

Solutions Exercices Prévisions (suite)

Exos: 1 ; 4 ; 5 + précisions

sur l'exo 18

1 - Les données

t	1	2	3	4	5	6	7	8	9
y _t	1	5	4	7	10	11	12	14	17

- Modèle: $y_t = \alpha + \beta X_t + \varepsilon_t, t = 1, \dots, 9$

(En remplaçant i par t)

NB: Les observations de X (qui sont dans l'énoncé) ne sont pas importantes pour répondre à la question (par besoin de X !)

Calculons l'autocorrélation échantillonnale, d'ordre 2.

Par définition, l'autocorrélation d'ordre k est donnée

par:

$$\hat{\rho}_y(k) = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$$

Pour $k=2$

$$\Rightarrow \hat{\rho}_y(2) = \frac{\sum_{t=1}^{T-2} (y_t - \bar{y})(y_{t+2} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$$

Statistiques descriptives :

$$\bar{y} = \frac{\sum_{t=1}^9 y_t}{9} = \frac{81}{9} = 9 ; \quad \sum_{t=1}^9 y_t^2 = 941$$

$$\begin{aligned} \sum_{t=1}^9 (y_t - \bar{y})^2 &= (1-9)^2 + (5-9)^2 + (4-9)^2 + (7-9)^2 \\ &\quad + (10-9)^2 + (11-9)^2 + (12-9)^2 \\ &\quad + (14-9)^2 + (17-9)^2 \end{aligned}$$

$$= 64 + 16 + 25 + 4 + 1 + 4 + 9 + 25 + 64$$

$$\Rightarrow \sum_{t=1}^9 (y_t - \bar{y})^2 = 212$$

$$\begin{aligned} \sum_{t=1}^{9-2} (y_t - \bar{y})(y_{t+2} - \bar{y}) &= (1-9)(4-9) + (5-9)(7-9) \\ &\quad + (4-9)(10-9) + (7-9)(11-9) + (10-9)(12-9) \\ &\quad + (11-9)(14-9) + (12-9)(17-9) \end{aligned}$$

$$= 40 + 8 - 5 - 4 + 3 + 10 + 24$$

$$\Rightarrow \sum_{t=1}^7 (y_t - \bar{y})(y_{t+2} - \bar{y}) = 76$$

et donc,

$$\hat{\rho}_y(2) = \frac{\sum_{t=1}^7 (y_t - \bar{y})(y_{t+2} - \bar{y})}{\sum_{t=1}^9 (y_t - \bar{y})^2} = \frac{76}{212} = 0.35849 \dots$$

$$\Rightarrow \hat{\rho}_y(2) \approx 0.36 \quad \Rightarrow \text{Réponse C}$$

(2)

[4]

Modèles Marche aléatoire

Sans dérive

$$Y_t = Y_{t-1} + \varepsilon_t$$

Avec dérive

$$Y_t = Y_{t-1} + d + \varepsilon_t$$

- On demande l'affirmation correcte!

① - Fausse. Le processus marche aléatoire (sans ou avec dérive) n'est pas stationnaire. Pour le voir il ~~fa~~ suffit de vérifier les conditions d'une stationnarité d'ordre deux. C'est-à-dire, on calcule les moments d'ordre un et deux: i.e:

$$E[Y_t] \text{ et } \text{Var}[Y_t].$$

Y_t est dite stationnaire au second ordre si les trois conditions suivantes sont vérifiées:

- $E[Y_t] = \text{cte}, \forall t$

- $\text{Var}[Y_t] = \text{cte}, \forall t$

- $\text{Cor}(Y_t, Y_{t+k}) = \text{cte}, \forall t, k \in \mathbb{Z}$

Si l'une des trois conditions n'est pas vérifiée la série n'est pas stationnaire.

