

# **PROJET D'ECONOMETRIE**

---

*Gaye Abdoulaye*

*Benkaddour Joudy*

*Nous ferons tous nos tests à 5% pour éviter toutes confusions.*

## **Modélisation univariée :**

### **Question 1 :**

*Choisissez au minimum deux séries temporelles en motivant ce choix. Les séries doivent être disponibles pour la même période et la même fréquence d'observation (annuelle, trimestrielle, mensuelle, hebdomadaire ou quotidienne). La période d'observation, la fréquence, les sources des données doivent être soigneusement reportées. La période d'observation doit s'achever au plus tôt en 2019 et comporter 50 observations environ lorsque c'est possible. Il est préférable de ne pas inclure l'année 2020 dans l'échantillon en raison du comportement atypique de beaucoup de séries économiques pendant cette année.*

Nous avons choisi comme première série temporelle, le taux de chômage trimestriel en France métropolitaine du 1er janvier 2003 au 31 décembre 2018. Il s'agit d'une série corrigée de variations saisonnières, donnée par l'INSEE.

Notre deuxième série temporelle est le nombre de création d'entreprise (autoentrepreneurs inclus) trimestriellement en France métropolitaine du 1er janvier 2003 au 31 décembre 2018. Il s'agit toujours d'une série corrigée de variations saisonnières, donnée par l'INSEE.

Nous avons donc un échantillon de 63 observations pour les deux séries.

L'objectif de ce projet est de comparer ces deux séries temporelles car nous pensons qu'un lien existe entre ces séries et qu'elles sont inversement corrélées. Nous avons eu l'intuition d'utiliser le cours de macroéconomie de cette année (notamment la partie sur le modèle de croissance de Schumpeter) qui prédit la relation entre ces deux variables. En effet, nous sommes amenés à penser que la création d'entreprises a un impact direct sur le taux de chômage de l'économie, qui pourra ensuite à son tour, impacter le PIB de cette même économie. Nous allons donc, avec nos données, chercher à expliquer le modèle de croissance de Schumpeter.

## Question 2 :

Faites et commentez les représentations graphiques et les autocorrélogrammes simples et partiels des séries.



### Taux de chômage (en %) au cours du temps :

Le taux de chômage connaît de nombreuses phases de hausse et de baisse depuis 2003. Parmi elles, une grosse phase de baisse pendant la période de crise financière/économique survenue en 2008. De plus, on peut relever la phase de forte hausse de 2008 à 2016 (au lendemain de la crise) suivi d'une phase de baisse du taux de chômage à partir de 2016, et ce, jusqu'à 2020. Le taux de chômage est ainsi soumis à d'importantes variations depuis 2003.

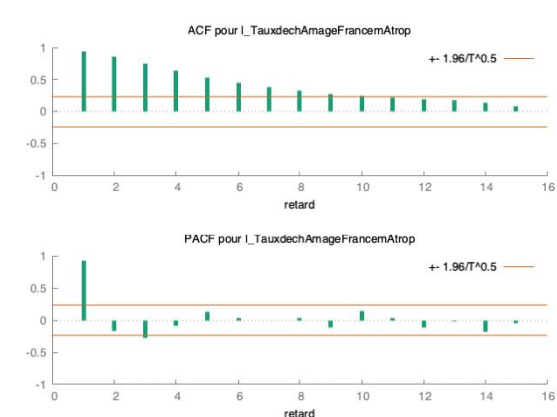
### Création d'entreprise au cours du temps :

La création d'entreprise suit une tendance globalement haussière depuis 2003. On constate, de plus, une forte hausse de la création d'entreprise après la crise de 2008 qui s'est maintenue jusqu'à ce jour. En effet, les économies (en France notamment) avaient besoin de se relancer après la crise, d'où cette explosion.

Par ailleurs, on peut noter à première vue, la relation négative existante entre le taux de chômage et le taux de création d'entreprise en France métropolitaine. En effet, comme dit dans l'introduction, il est licite de penser que plus d'entreprises présentes dans l'économie augmente le nombre de travailleurs donc diminue le taux de chômage.

Pour des raisons de praticité et afin de réduire l'hétéroscédasticité, nous allons nous intéresser au logarithme des séries temporelles.

### Autocorrélogramme du logarithme du taux de chômage en France métropolitaine :

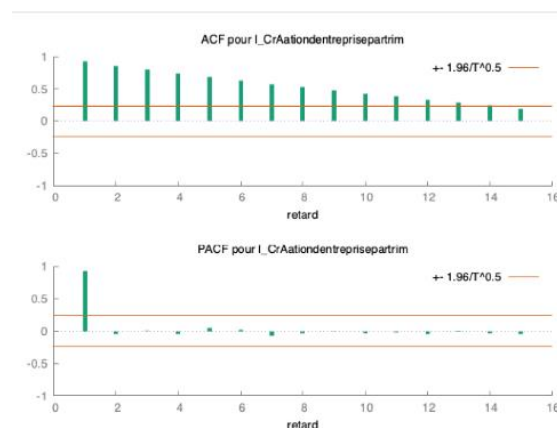


Fonction d'autocorrélation de l\_TauxdechAmeFrancemAtrop  
 \*\*\*, \*\*, \* indiquent une significativité au seuil de 1%, 5% et 10% sur la base d'écart-type  $1/T^{0.5}$

RETARD	ACF	PACF	Q	[p. critique]
1	0.9376 ***	0.9376 ***	62.4565	[0.000]
2	0.8589 ***	-0.1676	115.6552	[0.000]
3	0.7514 ***	-0.2701 **	156.9971	[0.000]
4	0.6357 ***	-0.0878	187.0576	[0.000]
5	0.5362 ***	0.1344	208.7804	[0.000]
6	0.4482 ***	0.0346	224.2060	[0.000]
7	0.3773 ***	0.0004	235.3147	[0.000]
8	0.3256 ***	0.0404	243.7257	[0.000]
9	0.2728 **	-0.1135	249.7310	[0.000]
10	0.2451 **	0.1451	254.6597	[0.000]
11	0.2233 *	0.0321	258.8233	[0.000]
12	0.1985	-0.1148	262.1728	[0.000]
13	0.1792	-0.0150	264.9524	[0.000]
14	0.1368	-0.1758	266.6028	[0.000]
15	0.0880	-0.0407	267.2983	[0.000]

Nous avons pris un retard de 12 pour vérifier les autocorrélations sur 3 ans. Nous voyons sur le graphe ACF que les autocorrélations diminuent au fil du temps tandis que sur le graphe PACF il n'y a que la première autocorrélation partielle qui est significative. Nous pouvons donc supposer que le taux de chômage suit un AR(1).

### Autocorrélogramme du logarithme du taux de création d'entreprises en France métropolitaine :



Fonction d'autocorrélation de l\_CrAationentreprisepartrim  
 \*\*\*, \*\*, \* indiquent une significativité au seuil de 1%, 5% et 10% sur la base d'écart-type  $1/T^{0.5}$

RETARD	ACF	PACF	Q	[p. critique]
1	0.9314 ***	0.9314 ***	61.6381	[0.000]
2	0.8620 ***	-0.0419	115.2327	[0.000]
3	0.7988 ***	0.0091	161.9553	[0.000]
4	0.7334 ***	-0.0513	201.9646	[0.000]
5	0.6790 ***	0.0467	236.7951	[0.000]
6	0.6319 ***	0.0208	267.4543	[0.000]
7	0.5784 ***	-0.0730	293.5595	[0.000]
8	0.5257 ***	-0.0247	315.4856	[0.000]
9	0.4770 ***	-0.0041	333.8416	[0.000]
10	0.4279 ***	-0.0266	348.8680	[0.000]
11	0.3806 ***	-0.0226	360.9678	[0.000]
12	0.3335 ***	-0.0384	370.4244	[0.000]
13	0.2888 **	-0.0097	377.6433	[0.000]
14	0.2442 **	-0.0325	382.9020	[0.000]
15	0.1990	-0.0409	386.4579	[0.000]

A première vue, cet autocorrélogramme ne semble pas si différent de celui du taux de chômage. Nous avons toujours plusieurs autocorrélations significatives mais décroissantes avec le retard mais aussi une seule autocorrélation partielle positive non nulle. On peut donc s'attendre à ce que la modélisation de la création d'entreprise suive un AR(1), ce qui reste à confirmer avec les prochains tests.

### Question 3) :

Appliquez la stratégie de test de racine unitaire à chaque série et indiquez la méthode de stationnarisation appropriée pour chacune d'elle.

Nous prendrons pour la suite des questions uniquement les logarithmes des séries car cela permet d'atténuer les variations.

