Chapitre III : Estimation, spécification et prévision d'un modèle ARMA

Yannick Le Pen

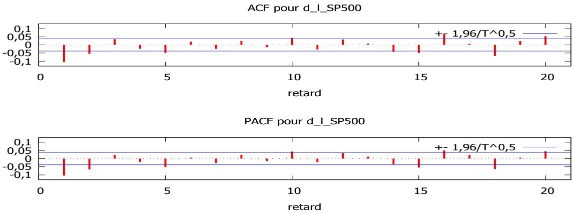
- Utilisation d'un modèle ARMA pour modéliser une série économique.
- Conditions d'application :
- Une série stationnaire
- cette série présente de l'autocorrélation ou de l'autocorrélation partielle.
- Dans le cas d'une série stationnaire sans autocorrélation ni autocorrélation partielle, il n'y a pas à chercher de représentation par un modèle ARMA.

Introduction : étapes de la procédure d'estimation

- Détermination du nombre de retards p et q à l'aide de l'autocorrélogramme de la série
- Estimation du modèle
- Tests de validation du modèle :
  1. test de significativité des coefficients et analyse des racines
  2. test sur les résidus
- Recherche d'un modèle alternatif si les conditions de validité du modèle estimé ne sont pas satisfaites.
- Prévision à l'aide d'un modèle ARMA

Exemple de l'indice S&P 500

- Indice S&P 500 pas stationnaire mais son taux de rendement stationnaire
- Modélisation du taux de rendement du S&P 500
- Autocorrélogramme du taux de rendement



- Autocorrélations (et les autocorrélations partielles) significatives aux ordres 1 et 2

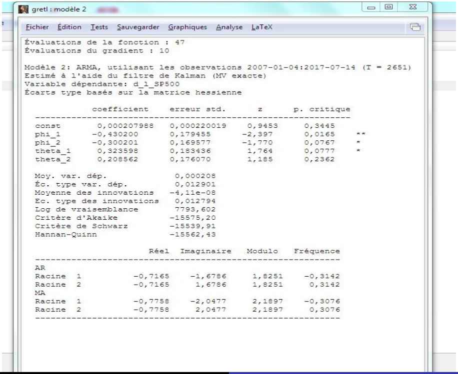
- On veut estimer un modèle ARMA(p,q) :

$$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

- Echantillon de T observations $\{x_1,\ x_2, ...xT\}$
- Les paramètres inconnus sont :
    - les nombres de retards p (composante AR) et q (composante MA)
    - la constance c
    - les p paramètres $\phi_1, \phi_2, ..., \phi_p$ de la partie AR
    - les q paramètres $\theta_1,\ \theta_2, ..., \theta_q$ de la partie MA
    - la variance σ_ϵ^2 du terme d'erreur ϵ_t.
- soit au final $1 + 1 + 1 + p + q + 1$ paramètres inconnus.

- Méthode d'estimation : estimateur du maximum de vraisemblance.
- Il n'existe pas de solution analytique et l'estimation des paramètres se fait par des méthodes de résolution numérique.
- Sous des conditions assez générales, cet estimateur suit une loi normale, ce qui implique que les tests statistiques standards sur les coefficients peuvent être appliqués.
- Dans le cas d'un processus AR(p) sans composante MA, l'estimateur du maximum de vraisemblance correspond à l'estimateur des MCO.

- Exemple du taux de rendement de l'indice S&P 500
- Autocorrélation et l'autocorrélation partielle jusqu'à l'ordre 2 pour r_{SP500}
- Ni à un modèle AR, ni à un modèle MA purs
- On commence par un ARMA(2,2).



gretl : modèle 2

Fichier Édition Tests Sauvegarder Graphiques Analyse LaTeX

```
Évaluations de la fonction : 47
Évaluations du gradient : 10

Modèle 2: ARMA, utilisant les observations 2007-01-04:2017-07-14 (T = 2651)
Estimé à l'aide du filtre de Kalman (MV exacte)
Variable dépendante: d_l_SP500
Écarts type basés sur la matrice hessienne

             coefficient    erreur std.      z       p. critique
  ---------------------------------------------------------------
  const       0,000207988   0,000220019    0,9453    0,3445
  phi_1      -0,430200      0,179455      -2,397     0,0165    **
  phi_2      -0,300201      0,169577      -1,770     0,0767    *
  theta_1     0,323598      0,183436       1,764     0,0777    *
  theta_2     0,208562      0,176070       1,185     0,2362

  Moy. var. dép.             0,000208
  Éc. type var. dép.         0,012901
  Moyenne des innovations   -4,11e-08
  Ec. type des innovations   0,012794
  Log de vraisemblance       7793,602
  Critère d'Akaike          -15575,20
  Critère de Schwarz        -15539,91
  Hannan-Quinn              -15562,43

                      Réel   Imaginaire    Modulo   Fréquence
  -----------------------------------------------------------
  AR
  Racine  1         -0,7165    -1,6786     1,8251    -0,3142
  Racine  2         -0,7165     1,6786     1,8251     0,3142
  MA
  Racine  1         -0,7758    -2,0477     2,1897    -0,3076
  Racine  2         -0,7758     2,0477     2,1897     0,3076
  -----------------------------------------------------------
```

