# TD4 : test de la moyenne 1 (cas symétrique)

## Rappel :

Test de l'hypothèse nulle $H_0$ contre l'hypothèse alternative $H_1$.
$H_0$ est l'hypothèse non coûteuse : on n'agit pas si elle n'est pas fausse, tout est normal.
$H_1$ est l'hypothèse coûteuse : on doit agir, le résultat non attendu.
2 types d'erreurs possibles : voir CM 2.
Cas pourcentage : $\hat{\sigma} = \sigma_0 = \sqrt{p_0(1-p_0)}$.

**Exercice 1.**

1. Trouver $z$ tel que $P(-z \leq \mathcal{N}(0,1) \leq z) = 0.93$.
2. Trouver $z$ tel que $P(-z \leq \mathcal{T}(6) \leq z) = 0.90$.
3. Trouver $z$ tel que $P(-z \leq \mathcal{N}(0,1) \leq z) = 0.99$.

## 1 Exercices

**Exercice 2.** Une entreprise de fabrication de matelas souhaite savoir si elle doit ajuster sa production de matelas. Lors des dernières années, l'usine produisait 60% de matelas durs et 40% de matelas mous. Sur les 300 derniers matelas vendus, il y a eu 140 matelas durs vendus.

1. Au risque 1%, l'usine doit-elle ajuster sa production ou non ?

**Correction.** Loi Bernoulli/pourcentage. On note $p$ la probabilité de vendre un matelas dur cette année.

$p_0 = 0.6$ la probabilité les années précédentes.

Comme $n \geq 30$, $np_0 = 180 \geq 5$ et $n(1-p_0) = 120 \geq 5$, on peut faire le test. Test statistique : 2 hypothèses :

— hypothèse nulle : $H_0 : p = p_0$ : le marché n'a pas changé *versus*
— hypothèse alternative : $H_1 : p \neq p_0$ : le marché a changé et donc penser à ajuster la production.

$\alpha = 1\%$, $\hat{p} = \frac{140}{300} = 0.467$ et $\sigma_0 = \sqrt{p_0(1-p_0)} = \sqrt{0.6 \times 0.4} = 0.49$. Calcul de la statistique :

$$Z = \frac{\hat{p} - p_0}{\sigma_0/\sqrt{n}} = \frac{0.467 - 0.6}{0.49/\sqrt{300}} = -4.7.$$

Zone de rejet : par le TCL, $Z \sim \mathcal{N}(0;1)$, la loi centrée réduite. Donc, on met 1% à gauche et à droite de manière équitable car $H_1 : p \neq p_0$ (on verra semaine prochaine des $p > p_0$ ou $p < p_0$). Faire dessin. On cherche $z$ tel que $P(-z \leq \mathcal{N}(0;1) \leq z) = 99\%$. Donc, par lecture dans la table, la zone de rejet est

$$R = ]-\infty; -2.576] \cap [2.576; +\infty[.$$

Comme $Z$ appartient à la zone de rejet $R$, on rejette $H_0$ et l'usine devrait ajuster sa production.

**Exercice 3.** Une coopérative agricole utilise une machine pour remplir ses sacs de pommes de terre : la machine est réglée afin que le poids des sacs soit en moyenne de 50kg. Si la machine est déréglée alors il procéder à un nouveau réglage. Pour contrôler le réglage de la machine, tous les 10 000 sacs, on mesure le poids de 50 sacs. Lors d'un contrôle, on a obtenu les résultats suivants :

| Poids | 47 | 48 | 49 | 50 | 51 | 52 | 53 |
|---|---|---|---|---|---|---|---|
| Effectifs | 8 | 9 | 4 | 12 | 8 | 4 | 5 |

1. Au risque 13%, la machine est-elle déréglée ?

**Correction.** Loi inconnue de moyenne $m$ inconnue, mais $n = 50 \geq 30$. Donc on peut faire un test.

La valeur de référence est $m_0 = 50$

Test statistique : 2 hypothèses :

— hypothèse nulle : $H_0 : m = m_0$ : la machine est bien réglée *versus*

— hypothèse alternative : $H_1 : m \neq m_0$ : la machine n'est pas bien réglée et on ne veut pas de sac trop rempli (perte) ni de sac trop peu rempli (les clients ne nous feront plus confiance).

