TD5 : test de la moyenne 2 (cas asymétrique)

Rappel :

Test de l'hypothèse nulle H_0 contre l'hypothèse alternative H_1.
Une seule forme pour $H_0 : m = m_0$.
Trois formes pour $H_1 : m \neq m_0$, $H_1 : m < m_0$ ou $H_1 : m > m_0$.

Exercice 1.
1. Trouver z tel que $P(\mathcal{N}(0,1) \leq z) = 0.93$.
2. Trouver z tel que $P(\mathcal{T}(16) \leq z) = 0.90$.
3. Trouver z tel que $P(z \leq \mathcal{N}(0,1)) = 0.99$.
4. Trouver z tel que $P(z \leq \mathcal{T}(9)) = 0.90$.

1 Exercices

Exercice 2. De 2006 à 2013, la consommation de lait en France était de 350 mille litres de lait par mois. Sur les 12 mois de l'année 2016, on observe que la consommation moyenne de lait par mois est de 367 mille litres avec un écart type de 24 mille litres. Les chiffres mensuels de consommation sont supposées suivre une loi gaussienne.

1. La hausse de consommation est-elle significative au risque 5% ?

Correction. La donnée étudiée est la consommation de lait en France par mois.

Le nombre de données est de $n = 12$ et elles suivent une loi gaussienne de moyenne m (inconnue), donc on sait faire.

La valeur de référence est $m_0 = 350$ qui correspond en milliers de litres au volume de lait consommé par mois en France.

On fait un test d'hypothèse avec
— $H_0 : m = m_0$: la consommation n'a pas changé *versus*
— $H_1 : m > m_0$: la consommation a augmenté.

On a $n = 12$, $\hat{m} = 367$ et $\hat{\sigma} = 24$. Donc, on obtient pour statistique

$$Z = \frac{350 - 367}{24/\sqrt{12}} = -2.45 \leq 0.$$

Comme $n = 12$ et que les données sont gaussiennes, Z suit une loi de Student à 11 degré de liberté. Comme $H_1 : m > m_0$, la zone de rejet est de la forme $Z_R =]-\infty, -t]$, où $t \geq 0$ (on rejette l'hypothèse H_0 si "Z est trop éloigné de 0"). Comme le risque est 5%, on cherche donc t tel que $P(\mathcal{T}(11) \leq -t) = 0.05$, donc $P(|\mathcal{T}(11)| \geq t) = 2 \times 0,05 = 0,1$. La lecture dans la table donne $t = 1.796$, donc la zone de rejet Z_R est

$$Z_R =]-\infty, -1.796].$$

Comme $-2.45 \in Z_R$, on rejette H_0 et la hausse de consommation est significative.

Exercice 3. Une entreprise de meubles, avec son fournisseur actuel, reçoit en moyenne 8% de planches inutilisables car trop abimées. L'entreprise décide de tester un autre fournisseur en espérant qu'il livre significativement moins de planches abimées. Lors de l'échantillon test, sur les 10 000 planches reçues, 650 ont été inutilisables.

1. Au risque 1%, cette entreprise doit-elle changer son fournisseur pour le nouveau ?

Correction. La donnée étudiée est le pourcentage de planches inutilisables.

Le nombre de données est de $n = 10\,000 \geq 30$ et on a $650 \geq 5$ planches inutilisables et $10000 - 650 = 9350 \geq 5$ utilisables, donc on sait faire.

La loi est une loi de Bernoulli (car pourcentage) de paramètre p (inconnu), le pourcentage de planches inutilisables chez le fournisseur testé.

La valeur de référence est $p_0 = 0.08$ qui correspond au pourcentage du fournisseur habituel.

On fait un test d'hypothèse avec

— $H_0 : p = p_0$: le nombre de planches inutilisables est comparable entre les 2 fournisseurs *versus*

— $H_1 : p < p_0$: le nombre de planches inutilisables à l'air significativement plus faible chez le second fournisseur.

On a $n = 10\,000$, $\hat{p} = 0.065$ et $\hat{\sigma} = \sqrt{p_0(1 - p_0)} = \sqrt{0.08 \times 0.92} = 0.27$. Donc, on obtient pour statistique

$$Z = \frac{0.08 - 0.065}{0.27/\sqrt{10000}} = 5.6.$$

Comme $n = 10000$ et que $n\hat{p} = 650 \geq 5$ et $n(1 - \hat{p}) = 9350$, Z suit une loi gaussienne. Comme le risque est 1% et $H_1 : p < p_0$, la lecture dans la table, pour un risque de 1%, donne que la zone de rejet Z_R est

$$Z_R = [2.326, +\infty[.$$

Comme $5.6 \in Z_R$, on rejette H_0 et il faut penser à prendre le nouveau fournisseur.