- Vérifier si le modèle estimé est conforme au modèle théorique d'un processus ARMA
- Test 1 : Analyse des racines
  - Vérifier que les racines de la partie AR sont de module strictement supérieur à 1
  - Si oui : la série modélisée est bien stationnaire
  - Si non : la série modélisée n'est pas stationnaire
- Exemple du taux de rendement du S&P 500
- le module des deux racines est égal à 1,8251
- taux de rendement est stationnaire

Test de Student des paramètres

- On vérifie la significativité des paramètres estimés avec la statistique de Student.
- En particulier, les coefficients des derniers retards doivent être significativement différents de 0, sinon, ils doivent être éliminés.
- Exemple du taux de rendement du S&P 500
  - la constante c n'est pas significative
  - le coefficient ϕ_1 est significatif à 5%
  - les coefficients ϕ_2 et θ_1 sont significatifs à 10%
  - le coefficient θ_2 n'est pas significatif
  - Ce modèle ARMA(2,2) n'est pas valide. On doit estimer un modèle avec moins de retards pour la partie MA.

- Après élimination du deuxième puis du premier retard de la composante MA, on estime le modèle AR(2) :

```
gretl : modèle 10
Fichier  Édition  Tests  Sauvegarder  Graphiques  Analyse  LaTeX

Évaluations de la fonction : 43
Évaluations du gradient : 6

Modèle 10: ARMA, utilisant les observations 2007-01-04:2017-07-14 (T = 2651)
Estimé à l'aide du filtre de Kalman (MV exacte)
Variable dépendante: d_l_SP500
Écarts type basés sur la matrice hessienne

             coefficient    erreur std.      z      p. critique
  --------------------------------------------------------------
  const       0,000207862   0,000211052    0,9849    0,3247
  phi_1      -0,111405      0,0193659     -5,753     8,79e-09    ***
  phi_2      -0,0665268     0,0193200     -3,443     0,0006      ***

  Moy. var. dép.            0,000208
  Éc. type var. dép.        0,012901
  Moyenne des innovations  -9,29e-08
  Ec. type des innovations  0,012799
  Log de vraisemblance      7792,442
  Critère d'Akaike         -15576,88
  Critère de Schwarz       -15553,35
  Hannan-Quinn             -15568,37

                      Réel  Imaginaire    Modulo   Fréquence
  -----------------------------------------------------------
  AR
  Racine  1         -0,8373    -3,7856    3,8771    -0,2846
  Racine  2         -0,8373     3,7856    3,8771     0,2846
  -----------------------------------------------------------
```

- ϕ_1 et ϕ_2 significatifs
- Racines de la partie AR sont de module supérieur à 1.

- Dans le modèle théorique, $\epsilon_t \sim BB(0, \sigma_\epsilon^2)$.
- Vérification que les résidus $\hat{\epsilon}_t$ se comportent comme un bruit blanc.
  1. tests de Ljung-Box d'absence d'autocorrélation
  2. tests d'homoscédasticité : test d'un effet ARCH
  3. test de l'hypothèse de normalité des résidus : test de Jarque et Bera

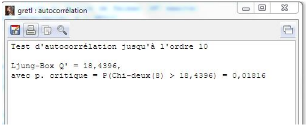
Tests de Ljung-Box d'absence d'autocorrélation des résidus

- On teste l'hypothèse d'absence d'autocorrélation des résidus jusqu'au rang k :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = ... = \rho_k = 0 \\ H_1 : \exists i \in \{1, 2, ..., k\} \ t.q. \ \rho_i \neq 0 \end{cases}$$

- ρ_i est l'autocorrélation des résidus $\hat{\epsilon}_t$.
- La statistique de test est $Q_k^* = T\,(T+2)\sum_{j=1}^{k} \frac{1}{T-j}\hat{\rho}_j^2$
- Sous H_0, $Q_k^* \rightarrow \chi^2(k-p-q)$
- On rejette H_0 si la statistique de test est supérieure au seuil critique choisi, ou si la probabilité critique est inférieure au risque de première espèce.
- Si l'on rejette H_0, on doit estimer un nouveau modèle, en augmentant le nombre de retards pour la partie AR ou la partie MA.