Nous allons effectuer la stratégie de test séquentielle de racines unitaires pour nos 2 modèles :

#### Test avec création d'entreprise :

```
k = 15: AIC = -89,6549
k = 14: AIC = -91,6299
k = 13: AIC = -93,2242
k = 12: AIC = -95,0525
k = 11: AIC = -97,0522
k = 10: AIC = -97,7798
k = 9: AIC = -99,6965
k = 8: AIC = -100,148
k = 7: AIC = -101,957
k = 6: AIC = -103,858
k = 5: AIC = -104,279
k = 4: AIC = -105,384
k = 3: AIC = -107,366
k = 2: AIC = -107,264
k = 1: AIC = -109,048
k = 0: AIC = -109,402
```

```
Test de Dickey-Fuller augmenté pour l_CrAationdentreprisepartrim
test à reculons à partir de 20 retards, suivant le critère AIC
taille de l'échantillon 67
hypothèse nulle de racine unitaire : a = 1

avec constante et tendance temporelle
avec 0 retards de (1-L)l_CrAationdentreprisepartrim
modèle: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
valeur estimée de (a - 1): -0,101288
statistique de test: tau ct(1) = -1,88412
p. critique asymptotique 0,6627
Coeff. d'autocorrélation du 1er ordre pour e: 0,157
```

Posons  $x_t$ : le logarithme de la variable création d'entreprise

Avec le critère AIC, nous pouvons conclure que le nombre de retard optimal pour nos prochains tests est 0, c'est notre  $p_{\max}$ .

Test Dickey-Fuller augmenté (test unilatéral) :

$$H_0 : \rho = 0 \Rightarrow x_t \text{ non stationnaire}$$
$$H_1 : \rho < 0 \Rightarrow x_t \text{ stationnaire}$$

Avec modèle 3 :  $\Delta x_t = \rho * x_{t-1} + b * t + c + \varepsilon_t$

Pour le tableau de valeur, comme nous avons 63 observations, nous prendrons par principe de prudence, les valeurs pour 50 observations.

$\alpha$	1%	5%	10%
$C(\alpha)$	-4,15	-3,45	-3,18

$$\tau = -1,88 > C(0,05) = -3,45$$

Ainsi on ne rejette pas  $H_0$ , avec ce modèle (3), le processus n'est pas stationnaire. Comme nous n'avons pas rejeté  $H_0$ , nous passons au test de la tendance.

### Test de la tendance (test bilatéral) :

$$H_0 : b=0$$

$$H_1 : b \neq 0$$

```
Régression de Dickey-Fuller
MCO, utilisant les observations 2003:2-2019:4 (T = 67)
Variable dépendante: d_l_CrAationdentreprisepartri
```

	coefficient	éc. type	t de Student	p. critique	
const	1,14206	0,594211	1,922	0,0591	*
l_CrAationdent~_1	-0,101288	0,0537589	-1,884	0,6627	
time	0,00162056	0,00101974	1,589	0,1169	

AIC: -156,118    BIC: -149,504    HQC: -153,501

D'après la p-value (=0.1169) et les valeurs sur le tableau ( $C(0,05) = 3,18 > |t_b| = 1,589$ ), on ne rejette pas l'hypothèse  $H_0$  : il n'y a donc pas de tendance déterministe dans ce modèle.

### Passons donc au test du modèle 2 :

```
k = 15: AIC = -87,1922
k = 14: AIC = -88,8921
k = 13: AIC = -90,8902
k = 12: AIC = -92,8353
k = 11: AIC = -94,3240
k = 10: AIC = -96,0953
k = 9: AIC = -98,0141
k = 8: AIC = -99,4416
k = 7: AIC = -101,439
k = 6: AIC = -102,940
k = 5: AIC = -104,141
k = 4: AIC = -104,412
k = 3: AIC = -106,320
k = 2: AIC = -107,136
k = 1: AIC = -108,497
k = 0: AIC = -109,565
```

Test de Dickey-Fuller augmenté pour l\_CrAationdentreprisepartrim  
test à reculons à partir de 15 retards, suivant le critère AIC  
taille de l'échantillon 67  
hypothèse nulle de racine unitaire : a = 1

test avec constante  
avec 0 retards de (1-L)l\_CrAationdentreprisepartrim  
modèle: (1-L)y = b0 + (a-1)\*y(-1) + e  
valeur estimée de (a - 1): -0,0253171  
statistique de test: tau\_c(1) = -1,01767  
**p. critique asymptotique 0,7492**  
Coeff. d'autocorrélation du 1er ordre pour e: 0,125

Régression de Dickey-Fuller  
MCO, utilisant les observations 2003:2-2019:4 (T = 67)  
Variable dépendante: d\_l\_CrAationdentreprisepartri

Avec le critère AIC, nous pouvons conclure que le nombre de retard optimal pour nos prochains tests est 0, c'est notre  $p_{\max}$ .

### Test Dickey-Fuller augmenté (test unilatéral) :

$$H_0 : \rho = 0 \Rightarrow x_t \text{ non stationnaire}$$

$$H_1 : \rho < 0 \Rightarrow x_t \text{ stationnaire}$$

Avec modèle 2 :  $\Delta x_t = \rho * x_{t-1} + c + \varepsilon_t$

$$\tau = -1,02 > C(0,05) = -2,93 \text{ aussi } p\text{-value} = 0,75 > 0,10$$

Ainsi on ne rejette pas  $H_0$ , avec ce modèle (2), le processus n'est pas stationnaire. Comme nous n'avons pas rejeté  $H_0$ , nous passons au test de la constante du modèle.

### Test de la constante (test bilatéral) :

$$H_0 : c=0$$

$$H_1 : c \neq 0$$

	coefficient	éc. type	t de Student	p. critique
const	0,314617	0,289674	1,086	0,2814
l_CrAationdent~_1	-0,0253171	0,0248776	-1,018	0,7492

AIC: -155,525    BIC: -151,116    HQC: -153,78

Comme la p-value du coefficient c vaut 28% > 10%, on ne rejette pas  $H_0$ , donc la constante c n'est pas significativement différent de 0.

Nous passons au modèle 1 :

```

k = 15: AIC = -86,3845
k = 14: AIC = -88,1516
k = 13: AIC = -90,1515
k = 12: AIC = -92,0990
k = 11: AIC = -93,5686
k = 10: AIC = -95,4133
k = 9: AIC = -97,2848
k = 8: AIC = -98,7709
k = 7: AIC = -100,771
k = 6: AIC = -102,177
k = 5: AIC = -103,516
k = 4: AIC = -103,503
k = 3: AIC = -105,309
k = 2: AIC = -106,494
k = 1: AIC = -107,532
k = 0: AIC = -108,877

Test de Dickey-Fuller augmenté pour l_CrAationdentreprisepartrim
test à reculons à partir de 15 retards, suivant le critère AIC
taille de l'échantillon 67
hypothèse nulle de racine unitaire : a = 1

test sans constante
avec 0 retards de (1-L)l_CrAationdentreprisepartrim
modèle: (1-L)y = (a-1)*y(-1) + e
valeur estimée de (a - 1): 0,00168922
statistique de test: tau nc(1) = 2,15235
p. critique asvmptotique 0,993
Coeff. d'autocorrélation du 1er ordre pour e: 0,118

```

Avec le critère AIC, nous pouvons conclure que le nombre de retard optimal pour nos prochains tests est 0, c'est notre  $p_{\max}$ .

Test Dickey-Fuller augmenté (test unilatéral) :

$$H_0 : \rho = 0 \Rightarrow x_t \text{ non stationnaire}$$

$$H_1 : \rho < 0 \Rightarrow x_t \text{ stationnaire}$$

Avec modèle 1 :  $\Delta x_t = \rho * x_{t-1} + \varepsilon_t$

$$\tau = 2,15 > C(0,05) = -1,95 \text{ aussi } p\text{-value} = 0,99 > 0,10$$

Ainsi on ne rejette pas  $H_0$ , avec ce modèle (1), le processus n'est pas stationnaire. Ainsi, l'organigramme nous indique que :

A l'issu de tous ces tests, le processus  $x_t \sim I(1)$  et peut s'écrire (donc sous la forme d'un bruit blanc) :  $\Delta x_t = \varepsilon_t$

Nous allons donc vérifier cela à l'aide de 2 tests :

Test Phillips-Perron :

```

Phillips-Perron unit-root test for ld_CrAationdentreprisepartrim, Bartlett bandwic
Z t = -6,63402 (p-value = 0,0000)

Test regression (OLS, dependent variable ld_CrAationdentreprisepartrim, T = 66):

-----
                coefficient   éc. type      z      p. critique
-----
ld_CrAationdentreprisepartrim(-1)   0,175050   0,122497  1,429    0,1530

Sample variance of residual      0,00587869
Estimated long-run error variance 0,00466083

```

L'intérêt de ce test est qu'il prend en compte l'autocorrélation et l'hétéroscédasticité

Pour ce test nous avons les mêmes règles de décision que le test ADF et l'on compare la statistique de test aux seuils critiques de la table Dickey-Fuller :

$$H_0 : \rho = 0 \Rightarrow y_t \text{ non stationnaire}$$

$$H_1 : \rho < 0 \Rightarrow y_t \text{ stationnaire}$$

Ainsi, nous voyons bien qu'avec une p-value de 0, nous rejetons l'hypothèse  $H_0$  et donc la série des  $\Delta x_t$  est bien stationnaire.

