

# SCI03 Automne 2014

## TP 2: Echantillonnage, estimation

### 1 Echantillonnage

1. Générez une réalisation d'un échantillon de taille  $n = 50$ , de variable aléatoire parente  $X$  suivant une loi continue uniforme sur l'intervalle  $[1, 3]$ .
2. Calculez la moyenne empirique  $\bar{x}$  et la variance empirique  $s^{*2}$  de cet échantillon. Comparez ces valeurs aux valeurs théoriques (espérance et variance de  $X$ ).
3. Répétez 100 fois les calculs précédents. Calculez la moyenne et la variance des 100 valeurs de  $\bar{x}$  obtenues. Comparez ces valeurs aux valeurs théoriques (espérance et variance de  $\bar{X}$ ).
4. Répétez les calculs précédents pour différentes valeurs de  $n$  entre 1 et  $10^4$ . Qu'observez vous ? Représentez graphiquement les résultats.
5. Pour  $n = 1, 10, 100$  et  $1000$ , faites un histogramme des 100 valeurs de  $\bar{x}$ . Qu'observez vous ? Superposez la courbe de densité d'une loi normale avec des paramètres convenablement choisis.
6. Répétez les calculs précédents (questions 1 à 4) pour  $X$  suivant cette fois une loi de Bernoulli de paramètre  $p = 0.3$ .

### 2 Estimation

#### 2.1 Données CEO

Le jeu de données `CEO.txt`, publié en 1993 par le magazine Forbes, contient des données relatives à l'âge et au salaire (en milliers de dollars) de 59 chefs d'entreprises de taille moyenne (chiffre d'affaire annuel entre 5 et 350 millions de dollars).

Calculez des intervalles de confiance sur l'âge et le salaire moyens dans la population totale, pour différents niveaux de confiance compris entre 90% et 99%. Qu'observez vous ?

Calculez des intervalles de confiance sur la proportion de chefs d'entreprise âgés de moins de 40 ans, puis de moins de 50 ans, pour différents niveaux de confiance compris entre 90% et 99%. Qu'observez vous ?

#### 2.2 Simulations

1. Générez une réalisation d'un échantillon de taille  $n = 20$ , de variable aléatoire parente  $X$  suivant une loi normale d'espérance  $\mu = 5$  et d'écart-type  $\sigma = 1$ . Calculez un intervalle de confiance sur  $\mu$  de niveau de confiance  $1 - \alpha = 0.95$ .
2. Répétez 1000 fois le calcul précédent. Combien de fois l'intervalle de confiance contient-il la valeur  $\mu = 5$ ? Comparez cette valeur à la valeur théorique et expliquez ce résultat.
3. Répétez les calculs précédents en prenant  $1 - \alpha = 0.99$ .
4. Répétez les calculs précédents (questions 1 à 4) en prenant  $n = 30$  et  $X$  suivant une loi uniforme sur  $[0, 2]$ .