Test de Ljung-Box appliqué au taux de rendement du S&P 500



- On voit que l'on accepte l'hypothèse nulle d'absence d'autocorrélation (de peu) pour un risque de première espèce de 1%.
- On rejette cette hypothèse pour un risque de 5% et 10%
- le modèle estimé passe (de peu) les critères de validation.

Test de l'hypothèse de normalité de Jarque et Bera

- On veut vérifier si les résidus $\hat{\epsilon}_t$ suivent une loi normale. Ce test est basé sur :
  - le skewness S : le coefficient d'asymétrie d'une distribution,
  - le kurtosis K : le coefficient d'aplatissement.

Test de Jarque et Bera : Interprétation du Skewness

- Le skewness mesure le degré d'asymétrie d'une distribution
- L'expression du Skewness est $S(X) = \frac{E(X-E(X))^3}{\sigma(X)^3}$
- Si $S = 0$, la distribution est symétrique autour de sa moyenne
- Si $S > 0$, la distibution est étalée vers la droite
- Si $S < 0$, la distribution est étalée vers la gauche.
- Si $X \sim N(0,1)$, la distribution de X est symétrique autour de 0 et $S(X) = 0$ et $K(X) = 3$ [1].

1. On définit aussi l'excès de kurtosis comme $K(X) - 3$.

Test de Jarque et Bera : interprétation du kurtosis

- Le kurtosis mesure le degré d'applatissement d'une distribution.
- L'expression du kurtosis est $K(X) = \frac{E(X-E(X))^4}{\sigma(X)^4}$
- Il est lié à la probabilité d'apparition des valeurs extrêmes
- Si $X \sim N(0,1)$, $K(X) = 3$.
- On définit aussi l'excès de kurtosis comme $K(X) - 3$.
- Si pour une variable Y, on a $K(Y) > 3$ alors la probabilité d'apparition de valeurs extrêmess pour Y est plus importante que dans le cas d'une loi normale.

Test de Jarque et Bera : Estimation de $S(X)$ et $K(X)$

Le skewness et le kurtosis sont estimés par :

$$\hat{S}(X) = \frac{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^3}{\left[\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^{3/2}}$$

$$\hat{K}(X) = \frac{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^4}{\left[\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2\right]^{2}}$$

Test de Jarque et Bera

- fondé sur la comparaison entre le skewness et le kurtosis estimé à partir d'un échantillon et les valeurs de référence de la loi normale.
- Hypothèses du test :

$$\begin{cases} H_0 : \ S(X) = 0 \ et \ K(X) = 3 \\ H_1 : S(X) \neq 0 \ ou \ K(X) \neq 3 \end{cases}$$

- La statistique du test notée LJB est une moyenne pondérée de $(S(X) - 0)^2$ et $(K(X) - 3)^2$.
- Sous H_0, la statistique de test suit la loi $LJB \rightarrow \chi^2(2)$.
- On rejettera H_0 si la statistique de test est supérieure au seuil critique ou si la probabilité critique est inférieure au risque de première espèce choisi.

Application au taux de rendement du S&P 500

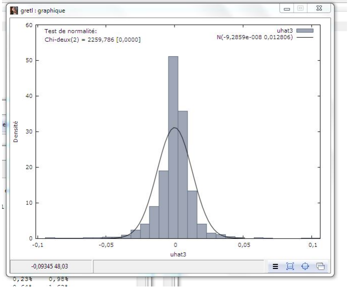
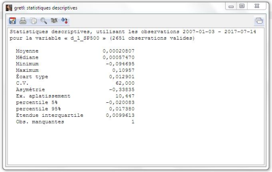


Figure – Résultat du test de normalité de Jarque et Bera

- La valeur de la statistique de test "chi-deux = 2257,786"
- La probabilité critique égale à 0.
- On rejette donc l'hypothèse de normalité des résidus.

