Université Joseph Fourier
L2/STA230

TP4 : Intervalles de confiance

L'objectif de ce TP est de construire un intervalle de confiance autour de paramètres estimés tels qu'une proportion, une moyenne, une variance. En effet l'estimateur de la proportion des bébés dont le poids est inférieur à 2,5Kg est une variable aléatoire. L'estimation varie avec l'échantillon. Aussi, il est utile de définir autour d'une valeur estimée un intervalle de confiance dans lequel la valeur recherchée a x% (généralement 95 %) de chances d'appartenir. Plus l'intervalle est court, plus la précision de votre estimation est grande, mais plus le niveau de confiance est faible.
Nous travaillons sur le fichier des BB.csv. Nous nous intéressons à la répartition des bébés suivant leur poids. Nous supposons que le poids suit une distribution gaussienne.
Afin de bien vous préparer aux tests statistiques, il est essentiel de savoir représenter sur un brouillon, la loi de votre estimateur et placer sur votre dessin, l'estimation, les bornes de l'intervalle de confiance ainsi que la probabilité α.

1. Estimez à partir des données brutes le poids moyen des bébés ainsi que la variance et la variance sans biais associées.
2. Quelle loi suit cet estimateur. Pourquoi ?
3. Construisez un intervalle de confiance pour différentes valeurs de α (1%, 5 %, 10%) autour de cette moyenne. Interprétez. *Utilisez la formule de votre cours*
4. Sélectionnez aléatoirement un sous échantillon des bébés, de taille 4 fois plus petite que la taille de l'échantillon initial, *en utilisant ALEA()*,
5. Calculez sur le sous échantillon précédemment construit le poids moyen ainsi que la variance et la variance sans biais associées. Construisez un intervalle de confiance de niveau $1-\alpha$ autour de la moyenne. On prendra les valeurs de $\alpha = 5\%,\ 1\%,\ 10\%,\ 20\%$. Comparez avec les intervalles précédemment calculés. En particulier que pensez vous de la précision de votre estimation ?
6. Estimez à partir des données brutes la variance des poids des bébés.
7. Quelle loi suit cet estimateur. Pourquoi ?
8. Construisez un intervalle de confiance autour de la variance des données brutes. Interprétez.
9. Estimez à partir des données brutes la proportion de bébés pesant moins de 2,5 Kg. Interprétez.
10. Construisez un intervalle de confiance de niveau $1 - \alpha$ autour de cette proportion. On prendra les valeurs de $\alpha = 5\%,\ 1\%,\ 10\%,\ 20\%$.
11. Calculez sur le sous echantillon ci dessus (cf question 4) la proportion de bébés dont le poids est inférieur à 2,5 Kg et construisez un intervalle de confiance de niveau $1-\alpha$ autour de cette proportion. On prendra les valeurs de $\alpha = 5\%,\ 1\%,\ 10\%,\ 20\%$. Comparer avec les intervalles précédemment calculés. En particulier que pensez vous de la précision de votre estimation ?