Université d'Orléans - Licence Economie et Gestion
Statistique Mathématique

C. Hurlin. Correction Examen Novembre 2011

Exercice 1 *Maximum de Vraisemblance. Barême 12 points*

Question 1 (2 points) On sait que les variables $\{X_1,..,X_N\}$ sont *i.i.d.* de même loi que X. Dès lors, on a :

$$\begin{aligned}\ln L\left(X_1,..,X_N;\sigma^2\right) &= \sum_{i=1}^{N}\ln f_X\left(X_i;\sigma^2\right) \\ &= \prod_{i=1}^{N}\left(-\frac{X_i^2}{2\sigma^2}\right)+\ln\left(X_i\right)-\ln\left(\sigma^2\right) \quad 0.5 \text{ point} \end{aligned} \tag{1}$$

Dès lors, la log-vraisemblance associée au N-échantillon $\{X_1,..,X_N\}$ s'écrit :

$$\ln L\left(X_1,..,X_N;\sigma^2\right)=-\frac{1}{2\sigma^2}\sum_{i=1}^{N}X_i^2+\sum_{i=1}^{N}\ln\left(X_i\right)-N\ln\left(\sigma^2\right) \quad (1.5 \text{ point}) \tag{2}$$

Question 2 (2 points) Soit $\widehat{\sigma}^2_{MV}$ l'estimateur du maximum de vraisemblance du paramètre σ^2. Ce dernier vérifie :

$$\widehat{\sigma}_{MV}=Arg\max L\left(X_1,..,X_N;\sigma\right)$$

On en déduit la CN :

$$\left.\frac{\partial\ln L\left(X_1,..,X_N;\sigma^2\right)}{\partial\sigma^2}\right|_{\widehat{\sigma}^2_{MV}}=\frac{1}{2\widehat{\sigma}^4_{MV}}\sum_{i=1}^{N}X_i^2-\frac{N}{\widehat{\sigma}^2_{MV}}=0 \quad (0.5 \text{ point}) \tag{3}$$

D'où l'on tire que :

$$\frac{1}{2\widehat{\sigma}^4_{MV}}\sum_{i=1}^{N}X_i^2=\frac{N}{\widehat{\sigma}^2_{MV}}$$

et finallement :

$$\widehat{\sigma}^2_{MV}=\frac{1}{2N}\sum_{i=1}^{N}X_i^2 \quad (0.5 \text{ point}) \tag{4}$$

Ohn vérifie la CS :

$$\begin{aligned}\left.\frac{\partial^2\ln L\left(X_1,..,X_N;\sigma^2\right)}{\partial\sigma^4}\right|_{\widehat{\sigma}^2_{MV}} &= -\frac{1}{\widehat{\sigma}^6_{MV}}\sum_{i=1}^{N}X_i^2+\frac{N}{\widehat{\sigma}^4_{MV}} \\ &= \frac{1}{\widehat{\sigma}^6_{MV}}\left(-\sum_{i=1}^{N}X_i^2+\widehat{\sigma}^2_{MV}N\right)\end{aligned} \tag{5}$$

Or, nous savons que $\widehat{\sigma}^2_{MV} = \frac{1}{2N}\sum_{i=1}^N X_i^2$. Dès lors, il vient :

$$\begin{aligned}\left.\frac{\partial^2 \ln L\left(X_1,..,X_N;\sigma^2\right)}{\partial\sigma^4}\right|_{\widehat{\sigma}^2_{MV}} &= \frac{1}{\widehat{\sigma}^6_{MV}}\left(-\sum_{i=1}^N X_i^2+\frac{1}{2}\sum_{i=1}^N X_i^2\right)\\ &= -\frac{1}{2\widehat{\sigma}^6_{MV}}\sum_{i=1}^N X_i^2<0 \quad \text{(1 point)} \qquad (6)\end{aligned}$$

Question 3 (1 point) On sait que

$$E\left(X\right)=\sigma\sqrt{\frac{\pi}{2}} \qquad V\left(X\right)=\left(\frac{4-\pi}{2}\right)\sigma^2 \qquad (7)$$