Application au taux de rendement du S&P 500



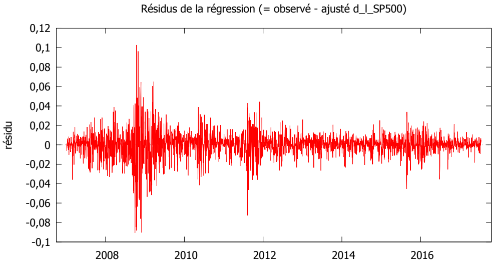
Statistiques descriptives, utilisant les observations 2007-01-03 - 2017-07-14
pour la variable « d_l_SP500 » (2651 observations valides)

Moyenne	0,00020807
Médiane	0,00057470
Minimum	-0,094695
Maximum	0,10957
Écart type	0,012901
C.V.	62,000
Asymétrie	-0,33835
Ex. aplatissement	10,447
percentile 5%	-0,020083
percentile 95%	0,017380
Etendue interquartile	0,0099613
Obs. manquantes	1

- $S = -0,33835$. La distribution du taux de rendement est légèrement étalée vers la gauche.
- La probabilité de rendements négatifs est plus élevée que celle de rendements positifs
- $K - 3 = 10,447$.
- Probabilité d'apparition de rendements extrêmes est plus importante que dans le cas d'une loi normale.

Test d'un effet ARCH

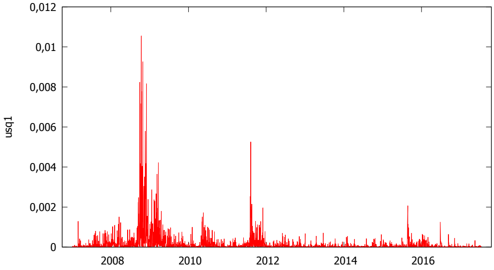
- On veut vérifier si la variance du résidu $\hat{\epsilon}_t$ ne dépend pas du temps.
- Une caractéristique des séries financières.



- Alternance de période de fortes et faibles fluctuations (alternance de "boom and bust".)
- On parle aussi de cluster de volatilité

Test d'un effet ARCH

- On utilise $\hat{\epsilon}_t^2$ comme une proxy de la volatilité :



- Alternance de périodes tranquilles et de périodes de forte volatilité .

Test d'un effet ARCH

- Le test d'un effet ARCH est fondé sur l'estimation de la régression :

$$\hat{\epsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\epsilon}_{t-1}^2 + ... + \alpha_q \hat{\epsilon}_{t-q}^2 + u_t$$

- Hypothèses du test :

$$\begin{cases} H_0 : \alpha_1 = \alpha_2 = ... = \alpha_q = 0 \Rightarrow pas\, d'\, effet\, ARCH \\ H_a : \alpha_1 \neq 0\, ou\, \alpha_2 \neq 0 ...\, ou\, \alpha_q \neq 0 \Rightarrow effet\, ARCH \end{cases}$$

  Statistique de test est :

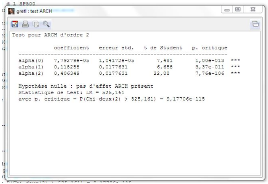
$$LM_{stat} = TR^2 \rightarrow \chi^2(q)$$

  où T est le nombre d'observations et R^2 le R^2 de la régression de $\hat{\epsilon}_t^2$ sur les $\{\hat{\epsilon}_{t-i}^2\}$.

- Le test d'un effet ARCH est le plus souvent appliqué à des séries de rendement d'actifs financiers.

Test d'un effet ARCH sur le taux de rendement du S&P 500

- On applique le test d'un effet ARCH avec 2 retards au taux de rendement du S&P 500.


gretl : test ARCH

Test pour ARCH d'ordre 2

	coefficient	erreur std.	t de Student	p. critique	
alpha(0)	7,79279e-05	1,04172e-05	7,481	1,00e-013	***
alpha(1)	0,118258	0,0177631	6,658	3,37e-011	***
alpha(2)	0,406349	0,0177631	22,88	7,76e-106	***

Hypothèse nulle : pas d'effet ARCH présent
Statistique de test: LM = 525,161
avec p. critique = P(Chi-deux(2) > 525,161) = 9,17706e-115

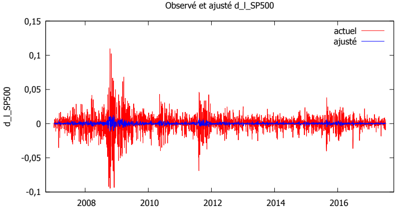
- Probabilité critique égale à 0 : rejet l'hypothèse d'absence d'effet ARCH.

