CPP-la prépa des INP 2 ème année
année 2015-2016

TP 1: Probabilités et statistiques

Lancer matlab. Aller dans l'onglet *file* puis *set path* et ajouter le chemin

```
C:\Program Files\MATLAB\R2012a\toolbox\shared\stixbox
```

Le but de cette manipulation est de charger certaines nouvelles commandes matlab comme par exemple la commande *histo* qui permet de tracer un histogramme normalisé contrairement à la commande classique *hist* de matlab. (Si ce n'est pas possible, il faudra calculer l'aire de chaque histogramme et multiplier la densité avec laquelle on veut comparer l'histogramme...)

Partie 1. Prise en main

Essayer les commandes

```
>> A= [1,2;3,4]
>> A^2;
>> A^2
>> A.^2
>> B=[1:10]
>> cumsum (B)
>> max(B)

>> x= [-4:0.01:4];
>> f= 1/ sqrt(2*pi) *exp(-x.^2 / 2);
>> F=0.01*cumsum(f);
>> plot(x,f)
>> plot(x,F)
>> help pnorm
>> y=pnorm(x,0,1);
>> plot(x,f,x,y)
```

On va maintenant s'intéresser d'un peu plus près aux probabilités.

```
>>U=rand(1,5)
>>U=rand(1,1000);
>>histo(U, 20, 0,1)
>>hold on
>>x=[0:0.01:1];
>>size(x)
>>f=ones(1,size(x))
>>plot(x,f,'r')
```

Dans la commande *histo*, le 20 ici désigne le nombre de classes de l'histogramme, 0 un paramètre de centrage des classes et 1 l'aire totale de l'histogramme. La commande *hold on* permet de continuer à tracer sur le même graphique. La commande *hold off* l'arrête.

Que fait la commande *rand*?

Compléter le programme suivant:

```
>>U=rand(1,5)
>>U=1 +3 * rand(1,1000);
>>histo(U, 20, 0,1)
>>hold on
>>x=[1:0.01:4];
>>size(x)
>>f= ...
>>plot(x,f,'r')
```

Partie 2. Simulation de lois de variables aléatoires Ouvrir un nouveau script matlab, que l'on sauvera sous le nom *expo.m* et taper dans le srcipt les commandes

```
function A=expo(N)
nc=floor(sqrt(N));  % nc désigne le nombre de classes de l'histogramme
U=rand(1,N);
A= -2*log(U);
histo(A,nc,0,1)
x_max=max(A);
hold on
x=[0:0.01:x_max];
f=1/2* exp(-1/2 *x);
plot(x,f)
hold off
```

Tester maintenant ce programme dans matlab pour $N = 100$, $N = 1000$ et $nc = 30$. Que fait ce programme? Justifier théoriquement.

Taper directement dans matlab les commmandes

```
A=[1,4,2,9,8]
B=sort(A)
C=[1:5]/5
stairs(B,C)
```

Ecrire un programme *expo2(N)* permettant de tracer sur un même graphe la fonction de répartition empirique ainsi que la fonction de répartition théorique des variables aléatoires de l'exemple précédent. On rappelle que la fonction de répartition empirique est donnée par

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{X_i \leq t}.$$

Soit $U_1, \ldots, U_k$ des variables aléatoires de loi uniforme sur $[0, 1]$. On construit alors $M = \max_{1 \leq i \leq k} U_i$. On veut maintenant simuler N variables aléatoires indépendantes de même loi que M et comparer l'histogramme avec la densité théorique. En TD, on a vu que la densité de M est donnée par

$$f_M(t) = kt^{k-1}\mathbf{1}_{[0,1]}(t).$$

Recopier et compléter le programme suivant dans l'éditeur associé à matlab.

```
function M=maxU(N,k)
nc=floor(sqrt(N));
M=zeros(1,N);
for i=1:N
A=rand(1,k);
M(1,i)= ... ;
end

histo(M,nc,0,1)
hold on
x=[0:0.01:1];
f=...;
plot(x,f)
hold off
```

