ECE 2 - Année 2018-2019
Lycée français de Vienne
Mathématiques - F. Gaunard
`http://frederic.gaunard.com`

T.P. n°7
Autour du Théorème Central Limite
Convergence(s) en loi

---

1 Rappel: Énoncé du Théorème central limite

Théorème 1. *Soit (X_n) une suite de v.a. indépendantes de même loi, d'espérance (commune) finie m et de variance (commune) finie (et non nulle) σ^2. Pour $n \in \mathbb{N}^*$, on note*

$$S_n = \sum_{k=1}^{n} X_k, \quad \overline{X}_n = \frac{S_n}{n}$$

et

$$S_n^* = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{S_n - nm}{\sigma\sqrt{n}}, = \frac{\sqrt{n}(\overline{X}_n - m)}{\sigma} = \frac{\overline{X}_n - E(\overline{X}_n)}{\sigma(\overline{X}_n)} = \overline{X}_n^*.$$

Alors,

$$S_n^* \xrightarrow{\mathcal{L}} Z \qquad \text{et} \qquad \overline{X}_n^* \xrightarrow{\mathcal{L}} Z, \qquad \text{où} \qquad Z \hookrightarrow \mathcal{N}(0,1).$$

En particulier, notant Φ la fonction de répartition de la loi normale $\mathcal{N}(0,1)$, on a, pour tous réels a, b avec $a < b$,

$$\lim_{n \to +\infty} P\left(a < S_n^* \leq b\right) = \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} \mathrm{d}t.$$

2 La loi normale sous `SciLab`

Pour toute la suite de ce TP, on aura besoin de la densité f (et de la fonction de répartition Φ) de la loi normale centrée réduite

$$f\ x \longmapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}, \qquad \Phi(t) = \int_{-\infty}^{t} f(x)\mathrm{d}x.$$

Sous `Scilab`, un calcul approché de la fonction de répartition Φ est implémenté par la fonction `cdfnor()` (**c**umulative **d**istribution **f**unction **nor**mal distribution). L'appel général est le suivant:

`[P, Q] = cdfnor("PQ", X, Mean, Std),` où

- `X` est la borne d'intégration;
- `Mean` est l'espérance choisie (la loi n'est pas forcément centrée)
- `Std` est l'écart-type choisi (la loi n'est pas forcément réduite)
- `P` et `Q` sont les résultats de l'appel, c'est la valeur approchée de

$$P = \int_{-\infty}^{x} f_{m,\sigma^2}(t)\mathrm{d}t, \quad Q = 1 - P = \int_{x}^{+\infty} f_{m,\sigma^2}(t)\mathrm{d}t.$$

Exercice 1. (densité et fonction de répartition de $\mathcal{N}(0,1)$)

(1) Écrire une fonction `y=densiteNorCR(x)` correspondant à la densité de la loi normale centrée réduite.
(2) Écrire une fonction `y=repartitionNorCR(x)` correspondant à la fonction de répartition de la loi normale centrée réduite.

Exercice 2. Dans une population, 40% des individus possèdent le caractère C. On considère un échantillon de 200 de ces individus. À l'aide du Théorème central limite et et la fonction `repartitionNorCR()` précédemment écrire, estimer la probabilité que la fréquence d'apparition du caractère C dans cet échantillon soit comprise entre 30% et 50%.

3 Simulation de la loi normale *via* la loi uniforme

Exercice 3. Soit (X_n) une suite de v.a. indépendantes de même loi uniforme $\mathcal{U}([0;1])$.

(1) Justifier que
$$\sqrt{\frac{12}{n}}\sum_{k=1}^{n} X_n - \sqrt{3n} \xrightarrow{\mathcal{L}} X, \qquad \text{où} \quad X \hookrightarrow \mathcal{N}(0,1).$$
(2) Utiliser cette observation pour compléter le programme `SciLab` suivant permettant de simuler la variable
$$S_n^* = \sqrt{\frac{12}{n}}\sum_{k=1}^{n} X_n - \sqrt{3n}$$

```
function y=normale(n)
    s=0;
    for k=1:n
        s=s+..........................
    end
    y=................................
endfunction
```

