Correction épreuve de Séries Temporelles Univariées du 5 janvier 2021

Gilbert Colletaz

6 janvier 2021

Vrai ou Faux ?

1. Vrai.
2. Faux.
3. Faux.
4. Faux.
5. Vrai.
6. Vrai.

Exercices

1. (a) La modélisation retenue suppose que $T_t = 33 + 1.90t$ représente un trend déterministe de long terme présent dans le niveau du PIB. En $t = 240$ ce trend vaut 489. Le PIB observé à cette date est supérieur à sa tendance de long terme, on serait donc plutôt en période de conjoncture favorable.

   (b) — La prévision est donnée par :

$$\begin{aligned} E_{240}[T_{243} + u_{243}] &= T_{243} + E_{240}[-1.26u_{242} - 1.11u_{241} - 0.61u_{240}] \\ &= T_{243} - 0.61E_{240}[u_{240}] \\ &= T_{243} - 0.61u_{240} \\ &= T_{243} - 0.61[PIB_{240} - T_{240}] \\ &= T_{243} - 0.61[506 - 489] \\ &= [33 + 1.90 \times 243] - 0.61 \times 17 \\ &= 494.7 - 10.37 \\ &= 484.33 \end{aligned}$$

   — L'erreur de prévision est égale à $PIB_{243} - E_{240}[PIB_{243}] = -1.26u_{242} - 1.11u_{241}$. Sa variance est de $(1.26^2 + 1.11^2)\sigma_u^2 \approx 2.82 * 120 = 338.4$. L'intervalle de confiance cherché est donc : $484.33 \pm 2\sqrt{338.4} \approx [448, 521]$.

2. (a) — pour les prévisions :
   — $E_{5/1/2021}[x_{6/1/2021}] = E_{5/1/2021}[0.6\, x_{5/1/2021} + 20 + u_{6/1/2021}] = 0.6 \times 88 + 20 = 72.0$ et,

— $E_{5/1/2021}[x_{7/1/2021}] = E_{5/1/2021}[0.6\,x_{6/1/2021} + 20 + u_{7/1/2021}] = 0.6 \times 72 + 20 = 63.2$

— pour les IC :

— l'erreur de prévision du 6/1/2021 est $u_{6/1/2021}$, l'IC est donc : $72 \pm 2\sqrt{36} = [60, 84]$

— l'erreur du 7/1/2021 est :

$$\begin{aligned} x_{7/1/2021} - E_{5/1/2021}[x_{7/1/2021}] &= 0.6\,x_{6/1/2021} + 20 + u_{7/1/2021} - E_{5/1/2021}[0.6\,x_{6/1/2021} - 20] \\ &= 0.6 \times [x_{6/1/2021} - E_{5/1/2021}[x_{6/1/2021}] + u_{7/1/2021} \\ &= 0.6 \times u_{6/1/2021} + u_{7/1/2021} \end{aligned}$$

Sa variance est donc $(1 + 0.6^2)\sigma_u^2 = 1.36 \times 36 = 48.96$. Ainsi, l'IC recherché est : $63.2 \pm 2\sqrt{48.96} \approx [49.2, 77.2]$.

(b) Ici il s'agit de faire une prévision pour un horizon de 365 jours et on sait d'une part que pour des horizons lointains les prévisions convergent vers l'espérance non conditionnelle, soit ici vers $\frac{20}{1-0.6} = 50$, et d'autre part que leur variance tend vers la variance non conditionnelle, ici $36/(1 - 0.6^2) = \frac{36}{0.64} = 56.25$. L'IC est donc : $50 \pm 2\sqrt{56.25} = [35, 65]$.

3. On observe une décroissance régulière sur les autocorrélations et une évolution beaucoup plus irrégulière des autocorrélations partielles ce qui pointe vers un processus AR. Il faut alors trouver l'ordre de l'AR, *i.e.* l'ordre à partir duquel les partielles s'annulent. Avec 100 observations on peut construire un IC autour de zéro égal à $\pm\frac{2}{10} = \pm 0.2$, ce qui conduit à rejeter la nullité des deux premières et ne pas rejeter la nullité des suivantes. On retiendrait un AR(2).

4. On peut déjà calculer les valeurs critiques afférentes aux tests DF ou ADF de Dickey-Fuller, soit pour un seuil de 5%, 150 observations et selon la table de McKinnon :

— pour τ : $-1.9393 - \frac{0.398}{150} = -1.94$

— pour $\tau_m u$: $-2.8621 - \frac{2.738}{150} - \frac{8.36}{150^2} = -2.88$

— pour τ_t : $-3.4126 - \frac{4.039}{150} - \frac{17.83}{150^2} = -3.44$

Les trois premières colonnes de la table s'intéressent à la présence d'une racine unitaire dans le niveau de la variable. Les trois dernières à la présence d'un trend stochastique dans la série des différences premières de la variable. On sait que ce test requiert l'orthogonalité des résidus qui est étudiée ici par le test de Ljung-Box qui sous H0 suit un $\chi 2$ dont le nombre de degrés de liberté est égal au nombre de corrélations utilisées pour sa construction, ici 8, diminué du nombre de coefficients estimés pour calculer les résidus empiriques non compris la constante. On obtient donc le tableau suivant :

régression afférente à	nombre de corrélations	nb de coefficients estimés hors constante			degrés de liberté			valeur critique à 5%		
		$p=0$	$p=1$	$p=2$	$p=0$	$p=1$	$p=2$	$p=0$	$p=1$	$p=2$
τ	8	1	2	3	7	6	5	14.07	12.59	11.07
$\tau\mu$	8	1	2	3	7	6	5	14.07	12.59	11.07
τ_t	8	2	3	4	6	5	4	12.59	11.07	9.49