Cas avec dérive

$t=1$, $Y_1 = Y_0 + d + \varepsilon_1$

$t=2$, $Y_2 = Y_0 + 2d + \varepsilon_1 + \varepsilon_2$

\vdots
 t , $Y_t = Y_0 + td + \sum_{i=1}^t \varepsilon_i$

(3)

- En prenant l'espérance de la dernière écriture de Y_t
On a :

$$\mathbb{E}[Y_t] = \mathbb{E}[Y_0] + \mathbb{E}[td] + \mathbb{E}\left[\sum_{i=1}^t \varepsilon_i\right]$$

(par linéarité de l'espérance mathématique)

$$\mathbb{E}[Y_t] = Y_0 + td + \sum_{i=1}^t \mathbb{E}[\varepsilon_i]$$

(par le fait que $Y_0 = 0^{\text{te}}$ et td aussi, $\forall t$)

$$\Rightarrow \mathbb{E}[Y_t] = Y_0 + td$$

(Puisque $\mathbb{E}[\varepsilon_i] = 0$, $\forall i$)

On voit que l'espérance de Y_t dépend du temps t .

Y_t n'est donc pas stationnaire,

\forall les deux autres conditions.

Cas sans dérive

$$\begin{array}{l} t=1 \\ t=2 \\ \vdots \\ t \end{array} \quad \begin{array}{l} Y_1 = Y_0 + \varepsilon_1 \\ Y_2 = Y_0 + \varepsilon_1 + \varepsilon_2 \\ \vdots \\ Y_t = Y_0 + \sum_{i=1}^t \varepsilon_i \end{array}$$

⇒ En prenant l'espérance de la dernière écriture de Y_t
on a :

$$\begin{aligned} \mathbb{E}[Y_t] &= \mathbb{E}[Y_0] + \mathbb{E}\left(\sum_{i=1}^t \varepsilon_i\right) \\ &= \mathbb{E}[Y_0] + \sum_{i=1}^t \mathbb{E}[\varepsilon_i] \\ &= Y_0 + 0. \end{aligned}$$

La première condition de la stationnarité du second ordre est vérifiée ici.

$$\begin{aligned} \Rightarrow \text{Var}[Y_t] &= \text{Var}\left(Y_0 + \sum_{i=1}^t \varepsilon_i\right) \\ &= \text{Var}\left(\sum_{i=1}^t \varepsilon_i\right) \end{aligned}$$

(Puisque Y_0 est ~~supplé~~ constant)

$$\text{Var}[Y_t] = \sum_{i=1}^t \text{Var}(\varepsilon_i)$$

(Puisque $\text{cor}(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$)

$$\Rightarrow \text{Var}(Y_t) = \sum_{i=1}^t \sigma_\varepsilon^2$$

(Puisque, par définition $\text{Var}(\varepsilon_i) = \text{cste}, \forall i$)

$$\Rightarrow \text{Var}(Y_t) = t \sigma_\varepsilon^2, \text{ dépend du temps } t$$

Donc Y_t non stationnaire, \forall la 3^{ème} condition.

② - Fausse.

Rappelons: ~~cas de Y_{N+1}~~ \hat{Y}_{N+1}

\hat{Y}_{N+1} = prévision de Y_{N+1}

- Pour un niveau de probabilité fixé α (par exemple), plus on avance dans le temps (c'est à dire, plus l'horizon de prévision augmente $h=1, 2, \dots$) plus les bornes de l'intervalle de probabilité augmente d'un facteur \sqrt{h} . Elles ne s'approchent donc pas des droites horizontales, pour un modèle de marche aléatoire avec dérive.

$${}^N Y_{N+h} = Y_N + h d.$$

③ - Vraie. ~~Ce cas peut se~~ c'est le cas pour une marche aléatoire sans dérive.

$${}^N Y_{N+h} = Y_N, \forall h$$

15] . Vraie.

Dans l'analyse des séries temporelles, on cherche à expliquer la série y_t à partir de ces propres retards (ou de observations passées : y_{t-1}, y_{t-2}, \dots). On suppose donc qu'il y aurait une corrélation non nulle entre y_t et l'une ou plus de ces observations passées.