### Test KPSS :

Hypothèse inversée par rapport aux tests de Dickey-Fuller :

$$H_0 : \sigma_\varepsilon = 0 \Rightarrow x_t \text{ stationnaire}$$

$$H_1 : \sigma_\varepsilon > 0 \Rightarrow x_t \text{ non stationnaire}$$

```

Régression KPSS
MCO, utilisant les observations 2003:2-2019:4 (T = 67)
Variable dépendante: ld_CrAationdentreprisepartrim

-----
                coefficient   éc. type      t de Student   p. critique
-----
const          0,0199717    0,00912855    2,188         0,0322   **

AIC: -156,466   BIC: -154,261   HQC: -155,594

Estimation robuste de la variance: 0,00463399
Somme des carrés des résidus récursifs: 2,20046

Test KPSS pour ld_CrAationdentreprisepartrim

T = 67
Paramètre du retard de troncation = 15
Statistique de test = 0,105781

                10%      5%      1%
Valeurs critiques: 0,350  0,462  0,729
P. critique > .10

```

Ainsi nous avons une p-value > 0.05 donc nous ne rejetons pas à 5% l'hypothèse nulle de stationnarité du modèle, de plus nous voyons que la statistique de test est inférieure à la valeur critique de 5%. Dès lors, selon le test de KPSS, la série  $\Delta x_t$  est stationnaire.

**Selon l'ensemble de nos tests, nous concluons alors que la série : différence première du logarithme du taux de création d'entreprise est une série stationnaire.**

### Test avec le taux de chômage :



k = 15: AIC = -211.433  
 k = 14: AIC = -212.761  
 k = 13: AIC = -214.581  
 k = 12: AIC = -215.645  
 k = 11: AIC = -215.669  
 k = 10: AIC = -215.124  
 k = 9: AIC = -217.122  
 k = 8: AIC = -218.881  
 k = 7: AIC = -220.875  
 k = 6: AIC = -222.818  
 k = 5: AIC = -224.649  
 k = 4: AIC = -225.063  
 k = 3: AIC = -223.453  
 k = 2: AIC = -223.874  
 k = 1: AIC = -219.715  
 k = 0: AIC = -216.854

Test de Dickey-Fuller augmenté pour  $\ln(\text{TauxdechAmageFrancemAtrop})$   
 test à reculons à partir de 15 retards, suivant le critère AIC  
 taille de l'échantillon 63  
 hypothèse nulle de racine unitaire :  $a = 1$

avec constante et tendance temporelle  
 avec 4 retards de  $(1-L)\ln(\text{TauxdechAmageFrancemAtrop})$   
 modèle:  $(1-L)y = b_0 + b_1t + (a-1)y(-1) + \dots + e$   
 valeur estimée de  $(a - 1)$ : -0.0674839  
 statistique de test:  $\tau_{ct}(1) = -1.35551$   
 p. critique asymptotique 0.8736  
 Coeff. d'autocorrélation du 1er ordre pour e: -0.018  
 différences retardées:  $F(4, 56) = 4.656 [0.0026]$

Régression du test de Dickey-Fuller augmenté  
 MCO, utilisant les observations 2004:2-2019:4 (T = 63)  
 Variable dépendante:  $d_{\ln(\text{TauxdechAmageFrancemAtrop})}$

Posons  $y_t$  : le logarithme de la variable taux de chômage

Avec le critère AIC, nous pouvons conclure que le nombre de retard optimal pour nos prochains tests est 4, c'est notre  $p_{\max}$ .

Test Dickey-Fuller augmenté (test unilatéral) :

$H_0 : \rho = 0 \Rightarrow y_t$  non stationnaire

$H_1 : \rho < 0 \Rightarrow y_t$  stationnaire

Avec modèle 3 :  $\Delta y_t = \rho * y_{t-1} + b * t + c + t$

$\tau = -1,35 > C(0,05) = -3,45$  de plus, la probabilité critique de ce test est de :  $0,8736 > 0,05$ .

Ainsi on ne rejette pas  $H_0$ , avec ce modèle (3), le processus n'est pas stationnaire. Comme nous n'avons pas rejeté  $H_0$ , nous passons au test de la constante :

Test de la tendance (test bilatéral) :

$H_0 : b=0$

$H_1 : b \neq 0$

	coefficient	éc. type	t de Student	p. critique	
const	0.143013	0.103193	1.386	0.1713	
$\ln(\text{TauxdechAmag})_1$	-0.0674839	0.0497849	-1.356	0.8736	
$d_{\ln(\text{TauxdechAmag})}_1$	0.194659	0.124261	1.567	0.1229	
$d_{\ln(\text{TauxdechAmag})}_2$	0.348666	0.124291	2.805	0.0069	***
$d_{\ln(\text{TauxdechAmag})}_3$	0.205973	0.128252	1.606	0.1139	
$d_{\ln(\text{TauxdechAmag})}_4$	-0.250487	0.132913	-1.885	0.0647	*
time	7.42199e-05	0.000238609	0.3111	0.7569	
AIC: -273.072 BIC: -258.071 HQC: -267.172					

D'après la p-value (=0.7569) et les valeurs sur le tableau ( $C(0,05) = 3,18 > |t_b| = 0,3111$ ), on ne rejette pas l'hypothèse  $H_0$  donc, on peut affirmer qu'à 95% il n'y a donc pas de tendance déterministe ( $b=0$ ) dans ce modèle.

Nous passons donc au test du modèle 2 :

k = 15: AIC = -213.378  
 k = 14: AIC = -214.761  
 k = 13: AIC = -216.556  
 k = 12: AIC = -217.606  
 k = 11: AIC = -217.370  
 k = 10: AIC = -216.190  
 k = 9: AIC = -218.092  
 k = 8: AIC = -220.027  
 k = 7: AIC = -221.959  
 k = 6: AIC = -223.957  
 k = 5: AIC = -225.476  
 k = 4: AIC = -226.514  
 k = 3: AIC = -225.418  
 k = 2: AIC = -225.659  
 k = 1: AIC = -220.594  
 k = 0: AIC = -216.261

Test de Dickey-Fuller augmenté pour l\_TauxdechAmageFrancemAtrop  
 test à reculons à partir de 15 retards, suivant le critère AIC  
 taille de l'échantillon 63  
 hypothèse nulle de racine unitaire : a = 1

test avec constante  
 avec 4 retards de (1-L)l\_TauxdechAmageFrancemAtrop  
 modèle: (1-L)y = b0 + (a-1)\*y(-1) + ... + e  
 valeur estimée de (a - 1): -0.0578184  
 statistique de test: tau c(1) = -1.49837  
 p. critique asyptotique 0.5347  
 Coeff. d'autocorrélation du 1er ordre pour e: -0.024  
 différences retardées: F(4, 57) = 4.851 [0.0019]

Avec le critère AIC, nous pouvons conclure que le nombre de retard optimal pour nos prochains tests est 4, c'est notre  $p_{\max}$ .

Test Dickey-Fuller augmenté (test unilatéral) :

$H_0 : \rho = 0 \Rightarrow y_t \text{ non stationnaire}$

$H_1 : \rho < 0 \Rightarrow y_t \text{ stationnaire}$

Avec modèle 2 :  $\Delta y_t = \rho * y_{t-1} + c + \varepsilon_t$

$\tau = -1,49837 > C(0,05) = -2,95$ . De plus, nous avons une p-value de 0,5347

Ainsi on ne rejette pas  $H_0$ , avec ce modèle (2), le processus n'est pas stationnaire. Comme nous n'avons pas rejeté  $H_0$ , nous passons au test de la constante.