Remarques sur les tests sur les résidus

On accorde pas la même importance à tous les tests sur les résidus.

1. Le test d'absence d'autocorrélation doit être validé. Si ce n'est pas le cas, il faut chercher un modèle ARMA alternatif
2. Si on rejette l'hypothèse de loi normale, on ne fait rien de plus
3. Si on rejette l'hypothèse d'absence d'effet ARCH, on pourra estimer un modèle de volatilité conditionnelle (voir le chapitre sur la modélisation GARCH).

- Le modèle AR(2) a passé les tests de validation
- Faible différence entre l' écart type des résidus (innovation) et celui de la variable dépendante est petite.
- Faible pouvoir explicatif du modèle (assez attendu pour une série financière).
- Conclusion confirmée par le graphique des valeurs observées et des valeurs ajustées :



- Utilisation du modèle validé pour faire de la prévision (à court terme).
- Informations nécessaires pour faire une prévision sont :
  1. un modèle de génération des données : la modèle que l'on a estimé et validé.
  2. les valeurs passées de x_t à la date t : $I_t = \{x_s, s \leq t\}$
  3. la date t à laquelle la prévision est faite
  4. le nombre de périodes dans le futur pour lesquelles la prévision est faite : l'horizon de la prévision h
- $x_t^a(h) =$ la valeur prévue en t de x_{t+h}.
- L'erreur de prévision est la différence entre la valeur observée et la valeur prévue $e_t(h) = x_{t+h} - x_t^a(h)$.
- Deux types de prévision :
  1. la prévision ponctuelle
  2. la prévision par intervalle de confiance

- On veut calculer des prévisions à la date t
- A cette date t, si l'on considère toutes les valeurs de x_{t+h}, on aura :
  - x_{t+h} connu si $h \leqslant 0$
  - x_{t+h} inconnu si $h > 0$ et remplacé par son prédicteur $x_t^a(h)$
- De la même façon, on aura :
  - ϵ_{t+h} connu si $h \leqslant 0$
  - ϵ_{t+h} inconnu si $\forall h > 0$ et remplacé par sa moyenne $E(\epsilon_t) = 0$
- On utilise ces formules pour calculer tous les prédicteurs.

Prévision ponctuelle dans le cas d'un *AR(1)*

- *AR(1)* stationnaire $x_t = \phi_1 x_{t-1} + \epsilon_t$
- En t, on veut prévoir x_{t+1} : prévision à l'horizon *h=1*
- On a $x_{t+1} = \phi_1 x_t + \epsilon_{t+1}$
- En t, x_t est connu et ϵ_{t+1} est inconnu et remplacé par 0.
- La prévision en t de x_{t+1} est :

$$x_t^a(1) = \phi_1 x_t$$

- $x_t^a(1)$ est la valeur ajustée du modèle
- ϵ_{t+1} est la variation de x entre t et $t+1$ non prévisible à la date t.

Prévision ponctuelle dans le cas d'un *AR(1)*

- L'erreur de prévision à l'horizon *h=1* est :

$$e_t(1) = x_{t+1} - x_t^a(1) = (\phi_1 x_t + \epsilon_{t+1}) - \phi_1 x_t = \epsilon_{t+1}$$

- L'erreur de prévision est nulle *en moyenne* :

$$E\left(e_t(1)\right) = E\left(\epsilon_{t+1}\right) = 0 \Rightarrow E(x_t^a(1)) = x_{t+1}$$

- Le prédicteur $x_t^a(1)$ est dit sans biais[2].
- La variance de l'erreur de prévision est $V\left(e_t(1)\right) = V\left(\epsilon_{t+1}\right) = \sigma_\epsilon^2$

2. Sous réserve que le modèle *AR(1)* soit le bon modèle.

Prévision à l'horizon h=2

- On applique les mêmes principes que précedemment
- $x_{t+2} = \phi_1 x_{t+1} + \epsilon_{t+1}$
- En t, x_{t+1} est inconnu et remplacé par son prédicteur $x_t^a(1) = \phi_1 x_t$
- En t, ϵ_{t+1} est inconnu et remplacé par son espérance $E(\epsilon_{t+1}) = 0$.
- En t, la prévision de x_{t+2} est donc :

$$x_t^a(2) = \phi_1 x_t^a(1) = \phi_1^2 x_t$$

Prévision à l'horizon *h=2*

- L'erreur de prévision à l'horizon *h=2* est :

$$
\begin{aligned}
e_t(2) &= x_{t+2} - x_t^a(2) \\
&= (\phi_1 x_{t+1} + \epsilon_{t+2}) - \phi_1 x_t^a(1) \\
&= \phi_1 \left(x_{t+1} - x_t^a(1)\right) + \epsilon_{t+2} \\
&= \phi_1 \epsilon_{t+1} + \epsilon_{t+2}
\end{aligned}
$$

