Université d'Orléans - Licence Economie et Gestion Statistique Mathématique

C. Hurlin. Correction de l'Examen Terminal Janvier 2008

Exercice 1 *Value-at-Risk et Tests Paramétriques. Barème : 12 points.*

Question 1 (1 point) Sachant que :

$$VaR_{(0.05)} = m + \sigma \ \Phi^{-1}(0.05) \tag{1}$$

On montre immédiatement que :

$$H_0 : \sigma^2 = \left[\frac{V_0 - m}{\Phi^{-1}(0.05)}\right] \tag{2}$$

$$H_1 : \sigma^2 \neq \left[\frac{V_0 - m}{\Phi^{-1}(0.05)}\right] \tag{3}$$

Remarque : -0.5 point, si le test porte sur σ et non σ^2.

Question 2 (1 point) On sait que sous l'hypothèse de normalité :

$$L(r_1, .., r_T, \sigma_0) = \left(\frac{1}{\sigma_0\sqrt{2\pi}}\right)^T \exp\left[-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{r_t - m}{\sigma_0}\right)^2\right] \quad \text{0.5 point} \tag{4}$$

$$L(r_1, .., r_T, \sigma_1) = \left(\frac{1}{\sigma_1\sqrt{2\pi}}\right)^T \exp\left[-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{r_t - m}{\sigma_1}\right)^2\right] \tag{5}$$

Il s'ensuit immédiatement que :

$$\frac{L(r_1, .., r_T, \sigma_0)}{L(r_1, .., r_T, \sigma_1)} = \left(\frac{\sigma_1}{\sigma_0}\right)^T \exp\left[\frac{1}{2\sigma_1^2}\sum_{t=1}^{T}(r_t - m)^2 - \frac{1}{2\sigma_0^2}\sum_{t=1}^{T}(r_t - m)^2\right] \tag{6}$$

Après simplification, il vient :

$$\frac{L(r_1, .., r_T, \sigma_0)}{L(r_1, .., r_T, \sigma_1)} = \left(\frac{\sigma_1}{\sigma_0}\right)^T \exp\left[\frac{1}{2}\left(\frac{\sigma_0^2 - \sigma_1^2}{\sigma_0^2\sigma_1^2}\right)\sum_{t=1}^{T}(r_t - m)^2\right] \quad \text{0.5 point} \tag{7}$$

Question 3 (1 point) D'après le lemme de Neyman Pearson, la région critique du test UPP de niveau α de l'hypothèse nulle $H_0 : \sigma^2 = \sigma_0^2$ contre l'hypothèse alternative $H_1 : \sigma^2 = \sigma_1^2$ est de la forme :

$$W = \left\{ r_1, .., r_T \mid \frac{L(r_1, .., r_T, \sigma_0)}{L(r_1, .., r_T, \sigma_1)} \leq K \right\} \quad \text{0.5 point} \tag{8}$$

où K est une constante déterminée par le risque de première espèce. Il s'ensuit que :

$$W = \left\{ r_1, .., r_T \mid \left(\frac{\sigma_1}{\sigma_0}\right)^T \exp\left[\frac{1}{2}\left(\frac{\sigma_0^2 - \sigma_1^2}{\sigma_0^2\sigma_1^2}\right)\sum_{t=1}^{T}(r_t - m)^2\right] \leq K \right\} \quad \text{0.5 point} \tag{9}$$

Il vient :

$$\left(\frac{\sigma_1}{\sigma_0}\right)^T \exp\left[\frac{1}{2}\left(\frac{\sigma_0^2-\sigma_1^2}{\sigma_0^2\sigma_1^2}\right)\sum_{t=1}^{T}(r_t-m)^2\right] \leq K \tag{10}$$

$$\Longleftrightarrow \frac{1}{2}\left(\frac{\sigma_0^2-\sigma_1^2}{\sigma_0^2\sigma_1^2}\right)\sum_{t=1}^{T}(r_t-m)^2 \leq K_1 \text{ avec } K_1=\ln\left[K.\left(\frac{\sigma_0}{\sigma 1}\right)^T\right] \tag{11}$$

$$\Longleftrightarrow \left(\sigma_0^2-\sigma_1^2\right)\sum_{t=1}^{T}(r_t-m)^2 \leq K_2 \text{ avec } K_2=2\sigma_0^2\sigma_1^2\ln\left[K.\left(\frac{\sigma_0}{\sigma 1}\right)^T\right] \tag{12}$$

Sachant que $\sigma_1 < \sigma_0$, il vient :

$$\frac{1}{T-1}\sum_{t=1}^{T}(r_t-m)^2 \leq A \quad \text{1 point} \tag{13}$$

avec $A = K_3/\left[\left(\sigma_0^2-\sigma_1^2\right)(T-1)\right]$. On a donc :

$$W=\left\{r_1,..,r_T \mid \frac{1}{(T-1)}\sum\nolimits_{t=1}^{T}(r_t-m)^2 \leq A\right\} \quad \text{0.5 point} \tag{14}$$