Niveau de risque : $\alpha = 13\%$.

Estimation de la moyenne : $\hat{m} = \frac{8\times47+\cdots+5\times53}{8+\cdots+5} = \frac{2485}{50} = 49.7$.

Estimation de l'écart-type : $\hat{\sigma} = \sqrt{\frac{8\times2.7^2+\cdots+5\times3.3^2}{50-1}} = \sqrt{\frac{176.5}{49}} \simeq 1.9$.

Calcul de la statistique : $Z = \frac{\hat{m}-m_0}{\hat{\sigma}/\sqrt{n}} = \frac{-0.3}{1.9/\sqrt{50}} \simeq -1.12$. Zone de rejet : par le TCL, $Z \sim \mathcal{N}(0;1)$, au risque 13%, pour $H_1 : m \neq m_0$, on trouve

$$R = ]-\infty; -1.515] \cap [1.515; +\infty[.$$

Comme $Z \notin R$, on ne rejette pas $H_0$ et on conclut que l'on n'a pas besoin de régler la machine.

**Exercice 4.** Une usine de médicament produit des dolipranes avec 1g de principe actif. Tous les 10000 comprimés produits, un test statistique sur 10 comprimés est réalisé. Lors du dernier test, on a trouvé les doses suivantes de principe actif dans les 10 comprimés testés :

1017mg, 998mg, 1027mg, 984mg, 1006mg, 1020mg, 1000mg, 994mg, 1011mg, 1001mg.

1. Au risque 5%, faut-il régler la machine ?
2. En supposant que la concentration suit une loi gaussienne, au risque 5%, faut-il régler la machine ?

**Correction.**

1. Loi inconnue de moyenne $m$ inconnue et $n = 10 < 30$. Donc on ne peut pas faire un test.
2. Loi gaussienne de moyenne $m$ inconnue et $n = 10 < 30$. On peut faire un test car les données sont supposées gaussiennes.

   La valeur de référence est $m_0 = 1000$mg.

   Test statistique : 2 hypothèses :

— hypothèse nulle : $H_0 : m = m_0$ : la machine est bien réglée *versus*
— hypothèse alternative : $H_1 : m \neq m_0$ : la machine n'est pas bien réglée (on ne veut ni sous-dosage, ni sur-dosage).

Niveau de risque : $\alpha = 5\%$.

Estimation de la moyenne : $\hat{m} = \frac{983+\cdots+999}{10} = \frac{10058}{10} = 1005.8$.

Estimation de l'écart-type : $\hat{\sigma} = \sqrt{\frac{(983-1005.8)^2+\cdots+(999-1005.8)^2}{10-1}} = \sqrt{\frac{1535.6}{9}} \simeq 13.1$.

Calcul de la statistique : $Z = \frac{\hat{m}-m_0}{\hat{\sigma}/\sqrt{n}} = \frac{5.8}{13.1/\sqrt{10}} \simeq 1.4$. Zone de rejet : comme les données sont gaussiennes, $Z \sim \mathcal{T}(n-1) = \mathcal{T}(9)$, au risque 5%, pour $H_1 : m \neq m_0$, on trouve

$$R = ]-\infty; -2.262] \cap [2.262; +\infty[.$$

Comme $Z \notin R$, on ne rejette pas $H_0$ et on conclut que l'on n'a pas besoin de régler la machine.

## 2 Exercices d'entrainement

**Exercice 5.** L'année $N-1$, 80% des ventes provenaient de la boutique en ligne d'une entreprise. Le service commercial de l'entreprise regarde les statistiques de ventes du premier trimestre de l'année $N$ afin de détecter une évolution des ventes. Les statistiques donnent que sur les 13240 ventes du premier trimestre, 9732 ont eu lieu en ligne.

Au risque 1%, y a-t-il eu une évolution des ventes ?

**Correction.** Loi Bernoulli/pourcentage. On note $p$ la proportion de ventes en ligne.

$p_0 = 0.8$ le pourcentage de ventes en ligne de l'année $N-1$.

Comme $n \geq 30$, $np_0 = 10592 \geq 5$ et $n(1-p_0) = 2648 \geq 5$, on peut faire le test. Test statistique : 2 hypothèses :

— hypothèse nulle : $H_0 : p = p_0$ : les ventes n'ont pas évolué depuis l'année $N-1$ *versus*
— hypothèse alternative : $H_1 : p \neq p_0$ : les ventes ont évolué.