Or puisque $V\left(X\right)=E\left(X^2\right)-E\left(X\right)^2$, on a :

$$\begin{aligned}E\left(X^2\right) &= V\left(X\right)+E\left(X\right)^2\\ &= \left(\frac{4-\pi}{2}\right)\sigma^2+\frac{\pi}{2}\sigma^2\\ &= 2\sigma^2 \quad \text{(1 point)} \qquad (8)\end{aligned}$$

Question 4 (1 point) Calculons $E\left(\widehat{\sigma}^2_{MV}\right)$:

$$E\left(\widehat{\sigma}^2_{MV}\right)=\frac{1}{2N}\sum_{i=1}^N E\left(X_i^2\right)=\frac{1}{2N}\sum_{i=1}^N E\left(X^2\right)=\frac{E\left(X^2\right)}{2} \quad \text{(0.5 point)} \qquad (9)$$

Sachant que $E\left(X_i^2\right)=2\sigma^2$, on a :

$$E\left(\widehat{\sigma}^2_{MV}\right)=\sigma^2 \quad \text{(0.5 point)} \qquad (10)$$

L'estimateur $\widehat{\sigma}^2_{MV}$ est sans biais.

Question 5 (1 point) On admet que $V\left(X^2\right)=4\sigma^4$. Dès lors, puisque les X_i sont *i.i.d.* de même loi que X, il vient :

$$V\left(\widehat{\sigma}^2_{MV}\right)=\frac{1}{4N^2}V\left(\sum_{i=1}^N\left(X_i^2\right)\right)=\frac{1}{4N^2}\sum_{i=1}^N V\left(X^2\right)=\frac{V\left(X^2\right)}{4N} \quad \text{0.5 point} \qquad (11)$$

D'où l'on tire que :

$$V\left(\widehat{\sigma}^2_{MV}\right)=\frac{\sigma^4}{N} \quad \text{0.5 point} \qquad (12)$$

Question 6 (1 point) Par conséquent :

$$\lim V\left(\widehat{\sigma}^2_{MV}\right)=\lim\frac{\sigma^4}{N}=0 \quad \text{lorsque } N\to\infty \quad \text{0.5 point} \qquad (13)$$

Sachant que par ailleurs $E\left(\widehat{\sigma}^2_{MV}\right)=\sigma^2$, on en déduit que $\widehat{\sigma}^2_{MV}$ est un estimateur convergent de σ^2 :

$$\widehat{\sigma}^2_{MV} \underset{N\to\infty}{\xrightarrow{p}} \sigma^2 \quad \text{0.5 point} \qquad (14)$$

L'estimateur $\widehat{\sigma}^2_{MV}$ est convergent en probabilité.

Question 7 (2 points) On admet que :

$$\frac{\partial^2 \ln L\left(X_1,..,X_N;\sigma^2\right)}{\partial\sigma^4}=-\frac{1}{2\sigma^6}\sum_{i=1}^{N}X_i^2 \tag{15}$$

Dès la borne la quantité d'information Fisher vaut :

$$\begin{aligned}
I_N\left(\sigma^2\right) &= E\left[-\frac{\partial^2 \ln L\left(X_1,..,X_N;\sigma^2\right)}{\partial\sigma^4}\right] \\
&= \frac{1}{2\sigma^6}\sum_{i=1}^{N}E\left(X_i^2\right) \\
&= \frac{N2\sigma^2}{2\sigma^6} \\
&= \frac{N}{\sigma^4}\quad \text{1 point}
\end{aligned} \tag{16}$$

Dès lors,

$$V\left(\widehat{\sigma}_{MV}^2\right)=\frac{1}{I_N\left(\sigma^2\right)}=\frac{\sigma^4}{N}\quad \text{0.5 point}$$

L'estimateur du maximum de vraisemblance $\widehat{\sigma}_{MV}^2$ est efficace au sens de la borne FDCR

Question 8 (2 points) D'après le TCL, la loi asymptotique de de la quantité $\frac{1}{N}\sum_{i=1}^{N}X_i^2$

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}X_i^2-E\left(X^2\right)\right)\underset{N\rightarrow\infty}{\overset{d}{\longrightarrow}}N\left(0,V\left(X^2\right)\right) \tag{17}$$

$$\Longleftrightarrow \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}X_i^2-2\sigma^2\right)\underset{N\rightarrow\infty}{\overset{d}{\longrightarrow}}N\left(0,4\sigma^4\right)\quad \text{(1 point)} \tag{18}$$

