

TP 1 : Tests paramétriques à un échantillon

Exercice 1

Télécharger sur moodle ou sur le commun la base de donnée `QI.dat`. Le QI de 30 enfants dyslexiques a été mesuré au moyen d'un test nommée WISC-3 par des psychologues. L'objectif est de comparer le QI moyen d'un enfant dyslexique par rapport au QI moyen de référence, par définition égale à 100 (cette valeur de référence a été obtenue à partir d'un très large échantillon d'enfants de la population générale).

1. Décrire la variable et procéder à une étude descriptive univariée de cette variable.
2. Au vu des résultats de statistique descriptive, le QI des enfants dyslexiques vous semblent-il différent du QI de référence ?
3. Proposer une méthodologie pour tester si le QI des enfants dyslexiques est différent du QI de référence.
 - (a) Les psychologues affirment que le QI d'un enfant se comporte comme une loi gaussienne. Cette hypothèse vous semble-t-elle valide ici ?
 - (b) Faire le test au niveau $\alpha = 5\%$. On construira le test « pas à pas » : en particulier, on précisera les hypothèses nulle et alternative, la statistique de test ainsi que sa loi sous (H_0). Donner la zone de rejet et conclure.
 - (c) En utilisant cette fois-ci la commande `t.test` de R, calculer la p-valeur de ce test. Expliquer comment cette p-valeur a été obtenue. Commenter
 - (d) Donner l'intervalle de confiance à 95%, puis à 90%, du QI moyen chez les enfants dyslexiques. Commenter.

Exercice 2

Une étude sur l'insuffisance pondérale des nouveaux nés à été menée en 1986 aux États-Unis (Baystate Medical Center, Springfield, Massachusetts) auprès de 189 mères. Les données sont disponibles sur moodle et sur le commun sous le nom de `birthwt.dat`. On qualifie dans la suite un bébé d'hypotrophe si son poids de naissance est inférieur à 2500 grammes.

Les variables de la base sont les suivantes :

- `low` : indicateur d'insuffisance pondérale (1 si poids inférieur à 2500 g, 0 sinon)
- `lwt` : poids de la mère au début de la grossesse (en livre)
- `ht` : hypertension pendant la grossesse (1 si hypertension, 0 sinon).
- `bwt` : poids du bébé en grammes.

1. Décrire les variables et procéder à une étude descriptive univariée de la base de données.
2. On s'intéresse dans un premier temps au poids des mères au début de leur grossesse. Une étude réalisée aux États-Unis a établi que le poids moyen d'une femme au début de sa grossesse est de 130 livres. On se demande alors si le poids moyen des femmes de l'étude est représentatif du poids moyen de 130 livres.

- (a) Quelle méthodologie descriptive proposez-vous pour répondre à la question ? Y-a-t'il des hypothèses de modélisation à vérifier au préalable ?
 - (b) Mettre en place un test pour répondre à la question précédente. On commencera par construire le test « pas à pas », en choisissant un risque de première espèce de 5%. On précisera les hypothèses nulle et alternative, la statistique de test ainsi que sa loi sous (H_0). Donner la zone de rejet et conclure.
 - (c) En utilisant cette fois-ci la commande `t.test` de R, calculer la p-valeur de ce test. Expliquer comment cette p-valeur a été obtenue. Commenter.
3. On s'intéresse à présent à la présence d'hypertension chez les femmes au cours de leur grossesse. La proportion de femmes touchées par l'hypertension en Europe et aux États-Unis est de 11%.
- (a) L'échantillon des 189 femmes de l'étude est-il représentatif de la population des femmes hypertendues au cours de leurs grossesses ? On procédera de la même manière qu'à la question 2.
 - (b) Répondre à la même question pour chacun des deux groupes de mères ayant accouché de bébés hypotrophes et n'ayant pas accouché de bébés hypotrophes. On pourra séparer les données en deux groupes selon les valeurs de la variable *low* pour répondre à cette question.
4. Il est admis que le poids moyen des bébés (nés à terme) à la naissance est de 3,3 kg. Tester si le poids à la naissance des bébés des mères de l'étude s'écarte ou non de ce poids de référence à la naissance.