

Feuille TD Numéro 3 - Les tests (suite)

Plan

- Données qualitatives : Ki2
- Données quantitatives : test de Kruskal et Wallis, Wilcoxon (Mann et Witney)
- Le cas des données appariées.

Tests non paramétriques

Contexte :

- Cas d'une distribution ne suivant pas la loi normale.
- Effectifs réduits et difficultés de valider l'égalité des variances.

Avantages :

- Pas de conditions sur la population dont est extrait l'échantillon.
- Les résultats sont généralement des probabilités exactes.
- Possibilités d'avoir un échantillon de taille inférieur à 6.
- Traitement des données qualitatives.

Inconvénients :

- Moins puissants qu'un test paramétrique si il peut être utilisé.
- Explications parfois moins accessibles dans la littérature.

Test du χ^2

- Cas d'une population composée de deux groupes : sujets malades et non malades.
- Variable explicative : sexe.
- Question : le statut malade/non malade est-il lié au sexe ?
- Hypothèse H_0 : pas de différence d'état malade/non malade en fonction du sexe.

Test du χ^2 - Pearson

- Ecriture du tableau de contingence : vérification des effectifs minimaux et calcul des fréquences théoriques attendues.
- Détermination du degré de liberté : $(\text{nombre de lignes} - 1) * (\text{nombre de colonnes} - 1)$
- Lire une table de χ^2 (donner en annexe)
- Interpréter la P _value -
- Remarque : si les données sont quantitatives, possibilités de créer des classes puis d'utiliser le χ^2 . Attention aux choix des classes et à leur effectif.

Test du Ki2 - mise en oeuvre sur R

- Saisir les données sous forme de matrice.
- Appel au test `chisq.test()`.
- Accès aux champs de résultats : `statistic`,
`parameter`, `p.value`, `observed`, `expected`

Test Wilcoxon

- Statistique de rang : s'affranchit de la différence de moyenne.
- Application : hauteur de 2 sortes d'arbres dans une forêt. Question : les hauteurs médianes de ces 2 peuplements sont-elles égales ?
- Variante : test de Mann-Whitney. Comparaison des 2 réponses pour l'exercice précédent.