Remarque : compter -0.5 point si les valeurs K_j ne sont pas explicitées. Cette région critique est indépendante de la valeur de σ sous H_1. Par conséquent la région critique du test UPP de l'hypothèse $H_0 : \sigma^2 = \sigma_0^2$ contre l'hypothèse alternative $H_1 : \sigma^2 < \sigma_0^2$, est identique à celle du test de l'hypothèse $H_0 : \sigma^2 = \sigma_0^2$ contre l'hypothèse alternative $H_1 : \sigma^2 = \sigma_1^2$ avec $\sigma_0^2 > \sigma_1^2$ (1 point). Remarque II : sachant que $\overline{X}_N$ est un estimateur convergent de m, on admet que pour N grand, la statistique $\frac{1}{(T-1)}\sum_{t=1}^{T}(r_t-m)^2$ converge (en loi) vers S_N^2.

Question 4 (1 point) On sait que par définition du risque I :

$$\alpha = \Pr\left(S_N^2 \leq A \mid H_0 \text{ vraie}\right) \quad \text{0.5 point} \tag{15}$$

Or, on sait que sous l'hypothèse nulle, la variable $S_N^2/\sigma_0^2 \times (T-1)$ suit un chi-deux à $T-1$ degrés de liberté. Soit $F(.)$ la fonction de répartition de la loi du chi-deux à $T-1$ degrés de liberté. On a donc :

$$\alpha = \Pr\left[\frac{S_N^2}{\sigma_0^2}(T-1) \leq \frac{A}{\sigma_0^2}(T-1)\right] \tag{16}$$

D'où :

$$\alpha = F\left[\frac{A}{\sigma_0^2}(T-1)\right] \tag{17}$$

ou encore :

$$A = F^{-1}(\alpha)\left(\frac{\sigma_0^2}{T-1}\right) \quad \text{0.5 point} \tag{18}$$

Question 5 (1 point) On sait que si $\alpha = 5\%$ et si $T = 250$, alors on a :

$$\begin{aligned} A &= F^{-1}(0.05)\times\left(\frac{4}{101-1}\right) && (19)\\ &= 77.93\times\left(\frac{4}{101-1}\right) && (20)\\ &= 3.11 \quad \text{(0.5 point)} && (21) \end{aligned}$$

Donc si $S_N^2 = 4.2$, pour un risque de première espèce de 5%, on ne peut pas rejetter l'hypothèse nulle selon laquelle $\sigma^2 = 4$ (0.5 point).

Question 6 (1 point) On considère à présent le test sur la VaR à 5% :

$$H_0: \ VaR_{(0.05)} = -15 \tag{22}$$

$$H_1: \ VaR_{(0.05)} < -15 \tag{23}$$

On sait que la VaR à 5% sous hypothèse de normalité vérifie la relation :

$$VaR_{(0.05)} = m + \sigma \ \Phi^{-1}(0.05) \tag{24}$$

Dès lors, si $VaR_{(0.05)} = -15$, on a (cf. question 1) :

$$VaR_{(0.05)} = -15 \Longleftrightarrow \sigma^2 = \left[\frac{V_0 - m}{\Phi^{-1}(0.05)}\right]^2 = \left[\frac{-15 + 11.71}{-1.645}\right]^2 = 4 \tag{25}$$

Par conséquent à m connu, tester $H_0: \ VaR_{(0.05)} = -15$ contre $H_1: \ VaR_{(0.05)} < -15$ est équivalent à tester $H_0: \sigma^2 = 4$ contre l'hypothèse alternative $H_1: \sigma^2 < 4$ (0.5 point). Or on sait que si $S_N^2 = 4.2$, pour un risque de première espèce de 5%, on ne peut pas rejetter l'hypothèse nulle selon laquelle $\sigma^2 = 4$. Donc, pour un risque de première espèce de 5%, on ne peut pas rejeter l'hypothèse que la VaR soit égale à -15 (0.5 point).