On en déduit la loi asymptotique de l'estimateur $\widehat{\sigma}_{MV}^2=\frac{1}{2N}\sum_{i=1}^{N}X_i^2$:

$$\sqrt{N}\left(\widehat{\sigma}_{MV}^2-\sigma^2\right)\underset{N\rightarrow\infty}{\overset{d}{\longrightarrow}}N\left(0,\sigma^4\right)\quad \text{(1 point)} \tag{19}$$

================

Exercice 2 *Intervalle de confiance. Barème : 4 points*

Question 1 (2 points) On sait que la moyenne empirique $\overline{X}_N=\sum_{i=1}^{N}X_i/N$ est un estimateur convergent de l'espérance m. On fait l'hypothèse que le taux de croissance du PIB est distribué selon une normale $N\left(m,\sigma^2\right)$. Dès lors, on sait que :

$$\frac{\overline{X}_N-m}{\sigma/\sqrt{N}}\sim N\left(0,1\right)\quad \text{(0.5 point)} \tag{20}$$

Par définition (0.5 point) :

$$\Pr\left[t_{\alpha/2} < \frac{\left(\overline{X}_N - m\right)}{\sigma/\sqrt{N}} < t_{1-\alpha/2}\right] = 1-\alpha \tag{21}$$

où $t_{\alpha/2}$ et $t_{1-\alpha/2}$ désignent respectivement les fractiles d'ordre $\alpha/2$ et $1-\alpha/2$ de la loi $N\left(0,1\right).$ On en déduit finalement (0.5 point):

$$IC_{1-\alpha} = \left[\overline{X}_N - \frac{\sigma}{\sqrt{N}}t_{1-\alpha/2}; \overline{X}_N - \frac{\sigma}{\sqrt{N}}t_{\alpha/2}\right] \tag{22}$$

Application Numérique : $\sigma = 1.580, \overline{X}_N = 1.20,\ N = 15,\ \alpha = 10\%$

$$t_{1-\alpha/2} = 1.6449 \quad t_{\alpha/2} = -1.6449 \tag{23}$$

D'om l'on tire que :

$$IC_{0.95} = \left[1.20 - \frac{1.580}{\sqrt{15}}1.96\ ;\ 1.20 + \frac{1.580}{\sqrt{15}}1.96\right] = [0.52; 1.87] \quad \text{(0.5 point)} \tag{24}$$

Question 2 (2 points) Soit S_N^2, la variance empirique corrigée telle que :

$$S_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}_N\right)^2 \tag{25}$$

On sait que

$$\frac{\overline{X}_N - m}{S_N/\sqrt{N}} \sim Student\left(N-1\right) \qquad \text{(0.5 point)} \tag{26}$$

Dès lors on a par définition :

$$\Pr\left[t_{\alpha/2} < \frac{\left(\overline{X}_N - m\right)}{S_N/\sqrt{N}} < t_{1-\alpha/2}\right] = 1-\alpha \tag{27}$$

où $t_{\alpha/2}$ et $t_{1-\alpha/2}$ désignent respectivement les fractiles d'ordre $\alpha/2$ et 1-$\alpha/2$ de la loi $Student\left(N-1\right).$ On en déduit finalement (0.5 point):

$$IC_{1-\alpha} = \left[\overline{X}_N - \frac{S_N}{\sqrt{N}}t_{1-\alpha/2}; \overline{X}_N - \frac{S_N}{\sqrt{N}}t_{\alpha/2}\right] \tag{28}$$

Application Numérique : $S_N == 1.580, \overline{X}_N = 1.20,\ N = 15,\ \alpha = 10\%$

$$t_{\alpha/2} = -1.7613 \qquad t_{1-\alpha/2} = 1.7613 \ \ \text{(0.5 point)} \tag{29}$$

D'où l'on tire que :

$$IC_{0.95} = \left[1.20 - \frac{1.580}{\sqrt{15}}1.7613\ ;\ 1.20 + \frac{1.580}{\sqrt{15}}1.7613\right] = [0.4815; 1.9185] \quad \text{(0.5 point)} \tag{30}$$

On vérifie que l'intervalle de confiance est alors beaucoup plus large que lorsque l'on suppose σ^2 connue puisque sur la base de l'espérance non conditionnelle, il y a 90% de chance que le taux de croissance soit compris entre 0.48% et 1.91%