- L'erreur de prévision est nulle en moyenne et le prédicteur $x_t^a(2)$ est sans biais : $E\left(e_t(2)\right) = 0 \Rightarrow E(x_t^a(2)) = x_{t+2}$
- La variance de l'erreur de prévision est $V\left(e_t(2)\right) = \sigma_\epsilon^2(1+\phi_1^2)$
- L'erreur de prévision est nulle en moyenne mais sa variance augmente avec l'horizon de la prévision.

Généralisation à un AR(p)

- On en déduit que pour $h > 1$, on aura $x_t^a(h) = \phi_1 x(h-1)^a = \phi_1^h x_t$
- Les prévisions faîtes en t pour tous les horizons $h > 0$ se calculent à partir des paramètres du modèles et des variables observées jusqu'en t
- Dans le cas d'un AR(1) stationnaire, on a $\mid \phi_1 \mid < 1$. On en déduit que $lim_{h \to +\infty} x_t^a(h) = 0 = E(x_t)$. A mesure que l'horizon de la prévision augmente, la prévision converge vers l'espérance inconditionnelle de x_t.

Prévisions statiques et prévisions dynamiques

- $x_t^a(2) = \phi_1^2 x_t$ est une prévision *dynamique* : on a un date t fixe et l'on fait varier l'horizon de la prévision.
- Les prévisions dynamiques se calculent de façon récursive : pour $h = 1$, puis $h = 2$, puis $h = 3, \ldots$
- Dans le cas de la prévision *statique*, l'horizon de la prévision est toujours $h = 1$. C'est la date t de la prévision qui varie :
  - $x_t^a(1) = \phi_1 x_t$ est la prévision statique de x_{t+1}.
  - $x_{t+1}^a(1) = \phi_1 x_t$ est la prévision statique de x_{t+2}.
- $x_t^a(1)$ est la valeur ajustée du modèle

Prévision à l'horizon *h=1* dans le cas d'un *MA(1)*

- $x_t = \epsilon_t + \theta_1 \epsilon_{t-1}$
- le mode de calcul des prévisions se généralise à un processus MA(q) quelconque
- En *t+1*, on a $x_{t+1} = \epsilon_{t+1} + \theta_1 \epsilon_t$
- En *t*, ϵ_t est connu
- En *t*, ϵ_{t+1} est inconnu et remplacé par son espérance $E(\epsilon_{t+1}) = 0$
- On en déduit la prévision en *t* de x_{t+1} :

$$x_t^a(1) = \theta_1 \epsilon_t$$

Prévision à l'horizon *h=1* dans le cas d'un *MA(1)*

- L'erreur de prévision est :

$$e_t(1) = x_{t+1} - x_t^a(1) = (\epsilon_{t+1} + \theta_1 \epsilon_t) - \theta_1 \epsilon_t = \epsilon_{t+1} \Rightarrow E\left(e_t(1)\right) = 0$$

- L'erreur de prévision est nulle en moyenne, le prédicteur est sans biais.
- La variance de l'erreur de prévision est $V(e_t(1)) = V(\epsilon_{t+1}) = \sigma_\epsilon^2$

Prévision à l'horizon $h=2$

- En $t+2$, on a $x_{t+2} = \epsilon_{t+2} + \theta_1 \epsilon_{t+1}$
- En t, ϵ_{t+1} et ϵ_{t+2} sont inconnus et remplacés par leur espérance $E(\epsilon_{t+1}) = 0$
- On en déduit la prévision en t de x_{t+2} :

$$x_t^a(2) = 0$$

- L'erreur de prévision est : $e_t(2) = x_{t+2} - x_t^a(2) = (\epsilon_{t+2} + \theta_1 \epsilon_{t+1})$.
- L'erreur de prévision est nulle en moyenne, le prédicteur est sans biais.
- La variance de l'erreur de prévision est : $V(e_t(2) = V(\epsilon_{t+2} + \theta_1 \epsilon_{t+1}) = \sigma_\epsilon^2 (1 + \theta_1^2)$
- La variance de l'erreur de prévision s'accroît avec l'horizon de la prévision.