Test de la constante (test bilatéral) :

$H_0 : c=0$

$H_1 : c \neq 0$

	coefficient	éc. type	t de Student	p. critique
const	0.124713	0.0841049	1.483	0.1436
l_TauxdechAmag~_1	-0.0578184	0.0385875	-1.498	0.5347
d_l_TauxdechAm~_1	0.191072	0.122741	1.557	0.1251
d_l_TauxdechAm~_2	0.341435	0.121126	2.819	0.0066 ***
d_l_TauxdechAm~_3	0.196292	0.123429	1.590	0.1173
d_l_TauxdechAm~_4	-0.263141	0.125528	-2.096	0.0405 **

AIC: -274.964    BIC: -262.105    HQC: -269.906

D'après la p-value (=0.1436) et les valeurs sur le tableau ( $C(0,05) = 2,89 > |t_b| = 1,483$ ), on ne rejette pas l'hypothèse  $H_0$  à 5%, il n'y a donc pas de constante ( $c=0$ ).

Nous passons donc au test du modèle 1 :

Avec le critère AIC, nous pouvons conclure que le nombre de retard optimal pour nos prochains tests est 4, c'est notre  $p_{\max}$ .

Test Dickey-Fuller augmenté (test unilatéral) :

$H_0 : \rho = 0 \Rightarrow y_t \text{ non stationnaire}$

$$H_1 : \rho < 0 \Rightarrow y_t \text{ stationnaire}$$

```
test sans constante
avec 4 retards de (1-L)\l_TauxdechAmageFrancemAtrop
modèle: (1-L)y = (a-1)*y(-1) + ... + e
valeur estimée de (a - 1): -0.000643581
statistique de test: tau_pc(1) = -0.422078
p. critique asymptotique 0.5314
Coef. d'autocorrélation du 1er ordre pour e: -0.042
différences retardées: F(4, 58) = 4.470 [0.0032]

Régression du test de Dickey-Fuller augmenté
MC0, utilisant les observations 2004:2-2019:4 (T = 63)
Variable dépendante: d_l_TauxdechAmageFrancemAtrop
```

Avec modèle 1 :  $\Delta y_t = \rho * y_{t-1} + \varepsilon_t$

$\tau = -0,422078 > C(0,05) = -1,95$ . De plus, nous avons une p-value de 0,5314.

Ainsi on ne rejette pas  $H_0$ , avec ce modèle (1), le processus n'est pas stationnaire.

A l'issue de tous ces tests, on en conclut donc que le processus  $y_t \sim I(1)$  et peut s'écrire (donc sous la forme d'un bruit blanc) :  $\Delta y_t = \varepsilon_t$

### Test Phillips-Perron pour confirmer (sans constante, ni tendance)

```
Phillips-Perron unit-root test for ld_TauxdechAmageFrancemAtrop, Bartlett bandwidth
Z_t = -6,4321 (p-value = 0,0000)

Test regression (OLS, dependent variable ld_TauxdechAmageFrancemAtrop, T = 66):
```

	coefficient	éc. type	z	p. critique	
ld_TauxdechAmageFrancemAtrop(-1)	0,248021	0,120701	2,055	0,0399	**

```
Sample variance of residual      0,000807967
Estimated long-run error variance 0,000974434
```

L'intérêt de ce test est qu'il prend en compte l'autocorrélation et l'hétéroscédasticité

Pour ce test nous avons les mêmes règles de décision que le test ADF et l'on compare la statistique de test aux seuils critiques de la table Dickey-Fuller :

$$H_0 : \rho = 0 \Rightarrow y_t \text{ non stationnaire}$$

$$H_1 : \rho < 0 \Rightarrow y_t \text{ stationnaire}$$

Ainsi, nous voyons bien qu'avec une p-value de 0, nous rejetons l'hypothèse  $H_0$  et donc la série des  $\Delta y_t$  est bien stationnaire.

### Test KPSS :

Hypothèse inversée par rapport à des tests de Dickey-Fuller :

$$H_0 : \sigma_\varepsilon = 0 \Rightarrow y_t \text{ stationnaire}$$

$$H_1 : \sigma_\varepsilon > 0 \Rightarrow y_t \text{ non stationnaire}$$

```

Régression KPSS
MCO, utilisant les observations 2003:2-2019:4 (T = 67)
Variable dépendante: d_l_TauxdechAmeFrancemAtrop

      coefficient      éc. type    t de Student    p. critique
-----
const      -0.000187743    0.00356123    -0.05272      0.9581

AIC: -282.6   BIC: -280.396   HQC: -281.728

Estimation robuste de la variance: 0.0013287
Somme des carrés des résidus récursifs: 1.30224

Test KPSS pour d_l_TauxdechAmeFrancemAtrop

T = 67
Paramètre du retard de troncation = 15
Statistique de test = 0.218332

      10%      5%      1%
Valeurs critiques: 0.350  0.462  0.729
P. critique > .10

```

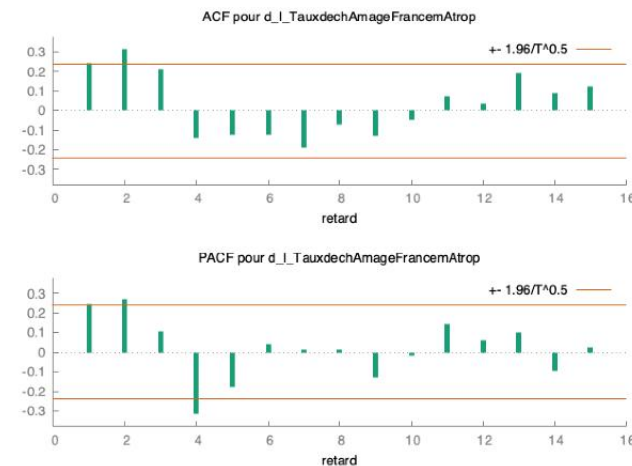
Ainsi nous avons une p-value  $> 0.10$  donc nous ne rejetons pas à 5% l'hypothèse nulle, de plus nous voyons que la statistique de test est inférieure à la valeur critique de 5%. Dès lors, selon le test de KPSS, la série  $\Delta y_t$  est stationnaire.

**Selon l'ensemble de nos tests, nous concluons alors que la série : différence première du logarithme du taux de chômage est une série stationnaire.**

#### Question 4) :

Pour l'une des séries, cherchez le modèle ARMA satisfaisant les conditions de validité.

Nous allons prendre comme série stationnarisée, la différence première du logarithme du taux de chômage, dont voici son autocorrélogramme :



Visuellement, nous devons nous attendre à un ARMA(4,2) car pour les ACF nous avons des autocorrélations significatives jusqu'à t-2, tandis que pour la partie PACF, nous avons les autocorrélations partielles jusqu'à t-4 qui sont significatives.

Détermination du meilleur modèle grâce à la fonction armax :

```
armax: successfully-l:p=0,q=0
```

Information Criteria of ARMAX(p,q) for ld_TauxdechAmageFrancemAt			
p, q	AIC	BIC	HQC
0, 0	-280,6004	-276,1911	-278,8556
0, 1	-283,3186	-278,9092	-281,5738
0, 2	-285,8906*	-283,0719*	-285,4010*
0, 3	-287,1913	-285,1193	-287,7123
0, 4	-287,2214	-285,9932	-286,9870
1, 0	-284,7516	-280,3422	-283,0068
1, 1	-285,5322	-278,9181	-282,9150
1, 2	-292,1778	-283,3590	-288,6882
1, 3	-294,0908	-283,0674	-289,7288
1, 4	-293,8098	-280,5816	-288,5754
1, 5	-291,8268	-276,3940	-285,7200
2, 0	-287,7884	-281,1743	-285,1712
2, 1	-286,0131	-277,1944	-282,5235
2, 2	-292,6518	-281,6284	-288,2898
2, 3	-292,2047	-278,9765	-286,9703
2, 4	-291,8279	-276,3950	-285,7210
2, 5	-289,8282	-272,1907	-282,8490
3, 0	-286,5832	-277,7645	-283,0936
3, 1	-286,2564	-275,2330	-281,8944
3, 2	-291,5474	-278,3193	-286,3130
3, 3	-291,0951	-275,6622	-284,9882
3, 4	-292,7129	-275,0753	-285,7337
3, 5	-288,0688	-268,2266	-280,2172
4, 0	-281,2782	-280,2501	-287,3112

Évaluations de la fonction : 45  
Évaluations du gradient : 17

Modèle 1: ARMA, utilisant les observations 2003:2-2019:4 (T = 67)  
Estimation par AS 197 (MV exact)  
Variable dépendante: ld\_TauxdechAmageFrancemAtrop  
Écart-types basés sur la matrice hessienne

	coefficient	éc. type	z	p. critique
const	-0,000870847	0,00603612	-0,1443	0,8853
theta_1	0,163556	0,116975	1,398	0,1620
theta_2	0,469492	0,113283	4,144	3,41e-05 ***
theta_3	0,385341	0,112073	3,438	0,0006 ***

Moyenne var. dép. -0,000188 Éc. type var. dép. 0,029150  
Moyenne innovations 0,000047 Éc. type innovations 0,024863  
R2 0,261636 R2 ajusté 0,238562  
Log de vraisemblance 151,9557 Critère d'Akaike -293,9115  
Critère de Schwarz -282,8880 Hannan-Quinn -289,5494

	Réel	Imaginaire	Modulo	Fréquence
MM				
Ordre 1	-1,7907	0,0000	1,7907	0,5000
Ordre 2	0,2861	-1,1693	1,2038	-0,2118
Ordre 3	0,2861	1,1693	1,2038	0,2118

La capture ci-dessus (à gauche) montre l'exécution de la fonction armax et selon cette dernière, le modèle idéal, ici, correspond à un MA (3), car c'est celui qui minimise l'ensemble des 3 critères retenus (AIC, BIC et HQ). Ce résultat contredit la prévision que l'on aurait pu faire à l'aide de l'autocorrélogramme, on pouvait s'y attendre car ceux-ci sont moins précis.

