

### Chapitre III. Test d'hypothèse

#### Objectifs :

- ✓ Principes générales du test d'hypothèse
- ✓ Région d'acceptation et de rejet d'un test

Dans la statistique inférentielle, si l'estimation permet la description ; le test d'hypothèse est processus probabiliste de décision sur toute la population à partir d'un (plusieurs) échantillon(s).

#### Jargon statistique

$H_0 / H_1$  :  $H_0$  : pas de différence : il n'y a pas d'impact « d'effet » du traitement dans la population .

Ex. on s'intéresse au phénomène biologique : les bœufs sont en danger au pâturage du début de printemps au hypo magnésie.

Un niveau sanguin bas de mg dans l'élevage à l'herbe contre celui à l'étable suggère un risque de la maladie.

$H_0$  : la moyenne de Mg plasmatique dans les 2 groupes est identique.

$H_1$  : les deux moyennes sont différentes.

il est possible d'émettre l'un des deux jugements :

\*\*  $H_0$  est vraie, donc  $H_1$  est fausse

\*\*  $H_1$  est vraie, donc  $H_0$  est fausse. On peut symboliser dans le tableau :

	Etat réalisé	
Jugement	$H_0$ réalisé	$H_1$ réalisé
$H_0$	Jugement correct	Jugement faux

H1	Jugement faux	Jugement correct
----	---------------	------------------

Parmi les deux hypothèses  $H_0$  et  $H_1$ , il en existe en général une dont le rejet à tort a des conséquences plus fâcheuses que pour l'autre. Il est donc normal de ne pas traiter  $H_0$  et  $H_1$  de façon symétrique. Ainsi, on peut commettre deux types d'erreur :

\*\* l'erreur de 1ère espèce qui est la probabilité de rejeter  $H_0$  alors que  $H_0$  est vraie ;

\*\* l'erreur de 2e espèce qui est la probabilité d'accepter  $H_0$  alors que  $H_1$  est vraie.

Pour relier maintenant le jugement porté à l'observation de la variable  $X$ , on opère ainsi :

— on dit que  $H_0$  est vraie si la valeur observée de  $X$ , soit  $x$ , se trouve dans un certain domaine  $\omega$ , appelé **région d'acceptation** de l'hypothèse  $H_0$  ;

— on dit que  $H_1$  est vraie si la valeur observée appartient à  $\bar{\omega}$ , appelé **région critique** ou région de rejet.

Pour choisir le domaine  $\omega$ , on impose en général deux conditions :

— que la probabilité de commettre l'erreur de première espèce soit égale à un seuil déterminé  $\alpha$  choisi a priori aussi faible qu'on le veut (souvent 5% en biologie) ;

— que la probabilité  $\beta$  de commettre l'erreur de deuxième espèce soit minimale. Il importe de noter en effet que la première condition ne suffit pas, sauf cas très particulier, à définir  $\omega$  de façon unique.

Il est possible maintenant de compléter le tableau précédent en indiquant les règles de jugement et les probabilités pour qu'il soit correct ou faux :

	état réalisé	
décision	$H_0$ est réalisée	$H_1$ est réalisée
$H_0(X \in \omega)$	jugement correct ( $1 - \alpha$ )	jugement faux ( $\beta$ )
$H_1(X \notin \omega)$	jugement faux ( $\alpha$ )	jugement correct ( $1 - \beta$ )

## Application Test hypothèse-

Cours : Usage du student. Comparaison entre moyenne (s)

Objectifs :

- ✓ *Savoir conditions et approches générales lors de l'utilisation du test t ;*
- ✓ *Tirer des conclusions avec t test.*

Nature des données

Si on s'intéresse :

1. Au profil métabolique (PB totale, albumine, calcium, phosphore...) chez un groupe de bœuf conduit ensemble. Comparer les moyennes à une valeur de référence (**un seul groupe**) ;
  2. A mesurer le stress du transport sur le bœuf (La cortisol sécrété par la glande adrénaline est la réponse de la situation de stress. En d'autre terme, on voudra comparer la moyenne plasmatique du cortisol (ng/ml) du bœuf durant le transport et en cas de repos. (**2 groupes indépendants**))
  3. A comparer les moyennes des deux groupes d'observations sur les mêmes individus (**2 groupes dépendantes**)
  4. Comparer plusieurs moyennes dans plusieurs groupes
- ✓ Il y a une seule variable d'intérêt : la concentration plasmatique de cortisol (ng / ml) ;
  - ✓ La variable est de nature numérique (ng / ml) ;
  - ✓ On assume que la variable est une distribution normale  $\sim N(\mu, \sigma^2)$ .