Généralisation à un MA(q)

- On a $x_t^a(h) = 0$ dès que $h > 1$ pour un processus *MA(1)*.
- Cette propriété se généralise à tout processus *MA(q)*. On a $x_t^a(h) = 0$ dès que $h > q$

Prévision ponctuelle l'horizon *h=1* le cas d'un ARMA(p,q)

- Application des principes vus dans le cas d'un AR(p) et d'un MA(q).
- Exemple d'un ARMA(1,1) stationnaire

$$x_t = \phi_1 x_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}$$

- Prévision à l'horizon $h = 1$
- $x_{t+1} = \phi_1 x_t + \epsilon_{t+1} + \theta_1 \epsilon_t \Rightarrow x_t^a(1) = \phi_1 x_t + \theta_1 \epsilon_t$
- L'erreur de prévision à l'horizon *h=1* est donc $) = \epsilon_{t+1}$.
- $E(e_t(1)) = 0$ et le prédicteur $x_t^a(1)$ est sans biais.
- La variance de l'erreur de prévision est $V(e_t(1)) = \sigma_\epsilon^2$

Prévision d'un ARMA(1,1) à l'horizon *h=2*

- $x_{t+2} = \phi_1 x_{t+1} + \epsilon_{t+2} + \theta_1 \epsilon_{t+1}$
- D'où : $x_t^a(2) = \phi_1 x_t^a(1) = \phi_1^2 x_t + \phi_1 \theta_1 \epsilon_t$
- L'erreur de prévision à l'horizon *h=2* est :

$$\begin{aligned} e_t(2) &= x_{t+2} - x_t^a(2) \\ &= (\phi_1 x_{t+1} + \epsilon_{t+2} + \theta_1 \epsilon_{t+1}) - \phi_1 x_t^a(1) \\ &= \phi_1 \left(x_{t+1} - x_t^a(1)\right) + \epsilon_{t+2} + \theta_1 \epsilon_{t+1} \\ &= \phi_1 \epsilon_{t+1} + \epsilon_{t+2} + \theta_1 \epsilon_{t+1} \end{aligned}$$

- L'erreur de prévision est nulle en moyenne.
- La variance de l'erreur de prévision est
  $V(e_t(2)) = \sigma_\epsilon^2 (1 + (\theta_1 + \phi_1)^2)$

Prévision d'un ARMA(1,1)

- $x_t^a(2)$ est une prévision dynamique.
- La prévision statique est $x_{t+1}^a(1) = \phi_1 x_t + \theta_1 \epsilon_t$ dans le cas de x_{t+2}.
- pour $h > 1$, on aura $x_t^a(h) = \phi_1 x_t^a(h-1)$
- pour $h > 1$, la composante *MA* n'apporte plus d'information au calcul de la prévision, dans le cas d'un processus *ARMA(1,1)*
- Se généralise à un ARMA(p,q)

- La prévision est calculée pour la variable stationnarisée
- Recoloration = calcul de la prévision pour la variable initiale non stationnarisé
- Cas d'un processus DS : $x_t = \Delta z_t$
  - La prévision en t de z_{t+1} sera : $z_t^a(1) = z_t + x_t^a(1) = z_t + \Delta_t^a z_t(1)$.
  - Pour les horizons de prévision $h > 1$, on aura $z_t^a(h) = z_t^a(h-1) + x_t^a(h) = z_t + \Delta_t^a z_t(h)$
- Cas d'un processus TS
  - $z_t = a + bt + x_t$ le processus TS
  - Le processus stationnarisé x_t est modélisé à l'aide d'un modèle ARMA.
  - La prévision en t de z_{t+h} sera : $z_t^a(h) = a + b(t+h) + x_t^a(h), \forall\, h > 0$.

