

MST - Logiciel R

TP n° 4

Tests d'hypothèses

Exercice n°1 – Test de conformité d'une moyenne

1. Importer le jeu de données `tempstrajet.txt` qui décrit le temps de trajet domicile-travail d'une personne selon deux itinéraires différents A et B.
2. Créer deux nouveaux jeux de données : un pour chacun des itinéraires A et B.
3. Fournir les statistiques descriptives des temps de trajet de l'itinéraire A ainsi qu'une représentation graphique de ces données.
4. Construire un intervalle de confiance au seuil 95% pour le temps de trajet moyen correspondant à l'itinéraire A. Deux possibilités :
 - en utilisant les formules du cours,
 - par le menu déroulant de Rcmdr : *Statistiques* → *Moyennes* → *t-test univarié*.
Que devient l'intervalle pour des niveaux de confiance de 90% ? de 99% ?
5. Peut-on accepter, au seuil de risque $\alpha = 5\%$, l'hypothèse selon laquelle le temps de trajet moyen μ selon l'itinéraire A est égal à 15 minutes ? Répondre de trois façons différentes (mais strictement équivalentes) :
 - en calculant la région de rejet de H_0 ,
 - par le calcul de la p -value,
 - sur la base de l'intervalle de confiance de μ .
6. Calculer le risque de deuxième espèce si l'on suppose que le vrai temps de trajet est en réalité égal à 16 minutes. En déduire la puissance du test.
7. De même, peut-on accepter l'hypothèse selon laquelle le temps de trajet moyen par l'itinéraire B est de 15 minutes ?

Exercice n°2 – Comparaison de deux moyennes, de deux variances

1. Sur la base du jeu de données global (les deux itinéraires), construire une représentation graphique permettant de comparer les temps de trajet des deux itinéraires. Fournir le temps de trajet moyen et l'écart type du temps de trajet pour chaque itinéraire.
2. Peut-on accepter l'hypothèse selon laquelle le temps de trajet moyen lié à l'itinéraire B est moins élevé que celui de l'itinéraire A ? Réaliser le test au seuil de risque de 5%, puis de 10% avec l'hypothèse de variances différentes.
3. Vérifier que le choix effectué lors de la question précédente (hypothèse d'inégalité des variances) est cohérent en réalisant un test F d'égalité des variances.

Exercice n°3 – Test de conformité d’une proportion

Parmi les étudiants du TP, dénombrer les gauchers et effectuer le test de l’hypothèse (au seuil de risque de 5%) selon laquelle la proportion de gauchers est égale à 15% (proportion estimée en France.)

Exercice n°4 – Test d’indépendance du χ^2

Charger le jeu de données `taillefum` et tester (au seuil de risque de 5%) les hypothèses suivantes :

- Le ratio filles/garçons est identique selon les années.
- Le statut de fumeur d’un étudiant est indépendant de celui de ses parents.
- Le fait de fumer pour un étudiant ne dépend pas de son sexe.