E2I4 année 2021-2022

Statistiques

Statistiques inférentielles : estimation

Exercice 1 : Simulation Loi de Bernouilli

À l'approche du second tour d'une élection présidentielle, on interroge n personnes au hasard et on note Xi = 1 si le i-ème individu se prononce pour le candidat A et Xi = 0 si c'est pour le candidat B.
Xi suit une loi de Bernoulli de paramètre p inconnu.

1. À l'aide de Python, choisir p au hasard dans l'intervalle [0; 1] (et pour le moment de regarder pas la valeur car nous allons essayer de l'estimer)

2. On va estimer le paramètre p à l'aide d'un échantillon. Créer un échantillon de taille n = 100 réalisations d'une variable aléatoire de Bernouilli à l'aide de la commande grand.

3. Donner une estimation de p. Comparer cette estimation ponctuelle avec la vraie valeur de p

4. Répéter le procédé pour 1000 échantillons chacun de taille n = 100, c'est à dire donner 1000 estimations du paramètre p. Pour cela, créer un vecteur F qui contient ces 1000 estimations.

5. Tracer l'histogramme de F. Qu'observez-vous ?

Exercice 2 : intervalle de confiance

Le but de cet exercice est que vous écriviez 2 programmes (ou fonctions) qui vous permettront de retourner à partir d'un échantillon de valeurs l'intervalle de confiance de la moyenne de la population (et de la proportion)

Les fonctions prennent en entrée un échantillon de valeurs, et un seuil de confiance (ou d'erreur)

Exercice 3 :
Créer un échantillon de taille n = 1000 contenant n réalisations d'une variable aléatoire qui suit une loi normale de moyenne μ et d'écart type σ.

Fixer μ et σ. Générer alors 100 valeurs issues de la loi normale et calculer une réalisation de l'intervalle de confiance de la moyenne μ.

Répéter cela 100 fois en testant chaque fois si l'intervalle de confiance contient ou pas μ