- La prévision est un intervalle de confiance
- Prend en compte la précision de la prévision
- Hypothèse sur ϵ_t : $\epsilon_t \sim N(0, \sigma_\epsilon^2)$
- $e_t(1) = \epsilon_{t+1}$. Il en résulte que $e_t(1) \sim N(0, \sigma_\epsilon^2) \Rightarrow \frac{e_t(1)}{\sigma_\epsilon}$
- expression d'un intervalle de confiance à *95%* pour les prévisions à l'horizon *h=1* :

$$
\begin{aligned}
& P \quad \left[-2 \leqslant \frac{e_t(1)}{\sigma_\epsilon} \leqslant 2\right] \\
\Rightarrow \quad & P \quad \left[-2\sigma_\epsilon \leqslant x_{t+1} - x_t^a(1) \leqslant 2\sigma_\epsilon\right] \\
\Rightarrow \quad & P \quad \left[x_t^a(1) - 2\sigma_\epsilon \leqslant x_{t+1} \leqslant x_t^a(1) + 2\sigma_\epsilon\right]
\end{aligned}
$$

- Prévision par intervalle de confiance à l'horizon h
- $e_t(h) = x_{t+h} - x_t^a(h) \sim N(0, \sigma^2(h))$
- Expression d'un intervalle de confiance à 95% pour la prévision est :

$$[x_t^a(h) - 2\sigma(h) \leqslant x_{t+1} \leqslant x_t^a(h) + 2\sigma(h)]$$

- $\sigma(h)$ est l'écart type de l'erreur de prévision à l'horizon h.
- $\sigma(h)$ diffère selon le modèle utilisé pour calculer les prévisions.

- Si l'on veut faire de la prévision dynamique sous Gretl, on doit modifier l'échantillon des observations. On décide de supprimer les 5 dernières observations.

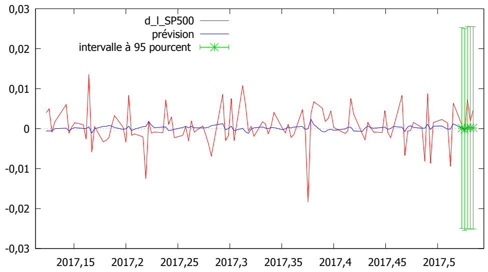


Figure – Représentation des prévisions dynamiques

Prévisions statiques

- On choisit maintenant de faire des prévisions statiques.

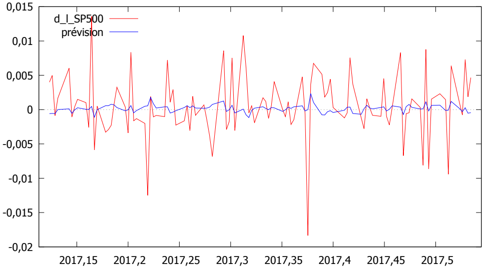


Figure – Représentation graphique des prévisions statiques

- Comme la qualité de la régression n'est pas très grande, les prévisions ne sont pas très proches des valeurs réalisées.

- Possibilité de plusieurs modèles valides pour une seule série
- Utilisation de critères de sélection
- Les critères standards :
- Fondés sur l'erreur de prévision à l'horizon *h=1*.
- Le meilleur modèle est celui qui minimise le critère de sélection choisi.
  - Erreur absolue moyenne : $MAE = \frac{1}{T}\sum_{t=1}^{T} \mid \hat{\epsilon}_t \mid$
  - Erreur quadratique moyenne : $MSE = \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_t^2$
  - RMSE (root mean square error) qui est l'écart type estimé de l'erreur de prévision : $RMSE = \sqrt{MSE}$

- Les critères d'information les plus utilisés
- Le critère d'information de Akaike (1969) :
  $AIC(p,q) = ln\left(\hat{\sigma}^2_{\hat{\epsilon}}\right) + \frac{2(p+q)}{T}$
- Le critère d'information BIC ou de Schwarz (1978) :
  $BIC(p,q) = SC(p,q) = ln\left(\hat{\sigma}^2_{\hat{\epsilon}}\right) + (p+q)\frac{ln(T)}{T}$
- Le critère d'information de Hannan et Quinn (1979) :
  $HQ(p,q) = ln\left(\hat{\sigma}^2_{\hat{\epsilon}}\right) + (p+q)\frac{2ln(ln(T))}{T}$
- Ces critères d'information représentent un arbitrage entre la précision du modèle et le nombre de paramètres à estimer.
- Le meilleur modèle est celui qui minimise le critère d'information retenu.

Propriétés des critères d'information

- quand ils sont appliqués à des *AR(p)* :
- le critère *AIC* surestime l'ordre de l'*AR*
- le critère *HQ* estimer l'ordre de l'*AR* de façon convergente
- le critère *SC* est convergent de façon presque sure
- si $T \geqslant 16$, on a $\hat{p}(SC) \leqslant \hat{p}(HQ) \leqslant \hat{p}(AIC)$