Nous allons donc faire la régression de ce modèle sur GRETL :

Tout d'abord, on vérifie bien que les modules des racines sont strictement supérieurs à 1 (voir tableau à droite), ce qui est bien le cas ici. De plus il ne peut pas y avoir de racine commune entre celles de MA et AR car il n'y a pas de composante AR dans ce modèle.

Dans ce modèle, nous voyons que la constante et  $\theta_1$  ne sont pas significatifs, donc on recommence l'estimation du modèle en supprimant ces deux variables.

```

Évaluations de la fonction : 43
Évaluations du gradient : 14

Modèle 3: ARMA, utilisant les observations 2003:2-2019:4 (T = 67)
Estimation par AS 197 (MV exact)
Variable dépendante: ld_TauxdechAmageFrancemAtrop
Écarts-types basés sur la matrice hessienne

```

	coefficient	éc. type	z	p. critique
theta_2	0,490128	0,115094	4,259	2,06e-05 ***
theta_3	0,290077	0,0983159	2,950	0,0032 ***

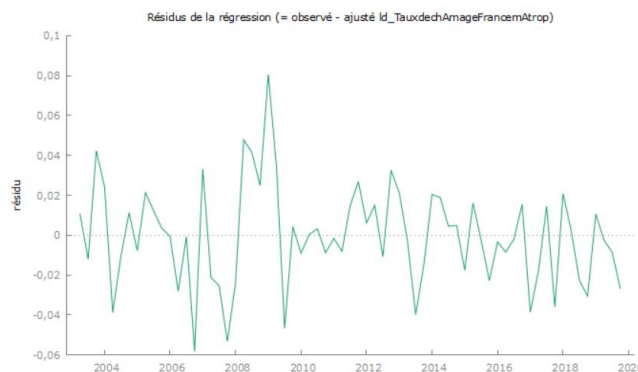
Moyenne var. dép.	-0,000188	Éc. type var. dép.	0,029150
Moyenne innovations	-0,000359	Éc. type innovations	0,025214
R2	0,241243	R2 ajusté	0,229569
Log de vraisemblance	151,0410	Critère d'Akaike	-296,0819
Critère de Schwarz	-289,4678	Hannan-Quinn	-293,4647

	Réel	Imaginaire	Modulo	Fréquence
MM				
Ordre 1	0,3184	-1,1749	1,2173	-0,2079
Ordre 2	0,3184	1,1749	1,2173	0,2079
Ordre 3	-2,3265	0,0000	2,3265	0,5000

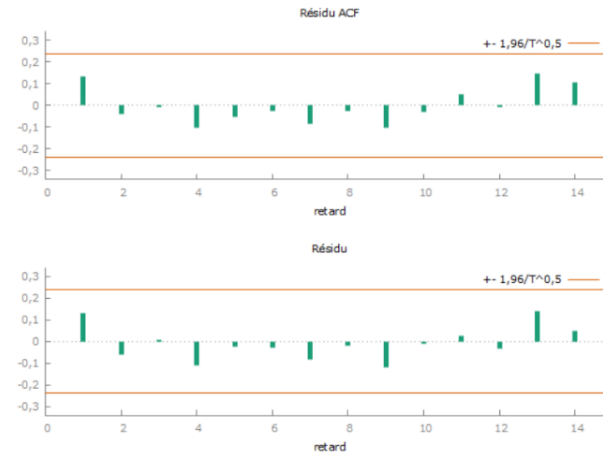
Dans la capture d'écran ci-dessus, l'ensemble des variables sont significatives, le modèle est donc satisfaisant. De plus les modules des racines restent strictement supérieurs à 1.

Concentrons-nous sur les résidus à présent :



Le graphique ci-dessus montre que malgré certaines grandes variations (notamment durant la période de crise 2008-2009), nous observons des oscillations autour de 0 et la volatilité reste globalement constante au cours du temps. Les résidus de la régression peuvent sans doute être modélisés par un bruit blanc, à confirmer dans la suite.

Le corrélogramme des résidus (ci-dessous) est bien caractéristique d'un bruit blanc : aucune autocorrélation et autocorrélation partielle significativement différente de 0 n'est à déclarer.



Une fois les résidus observés, passons dès à présent aux tests sur les résidus.

### 1. Test de Ljung-Box : Autocorrélation des résidus

Ce test repose sur les hypothèses suivantes :

$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0$  (absence d'autocorrélation des résidus)

$H_1 : \exists i \in \{1, 2, \dots, k\}$  tel que  $\rho_i \neq 0$  (présence d'autocorrélation des résidus)

Test d'autocorrélation jusqu'à l'ordre 15

Ljung-Box  $Q' = 8,36022$ ,

avec p. critique =  $P(\text{Chi-deux}(13) > 8,36022) = 0,8194$

D'après les résultats de ce test, la probabilité critique est de 0,8194 avec  $Q = 8,36022$ ,  $Q$  suivant une loi de chi-deux à 12 degrés de liberté.

Nous ne rejetons donc pas l'hypothèse nulle d'absence d'autocorrélation des résidus. Ainsi, nous pouvons en conclure, à 95%, qu'il n'y a pas d'autocorrélation entre les résidus du modèle.

### 2. Test de normalité des résidus (Jarque et Bera)

Les hypothèses de ce test sont comme suit :

$H_0 : S(X) = 0$  et  $K(X) = 3$  (la modélisation des résidus suit une loi normale)

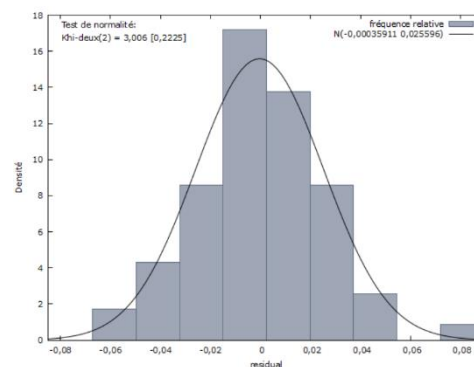
$H_1 : S(X) \neq 0$  ou  $K(X) \neq 3$  (la modélisation des résidus ne suit pas une loi normale)



Distribution des fréquences pour residual, obs 2-68  
nombre de classes = 9, moyenne -0,000359107, éc. type = 0,0255962

intervalle	centre	fréquence	rel.	cum.
< -0,049657	-0,058342	2	2,99%	2,99% *
-0,049657 - -0,032288	-0,040973	5	7,46%	10,45% **
-0,032288 - -0,014918	-0,023603	10	14,93%	25,37% *****
-0,014918 - 0,0024518	-0,0062331	20	29,85%	55,22% *****
0,0024518 - 0,019821	0,011137	16	23,88%	79,10% *****
0,019821 - 0,037191	0,028506	10	14,93%	94,03% *****
0,037191 - 0,054561	0,045876	3	4,48%	98,51% *
0,054561 - 0,071931	0,063246	0	0,00%	98,51%
>= 0,071931	0,080616	1	1,49%	100,00%

Test de l'hypothèse nulle de normalité de la distribution:  
Khi-deux(2) = 3,006 avec p. critique 0,22246



La probabilité critique de ce test est de  $0,22 > 0,05$ , donc nous ne rejetons pas l'hypothèse nulle de normalité des erreurs. Nous pouvons aussi observer que la distribution des erreurs semble avoir la forme d'une loi normale, réconfortant ainsi le test réalisé par Gretl.