Précisions sur lexo 18

$T = 100$

On donne 15 premiers fonctions d'auto-corrélation.

On donne la statistique de Box-pierce:

$$Q = T \sum_{k=1}^{15} \hat{r}_Y^2(k) = 9.38$$

(A) . Vraie.

C'est plutôt une conséquence du test de Bartlett qui teste l'hypothèse H_0 : si y_t a été généré par un processus bruit blanc, alors $\hat{r}_Y(k) \sim N(0, \frac{1}{T})$

i.e: $E[\hat{r}_Y(k)] = 0, \text{Var}(\hat{r}_Y(k)) = \frac{1}{T} = \frac{1}{100}$

$$\Rightarrow \hat{\sigma}_{\hat{r}_Y(k)} = \sqrt{\frac{1}{100}} = 0.1$$

(B) - Fausse.

C'est toujours le test de Bartlett tout-il s'agit.
Ce test est basé sur l'équation suivante:

Soit $\{y_t\}$ le processus temporelle:

$$y_t = \rho_k y_{t-k}, \quad k = 1, \dots, K.$$

$$H_0: \rho_Y(k) = 0, \quad k = 1, \dots, K.$$

$$H_1: \rho_Y(k) \neq 0, \quad k = 1, \dots, K.$$

On teste donc chaque coefficient individuellement
et non de façon simultanée (i.e. $\rho_Y(1) = \dots = \rho_Y(K) = 0$!).

Il vient, si H_0 est rejetée, cela ne veut
pas dire que le processus $\{y_t\}$ n'est pas
généralisé par un Bruit Blanc. Ça voudra
simplement dire que l'autocorrélation est nulle
entre y_t et y_{t-k} pour le k en question.

(C) - Fausse. Pour les mêmes raisons que dans (B).

(D) - Fausse. Dans ce cas-ci, il s'agit bien du
test de Box-Pierce. On a $\chi^2(15)$ où
15 est le degré de liberté (ici).

15 correspond au nombre de $\rho_Y(k)$ que l'on
test simultanément.

1.e:

Box-pièce:

~~$$y_t = \sum_{k=1}^K p_k y_{t-k} + \dots$$~~

$$y_t = \sum_{k=1}^K p_k y_{t-k} + \dots$$

$$H_0: p_1 = p_2 = \dots = p_k = 0$$

$$H_1: \exists k, p_k \neq 0.$$

ici: $k=15$ et non 85.

(E)

Fausse. Rien n'est sûr qu'à 95% que les coefficients d'autocorrélation sont tous nuls!

Puisque, si on prend $\alpha = 5\%$, la probabilité critique ou conditionnelle ~~est~~ à l'hypothèse H_0

1.e: $\alpha = P[\text{Rejette } H_0 \mid H_0 \text{ vraie}] = 5\%$

$$\alpha = P[\varphi > \chi^2(15) \mid p_1 = p_2 = \dots = p_{15} = 0]$$

il vient,

$$1 - \alpha = P[\text{Accepte } H_0 \mid H_0 \text{ vraie}] = 95\%$$

$$1 - \alpha = P[\varphi < \chi^2(15) \mid p_1 = p_2 = \dots = p_{15} = 0]$$

\Rightarrow la probabilité que φ soit plus petite que α de $\chi^2(15)$ sachant que $p_1 = p_2 = \dots = p_{15} = 0$ est de 95%.

Où, l'affirmation dans E est :

$$P \{ p_1 = p_2 = \dots = p_{15} = 0 \mid Q < \chi^2(15) \} = 95\%$$

$$P \{ H_0 \text{ vraie} \mid \text{Accepte } H_0 \}$$

ie: la probabilité que les p_k , $k=1, \dots, 15$ soient tous simultanément égaux à 0, sachant qu'on a accepté l'hypothèse nulle, est de 95%.

Ce qui n'est pas sûr, puisque $P \{ H_0 \text{ vraie} \mid \text{Accepte } H_0 \}$ n'est pas connue.