
Colles série 6 : Tests paramétriques

Sujet 1 : Charger la base de données `proteine.csv` et affecter à `x` l'échantillon des poids molaires dans la sous-population des individus de règne Metazoa.

1. Quelle est la nature de la variable étudiée ? Représenter sa distribution par un graphique adapté et y ajouter la densité d'une loi normale dont les paramètres seront choisis en fonction de l'échantillon. Donner les statistiques descriptives usuelles de l'échantillon.
2. On sait que dans le règne Viridiplantae le poids molaire est en moyenne de 20500. En supposant que le poids molaire dans le règne Metazoa suit une loi normale, proposer un test permettant de savoir si en moyenne le poids molaire chez les Metazoa dépasse celui connu chez les Viridiplantae. Donner l'expression littérale de la conclusion du test proposé pour un niveau de 3%.

Sujet 2 : Charger la base de données `proteine.csv` et affecter à `x` l'échantillon des poids molaires dans la sous-population des individus de règne Metazoa.

1. Quelle est la nature de la variable étudiée ? Représenter sa distribution par un graphique adapté et y ajouter la densité d'une loi normale dont les paramètres seront choisis en fonction de l'échantillon. Donner les statistiques descriptives usuelles de l'échantillon.
2. On sait que dans le règne Viridiplantae le poids molaire est aléatoire et d'écart-type connu $\sigma_0 = 12125$. En supposant que le poids molaire dans le règne Metazoa suit une loi normale, proposer un test permettant de savoir si la variabilité du poids molaire chez les Metazoa dépasse celle connue chez les Viridiplantae (calculer la p-valeur du test). Donner l'expression littérale de la conclusion du test proposé pour un niveau de 0.1%.

Sujet 3 : Charger la base de données `proteine.csv` et affecter à `x` l'échantillon de la variable `isoelec` dans le genre "Glycine".

1. Soit p la probabilité que `isoelec` dépasse 6 (> 6) dans la sous population des Glycines. Estimer p et en calculer un intervalle de confiance de niveau 99%.
2. On sait que dans la population tous genres confondus la probabilité de dépasser 6 vaut 0.7. Cette probabilité diffère-t-elle lorsque l'on considère la sous-population des Glycines ? Faire le test asymptotique, vu en cours pour répondre à cette question (soit à la main, soit avec `prop.test` en utilisant l'option adaptée). Conclure pour $\alpha = 1\%$.

Sujet 4 : Charger la base de données `proteine.csv` et affecter à `x` l'échantillon de la variable `isoelec` dans le genre "Arabidopsis".

1. Soit p la probabilité que `isoelec` dépasse 6 (> 6) dans la sous population des Arabidopsis. Estimer p et en calculer un intervalle de confiance de niveau 99%.
2. On sait que dans la population tous genres confondus la probabilité de dépasser 6 vaut 0.7. Proposer un test permettant de savoir si cette probabilité est plus faible dans la sous-population des Arabidopsis ? Faire le test asymptotique, vu en cours pour répondre à cette question (soit à la main, soit avec `prop.test` en utilisant l'option adaptée). Conclure pour $\alpha = 1\%$.