### 3. Test d'effet ARCH sur l'hétéroscédasticité des erreurs :

Les hypothèses de ce test sont les suivantes :

$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0 \Rightarrow$  pas d'effet ARCH (traduit l'absence d'hétéroscédasticité des erreurs du modèle)

$H_1 : \alpha_1 \neq 0$  ou  $\alpha_2 \neq 0 \dots$  ou  $\alpha_q \neq 0 \Rightarrow$  effet ARCH

Test pour ARCH d'ordre 15

	coefficient	éc. type	t de Student	p. critique
alpha(0)	0,000382398	0,000292696	1,306	0,1997
alpha(1)	0,0916545	0,150082	0,6107	0,5452
alpha(2)	0,263786	0,149030	1,770	0,0852 *
alpha(3)	0,208601	0,156104	1,336	0,1898
alpha(4)	-0,0205987	0,156189	-0,1319	0,8958
alpha(5)	0,0855891	0,158621	0,5396	0,5928
alpha(6)	-0,0908450	0,158941	-0,5716	0,5712
alpha(7)	-0,0924959	0,155975	-0,5930	0,5569
alpha(8)	-0,0712989	0,155233	-0,4593	0,6488
alpha(9)	0,259223	0,154682	1,676	0,1024
alpha(10)	-0,0441330	0,159442	-0,2768	0,7835
alpha(11)	0,0399264	0,157566	0,2534	0,8014
alpha(12)	-0,219125	0,160682	-1,364	0,1811
alpha(13)	-0,111997	0,158442	-0,7069	0,4842
alpha(14)	-0,0319956	0,151808	-0,2108	0,8343
alpha(15)	0,102190	0,152132	0,6717	0,5061

Hypothèse nulle: pas d'effet ARCH présent  
Statistique de test: LM = 15,2708  
avec p. critique = P(Khi-deux(15) > 15,2708) = 0,432081

Nous voyons par la capture d'écran ci-dessus que les coefficients des valeurs passées des résidus n'influent pas sur la variance de l'erreur (car ils ne sont pas significativement différents de 0).

Ainsi, la variance du résidu ne dépend pas du temps dès lors, une variance du résidu (élevée ou faible) à la date  $t$ , a très peu de chance de se propager à la date  $t+h$ .

**L'ensemble des tests réalisés sur les résidus satisfait l'hypothèse que ceux-ci sont un bruit blanc et ceux-ci suivent une loi normale centrée en 0 et de variance  $\sigma_e^2$ .**

**Pour conclure, la différence première du logarithme du taux de chômage, validant l'ensemble des tests de validation, suit bien un MA(3), les résidus étant caractérisés par un bruit blanc suivant une loi normale de moyenne nulle.**



### Question 5) :

Calculez les prévisions de la série modélisée sur les horizons de 1 à 3 en détaillant la méthode de calcul. Si nécessaire, calculez les prévisions de la série initiale non stationnarisée pour les horizons 1 à 3.

Nous pouvons désormais à l'aide de notre modèle MA(3), pouvoir faire des prévisions.

Gretl nous montre ceci :

Pour l'intervalle de confiance de 95%,  $z(0,025) = 1,96$

	ld_TauxdechAmage~	prédiction	éc. type	intervalle de 95%
2011:4	0,022473	-0,004411		
2012:1	0,011050	0,004924		
2012:2	0,032435	0,017468		
2012:3	0,000000	0,010801		
2012:4	0,041673	0,009113		
2013:1	0,020203	-0,000952		
2013:2	0,009950	0,012825		
2013:3	-0,020001	0,019814		
2013:4	-0,010152	0,004727		
2014:1	0,000000	-0,020348		
2014:2	0,000000	-0,018842		
2014:3	0,010152	0,005657		
2014:4	0,020001	0,015138		
2015:1	-0,009950	0,007669		
2015:2	0,019803	0,003688		
2015:3	-0,009852	-0,007225		
2015:4	-0,020001	0,002787		
2016:1	0,000000	0,003387		
2016:2	-0,020409	-0,011931		
2016:3	-0,010363	-0,008270		
2016:4	0,010363	-0,005138		
2017:1	-0,042111	-0,003485		
2017:2	-0,010811	0,006990		
2017:3	0,000000	-0,014436		
2017:4	-0,055880	-0,019930		
2018:1	0,022728	0,001912		
2018:2	-0,011300	-0,013433		
2018:3	-0,022990	-0,000226		
2018:4	-0,023530	0,007084		
2019:1	0,000000	-0,010538		
2019:2	-0,024098	-0,021608		
2019:3	-0,012270	-0,003715		
2019:4	-0,025001	0,001837		
2020:1		-0,004915	0,025214	-0,054333 - 0,044503
2020:2		-0,015636	0,025214	-0,065054 - 0,033783
2020:3		-0,007785	0,028080	-0,062820 - 0,047250



Nous pouvons faire une prévision uniquement sur 3 périodes car notre modèle est un MA(3), les prévisions futures à l'horizon  $h > 3$  ne seront que des prévisions nulles.

Nous voyons sur le graphique que l'écart-type de l'erreur augmente bien avec le temps.

Pour expliquer les résultats sur GRETL, il nous faut partir du modèle :

$y_t = \varepsilon_t + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3}$  avec ( $\theta_2 = 0,490128$  et  $\theta_3 = 0,290077$ ), le MA(3) retenu.

Avec une prévision dynamique, on peut déterminer  $y_{t+1}$  en fonction des résidus au temps  $t-1$  et  $t-2$  et on peut déterminer les résidus avec :

$\varepsilon_t = \text{Valeur observé en } t - \text{Valeur selon modèle en } t.$

$\varepsilon_{2019:4} = -0,026838$	$\varepsilon_{2019:3} = -0,008555$	$\varepsilon_{2019:2} = -0,00249$
------------------------------------	------------------------------------	-----------------------------------

Comme les résidus futurs ne sont pas connus, leur meilleure estimation est leur moyenne (donc 0) on obtient alors :

$$y_{2020:1} = \varepsilon_{2020:1} + \theta_2 \varepsilon_{2019:3} + \theta_3 \varepsilon_{2019:2} = 0 + 0,490128 * (-0,008555) + 0,290077 * (-0,00249) = -0,004915$$

$$y_{2020:2} = \varepsilon_{2020:2} + \theta_2 \varepsilon_{2019:4} + \theta_3 \varepsilon_{2019:3} = 0 + 0,490128 * (-0,026838) + 0,290077 * (-0,008555) = -0,015636$$

$$y_{2020:3} = \varepsilon_{2020:3} + \theta_2 \varepsilon_{2020:1} + \theta_3 \varepsilon_{2019:4} = 0 + 0 + 0,290077 * (-0,026838) = -0,007785$$

Le problème étant qu'ici  $y_t$  représente la différence du logarithme de série du taux de chômage, donc pour prévoir la véritable valeur du taux de chômage, il faut recolorer notre série  $y_t$  :

Soit :

- $y_t$  : la série stationnaire
- $z't$  : le logarithme de la série initiale (le logarithme taux de chômage)
- $z_t$  : la série initiale (le taux de chômage)

On a donc  $z_t(h) = \exp(z't + y_t(h))$

Dès lors,

$$z_t(1) = \exp(2,066863 - 0.004915) = 7,8612\%$$

$$z_t(2) = 7,7774\%$$

$$z_t(3) = 7,83873\%$$

*Ce sont donc nos prévisions pour le taux de chômage pour les 3 prochains trimestres selon notre modèle.*

## Modélisation multivariée :

### Question 6) :

Estimez le modèle VAR ayant le nombre de retards optimal. La composante déterministe de ce VAR devra être cohérente avec les conclusions de la question 4. Commentez les résultats des estimations.

Nous devons déterminer le nombre de retard optimal pour la VAR.

Nous obtenons ceci :

Système VAR, ordre des retards max 5

Les astérisques indiquent les meilleures (donc les plus faibles) valeurs des critères d'information suivants, AIC = critère d'Akaike, BIC = critère bayésien de Schwartz et HQC = critère d'Hannan-Quinn.

retards	log-vrais.	p(LR)	AIC	BIC	HQC
1	209,74179		-6,636832	-6,499598*	-6,582950
2	216,68074	0,00770	-6,731637	-6,457168	-6,623873*
3	218,05775	0,59979	-6,647024	-6,235321	-6,485379
4	224,12411	0,01639	-6,713681	-6,164743	-6,498154
5	231,42763	0,00559	-6,820246*	-6,134074	-6,550838

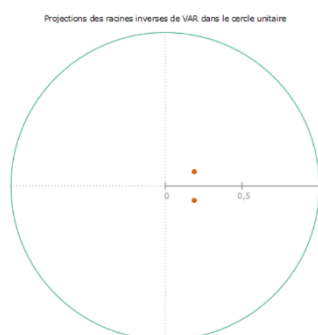
Les 3 critères d'informations diffèrent, nous décidons de suivre le BIC qui nous indique que le retard maximum optimal est 1. En effet, cela a pour objectif de réduire la taille du modèle, nous permettant d'optimiser et de faciliter les calculs.