Exercice 4. Dans un petit aéroport de province, il y avait en moyenne 36 mouvements (départs ou arrivées) d'avions par jour avant la crise du Covid. Sur le dernier trimestre de l'année 2022, on a observé le tableau suivant du nombre de mouvement

Mouvement journalier	24	32	34	46
Nombre de jours	28	20	24	20

1. Calculer la moyenne estimée $\hat{m}$ et l'écart-type estimé $\hat{\sigma}$.
2. Au risque 2%, le mouvement journalier moyen des avions dans cet aéroport a-t-il connu une baisse significative ?

Correction.

1. On a $n = 28 + 20 + 24 + 20 = 92$ jours. Donc la moyenne estimée est

$$\hat{m} = \frac{28 \times 24 + 20 \times 32 + 24 \times 34 + 20 \times 46}{92} = \frac{3048}{92} \simeq 33.1$$

et l'écart-type estimé est

$$\begin{aligned}\hat{\sigma} &= \sqrt{\frac{28 \times (24-33.1)^2 + 20 \times (32-33.1)^2 + 24 \times (34-33.1)^2 + 20 \times (46-33.1)^2}{91}} \\ &= \sqrt{\frac{28 \times 9.1^2 + 20 \times 1.1^2 + 24 \times 0.9^2 + 20 \times 12.9^2}{91}} \\ &= \sqrt{\frac{5690}{91}} \simeq 8.1.\end{aligned}$$

2. On étudie le nombre de mouvements d'avions dans un aéroport. Il suit une loi inconnue de moyenne m (inconnue). La valeur de référence (d'avant Covid) est $m_0 = 36$. On fait le test statistique suivant :
   — $H_0 : m = m_0$: c'est comme avant *versus*
   — $H_1 : m < m_0$: il y a une baisse significative du nombre d'avions dans l'aéroport.
   On a $n = 92$, $\hat{m} = 33.1$ et $\hat{\sigma} = 8.1$. On construit la statistique
   $$Z = \frac{36 - 33.1}{8.1/\sqrt{92}} = 3.43.$$
   Comme $n \geq 30$, par le TCL, Z suit une loi gaussienne avec $H_1 : m < m_0$, la zone de rejet Z_R est donc
   $$Z_R = [2.05, +\infty[.$$
   Comme $3.43 \in Z_R$, on rejette H_0 et la baisse du mouvement moyen semble significative.

2 Exercices d'entrainement

Attention : dans les exercices suivants, il y a des symétriques (cf TD4) $H_1 : m \neq m_0$ et des asymétriques $H_1 : m < m_0$ ou $H_1 : m > m_0$. Dans la correction, il est marqué les hypothèses H_1 attendues.

Exercice 5 (Sujet 2021-2022)**.** Une boutique en ligne teste un nouveau design pour son site web afin d'améliorer son taux de conversion (pourcentage de visiteurs qui réalisent un achat). Sur le premier mois de test, 2965 sur les 95634 visiteurs du site ont effectué un achat.

Sachant que le taux de conversion avec l'ancien design était de 2.65%. Au risque 1%, gardez-vous le nouveau design ?

Correction. La loi est une loi de Bernoulli (car pourcentage) de paramètre p (inconnu).

La valeur de référence est $p_0 = 2.65\%$. On fait un test d'hypothèse avec
— $H_0 : p = p_0$, aucun changement *versus*
— $H_1 : p > p_0$, le design est meilleur.
On a $n = 10\,000$, $\hat{p} = 0.031$ et $\hat{\sigma} = \sqrt{p_0(1-p_0)} = \sqrt{0.0265 \times 0.9735} = 0.16$. Donc, on obtient pour statistique

$$Z = \frac{0.0265 - 0.032}{0.16/\sqrt{95634}} = -8,698.$$

Comme $n = 95634 \geq 30$, $2965 \geq 5$ et $95634 - 2965 = 92669 \geq 5$, Z suit une loi gaussienne. Comme le risque est 1% et $H_1 : p > p_0$, la lecture dans la table, pour un risque de 1%, donne que la zone de rejet Z_R est

$$Z_R =]-\infty, -2.326].$$

Comme $Z \in Z_R$, on rejette H_0, il faudrait garder le nouveau design.

Exercice 6 (Rattrapage 2021-2022)**.** Après avoir regardé l'émission "Cauchemar en cuisine", un restaurant se dit qu'il serait temps qu'il change la décoration de sa salle. Sur les 3 mois après avoir refait la décoration de sa salle (soit 70 midis travaillés), il fait en moyenne 45.3 couverts par midi avec un écart-type de 6.4 couverts.

Sachant qu'avant redécoration, il faisait en moyenne 41.1 couverts par midi. Au risque 5%, a-t-il eu raison de redécorer ?

Correction. On a $n = 70 \geq 30$ et $m_0 = 41.1$. On teste $H_0 : m = m_0$, aucun changement *versus* $H_1 : m > m_0$, la décoration a un effet positif. La statistique est

$$Z = \frac{41.1 - 45.3}{6.4/\sqrt{70}} = -5.49.$$

Comme le risque est 5% et $H_1 : m > m_0$, la zone de rejet est

$$Z_R =]-\infty, -1.645].$$

Comme $Z \in Z_R$, on rejette H_0, le restaurant a bien fait de redécorer sa salle.