On veut enfin simuler des variables aléatoires indépendantes de loi normale. Ceci est fait dans matlab avec la commande *randn*. (Pour plus d'informations, ne pas hésiter à taper *help randn*...). Ecrire un programme que l'on apperlera *loinormale(N)* pour simuler N variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$ et comparer l'histogramme obtenu à la densité théorique. On rappelle que la densité de la loi normale centrée réduite est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

On peut se baser sur le bout programme suivant

```
function Z=loinormale(N)

nc=floor(sqrt(N));
```

```
Z=...;
histo(Z,nc,0,1)
hold on
x=[-4:0.01:4];
f=...;
plot(x,f)
hold off
```

Partie 3. Illustration de la loi forte des grands nombres
Enoncé de la loi forte des grands nombres:
Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires et de même loi, intégrable. On note $m = \mathbb{E}[X_1]$ et on pose $S_n = X_1 + \cdots + X_n$. Alors, avec probabilité 1,

$$\frac{S_n}{n} \longrightarrow_{n\to+\infty} m.$$

On va maintenant essayer d'illustrer la loi forte des grands nombres avec des variables aléatoires de loi exponentielle $\mathcal{E}(\lambda)$. On rappelle que si $X \sim \mathcal{E}(\lambda)$, alors $\mathbb{E}[X] = \frac{1}{\lambda}$. Pour illustrer la loi forte des grands nombres, il nous faut construire la suite de valeurs $a_n = \frac{\sum_{i=1}^{n} X_i}{n}$. Pour cela on crée un vecteur $X = (X_1, \ldots, X_n)$ avec des variables aléatoires indépendantes de loi $\mathcal{E}(2)$; puis, avec la commande *cumsum*, on crée le vecteur $S = (X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_n)$. Il ne reste plus qu'à diviser coordonnée par coordonnées le vecteur S par le vecteur $C = (1, \ldots, n)$ avec la commande "./". Enfin, il fait faire tracer la suite (a_n) ainsi que la valeur limite théorique.

Recopier et compléter le programme suivant dans l'éditeur associé à matlab.

```
function A=LFGN(n,lambda)
U=rand(1,n);
X=...;
B=...;
C=[1:n];
A=...;
plot(C,A,'b',C,..., 'r')
```

Tester ce programme pour différentes valeurs de λ différentes tailles d'échantillon n.

Partie 4. Illustration du théorème central limite

Enoncé du théorème central limite:
Soit $(X_i)_{i\geq 1}$ une suite des variables aléatoires de même loi, de carré intégrable. On note $m = \mathbb{E}[X_1]$, $\sigma^2 = Var(X_1)$ et on pose $S_n = X_1 + \cdots + X_n$. Alors, quand $n \to \infty$,

$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - m\right) \longrightarrow^{Loi} \mathcal{N}(0,1).$$

On va maintenant illustrer le théorème central limite. Pour cela, il faut se fixer n suffiamment grand et construire un grand nombre N de variables aléatoires indépendantes Y_i s'écrivant chacune:

$$Y_i = \frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - m\right),\ 1 \leq i \leq N.$$

Puis il faudra comparer l'histogramme de ces N variables aléatoires Y_i avec la densité de la loi gaussienne.

On va prendre ici avec des variables aléatoires de loi uniforme sur $[0,1]$. On rappelle que $\mathbb{E}[X] = \frac{1}{2}$ et $\mathrm{Var}(X) = \frac{1}{12}$. Compléter le programme suivant:

```
 function Z=TLC(n,N)
Y=zeros(1,N);
nc=floor(sqrt(N));
for i=1:N
U=rand(1,n);
s=sum(U);
Y(1,i)= ... ;
end;
histo(Y,nc,0,1)
hold on
x=[-4:0.01:4];
f=...;
plot(x,f)
hold off
```

Illustrer le théorème central limite pour des variables aléatoires de loi exponentielle.