Le nombre de retards optimal choisit, nous pouvons procéder à l'estimation optimal de notre modèle VAR.

### Estimation du VAR :

Avant toutes choses, nous allons vérifier que notre VAR (1) est bien stationnaire. Pour cela, nous allons représenter le graphique des racines inverses de notre modèle :

Nous observons bien que les modules des inverse des racines de notre VAR estimé sont inférieurs à 1, nous avons donc un modèle VAR(1) stationnaire.



Equation 1: ld\_TauxdechAmageFrancemAtrop

	coefficient	éc. type	t de Student	p. critique
ld_TauxdechAma~_1	0.285628	0.133118	2.146	0.0357 **
ld_CrAationden~_1	-0.0342121	0.0500850	-0.6831	0.4970

Moyenne var. dép. -0.000379 Éc. type var. dép. 0.029331  
Somme carrés résidus 0.052138 Éc. type régression 0.028542  
R2 0.067793 R2 ajusté 0.053227  
F(2, 64) 2.327141 P. critique (F) 0.105778  
rho -0.084572 Durbin-Watson 2.155051

Tests de Fisher d'absence de restriction:

Tous les retards de ld_TauxdechAmage~	F(1, 64) =	4.6039 [0.0357]
Tous les retards de ld_CrAationden~	F(1, 64) =	0.46660 [0.4970]
Toutes les variables, retard 1	F(2, 64) =	2.3271 [0.1058]

### Significativité des coefficients :

#### **Équation du taux de chômage en France métropolitaine :**

Les p-values des tests de significativité des coefficients indiquent que le coefficient du premier retard du log-différencié de la création d'entreprise (ld\_Creationentreprise) est non significatif. En effet, la p-value de ce test (0,4970) est supérieure à notre seuil de test de première espèce (5%).

En revanche, nous observons, (par une p-value égale à  $0,0377 < 0,05$ ), que le coefficient du premier retard du log-différencié du taux de chômage (ld\_tauxdechomage) est quant à lui significativement différent de 0 (rejet de l'hypothèse nulle de nullité de ce coefficient).

Commentaire : Le taux de chômage en France métropolitaine dépend positivement de sa valeur en t-1. En effet, cela n'est pas surprenant, puisqu'il est légitime de penser que le gouvernement ainsi que les banques centrales ajustent leur prévisions sur l'inflation ainsi que sur le taux de chômage passé pour essayer de contrôler le niveau du taux de chômage. Nous retrouvons ici la notion de la courbe de Philips vu l'année dernière en macroéconomie.

En revanche, il est surprenant de constater que le taux de chômage à la date t ne dépend pas de la valeur du taux de création d'entreprise à la date t-1. Puisqu'il serait tentant de croire que plus il y a d'entreprises, plus il y a de travailleurs et donc moins de chômage. Cependant, lorsqu'on y réfléchit bien, les données sur la création d'entreprises en France englobent même les créations d'auto-entreprises. Or, les auto-entrepreneurs ne permettent pas nécessairement d'avoir une influence notable sur le taux de chômage alors en vigueur.

Equation 2: ld\_CrAationentreprisepartrim

	coefficient	éc. type	t de Student	p. critique
ld_TauxdechAma~_1	0.537836	0.354051	1.519	0.1337
ld_CrAationden~_1	0.0913591	0.133210	0.6858	0.4953
Moyenne var. dép.	0.019774	Éc. type var. dép.	0.075275	
Somme carrés résidus	0.368817	Éc. type régression	0.075913	
R2	0.064202	R2 ajusté	0.049580	
F(2, 64)	2.195397	P. critique (F)	0.119628	
rho	0.022849	Durbin-Watson	1.941106	

Tests de Fisher d'absence de restriction:

Tous les retards de ld_TauxdechAma~	F(1, 64) =	2.3076 [0.1337]
Tous les retards de ld_CrAationdenr~	F(1, 64) =	0.47036 [0.4953]
Toutes les variables, retard 1	F(2, 64) =	2.1954 [0.1196]

#### **Équation de la création d'entreprise en France métropolitaine :**

Nous observons cette fois que le coefficient du premier retard du log de la création d'entreprise (ld\_Creationentreprise) est non significatif au seuil de 5%. De plus, nous observons aussi que le coefficient du premier retard du log du taux de chômage (ld\_tauxdechomage) est non significatif au seuil critique de notre test (de 5%) car sa p-value( $0,1337 > 0,05$ ).

Commentaire : La création d'entreprises en France métropolitaine à la date t ne dépend donc pas du taux de chômage à la date t-1, et pas non plus de la création d'entreprises à la date t-1. Cela n'est pas surprenant puisque la présence de profession libérale (d'auto-entrepreneurs) compromet la vraisemblance du modèle. En effet, comme dit dans le commentaire précédent, il n'est pas possible, ni concret de relier les notions d'auto-entrepreneurs et de taux de chômage dans l'économie.

### **Question 7) :**

Déterminez les relations de causalité entre les variables et commentez les résultats obtenus.

Nous allons mettre en place un test de causalité de Granger pour observer les relations de causalité entre nos variables, le test est représenté ci-dessous :

Hypothèse de test d'absence de causalité de Granger :

$$H_0 : a_{12}^l = 0 \Rightarrow x_{2t} \text{ ne cause pas } x_{1t}$$

$$H_1 : a_{12}^l \neq 0 \Rightarrow x_{2t} \text{ cause } x_{1t}.$$

La statistique de test utilisée pour réaliser ce test sera la statistique de Fisher, que l'on peut voir dans les captures d'écran ci-dessous :

**Équation du taux de chômage en France métropolitaine :**

Tests de Fisher d'absence de restriction:

Tous les retards de ld_IxdeChomage	F(1, 64) =	4,6039	[0,0357]
Tous les retards de ld_CreaEntr	F(1, 64) =	0,46660	[0,4970]
Toutes les variables, retard 1	F(2, 64) =	2,3271	[0,1058]

Etant donné la p-value du test de 0,4970 qui est supérieure au seuil critique de notre test (5%), nous ne rejetons pas l'hypothèse nulle selon laquelle le taux de création d'entreprise à la date t-1 n'affecte pas le taux de chômage à la date t.

**Équation de la création d'entreprise en France métropolitaine :**

Tests de Fisher d'absence de restriction:

Tous les retards de ld_IxdeChomage	F(1, 64) =	2,3076	[0,1337]
Tous les retards de ld_CreaEntr	F(1, 64) =	0,47036	[0,4953]
Toutes les variables, retard 1	F(2, 64) =	2,1954	[0,1196]