================

Exercice 3 *Estimation. Barème : 6 points*

Question 1 (2 points) Calculons $E\left(\widehat{\theta}_1\right)$ et $E\left(\widehat{\theta}_2\right)$:

$$E\left(\widehat{\theta}_1\right)=\frac{1}{N}\sum_{i=1}^{N}E\left(X_i\right)=\frac{1}{N}\sum_{i=1}^{N}E\left(X\right)=\theta \qquad \text{(0.5 point)} \tag{31}$$

$$E\left(\widehat{\theta}_2\right)=\frac{1}{N}\sum_{i=1}^{N}E\left(X_i^2\right)=\frac{1}{N}\sum_{i=1}^{N}E\left(X^2\right) \tag{32}$$

Or on sait que

$$\begin{aligned} E\left(X^2\right) &= V\left(X\right)+E\left(X\right)^2 && (33)\\ &= \theta-\theta^2+\theta^2 && (34)\\ &= \theta \quad \text{(1 point)} && (35) \end{aligned}$$

Il vient donc :

$$E\left(\widehat{\theta}_2\right)=\theta \qquad \text{(0.5 point)} \tag{36}$$

Les deux estimateurs sont sans biais.

Question 2 (1 point) Calculons $V\left(\widehat{\theta}_1\right)$ et $V\left(\widehat{\theta}_2\right)$:

$$V\left(\widehat{\theta}_1\right)=\frac{1}{N^2}\sum_{i=1}^{N}V\left(X_i\right)=\frac{1}{N^2}\sum_{i=1}^{N}V\left(X\right)=\frac{\theta-\theta^2}{N} \qquad \text{(0.5 point)} \tag{37}$$

$$V\left(\widehat{\theta}_2\right)=\frac{1}{N^2}\sum_{i=1}^{N}V\left(X_i^2\right)=\frac{1}{N^2}\sum_{i=1}^{N}V\left(X^2\right)=\frac{2\theta^2-2\theta^4}{N} \qquad \text{(0.5 point)} \tag{38}$$

Dans les deux cas, on a :

$$\lim_{N\rightarrow\infty}V\left(\widehat{\theta}_1\right)=\lim_{N\rightarrow\infty}V\left(\widehat{\theta}_2\right)=0 \tag{39}$$

Les deux estimateurs (sans biais) sont donc convergents au sens de la convergence en probabilité.

Question 3 (1 point) On sait que

$$V\left(\widehat{\theta}_1\right)=\frac{\theta-\theta^2}{N} \tag{40}$$

$$V\left(\widehat{\theta}_2\right)=\frac{2\theta^2-2\theta^4}{N} \tag{41}$$

Dès lors :

$$\frac{V\left(\widehat{\theta}_1\right)}{V\left(\widehat{\theta}_2\right)}=\frac{\theta\left(1-\theta\right)}{2\theta^2\left(1-\theta^2\right)}=\frac{1}{2\theta\left(1+\theta\right)} \tag{42}$$

On ne peut savoir si $V\left(\widehat{\theta}_1\right)>V\left(\widehat{\theta}_2\right)$ ou si $V\left(\widehat{\theta}_1\right)<V\left(\widehat{\theta}_2\right)$ puisque cette quantité dépend de la vraie valeur de θ qui est inconnue.

Question 4 (2 points) D'après le TCL, on a :

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}X_i - E\left(X\right)\right) \underset{N\to\infty}{\overset{d}{\longrightarrow}} N\left(0, V\left(X\right)\right) \tag{43}$$

Donc

$$\sqrt{N}\left(\widehat{\theta}_1 - \alpha\right) \underset{N\to\infty}{\overset{d}{\longrightarrow}} N\left(0, \theta - \theta^2\right) \quad \text{1 point} \tag{44}$$

De la même façon :

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}X_i^2 - E\left(X^2\right)\right) \underset{N\to\infty}{\overset{d}{\longrightarrow}} N\left(0, V\left(X^2\right)\right) \tag{45}$$

Donc

$$\sqrt{N}\left(\widehat{\theta}_2 - \alpha\right) \underset{N\to\infty}{\overset{d}{\longrightarrow}} N\left(0, 2\theta\left(1+\theta\right)\right) \quad \text{1 point} \tag{46}$$