(3) Compléter le programme suivant pour représenter simultanément le diagramme obtenu avec un échantillon de taille $N = 1000$ et $n = 12$ ainsi que la densité de la loi normale $\mathcal{N}(0;1)$.

```
n=12;
//simulation d'un échantillon de S_n*
S=zeros(1,1000);
for k=1:1000
    ................................
end

//répartition des effectifs par classe, calcul des fréquences et diagramme
classes=-10:.2:10
nbc=.....................; //nombre de classes
[i, eff]=dsearch(S,classes);
freq=...............................
bar(................., ..................)

//courbe de la densité de la loi normale centrée réduite
Y=feval(..............)
plot2d(........, ......., style=5) // style=5 correspond à une courbe rouge
```

(4) Reprendre le programme suivant afin de comparer la fonction de répartition et le diagramme des effectifs cumulés.

4 Annales de concours (et autres exercices) sur le sujet

On propose un regroupement d'exercices dont les questions sous `SciLab` font référence aux notions du Chapitre 11 (Convergences et approximations), comme l'utilisation du Théorème Central Limite, de la loi faible des grands nombres, etc...

Exercice 4. (D'après **EDHEC 2015**)
Trois personnes, notées A, B et C entrent simultanément dans une agence bancaire disposant de deux guichets. Les clients A et B occupent simultanément à l'instant 0 les deux guichets tandis que C attend que l'un des deux guichets se libère pour se faire servir.

On suppose que :

- Les durées de passage au guichet des trois personnes A, B et C sont mesurées en heures et on suppose que ce sont des variables aléatoires indépendantes, notées respectivement X Y et Z, et suivant toutes la loi uniforme sur $[0,1[$.
- La durée du changement de personne à un guichet est négligeable.

(1) On pose $U=\min(X,Y)$ et $V=\max(X,Y)$ et on admet que U et V sont des variables aléatoires.
 (a) Montrer que la fonction de répartition F_U de U est définie par :
$$F_U(x)=\begin{cases} 0, & \text{si } x<0\\ 2x-x^2, & \text{si } 0\leqslant x\leqslant 1\\ 1, & \text{si } x>1\end{cases}$$
 (b) En déduire que U est une variable aléatoire à densité et donner f_U de U.
 (c) Déterminer l'espérance et la variance de U.

(2) On note T le temps total passé par C dans l'agence bancaire.
 (a) Exprimer T en fonction de certaines variables précédentes.
 (b) En déduire $E(T)$ et $V(T)$.

(3) (a) On rappelle que, si `a` et `b` sont deux vecteurs lignes de taille n, les commandes `m=min(a,b)` et `M=max(a,b)` renvoient les vecteurs `m` et `M`, de même taille que `a` et `b`, et tels que, pour tout i de $[\![1,n]\!]$, on ait : `m(i)=min(a(i),b(i))` et `M(i)=max(a(i),b(i))`.
 On rappelle également que `grand(1,n,'unf',0,1)` simule n variables aléatoires indépendantes suivant la loi uniforme sur $[0,1[$.
 Complèter les commandes `Scilab` suivantes pour qu'elles permettent de simuler n fois les variables aléatoires U, V et T, pour n entré par l'utilisateur :

```
n=input('n=?')
x=grand(1,n,'unf',0,1)
y=grand(1,n,'unf',0,1)
z=grand(1,n,'unf',0,1)
u= ------ ; disp (u, 'u=')
v= ------ ; disp (v, 'v=')
t= ------ ; disp (t, 't=')
```

 (b) Que représente l'événement $(T\geqslant V)$?
 (c) On souhaite déterminer une valeur approchée de la probabilité $P(T\geqslant V)$, notée p, en simulant un grand nombre de fois le passage des clients A, B et C aux guichets.
 Compléter les commandes `p= ----- ; disp(p, 'p=')` pour que, placées sous les commandes écrite à la question 3.(a), elles permettent d'obtenir une valeur approchée de p.
 (d) Lors de plusieurs essais des commandes ci-dessus, avec $n=10000$, la réponse donnée par `Scilab` est comprise entre 0.66 et 0.67.
 Que peut-on conjecturer quant à la valeur exacte de p ?

