ECE 2 - Année 2017-2018
Lycée français de Vienne
Mathématiques - F. Gaunard
http://frederic.gaunard.com

## Séances de Révisions

*(*) Facile (**) Classique/intermédiaire (***) Difficile*

# Séance 5. Intervalles de confiance & Estimation

Dans toute cette partie $\Phi$ désigne la fonction de répartition de la loi normale $\mathcal{N}(0,1)$.

**Exercice 1.** (**) Soit $X$ une variable aléatoire d'espérance $m$ et de variance $\sigma^2$. On considère un $n-$échantillon $(X_1, X_2, ..., X_n)$ de $X$ et on note

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

(1) Expliquer pourquoi l'intervalle

$$\left[\overline{X}_n - \frac{\sigma}{\sqrt{n\alpha}};\ \overline{X}_n + \frac{\sigma}{\sqrt{n\alpha}}\right]$$

fournit un intervalle de confiance (exact) pour $m$ au risque $\alpha$.

(2) Expliquer pourquoi, en notant $u_\alpha = \Phi^{-1}(1-\alpha/2)$, l'intervalle

$$\left[\overline{X}_n - \frac{\sigma}{\sqrt{n}}u_\alpha;\ \overline{X}_n + \frac{\sigma}{\sqrt{n}}u_\alpha\right]$$

fournit un intervalle de confiance asymptotique pour $m$ au risque $\alpha$.

(3) **Application.** On suppose que $X \hookrightarrow \mathcal{B}(p)$ et on cherche à estimer $p$ (qu'on ne connait pas). On ne connait donc *a fortiori* pas non plus la variance. Expliquer comment obtenir des intervalles de confiance (exact et asymptotique) pour $m$ (au risque $\alpha$) qui ne dépendent pas de $\sigma$.

**Exercice 2.** Dans le cadre d'une étude sur la santé au travail, on a interrogé au hasard 500 salariés de différents secteurs et de différentes régions de France. 145 d'entre eux déclarent avoir déjà subi un harcèlement moral au travail.

(1) Donner une estimation ponctuelle de la proportion de salariés ayant déjà subi un harcèlement moral au travail.
(2) Donner une estimation de cette proportion par un intervalle de confiance à 90%.
(3) Si avec les mêmes données on calculait un intervalle de confiance à 95%, serait-il plus grand ou plus petit que celui trouvé à la question précédente ? (justifier sans calcul.)

**Exercice 3.** José et Josette jouent au poker. Chacun tire ainsi deux cartes dans un jeu de 52. Sur 100 parties, Josette a tiré 25 fois une paire.

(1) Calculer la probabilité d'obtenir une paire lors d'un tirage de deux cartes.
(2) En utilisant l'inégalité de Bienaymé-Tchebychev, peut-on dire que Josette triche au risque $\alpha = 0.05$?
(3) En utilisant le théorème central limite, peut-on dire que Josette triche au risque $\alpha = 0.05$?

**Exercice 4.** (*) Soient $\theta$ un paramètre réel strictement positif et $X$ une v.a. de densité $f$ donnée par

$$f(x) = \begin{cases} \frac{2x}{\theta^2}, & \text{si } x \in ]0;\theta] \\ 0, & \text{sinon} \end{cases}$$

Enfin, on considère un $n-$échantillon $(X_1, X_2, .., X_n)$ de $X$.

(1) Calculer l'espérance et la variance de $X$. Déterminer la fonction de répartition $F$ de $X$.
(2) On notant $\overline{X}_n$ la moyenne empirique de l'échantillon, déterminer un estimateur sans biais $T_n$ de $\theta$ de la forme $T_n = c\overline{X}_n$. Préciser son risque quadratique.
(3) On pose $M_n = \max(X_1, ..., X_n)$. Déterminer la fonction de répartition $G_n$ et une densité $g_n$ de $M_n$.
(4) En déduire un estimateur $W_n$, à partir de $M_n$, sans biais de $\theta$. Calculer son risque quadratique.
(5) Entre $W_n$ et $T_n$, quel estimateur choisir?
(6) Soit $\alpha$ un réel compris strictement entre 0 et 1.
   (a) Montrer qu'il existe deux réels $a, b$ strictement compris entre 0 et 1 tels que
   $$P(M_n \leq a\theta) \leq \frac{\alpha}{2}, \quad \text{et} \quad P(b\theta \leq M_n \leq \theta) = \frac{\alpha}{2}.$$
   (b) En déduire un intervalle de confiance pour $\theta$ au risque $1 - \alpha$.

