TP noté

Exercice 1 Une généralisation de la loi géométrique **11 points**

Soit p dans l'intervalle $]0\,;1[$. On considère une variable aléatoire X de loi de Bernoulli de paramètre p et X_1, X_2, ..., X_n, ... une suite de variables aléatoires indépendantes ayant la même loi que X. Si $i \geq 1$, on dit qu'il y a un succès au temps i lorsque $X_i = 1$.

Soit k un entier supérieur ou égal à 1. La variable aléatoire T_k est la valeur du premier instant où il y a eu exactement k succès.

Par exemple, si les premiers tirages sont :

$$X_1 = 0,\ X_2 = 1,\ X_3 = 0,\ X_4 = 1,\ X_5 = 1,\ X_6 = 0,\ X_7 = 1,$$

alors, $T_1 = 2$, $T_2 = 4$, $T_3 = 5$ et $T_4 = 7$.

1. Quelle instruction peut-on entrer dans Matlab ou Octave pour obtenir un nombre aléatoire égal à 1 avec probabilité p et à 0 avec probabilité $1 - p$?
   Solution: En supposant qu'un variable `p` de type `double` a déjà été affectée et contient un réel entre 0 et 1, il suffit d'entrer la commande `rand() <= p`.
2. Écrire dans un fichier de fonction `simuler_Tk.m` la fonction `simuler_Tk(p,k)` dont la valeur de retour est un nombre aléatoire qui a la même loi que T_k lorsque le paramètre des variables de Bernoulli indépendantes X_1, X_2, ... est égal à `p`.
   Solution: La fonction suivante répond aux questions 2 et 3.

```
function res = simuler_Tk(p, k, N)
  res = zeros(1, N);
  for i = 1:N
    nb_succes = 0;
    temps = 0;
    while nb_succes < k
      X = (rand() <= p);
      temps = temps + 1;
      if X % succes
        nb_succes = nb_succes + 1;
      end %if
    end %while
    res(i) = temps;
  end
end
```

3. Modifier cette fonction de manière à faire que `simuler_Tk(p,k,N)` retourne un vecteur ligne de taille `N` dont toutes les composantes sont des réalisations indépendantes ayant la même loi que T_k.
4. À l'aide de cette fonction, écrire, dans un script `loi_Tk.m` des instructions permettant d'afficher une valeur approchée de l'espérance de T_3 pour $p = 0,4$, à partir d'un échantillon de taille `N = 2000`. *Vous pourrez comparer avec la valeur théorique qui vaut* $3/0,4$.
   Solution: Réponse aux questions 4 et 5.

```
% Exercice sur la generalisation de la loi geometrique
1;
k = 3; p = 0.4; % paramètres
N = 2000; %taille de l'échantillon
echantillon = simuler_Tk(p, k, N);
m = mean(echantillon);
disp("L'espérance de T_k vaut environ : "); disp(m);
disp("Espérance exacte : "); disp( k / p);
```

```
freq_6 = sum( echantillon == 6 )/N;
disp("fréquence empirique de 6 : "); disp(freq_6);
disp("probabilité d'obtenir 6 : ");
proba = nchoosek(6-1, k-1) * p^k * (1-p)^(6-k);
disp(proba);
```

5. Compléter ce même script pour afficher une valeur approchée de la probabilité que T_3 soit égal à 6, toujours pour $p = 0,4$, à l'aide du même échantillon que précédemment. *Remarque : la valeur théorique est ici* $10p^3(1-p)^3$.

Exercice 2 Méthode de rejet **5 points**

On considère la fonction

$$f(x) = \begin{cases} \frac{1}{2}x + \frac{1}{2} & \text{si } x \text{ est dans } [-1\,;1]\,; \\ 0 & \text{sinon.} \end{cases}$$

On admet que f est bien une densité de probabilité. En utilisant la méthode de rejet, écrire une fonction `rejet()` sans argument qui retourne un nombre aléatoire dont la loi est à densité f.

```
function X = rejet()
  while true
    Y = -1 + 2*rand(); %uniforme sur [-1,1]
    U = rand();
    % ici c = 1/2 et g constante = 2 donc c*g(Y)=1 toujours
    if ( U <= (Y+1)/2 ) %on accepte
      X = Y;
      break;
    end
  end
end
```

Et si on veut tester, on peut écrire le fichier `test_rejet.m` suivant.

```
1;
N = 1000;
ech = zeros(1,N);
for i = 1:N
  ech(i) = rejet();
end
hist(ech,20);
```