Question 7 (1 point) On sait que :

$$\begin{aligned} P(\sigma) &= \Pr\left(S_N^2 \leq A \mid H_1 \text{ vraie}\right) \\ &= \Pr\left(S_N^2 \leq A \mid \sigma^2 < \sigma_0^2\right) \quad \text{0.5 point} \qquad (26) \\ &= \Pr\left(\frac{S_N^2}{\sigma^2}(T-1) \leq \frac{A}{\sigma^2}(T-1) \;\middle|\; \sigma^2 < \sigma_0^2\right) \qquad (27) \end{aligned}$$

Pour un niveau de risque I de $\alpha\%$, on sait que :

$$A = F^{-1}(\alpha) \ \left(\frac{\sigma_0^2}{T-1}\right) \tag{28}$$

Dès lors, on a :

$$\begin{aligned} P(\sigma) &= F\left[F^{-1}(\alpha) \ \left(\frac{\sigma_0^2}{T-1}\right)\frac{(T-1)}{\sigma^2}\right] \quad \forall \sigma < \sigma_0^2 \qquad (29) \\ &= F\left[F^{-1}(\alpha) \ \left(\frac{\sigma_0^2}{\sigma^2}\right)\right] \qquad (30) \end{aligned}$$

où $F(.)$ désigne la fonction de répartition de la loi du chi-deux à $T-1$ degrés de liberté.

Question 8 (2 points) La région d'acceptation (RA) de ce test fondée correspond à l'intersection des deux RA des deux tests unilatéraux $H_0: \sigma = \sigma_0^2$ contre $H_1: \sigma < \sigma_0^2$ et $H_0: \sigma = \sigma_0^2$ contre $H_1: \sigma > \sigma_0^2$, c'est à dire (0.5 point) :

$$W = \left\{r_1, .., r_T \mid S_N^2 \leq A\right\} \tag{31}$$

$$H_0: \sigma = \sigma_0^2 \text{ contre } H_1: \sigma < \sigma_0^2: \overline{W} = \left\{r_1, .., r_T \mid S_N^2 \geq A\right\} \tag{32}$$

$$H_0: \sigma = \sigma_0^2 \text{ contre } H_1: \sigma > \sigma_0^2: \overline{W} = \left\{r_1, .., r_T \mid S_N^2 \leq B\right\} \tag{33}$$

Toutefois les seuils A et B doivent être déterminés pour un niveau de risque de $\alpha/2\%$ afin de garantir un risque de première espèce de $\alpha\%$ pour le test bilatéral. Soient (1 point) :

$$A = F^{-1}\left(\frac{\alpha}{2}\right)\left(\frac{\sigma_0^2}{T-1}\right) \tag{34}$$

$$B = F^{-1}\left(1-\frac{\alpha}{2}\right)\left(\frac{\sigma_0^2}{T-1}\right) \tag{35}$$

Dès lors, l'intersection des deux RA nous donne pour le test bilatéral une RA définie par :

$$\overline{W} = \left\{r_1, .., r_T \mid A < S_N^2 < B\right\}$$

ou encore

$$\overline{W} = \left\{r_1, .., r_T \mid F^{-1}\left(\frac{\alpha}{2}\right)\left(\frac{\sigma_0^2}{T-1}\right) < S_N^2 < F^{-1}\left(1-\frac{\alpha}{2}\right)\left(\frac{\sigma_0^2}{T-1}\right)\right\} \text{ 0.5 point} \tag{36}$$

Question 9 (1.5 points) Commençons par définir la RA du test bilatéral sur σ^2 :

$$\overline{W} = \left\{r_1, .., r_T \mid F^{-1}\left(\frac{0.05}{2}\right)\left(\frac{4}{101-1}\right) < S_N^2 < F^{-1}\left(1-\frac{0.05}{2}\right)\left(\frac{4}{101-1}\right)\right\} \tag{37}$$

$$\overline{W} = \left\{r_1, .., r_T \mid 2.96 < S_N^2 < 5.18\right\} \text{ 0.5 point} \tag{38}$$

Dès lors, si $S_N^2 = 4.2$, on est dans la région d'acceptation pour un niveau de risque de première espèce de 5 %. Pour un risque de première espèce de 5%, on ne peut pas rejetter l'hypothèse nulle selon laquelle $\sigma^2 = 4$ (0.5 point). Par ailleurs, à m connu, tester $H_0: \ VaR_{(0.05)} = -15$ contre $H_1: \ VaR_{(0.05)} < -15$ est équivalent à tester $H_0 : \sigma^2 = 4$ contre l'hypothèse alternative $H_1 : \sigma^2 < 4$. Donc, pour un risque de première espèce de 5%, on ne peut pas rejeter l'hypothèse que la VaR soit égale à -15 (1 point).