**Exercice 5.** (**D'après **ESSEC II 2016**) Soit $X$ d'espérance $m$ et de variance $\sigma^2$. On considère alors un $n-$échantillon $(X_1, ..., X_n)$ de $X$ et on pose $T_n = X_1 + X_2 + ... + X_n$. On suppose que $\sigma^2$ est connue, mais pas $m$.

(1) Déterminer $E(T_n)$ et $V(T_n)$.
(2) Montrer que, pour tout $\varepsilon > 0$
$$P\left(|T_n - nm| > n\varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon}.$$
(3) Quel est le risque de l'intervalle de confiance
$$\left[\frac{T_n}{n} - \varepsilon; \ \frac{T_n}{n} + \varepsilon\right]$$
pour $m$?

**Exercice 6.** (**) Soit $(X_n)$ une suite de v.a.i.i.d de loi $\mathcal{B}(p^2)$. On note $\overline{X}_n$ la moyenne empirique des $X_i$ et on cherche à estimer $p$.

(1) Montrer que $E(\overline{X}_n^2) - E\left(\sqrt{\overline{X}_n}\right)^2 > 0$.
(2) En déduire que l'estimateur $T_n = \sqrt{\overline{X}_n}$ n'est pas sans biais.
(3) Montrer que, pour tous $x, y \in [0;1]$, $\left|\sqrt{y} - \sqrt{x}\right| \leq \sqrt{|y - x|}$.
(4) Soit $\varepsilon > 0$. À l'aide de la question précédente, montrer que
$$P\left(|T_n - p| > \varepsilon\right) \leq \frac{p^2(1 - p^2)}{n\varepsilon^4}.$$
(5) En déduire un intervalle de confiance au risque $\alpha$ pour $p$, basé sur $T_n$.

# Séance 6. `SciLab` pour les nuls

Attention, la suite d'exercices et les instructions `SciLab` nécessaires à leur résolution ne constitue en aucun cas une liste exhaustive des compétences à maîtriser en informatique.

### 1 - Calcul d'un entier $n$ qui vérifie une condition

**Exercice 7.** (*D'après **EDHEC 2016**) On considère une suite $(u_n)$ qui vérifie, pour tout $n \in \mathbb{N}$

$$e^{-\sqrt{u_n}} \leq u_n - n \leq e^{-\sqrt{n}}.$$

(1) Compléter les commandes `SciLab` suivantes qui permettent de calculer et d'afficher un entier $n$ pour lequel $u_n - n$ est inférieur ou égal à $10^{-4}$.

```
n=0;
while ...........
    n=...........
end
disp(n)
```

(2) Le Script ci-dessus affiche l'une des trois valeurs $n = 55, n = 70$ ou $n = 85$. En prenant $2,3$ comme valeur approchée de $\ln(10)$, déterminer laquelle.

**Exercice 8.** (*D'après **EML 2016**) On considère la fonction $f$ définie sur $\mathbb{R}_+$ par

$$f(x) = \begin{cases} x^2 - x\ln(x), & \text{si } x \neq 0 \\ 0, & \text{si } x = 0 \end{cases}$$

et la suite $(u_n)$ définie par la relation de récurrence

$$u_0 = \frac{1}{2}, \qquad u_{n+1} = f(u_n).$$

(1) *Démontrer que $(u_n)$ est croissante et converge vers* $1$.
(2) Écrire un programme sous `SciLab` qui calculer et affiche un entier $N$ tel que

$$1 - u_N < 10^{-4}.$$

**Exercice 9.** (*D'après **EML 2015**) On considère la fonction $f$ définie sur $\mathbb{R}$ par $f(x) = x^3 e^x$.

(1) Montrer la convergence de la série $\sum \frac{1}{f(n)}$. On note $S$ sa somme.
(2) Montrer que, pour tout $n \in \mathbb{N}^*$,

$$\left| S - \sum_{k=1}^{n} \frac{1}{f(k)} \right| \leq \frac{1}{(e-1)e^n}.$$

(3) En déduire l'écriture d'un programme sous `SciLab` qui calcule une valeur approchée de $S$ à $10^{-5}$ près.