Exercice 3 Méthode de l'inverse de la fonction de répartition **5points points**

On considère la fonction

$$f(x) = \begin{cases} \frac{1}{x\ln(2)} & \text{si } x \text{ est dans } [1\,;2]\,; \\ 0 & \text{sinon.} \end{cases}$$

On admet que f est bien une densité de probabilité et que sa fonction de répartition est la fonction F donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 1\,; \\ \ln(x)/\ln(2) & \text{si } 1 \leq x < 2\,; \\ 1 & \text{si } x \geq 2. \end{cases}$$

1. ✐Résoudre, pour u dans $[0\,;1]$, l'équation d'inconnue x dans $[1\,;2]$:

$$F(x) = u.$$

Solution: Pour tout u dans $[0\,;1]$ et x dans $[1\,;2]$, on a

$$F(x) = u \iff \ln(x) = \ln(2)u \iff x = \exp(\ln(2)u) \iff x = 2^u.$$

La solution $x = \exp(\ln(2)u)$ était aussi acceptée.

2. À l'aide de la méthode de l'inverse de la fonction de répartition, écrire une fonction Matlab/Octave `inv_repartition(N)` qui retourne un N nombres aléatoires indépendants dont la loi est à densité f.
Solution:

```
function X = inv_rep(N)
  X = 2.^(rand(1,N));
end
```

Pour tester on pouvait se servir du script suivant :

```
% besoin du fichier densite_empirique.m pour fonctionner
1;
clf
densite_empirique(inv_rep(10000), 50);
hold on
X = linspace(1,2,1000);
Y = 1 ./ (X .* log(2));
plot(X,Y,'r');
hold off
```

Exercice 4 Méthode de Monte-Carlo

1. À l'aide de la méthode de Monte-Carlo, donner une valeur approché de l'intégrale

$$\int_0^1 \sin(\pi x)\,\mathrm{d}x.$$

On utilisera un échantillon de taille 10000. *Remarque : la valeur exacte de cette intégrale se calcule facilement et vaut* $2/\pi$. On écrira un script `monte_carlo1.m` pour résoudre l'exercice.
Solution: Si X_1, X_2, ... sont des variables aléatoires indépendantes de loi commune la loi uniforme sur $[0\,;1]$, alors l'intégrale à calculer vaut $\mathbb{E}\left[\sin(\pi X_1)\right]$. D'après la loi des grands nombres, on a la convergence presque sûre

$$\mathbb{E}\left[\sin(\pi X_1)\right] = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \sin\left(\pi X_i\right).$$

On en déduit le script suivant pour calculer une valeure approchée par la méthode de Monte-Carlo.

```
1;
N = 10000; % taille de l'échantillon
ech = rand(1, N);
X = sin( pi * ech );
m = mean(X);
disp("Valeur approchée de l'intégrale : ");
disp(m)
disp("Valeur numérique (presque) exacte :");
disp( 2/pi );
```

2. En utilisant la méthode de Monte-Carlo, écrire un script `monte_carlo2.m` qui permet d'obtenir, à l'aide d'un échantillon de taille 10000, une valeur approchée de l'intégrale

$$\int_{-\infty}^{+\infty} x^2 e^{-x^2/2}\,\mathrm{d}x.$$

Remarque : la valeur exacte de cette intégrale est $\sqrt{2\pi}$.
Solution: Ici, il fallait (un peu) travailler l'intégrale de façon à faire apparaître la densité de la loi normale centrée réduite.

$$\int_{-\infty}^{+\infty} x^2 e^{-x^2/2}\,\mathrm{d}x = \int_{-\infty}^{+\infty} x^2 \sqrt{2\pi}\frac{e^{-x^2/2}}{\sqrt{2\pi}} = \mathbb{E}\left[N^2\sqrt{2\pi}\right],$$

où N est une variable aléatoire de loi normale centrée réduite. On en déduit, avec les mêmes arguments que précédemmment, le script suivant pour donner une valeur approchée de l'intégrale.

```
N = 10000; % taille de l'échantillon
ech_norm = randn(1, N);
Y = sqrt(2*pi) * ech_norm .^2;
m = mean(Y);
disp("Valeur approchée de l'intégrale : ");
disp(m);
disp("Valeur (presque) exacte : ");
disp( sqrt(2*pi) );
```