Exercice 2 *Maximum de Vraisemblance. Barème : 5 points.*

Question 1 (1 point) La vraisemblance de l'échantillon $\{X_1, .., X_N\}$ est définie par :

$$L(X_1, .., X_N; \theta) = \prod_{i=1}^N f_X(X_i) \tag{39}$$

D'où :

$$L(X_1, .., X_N; \theta) = \theta^N \exp\left(-\theta \sum_{i=1}^N X_i\right) \text{ 0.5 point} \tag{40}$$

On en déduit la log-vraisemblance :

$$\ln L(X_1, .., X_N; \theta) = N \ln(\theta) - \theta \sum_{i=1}^N X_i \text{ 0.5 point} \tag{41}$$

Question 2 (1 point) Soit $\widehat{\theta}$ l'estimateur du maximume de vraisemblance du paramètre θ. Cet estimateur vérifie la CN suivante :

$$\left.\frac{\partial \ln L\left(X_1,..,X_N;\theta\right)}{\partial \theta}\right|_{\widehat{\theta}}=0 \Longleftrightarrow N\frac{1}{\widehat{\theta}}-\sum_{i=1}^{N}X_i=0 \tag{42}$$

D'où l'on tire que :

$$\widehat{\theta}=\frac{1}{\overline{X}_N} \quad \text{0.5 point} \tag{43}$$

où $\overline{X}_N=(1/N)\sum_{i=1}^{N}X_i$ désigne la moyenne empirique. On vérifie qu'il s'agit bien d'un maximum puisque :

$$\left.\frac{\partial^2 \ln L\left(X_1,..,X_N;\theta\right)}{\partial \theta^2}\right|_{\widehat{\theta}}=0 \Longleftrightarrow -N\frac{1}{\widehat{\theta}^2}<0 \quad \text{0.5 point} \tag{44}$$

Question 3 (1 point) On montre tout d'abord que :

$$E\left(\overline{X}_N\right)=\frac{1}{N}\sum_{i=1}^{N}E\left(X_i\right)=E\left(X\right)=\frac{1}{\theta} \tag{45}$$

Donc $\overline{X}_N$ est un estimateur sans biais de $1/\theta$ (0.5 point). De plus, on sait que sous l'hypothèse d'indépendance des variables X_i :

$$V\left(\overline{X}_N\right)=\frac{1}{N^2}\sum_{i=1}^{N}V\left(X_i\right)=\frac{V\left(X\right)}{N}=\frac{1}{N\theta^2} \tag{46}$$

Dès lors :

$$\lim_{N\to\infty}V\left(\overline{X}_N\right)=0 \tag{47}$$

Cela implique que $\overline{X}_N$ est un estimateur convergent (en probabilité) de $1/\theta$ (0.5 point) :

$$\overline{X}_N \underset{N\to\infty}{\overset{p}{\longrightarrow}} \frac{1}{\theta} \tag{48}$$

Question 4 (1 point) Soit $g\left(Y\right)=1/Y$, cette fonction est continue en $1/\theta$, donc on obtient immédiatement que :

$$\widehat{\theta}=g\left(\overline{X}_N\right)=\frac{1}{\overline{X}_N} \tag{49}$$

converge vers $g\left(1/\theta\right)=\theta$. Donc :

$$\widehat{\theta} \underset{N\to\infty}{\overset{p}{\longrightarrow}} \theta \tag{50}$$

L'estimateur du maximum de vraisemblance $\widehat{\theta}$ est un estimateur convergent de θ.

Question 5 (1 point) Déterminons la borne FDCR, et pour cela déterminons la quantité d'information de Fisher :

$$I_N\left(\theta\right)=-E\left[\left.\frac{\partial^2 L\left(X_1,..,X_N;\theta\right)}{\partial \theta^2}\right|_{\theta}\right]=E\left(\frac{N}{\theta^2}\right)=\frac{N}{\theta^2} \tag{51}$$

On vérifie bien que :

$$V\left(\widehat{\theta}\right)=\frac{1}{I_N\left(\theta\right)} \tag{52}$$

L'estimateur du maximum de vraisemblance est un estimateur efficace au sens de la borne FDCR

=======

Exercice 3 *Test d'Indépendance : Application à l'Etude de la Relation Salaire / Diplôme. Barème : 4 points.*

Tableau des effectifs empiriques :

X \ Y	Bac	BEP-CAP	Sixième	Total
600-750	100	270	30	400
750-900	60	130	10	200
900-1650	40	50	10	100
Total	200	450	50	700

Tableau des effectifs théoriques (1.5 points) :

X \ Y	Bac	BEP-CAP	Sixième	Total
600-750	114.28	257.14	28.57	400
750-900	57.14	128.57	14.28	200
900-1650	28.57	64.28	7.14	100
Total	200	450	50	700

Distance du chi-deux :

$$D = \sum_i \sum_j \frac{(n_{ij} - N \times p_{ij})^2}{N \times p_{ij}} = 12.83 \quad 0.5 \text{ point} \tag{53}$$

Sous l'hypothèse d'independance, cette statistique suit un chi-deux à r degrés de liberté où

$$r = (3-1)(3-1) = 4 \quad 0.5 \text{ point} \tag{54}$$

Le fractile à 95% d'un chi-deux à 4 dégrés de liberté est égal à 9.49, donc pour un risque de première espèce de 5% on rejette l'hypothèse nulle d'indépendance entre le salaire et le diplôme (0.5 point).