**Exercice 10.** (*D'après **EML 2017**) On considère la fonction $f$ définie sur $\mathbb{R}_+^*$ par $f(x) = e^x - e\ln(x)$ et la suite $(u_n)$ définie par la relation de récurrence

$$u_0 = 2, \qquad u_{n+1} = f(u_n).$$

(1) *Montrer que $(u_n)$ est croissante et diverge vers $+\infty$.*
(2) Écrire un programme sous `SciLab` qui, étant donné un réel $A$ rentré par l'utilisateur, calcule et affiche un entier naturel $N$ tel que $u_n \geq A$.

## 2 - Simulation de v.a.r

**Exercice 11.** (*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 $n$.

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
```

**Exercice 12.** (*D'après **EML 2017**) On considère une urne contenant initialement une boule bleue et deux boules rouges. On effectue, dans cette urne, des tirages successifs de la façon suivante: on pioche une boule au hasard, on note sa couleur, puis on la replace dans l'urne en ajoutant une boule de la couleur de celle qui vient d'être obtenue.

Pour tout $k \in \mathbb{N}^*$, on note $B_k$ : "la boule tirée au $k-$ième tirage est bleue" et $R_k$ l'évènement: "la boule tirée au $k-$ième tirage est rouge".

(1) Recopier et compléter la fonction suivante afin qu'elle simule l'expérience étudiée et renvoie le nombre de boules rouges obtenues lors des $n$ premiers tirages, $n$ étant l'entier entré en argument.

```
function s=EML_17(n)
    b=1; //b représente le nombre de boules bleues dans l'urne
    r=2; //r représente le nombre de boules rouges dans l'urne
    s=0; //s représente le nombre de boules rouges obtenues en n tirages
    for k=1:n
        x=rand();
        if ............................. then
            ..................................
        else
            ..................................
        end
    end
endfunction
```

(2) L'exécution du programme ci-dessous renvoie `6.657`. Comment interpréter ce résultat?

```
n=0;
m=0;
for k=1:1000
    m=m+EML_17(n);
end
disp(m/1000)
```

**Exercice 13.** (*D'après **EML 2015**) On considère une v.a. $U$ telle que $U \hookrightarrow \mathcal{U}([0;1])$ et on introduit la v.a. $V$ définie par

$$V = -\frac{1}{\lambda}\ln(1-U),$$

où $\lambda$ est un réel strictement positif.

(1) Montrer que $V \hookrightarrow \mathcal{E}(\lambda)$.

(2) En déduire l'écriture d'une fonction `SciLab`, notée `y=loi_expo(lambda)` qui, prenant en argument un paramètre `lambda` renvoie une simulation de la loi exponentielle $\mathcal{E}($`lambda`$)$.

### 3 - Valeur approchée d'une espérance

**Exercice 14.** (Extrait de **HEC 2015**) On considère une v.a. $X \hookrightarrow \mathcal{E}(\lambda)$ (où $\lambda > 0$) et $F$ la fonction définie sur $\mathbb{R}$ par

$$F(x) = \exp\left(e^{-\lambda x}\right).$$

On *peut montrer* que $F$ est la fonction de répartition d'une v.a. $T$ (loi de Gumbel de paramètre $\lambda$). On peut aussi montrer que

$$Z = -\frac{1}{\lambda}\ln(\lambda X)$$

suit la même loi que $T$.

(1) On considère $\lambda = 1$. Écrire en `Scilab` les commandes qui permettent de simuler la loi de $T$.

(2) Écrire en `Scilab`, par la méthode de Monte-Carlo, les commandes qui permettent de renvoyer une valeur numérique approchée de l'intégrale

$$\int_0^{+\infty} e^{-t}\,\ln t\,dt.$$

### 4 - Représentations graphiques

**Exercice 15.** (*D'après **ECRICOME 2015**) On s'intéresse à la suite récurrente $(u_n)$ définie par

$$u_0 = 1, \qquad u_{n+1} = 1-\exp(-u_n).$$

(1) Compléter le programme ci-dessous permettant le calcul des 100 premiers éléments de $(u_n)$

```
U=zeros(1, 100)
U(1)=...............
for n=1: ...............
    U(n+1)=..........
end
plot2d(.........., U, -1)
```

(2) On modifie le programme précédent en remplaçant la dernière ligne par les instructions suivantes

```
X=1:100
S=cumsum(U)
Y=log(X)
plot2d(X, S, -1)
plot2d(X, Y)
```

(3) Que représente le vecteur `S`? Quelle conjecture pouvez-vous émettre sur la nature de la série terme général $u_n$?