### 1. Un seul groupe-t test

La moyenne observée dans une série de données prend une valeur de référence.

Ex. Comparer le gain de poids moyen d'un lot de taurillons avec le gain usuel de l'exploitation.

### 1.1. Conditions :

- ✓ L'échantillon soit représentatif de la population ;
- ✓ Les données observées proviennent d'une population normalement distribuée à cette variable d'intérêt (vérifier la normalité par un histogramme :

Si non : soit

\*\* Il faudra la normalisé log.... ;

\*\* alternativement, on utilise des tests non paramétriques appropriés tels que : Sign test,

Kolmogorov-Smirnov test, ou runs test.

### 1.1. Approche générale t test – un seule groupe

1. Spécifier  $H_0 : m = \mu$  et  $H_1 : m \neq \mu$  (test bilatéral) ;
2. Collecter les données puis les visualiser pour vérifier la normalité (histogramme) ;
3. Calculer le test statistique t :  $t = \frac{m - \mu}{\frac{s}{\sqrt{n}}}$  (ddl= n-1) ;
4. Obtenir p-value (référer |t calculé| à la table de distribution t) ;
5. Utiliser P-value pour juger la consistance des données observées avec  $H_0$  (on rejette  $H_0$  lorsque P-value < 0.05) ;
6. Spécifier IC de moyenne qui renforce le résultat trouvé

IC 95% :  $m \pm t(0.05, \text{ddl}) \frac{s}{\sqrt{n}}$  qui donne l'aire totale de 0.05

### Application :

Le tableau suivant présente les GMQ (g/j) d'un échantillon aléatoire représentatif de 36 taurillons d'un lot d'une ferme d'élevage. Si le contrôle de la performance de cette race indique une valeur standard de GMQ de l'ordre de 607 g/j que vous pouvez conclure de cette expérience ?

577 596 594 612 600 584 618 627 588 601 606 559 615 607 608 591 565 586  
621 623 598 602 581 631 570 595 603 605 616 574 578 600 596 619 636 589  
Mean,  $\bar{x} = 599.194$  g  
Standard deviation,  $s = 18.656$  g  
Standard error,  $s/\sqrt{n} = 18.66/\sqrt{36} = 3.109$  g

(Est ce que ces données sont consistantes avec la moyenne référentielle de la race 607 g/j) ?

### Solution

1.  $H_0 : 599.194 = 607$
2. **Histogramme** visualise une distribution approximativement normale
3.  $T = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{599.194 - 607}{\frac{18.656}{\sqrt{35}}} = -2.51$  (ddl=35)
4. **P-value = 0.017** : (sous  $H_0$  vraie on a qu' au moins de 2% de chance d'avoir un GMQ < 599.194) **P-value < 0.05 on rejette  $H_0$** . Donc les valeurs observées sont non-consistantes avec la valeur référence 607 g/J.
5. **IC 95% de le GMQ réelle est :**  
 $599.194 \pm 2.03(3.109) = (592.88, 605.51)$  g cet intervalle ne contient pas 607 g et on peut conclure que l' échantillon n' appartient pas à la population avec un GMQ de 607 g/j. l' échantillon présente un GMQ significativement inférieur à la valeur de référence 607 g/j.

## 2. Deux groupes indépendants-t test

### Conditions :

- ✓ **Variable** d'intérêt soit numérique et normalement distribué ;
- ✓ **2 groupes** : indépendants, représentatifs de la population ;
- ✓ **2 échantillons homoscedastique** :  $s_1^2 = s_2^2$  (**homogénéité des variances**)  
vérification est importante par œil ou utilisation des test comme **F- test**;

### Levene's test

(cas inverse : hétéroscedasticité :

\*\* on utilise le t-test modifié  $t = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

\*\* ou des tests non paramétrique comme Wilcoxon rank sum test ou Mann-Whitney  $U$  test ).