Les chiffres sont les mêmes que l'expliquée soit Δx_t ou $\Delta^2 x_t$. Au final, seule la statistique de 16.45 obtenue avec $p = 0$ permet le rejet de l'hypothèse nulle lorsque l'on travaille sur le niveau de la variable, soulignant ainsi la nécessité d'utiliser le test ADF plutôt que DF. Dès lors que l'on passe sur ADF, on ne peut pas rejeter la nullité des corrélations résiduelles que ce soit avec Δx_t ou $\Delta^2 x_t$.

On peut maintenant regarder les résultats de ces tests ADF et remarquer que les statistiques de Dickey-Fuller sont, quelle que soit la version du test, toutes supérieures à leur valeur critique lorsque l'expliquée est Δx_t (i.e. on ne rejette pas la nulle) et inférieures à celle-ci lorsqu'on passe à $\Delta^2 x_t$ (i.e. on rejette la nulle). En conséquence, on est amené à admettre la présence d'un seul trend stochastique dans le niveau de la variable étudiée qui serait donc I(1).

5. Sur données journalières on suppose que la valeur d'une variable un jour quelconque dépend de la valeur qu'elle avait 7 jours auparavant et d'un bruit blanc u de variance σ_u^2. Soit donc le processus correspondant :

$$(1 - \phi_1 L^7)x_t = u_t$$

On vous demande :

(a) On est face à un polynôme de degré 7, il a au plus 7 racines. Si on le factorise sur ces racines ω_i, il vient $(1-\phi_1 L^7) = (1-\omega_1^{-1}L)(1-\omega_2^{-1}L)\ldots(1-\omega_7^{-1}L)$ et on sait que chacun de ces polynômes élémentaires est inversible si $\|\omega_i\| > 1$. Pour une quelconque de ces racines on doit vérifier $(1-\phi_1^7\omega_i) = 0$ ou encore $\phi_1^7\omega_i = 1$. Compte-tenu de la condition imposée à la norme de ω_i, on doit vérifier $|\phi_1^7| < 1$ et donc $|\phi_1| < 1$. Sous cette condition il vient aisément :

$$x_t = \sum_{i=0}^{\infty} \phi_1^i L^{7i} u_{t-i} = u_t + \phi_1 u_{t-7} + \phi_1^2 u_{t-14} + \ldots$$

(b) Le processus $x_t = \phi_1 x_{t-7} + u_t$ étant supposé stationnaire et u_t orthogonal à x_{t-7}, on a : $\gamma_0 = \phi_1^2\gamma_0 + \sigma_u^2$ et donc $\gamma_0 = \frac{\sigma_u^2}{1-\phi_1^2}$.

(c) Partant de l'écriture MA, on remarque que les ensembles de u qui définissent respectivement x_t, x_{t-1} et $x_t - 8$ sont d'intersection nulle. En conséquence $\rho_1 = \rho_8 = 0$. Par ailleurs :

$$\gamma_7 = E(x_t x_{t-7}) = E(\phi_1 x_{t-7}^2 + u_t x_{t-7}) = \phi_1\gamma_0, \text{ et}$$
$$\gamma_{14} = E(x_t x_{t-14}) = E(\phi_1 x_{t-7}x_{t-14} + u_t x_{t-14}) = \phi_1\gamma_7$$

et donc : $\rho_7 = \phi_1$ et $\rho_{14} = \phi_1^2$.

(d) — $\phi_{11} = 0$ car $\phi_{11} = \rho_1$ et on a montré qu'ici $\rho_1 = 0$
— on identifie les coefficients de $x_t = \phi_{71}x_{t-1} + \phi_{72}x_{t-2} + \ldots + \phi_{77}x_{t-7} + v_t$ à ceux du modèle vrai : $x_t = \phi_1 x_{t-1} + u_t$, il vient $\phi_{77} = \phi_1$ (et aussi $\phi_{7i} = 0, i = 1, \ldots 6$).
— le processus étant un AR(7), les autocorrélations partielles sont nulles au-delà de 7 et donc : $\phi_{88} = \phi_{14,14} = 0$.

6. — $y_t = 0.8y_{t-1} + 10 + u_t$: c'est un processus stationnaire dont les observations oscilleraient autour de la constante 10/(1-0.8)=50. Ce n'est clairement pas le cas, elles oscillent autour de zéro.
— $y_t = 1.8y_{t-1} + u_t$: c'est un processus explosif et les valeurs absolues de ses observations tendraient vers l'infini. Ce n'est pas du tout vérifié.
— restent donc $y_t = 0.8y_{t-1} + 5\,u_t$ et $y_t = 0.8y_{t-1} + 0.5\,u_t$ qui sont des processus centrés sur zéro. La variance non conditionnelle de y est de $25/(1-0.8^2) \approx 69.4$ pour le premier et $0.25/(1-0.8^2) \approx 0.69$ pour le second. Ainsi, 95% des observations devraient se trouver dans l'intervalle ±16.7 pour l'un et ±1.67 pour l'autre. A l'évidence, seul le premier processus peut respecter cette condition. C'est donc lui qui a été simulé.