Exercice 7. Au cours de ces dernières années, une entreprise viticole a observé que la répartition des ventes de vins rouge et blanc était de 70% pour le vin rouge. Sur les 4884 dernières bouteilles vendus, il y a eu 2712 bouteilles de vin rouge vendues.

Au risque 1%, le marché de la vente de vin a-t-il changé ?

Correction. La loi est une loi de Bernoulli (car pourcentage) de paramètre p (inconnu).

La valeur de référence est $p_0 = 70\%$. On fait un test d'hypothèse avec

— $H_0 : p = p_0$, aucun changement *versus*

— $H_1 : p \neq p_0$, le marché de la vente de vin a changé.

On a $n = 4884$, $\hat{p} = 0.56$ et $\hat{\sigma} = \sqrt{p_0(1 - p_0)} = \sqrt{0.7 \times 0.3} = 0.46$. Donc, on obtient pour statistique

$$Z = \frac{0.7 - 0.56}{0.46/\sqrt{4884}} = 21.27.$$

Comme $n = 4884 \geq 30$, $2712 \geq 5$ et $4884 - 2712 = 2172 \geq 5$, Z suit une loi gaussienne. Comme le risque est 1% et $H_1 : p \neq p_0$, la lecture dans la table, pour un risque de 1%, donne que la zone de rejet Z_R est

$$Z_R =]-\infty, -2.576] \cup [2.576, +\infty[.$$

Comme $Z \in Z_R$, on rejette H_0, le marché du vin a bien changé.

Exercice 8. Une association sportive locale avait en moyenne 249.3 licenciés par an. Depuis 5 ans, elle organise annuellement un repas associatif afin de mieux se faire connaitre dans la ville. Ainsi, sur ces 5 dernières années, le nombre moyen de licenciés par an est de 265.6 avec un écart-type de 17.4. On suppose les données gaussiennes.

Au risque 5%, déterminer si le repas annuel permet ou non d'attirer plus de licenciés.

Correction. Le nombre de données est de $n = 5$ et elles suivent une loi gaussienne de moyenne m (inconnue).

La valeur de référence est $m_0 = 249.3$.

On fait un test d'hypothèse avec

— $H_0 : m = m_0$: le repas annuel n'attire pas plus de licenciés *versus*

— $H_1 : m > m_0$: le repas a bien un effet positif.

On a $n = 5$, $\hat{m} = 265.6$ et $\hat{\sigma} = 17.4$. Donc, on obtient pour statistique

$$Z = \frac{249.3 - 265.6}{17.4/\sqrt{5}} = -2.09.$$

Comme $n = 5$ et que les données sont gaussiennes, Z suit une loi de Student à 4 degrés de liberté. Comme le risque est 5% et $H_1 : m > m_0$, la zone de rejet Z_R est

$$Z_R =]-\infty, -2.132].$$

Comme $Z \notin Z_R$, on ne rejette pas H_0, le repas annuel n'a pas un effet positif significatif.

Exercice 9. La direction centrale d'une chaîne d'hypermarchés spécifie que le temps d'attente en caisse en heure pleine ne doit pas dépasser 15 minutes en moyenne. Si cette norme n'est pas respectée, le directeur du magasin doit alors procéder à l'engagement de personnel supplémentaire. Cette décision est prise à partir des statistiques mensuelles des magasins.

Sur un échantillon de 135 clients, les statistiques du magasin de Fort-De-France font apparaître un temps d'attente moyen de 17.4 minutes en heure pleine, avec un écart-type de 2.3 minutes.

Au risque 5%, quelle doit être la décision de la direction du magasin ?

Correction. On a $n = 135 \geq 30$, $m_0 = 15$. On teste $H_0 : m = m_0$, aucun changement *versus* $H_1 : m > m_0$, il faut engager quelqu'un.

La statistique est

$$Z = \frac{15 - 17.4}{2.3/\sqrt{135}} = -12.12.$$

Comme le risque est 5% et $H_1 : m > m_0$, la zone de rejet est

$$Z_R =]-\infty, -1.645].$$

Comme $Z \in Z_R$, on rejette H_0 et il faut engager du personnel supplémentaire.

Exercice 10. Une usine produit des tablettes de chocolats de 100g en moyenne. Les 50 dernières tablettes produites pèsent en moyenne 98.5g pour un écart-type de 5.2g.

Au risque 1%, la machine qui produit les tablettes de chocolat est-elle déréglée ?

Correction. On a $n = 50 \geq 30$, $m_0 = 100$. On teste $H_0 : m = m_0$, la machine va bien *versus* $H_1 : m \neq m_0$, la machine est déréglée.

La statistique est

$$Z = \frac{100 - 98.5}{5.2/\sqrt{50}} = 2.04.$$

Comme le risque est 1% et $H_1 : m \neq m_0$, la zone de rejet est

$$Z_R =]-\infty, -2.576] \cup [2.576, +\infty[.$$

Comme $Z \notin Z_R$, on ne rejette pas H_0 et il n'y a pas besoin de régler la machine.