### Approches générales 2 groupes indépendants et variances égales :

1.  $H_0 : m_1 = m_2 = \mu$

2. **Collecter** les données puis les visualiser pour vérifier la normalité (boite à moustache)  $m_1 - m_2$

$$Test_2 = \frac{\bar{m}_1 - \bar{m}_2}{\sqrt{s^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

3.

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (\text{ddl} = n_1 + n_2 - 2)$$

4. **Obtenir p-value** (référer  $|t \text{ calculé}|$  à la table de distribution t) ;  
 5. **Utiliser P-value** pour juger la consistance des données observées avec H0 (on rejette Ho lorsque P-value < 0.05) ;  
 6. **Spécifier IC de moyenne** qui renforce le résultat trouvé

$$m_1 - m_2 \pm t(\text{ddl}, \alpha) \sqrt{s^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$$

### Application

Afin d'étudier l'influence de flushing pendant la lutte sur le gain de poids vif chez les brebis. Nous constituons 2 groupes indépendants (24 femelles suralimentées pendant la lutte et 30 femelles témoins )

Controls			Flushed		
62.5	63.9	69.2	70.7	67.8	69.8
66.8	65.7	62.6	71.8	66.8	68.1
69.5	67.2	61.1	64.9	67.0	66.0
64.1	65.2	61.8	68.2	67.1	69.4
65.3	63.5	69.6	69.4	67.6	69.8
65.6	65.3	71.1	64.4	66.1	67.9
66.4	65.1	67.0	66.9	62.7	66.2
66.1	64.8	67.5	69.4	64.6	64.2
68.6	67.4	68.2			
62.5	66.0	63.6			

Spss output :

\*\*Homogénéité de variance : p=0.62 :la différence entre variances est non significative au seuil de 5%.

\*\* t = m1- m2/ SEM=2.4

$$SEM = s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad S = \sqrt{\frac{(n_1-1)s^2_1 + (n_2-1)s^2_2}{n_1+n_2-2}}$$

- P-value = 0.018 < 0.05 : la chance d'obtenir une différence de moyenne d'au moins 1,59 kg n'est que de 1,8% si l'hypothèse nulle est vraie (on rejette H<sub>0</sub> : les moyennes de poids sont significativement différentes. Les brebis subissent le flushing sont 1,59 kg plus lourdes que celles témoins.
- IC 95% : 0,20 à 2,91 kg : 0 n'appartient pas à cet intervalle.

### 3. Groupes dépendants (appariés)-t test

Analyse des différences entre composantes des paires des observations (avant traitement, après traitement) chez les mêmes individus de l'échantillon représentatif de la population.

**Conditions :**

- ✓ Les différences des paires sont normalement distribuées si non on fait appel à la normalisation par processus de transformation ou utilisation des tests non paramétriques appropriés comme **Wilcoxon signed rank test**.

**Approche générale**

1. H<sub>0</sub> = la moyenne de différences entre les paires observées = 0 ;
2. Vérifier la normalité des différences ;

3. Test :  $t = \frac{\bar{d}}{SE(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}}$       ddl = n-1 ; s<sub>d</sub> : écart type des différences

4. P-value

5. IC 95%

$$\bar{d} \pm t_{0.05} SE(\bar{d}) = \bar{d} \pm t_{0.05} (s_d/\sqrt{n})$$

**Application :**

Sue des chiens diabétiques dépendantes de l'insuline, nous intéressons à étudier l'effet d'un régime alimentaire (avec 2 modalités assez fibreux HF, moins fibreux LF) sur la glycémie. Après une phase d'adaptation de 2 mois, l'essai proprement dit était en 6 mois.

Preprandial serum glucose levels (mmol/l) in dogs with insulin-dependent diabetes mellitus fed a low- and high-fibre diet (based on summary data from Nelson *et al.*, 1998).

Dog	Low-fibre diet (LF)	High-fibre diet (HF)
1	9.44	9.28
2	17.61	8.67
3	8.89	6.28
4	16.94	12.67
5	10.39	6.67
6	11.78	7.28
7	15.06	15.39
8	7.06	5.61
9	19.56	11.94
10	8.22	5.11
11	23.17	17.33

HICHER Azzeddine-UMKB-2020