte l'hypothèse $H_0$, il ne faut *jamais* dire que c'est avec un risque d'erreur $\alpha$. Le risque $\alpha$ est le risque de rejeter à tort et ne mesure rien dans le cas où on accepte.