Université de Pau et des Pays de l'Adour **Année 2019-2020**
Département de Mathématiques
M1-Génie Pétrolier - Analyse de données

TP-5 : Théorème central limite

Le but est d'illustrer le théorème central limite.

1 Variable aléatoire discrète

Une variable aléatoire X peut prendre les valeurs allant de 1 à 6. Le vecteur des probabilités associées est

$$proba = (0.32, 0.05, 0.12, 0.15, 0.18, 0.18).$$

Cette variable aléatoire représente le nombre d'ouvriers présents un jour donné sur une plate-forme pétrolière.

1. Calculer l'espérance et l'écart type de la variable aléatoire X.

2. À l'aide de la fonction `sample`, générer $n = 50$ réalisations de la variable aléatoire X.

3. Effectuer un diagramme en bâtons des 50 réalisations. Comparer les fréquaneces empiriques aux probabilités. Que se passe-t-il si on augmente n ?

4. Comparer la moyenne empirique $\bar{x}$ et la moyenne théorique $\mu = \mathbb{E}(X)$. Que se passe-t-il lorsque l'on augmente la taille n de l'échantillon ? Existe-t-il un théorème qui justifie ce résultat ?

5. Lancer plusieurs fois votre programme (`Ctrl+Alt+R`) et constater que vous n'obtenez pas exactement la même valeur pour $\bar{x}$: $\bar{x}$ est une variable aléatoire.

6. Générer maintenant $m = 20$ valeurs de $\bar{x}$. Appeler Xbarv le vecteur des m valeurs. Pour chaque valeur de $\bar{x}$ on utilisera un échantillon de $n = 5$ réalisations de la variable aléatoire X.

7. Tracer un histogramme de $Xbarv$. Que devient l'histogramme lorsque l'on augmente m ?

8. Que devient l'histogramme lorsque l'on augmente n ?

9. D'après le théorème central limite, nous savons que $\bar{X}$ suit une loi normale de moyenne $\mu = \mathbb{E}(X)$ et d'écart type $\frac{\sigma}{\sqrt{n}}$. Sur le graphique de l'histogramme, ajouter à l'aide de la commande `lines`, la densité de la loi normale (commande `dnorm`) de moyenne μ et d'écart type $\frac{\sigma}{\sqrt{n}}$.

10. Faire varier n et observer le résultat.

2 Variable aléatoire continue

Nous allons maintenant illustrer le théorème central limite pour le cas où X suit une loi exponentielle de paramètre $\lambda = \frac{1}{3}$.

1. À l'aide de la fonction `rexp`, générer $n = 50$ réalisations de la variable aléatoire X.
2. Comparer la moyenne empirique $\bar{X}$ et la moyenne théorique $\mu = \mathbb{E}(X)$.
3. Lancer plusieurs fois votre programme et constater que vous n'obtenez pas exactement la même valeur pour $\bar{X}$: $\bar{X}$ est une variable aléatoire.
4. Générer maintenant $m = 20$ valeurs de $\bar{X}$. Appeler $Xbarv$ le vecteur des m valeurs. Pour chaque valeur de $\bar{X}$ on utilisera un échantillon de $n = 50$ réalisations de la variable aléatoire X.
5. Tracer un histogramme de $Xbarv$. Que devient l'histogramme lorsque l'on augmente n ?
6. D'après le théorème central limite, nous savons que $\bar{X}$ suit une loi normale de moyenne $\mu = \mathbb{E}(X)$ et d'écart type $\frac{\sigma}{\sqrt{n}}$. Sur le graphique de l'histogramme, ajouter à l'aide de la commande `curve`, la densité de la loi normale de moyenne μ et d'écart type $\frac{\sigma}{\sqrt{n}}$.