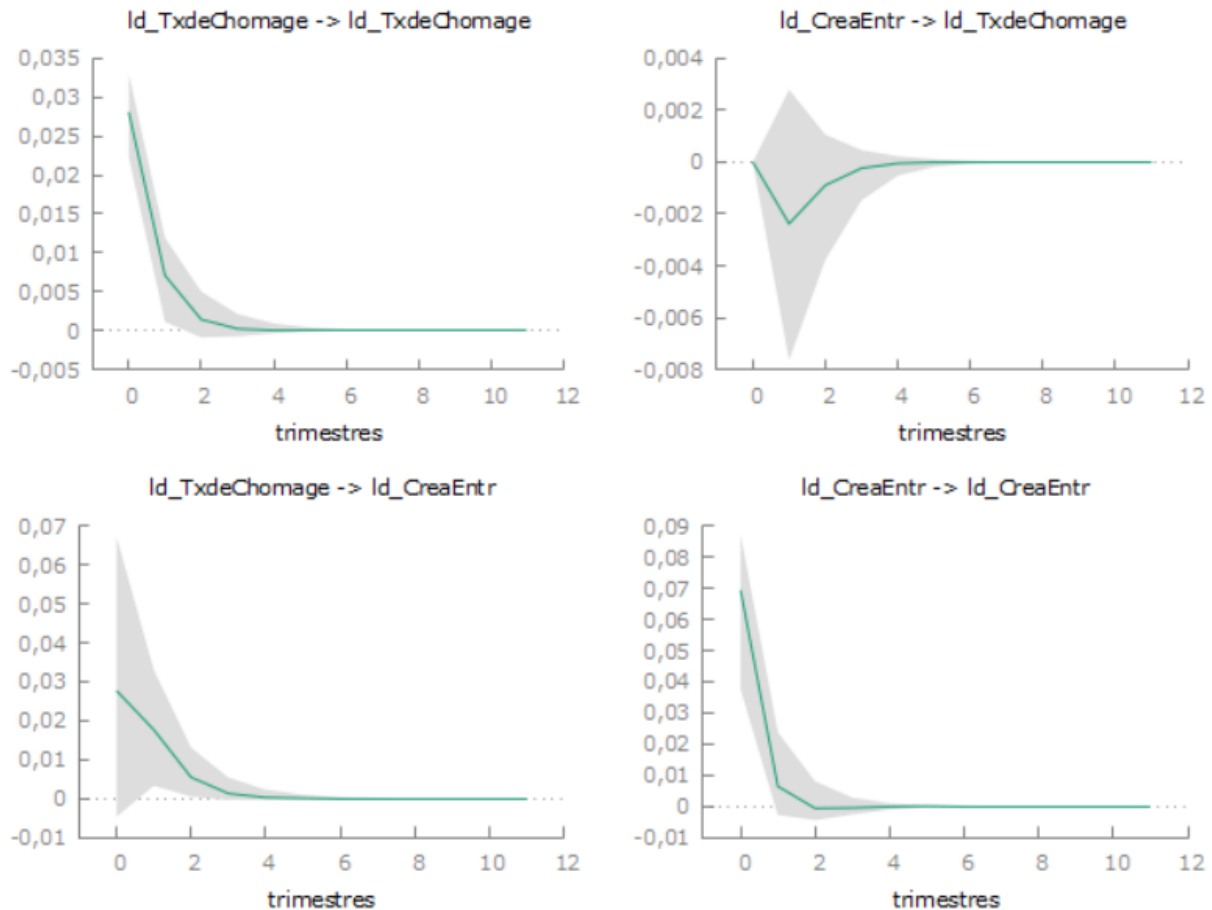
Etant donné la p-value du test de 0,1337 qui est supérieure au seuil critique de notre test (5%), nous ne rejetons pas l'hypothèse nulle selon laquelle le taux de chômage à la date t-1 n'affecte pas la création d'entreprise à la date t.

Le test de causalité au sens de Granger nous permet alors de conclure que les variables ne sont pas les causes l'une de l'autre.

### Question 8) :

En vous appuyant sur les conclusions des tests de causalité, effectuez et commentez une analyse impulsion-réponse des chocs sur les différentes variables du VAR. Vous justifierez votre choix pour l'ordre des variables dans la décomposition de Choleski.

#### Analyse impulsion-réponse :



Nous utilisons la décomposition de Choleski pour orthogonaliser les deux chocs. Étant donné que l'ordre des variables dans cette décomposition peut impacter l'analyse d'impulsion réponse, nous avons décidé de placer en première position la variable stationnarisée du taux de chômage, susceptible d'avoir une grande influence sur le nombre de création d'entreprises en France (d'après le test de causalité au sens de Granger, du fait que ce soit la plus petite des deux p-value vu précédemment).

Nous avons décidé d'étudier la réponse des variables face aux impulsions sur les 12 prochains trimestres (sur les 4 années à venir, qui est un horizon plausible pour étudier l'impact des chocs). En effet, nous voyons bien sur l'ensemble des 4 graphiques que cela suffit pour que les variables atteignent leurs valeurs d'équilibre (satisfaisant ainsi les conditions de stationnarité de notre VAR (1) vérifiées à la question 6).

Les graphiques de gauche correspondent à la réaction à un choc (à  $t=0$ ) sur le taux de chômage en France métropolitaine, tandis que les graphiques de droite correspondent à la réaction à un choc (à  $t=0$ ) sur le nombre de création d'entreprises en France.

#### Analyse des deux chocs :

- **Choc sur le taux de chômage :**

Réponse du taux de chômage : l'effet de ce choc décroît exponentiellement au cours du temps : Au début positif (à  $t=0$ ), l'effet se compense très rapidement, devenant nul à partir du troisième trimestre, devenant immédiatement après non-significatif. Le taux de chômage a su trouver sa valeur d'équilibre après seulement 3 trimestres (en 9 mois, donc moins d'une année).

Réponse du nombre de créations d'entreprises : L'effet de ce choc est très semblable au précédent, une décroissance significative. Bien que l'impact soit significatif à  $t=0$ , l'impact baisse de plus en plus au fur et à mesure des trimestres jusqu'à atteindre 0 à partir du troisième trimestre. Là aussi, la réponse au choc est immédiate, la création d'entreprises a su retrouver sa valeur d'équilibre après seulement 3 trimestres.

- **Choc sur la création d'entreprises en France métropolitaine :**

Réponse du taux de chômage : L'impact du choc connaît une rapide oscillation. En effet, du début du choc au troisième trimestre, celui-ci est négatif, avant de se reprocher de 0 au bout de, là aussi, du troisième trimestre. Après le troisième trimestre, son impact restera nul.

Réponse du nombre de création d'entreprises : L'effet du choc décroît immédiatement après le choc à  $t=0$ . Avec au début, une valeur proche de 0,07, le choc diminue fortement de telle sorte, que seulement à partir du deuxième trimestre après le choc, son impact est quasiment nul.

### **Conclusion :**

Un choc sur le taux de chômage ou le nombre de créations d'entreprises n'a que très peu défaut sur l'autre variable. Puisqu'en effet, le temps nécessaire pour que l'impact du choc ne soit plus ressenti est très bref (au sens de l'horizon temporel de nos variables), puisqu'il tourne autour de deux à trois trimestres.

L'analyse des impulsions-réponses nous permet ainsi de confirmer nos précédentes affirmations sur le test de causalité au sens de Granger. En effet, puisque les variables n'ont pas du tout d'influence l'une sur l'autre, il est logique qu'elle ne soit que très peu (voire pas du tout) impacté lorsque l'une d'entre elles subit un choc.

### Question 9) :

Faites le test de cointégration de Johansen si les séries initiales sont intégrées d'ordre 1. Indiquez si les résultats du test valident le modèle VAR. Proposez un modèle alternatif si vous le jugez nécessaire.

Nous avons dans notre cas deux séries initiales I(1) sans tendance ni constante, nous pouvons donc faire le test de cointégration de Johansen. Ce test permet de déterminer le nombre de relation de cointégration linéairement indépendantes et de les déterminer.

Nous avons fait le test avec une constante restreinte car nos séries n'ont pas de tendance et avec 1 retard comme précédemment.

```
Test de Johansen:
Nombre d'équations = 2
Ordre des retards = 1
Période d'estimation: 2003:2 - 2019:4 (T = 67)
```

Cas 2 : constante restreinte

Log de vraisemblance = 419,983 (constante incluse: 229,845)

Rang	Valeur propre	Test de la Trace	p. critique	Test Lmax	p. critique
0	0,11687	9,6608 [0,6771]	8,3272 [0,5204]		
1	0,019707	1,3335 [0,8891]	1,3335 [0,8882]		

Corrigé pour la taille de l'échantillon (ddl = 64)

Rang	Test de la Trace	p. critique
0	9,6608 [0,6904]	
1	1,3335 [0,8887]	

Nous allons faire une résolution par étape :

Pour le 1<sup>er</sup> test de la trace nous avons :

$H_0(0): \text{rang}(\Pi) = 0$  donc il n'existe pas de relations de cointégration  
 $H_a(0): \text{rang}(\Pi) > 0$

Nous obtenons une p-value de 0,67 > 0,05, ainsi nous ne rejetons pas l'hypothèse nulle, nous concluons qu'il n'y a pas de relation de cointégration car le rang de la matrice  $\Pi$ .

Nous arrêtons nos tests ici. Il n'existe pas de combinaison linéaire de nos deux séries qui forme une série stationnaire. On pouvait s'attendre à ce résultat car nous n'avons pas de lien de causalité entre les variables.

### Conclusion du projet :

L'ensemble des tests réalisés dans le cadre de notre projet permet ainsi d'infirmer, partiellement le fait que le taux de chômage ait un impact assez significatif sur le nombre de créations d'entreprises (au moins en France métropolitaine). En effet, le fait d'inclure les auto-entrepreneurs dans la variable de création d'entreprises « fausse » le modèle, puisque le modèle ainsi obtenu ne nous a pas permis de confirmer le modèle de croissance de Schumpeter, qui a pourtant prévalu pendant de longues années. En effet, les résultats des tests de causalité de Granger ainsi que de l'analyse impulsion réponses sont à la défaveur de la théorie avancée par Schumpeter.