TP 2

Exercice 1 Calcul de lois discrètes : les fonctions de répartition

1. Soit X une variable aléatoire de loi binomiale $B(30, 1/10)$. Calculer $P[X = 2]$ puis $P[X \leqslant 4]$, en effectuant un calcul exact puis en utilisant la loi de Poisson de paramètre np.
2. Soit Y une variable aléatoire de loi binomiale $B(n = 50, p = 1/5)$. Calculer $P[2 \leqslant Y \leqslant 6]$ puis $P[Y > 4]$, en effectuant un calcul exact puis en utilisant la loi normale. Pour une approximation plus précise, on peut utiliser une correction en continuité :
$$P[X \leqslant k] \backsimeq \int_{-\infty}^{k+1/2} f(x)dx$$
où f est la densité de la loi normale $N(np, np(1-p))$.
Nota bene : la commande pour la fonction de répartition de la loi binomiale est `cdfbin`

Exercice 2 Estimation

1. On dispose d'un échantillon simulé de taille $n = 50$ de variables aléatoires iid de loi normale $N(m = 5, \sigma^2 = 2)$. Après avoir simulé cet échantillon, déterminez à partir de ces observations uniquement, grâce au théorème central limite, un intervalle qui contienne m avec probabilité $1 - \alpha = 95\%$. Un tel intervalle est appelé intervalle de confiance, de niveau de confiance $1 - \alpha$.
Refaire le calcul avec un niveau de confiance $1 - \alpha = 98\%$. Au vu des observations toujours, estimer quelle valeur de n choisir pour obtenir un intervalle de confiance de niveau de confiance $1 - \alpha = 95\%$ et qui soit de largeur 0.2.
2. Considérons une v.a. X de loi uniforme sur l'intervalle $[0, \theta]$. On cherche à estimer le paramètre θ à partir de l'observation d'un échantillon $(X_1, ..., X_n)$, n-uplet de v.a. iid de même loi que X. Posons :
$$\hat{\theta}_n^{(1)} = \max_i X_i \qquad \text{et} \qquad \hat{\theta}_n^{(2)} = \frac{2}{n}\sum_i X_i$$
   (a) À partir d'un échantillon de taille $N = 100$, tracer les deux suites $\left(\hat{\theta}_n^{(1)}\right)_{n=1,...,N}$ et $\left(\hat{\theta}_n^{(2)}\right)_{n=1,...,N}$. Quel est l'estimateur qui converge le plus rapidement ?
   (b) On définit le biais d'un estimateur $\hat{\theta}$ par
$$Biais = E[\hat{\theta}] - \theta$$
   Quel est le biais de $\hat{\theta}_n^{(2)}$?
   (c) On peut estimer $E\left[\hat{\theta}_n^{(1)}\right]$ en calculant la moyenne empirique de $N = 200$ réalisations de $\hat{\theta}_n^{(1)}$ en vertu de la loi des grands nombres.
   Pour chaque $n = 10, 20, ..., 100$, estimer le biais de $\hat{\theta}_n^{(1)}$ en simulant $N = 200$ échantillons qui fourniront $N = 200$ réalisations de $\hat{\theta}_n^{(1)}$. Mener le calcul théorique et tracer les deux fonctions *Biais* obtenues (la fonction exacte et la fonction estimée).
   (d) La précision d'un estimateur est donné par l'erreur quadratique moyenne :
$$EQM = E\left[\left(\hat{\theta} - \theta\right)^2\right]$$
   Comme pour le biais, calculer l'erreur quadratique moyenne pour les deux estimateurs, puis estimer-la pour plusieurs valeurs de n.

Exercice 3 Calcul d'intégrales par Monte Carlo

Dans certains cas, les méthodes numériques classiques de calcul d'intégrales sont en échec (fonction pas régulière, grande dimension). Dans ces cas, on peut recourir au calcul de Monte Carlo qui consiste à écrire l'intégrale comme l'espérance d'une variable aléatoire. La loi des grands nombres fait le reste.

1. On cherche à calculer l'aire du domaine du plan

$$D = \{(x, y) \in [-10, 10]^2 : y^3 - x * y + 2x^2 \exp(-x^2/100) \cos(y) < 1\}$$

   Écrire cette aire comme l'espérance de $f(X, Y)$ où X et Y sont des v.a. indépendantes de loi uniforme sur $U[-10, 10]$. Utiliser la loi des grands nombres pour estimer cette espérance par une moyenne empirique puis utiliser le théorème central limite pour évaluer la précision de cette estimation.

2. Estimer le volume de la sphère d-dimensionnelle. On pourra choisir $d = 1, 2, 3, 5, 10, 20$. Simuler des nombres uniformément répartis sur [-1,1], puis des réalisations d'une loi normale pour améliorer la précision en grande dimension.

Exercice 4 Ruine du joueur

Deux joueurs A et B jouent à Pile ou Face, Face arrivant avec une probabilité 1/2. On suppose que les résultats successifs des lancers sont iid. Soient X_n la fortune en euros de A à l'instant n, a et b les fortunes initiales de A et B respectivement. Soit $c = a + b$, $q = 1 - p$. Le jeu s'arrête quand un des joueurs est ruiné, à l'instant noté T. On s'intéresse à la probabilité que A gagne :

$$u(a) = P[X_T = c | X_0 = a]$$

Simuler tout d'abord une partie jusqu'à la ruine de l'un des joueurs (on peut montrer qu'elle arrive presque sûrement).
Estimer la probabilité $u(a)$ pour diverses valeurs de a, à partir de la simulation de nombreuses parties.
Rendre le jeu déséquilibré en faisant de sorte que 'pile' arrive plus souvent et regarder ce qui se passe.