$\alpha = 1\%$, $\hat{p} = \frac{9732}{13240} = 0.735$ et $\sigma_0 = \sqrt{p_0(1-p_0)} = \sqrt{0.8 \times 0.2} = 0.4$. Calcul de la statistique :

$$Z = \frac{\hat{p} - p_0}{\sigma_0/\sqrt{n}} = \frac{0.735 - 0.8}{0.4/\sqrt{13240}} = -18.7.$$

Zone de rejet : par le TCL, $Z \sim \mathcal{N}(0;1)$, par lecture dans la table, la zone de rejet est

$$R = ]-\infty; -2.576] \cap [2.576; +\infty[.$$

Comme $Z$ appartient à la zone de rejet $R$, on rejette $H_0$ et on en déduit qu'il y a bien eu une évolution significative des ventes.

**Exercice 6.** Dans une usine d'embouteillage, une machine remplit des bouteilles d'1L de jus de clémentine. Toutes les 10 000 bouteilles, un contrôle qualité au risque 10% est effectué sur un échantillon de 10 bouteilles : la contenance est mesurée. On sait d'expérience que le remplissage suit une loi normale. Lors du dernier contrôle qualité, on a obtenu le volume suivant pour les 10 bouteilles :

0.991L, 0.993L, 0.995L, 0.974L, 0, 976L, 1.008L, 0.973L, 0.974L, 0.984L, 1.002L.

1. Qu'en déduire, à partir de ce test de contrôle qualité, sur la machine ?

Suite à ce premier test, s'il s'avère "raté" par la machine, un deuxième test plus précis est réalisé. Ce second test consiste en la mesure de la contenance de 100 bouteilles et le risque pris est de 1%. Le test donne une moyenne de 0.985L et un écart-type de 0.012L.

2. Suite à ce second test, faut-il arrêter la chaine de production et régler la machine ?

**Correction.**

1. Loi gaussienne (normale) de moyenne $m$ inconnue et $n = 10 < 30$. On peut faire un test car les données sont supposées gaussiennes.

   La valeur de référence est $m_0 = 1$L.

   Test statistique : 2 hypothèses :
   - hypothèse nulle : $H_0 : m = m_0$ : la machine est bien réglée *versus*
   - hypothèse alternative : $H_1 : m \neq m_0$ : la machine n'est pas bien réglée (on ne veut remplir les bouteilles ni trop, ni pas assez).

   Niveau de risque : $\alpha = 10\%$.

   Estimation de la moyenne : $\hat{m} = \frac{0.991+\cdots+1.002}{10} = \frac{9.87}{10} = 0.987$.

   Estimation de l'écart-type : $\hat{\sigma} = \sqrt{\frac{(0.991-0.987)^2+\cdots+(1.002-0.987)^2}{10-1}} = \sqrt{\frac{0.1446}{9}} \simeq 0.0127$.

   Calcul de la statistique : $Z = \frac{\hat{m}-m_0}{\hat{\sigma}/\sqrt{n}} = \frac{-0.013}{0.0127/\sqrt{10}} \simeq -3.24$.

   Zone de rejet : comme les données sont gaussiennes, $Z \sim \mathcal{T}(n-1) = \mathcal{T}(9)$, au risque 10%, pour $H_1 : m \neq m_0$, on trouve
   $$R = ]-\infty; -1.833] \cap [1.833; +\infty[.$$
   Comme $Z \in R$, on rejette $H_0$ et on conclut qu'il faut régler la machine.

2. On a encore des données gaussiennes, mais cette fois-ci, $n = 100 > 30$, on peut donc considérer que $\mathcal{T}(n-1) \sim \mathcal{T}(\infty) \sim \mathcal{N}(0;1)$.

   Calcul de la nouvelle statistique : $Z = \frac{\hat{m}-m_0}{\hat{\sigma}/\sqrt{n}} = \frac{-0.015}{0.012/\sqrt{100}} \simeq -12.5$.

   Zone de rejet : au risque 1%, pour $H_1 : m \neq m_0$, on trouve
   $$R = ]-\infty; -2.576] \cap [2.576; +\infty[.$$
   Comme $Z \in R$, on rejette $H_0$ et on conclut qu'il faut régler la machine.