Exercice 5. (Retour sur **EDHEC 2013**) On a pu voir en classe que, si (U_n) est une suite de v.a. indépendantes de même loi uniforme $\mathcal{U}([0;1])$, alors,

$$I_n = \min(U_1, ..., U_n) \xrightarrow{\mathcal{L}} 0.$$

Le programme suivant vise à illustrer cette observation. Recopier et compléter les instructions.

```
//simulation de la variable I_n
function y=edhec2013(n)
    U=zeros(1,n);
    for k=1:n
        U(k)=..................
    end
    y=............................
endfunction

//fréquence de la valeur obtenue par I_n sur un échantillon de taille 100
function y=freq_I(n)
    U=zeros(1,100)
    for k=1:100
        U(k)=..................
    end
    y=.............................
endfunction

I=feval([1:10:1000], freq_I)
plot2d([1:10:1000], I, style=5)
```

Exercice 6. (D'après **ECRICOME 2017**) Soit n un entier naturel non nul.
On effectue une série illimité de tirages d'une boule avec remise dans une urne contenant n boules numérotées de 1 à n. Pour tout entier naturel k non nul, on note X_k la variable aléatoire égale au numéro de la boule obtenue au k-ième tirage.
Pour tout entier naturel k non nul, on note S_k la somme des numéros des boules obtenues lors des k premiers tirages :

$$S_k = \sum_{i=1}^{k} X_i.$$

On considère enfin la variable aléatoire T_n égale au nombre de tirages nécessaires pour que, pour la première fois, la somme des numéros des boules obtenues soit supérieure ou égale à n.

On fait varier l'entier n et on étudie la convergence en loi de la suite de variable $(T_n)_{n\geq 1}$ obtenue.

(1) Compléter la fonction ci-dessous, qui prend en argument le nombre n de boules contenues dans l'urne, afin qu'elle simule la variable aléatoire T_n:

```
function y=T(n)
    S = ..........
    y = ..........
    while ..........
        tirage = grand(1,1,'uin',1,n)
        S = S + tirage
        y = ..........
    end
endfunction
```

```
function y=freqT(n)
    y = zeros(1,n)
    for i=1:100000
        k = T(n)
        y(k) = y(k)+1
    end
    y = y/100000
endfunction

function y=loitheoY(n)
    y = zeros(1,n)
    for k=1:n
        y(k) = (k-1)/prod(1:k)
    end
endfunction

clf
n = input('n=?')
plot2d(loitheoY(6),style=-2)
x = freqT(n)
bar(x(1:5))
```

(2) Comment interpréter la figure obtenue?

Exercice 7. (Loi du khi-deux) Soient $X_1, X_2, ..., X_n$ des v.a. indépendantes suivant une loi normale $\mathcal{N}(0,1)$. On dit que la variable T_{n-1} définie par

$$T_{n-1} = \sum_{k=1}^{n} X_k^2$$

suit la loi du khi-deux à n degrés de liberté, ce qu'on note $T_{n-1} \hookrightarrow \chi^2_{n-1}$.

(1) Écrire une fonction `y=khi_deux(n)` simulant la loi du khi-deux à n degrés de liberté. Stocker 10000 réalisations pour $n = 6$ dans un vecteur `echantillon`.

(2) Donner une estimation du réel t tel que $P(T_5 < t) = 0,95$.

☞ On commencera par taper l'instruction `echantillon=gsort(echantillon, 'g', 'i')`.

(3) On considère une urne contenant n jetons numérotés 1 à n. On tire k jetons un par un avec remise. On note, pour tout i entre 1 et n, N_i le nombre de jetons piochés portant le numéro i. On pose ensuite

$$X_k = \frac{n}{k} \sum_{i=1}^{n} \left(N_i - \frac{k}{n} \right)^2 .$$

☞ On peut montrer que $X_k \xrightarrow{\mathcal{L}} \chi^2_{n-1}$.

(a) Écrire une fonction `N=jeton(n,k)` renvoyant un vecteur N dont les composantes sont $N_1, N_2, ..., N_n$.

(b) On suppose $k = 6$. Utiliser la fonction précédente pour calculer X_{10000}.