# Validation de modèles ARIMA *

Jean-Marie Dufour †
Université de Montréal

Première version: Septembre 2000
Révisions: Mars 2002
Cette version: 14 mars 2002
Compilé: 14 mars 2002, 11:44am

* Cette recherche a bénéficié du support financier de la Chaire de recherche du Canada en économétrie, du Conseil des Arts du Canada (Bourse Killam), du Conseil de recherche en sciences humaines du Canada, du Conseil de recherche en sciences naturelles et en génie du Canada, du Réseau canadien de centres d'excellence (projet MITACS) et du Fonds FCAR du Québec.

† L'auteur est titulaire de la Chaire de recherche du Canada en économétrie. Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Centre de recherche et développement en économique (C.R.D.E.), et Département de sciences économiques, Université de Montréal. Adresse postale: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre Ville, Montréal, Québec, Canada H3C 3J7. TEL: (514) 343 2400; FAX: (514) 343 5831; courriel: jean.marie.dufour@umontreal.ca. Page Web: http://www.fas.umontreal.ca/SCECO/Dufour .

# Table des matières

**1. Problème** **1**

**2. Corrélogramme d'un bruit blanc** **1**

**3. Corrélogramme des résidus** **2**

**4. Statistique de Ljung-Box** **2**

# 1. Problème

$$X_t \sim ARIMA\,(p,d,q)$$
$$\varphi\,(B)\,\nabla^d X_t = \varphi_0 + \Theta\,(B)\,a_t \ .$$

Après estimation des paramètres on s'attend à ce que les résidus $\hat{a}_t$ soient à peu près un $BB$.

Critère de succès d'un modèle ARIMA : réduction de la série à un $BB$.

Examinons d'abord comment on pourrait tester si $a, \ \ldots \ , \ a_N$ est un bruit blanc.

# 2. Corrélogramme d'un bruit blanc

Soit

$$a_1, ..., a_N \sim BB\left(0, \sigma_a^2\right) \ . \tag{2.1}$$

Alors

$$r_k\,(a) = \sum_{t=1}^{N-k} a_t a_{t+k} \,/\, \sum_{t=1}^{N} a_t^2 \tag{2.2}$$

est un estimateur de $E\,(a_t a_{t+k})\,/E(a_t^2)$ .

Pour $N$ grand,

$$r_k(a) \ \sim \ N\left[0, \frac{1}{N}\right] \tag{2.3}$$

$$\frac{r_k\,(a)}{1/\sqrt{N}} \ \sim \ N\,[0,1] \ . \tag{2.4}$$

De plus, on peut montrer que $r_k\,(a)\,, k = 1, ..., K$, où $K < N$, sont indépendants. D'où :

$$Q\,(r) = \sum_{k=1}^{K}\left[\frac{r_k(a)}{1/\sqrt{N}}\right]^2 = N\sum_{k=1}^{K} r_k\,(a)^2 \sim \chi^2\,(K) \ .$$

On peut donc tester si $a_1, ..., a_N$ est effectivement un $BB$.

## 3. Corrélogramme des résidus

Au lieu de $a_1, ..., a_N$, on possède $\hat{a}_1, ..., \hat{a}_N$ . On veut tester

$$H_0 : \text{le modèle ARIMA } (p, d, q) \text{ est adéquat.} \tag{3.1}$$

Examinons les autocorrélations :

$$r_k(\hat{a}) \ , \quad k = 1, \ldots, K.$$

Pour $N$ grand

$$\sqrt{N}\, r_k(\hat{a}) \overset{a}{\sim} N[0, 1] \ , \quad k = 1, \ldots, K.$$

mais ne sont pas indépendants.

Toutefois, on peut montrer [Box et Pierce (1970)] que

$$\hat{r} \simeq (I - D)\, r$$

où

$$r = \begin{pmatrix} r_1(a) \\ \vdots \\ r_K(a) \end{pmatrix} \ , \quad \hat{r} = \begin{pmatrix} r_1(\hat{a}) \\ \vdots \\ r_K(\hat{a}) \end{pmatrix} \tag{3.2}$$

et

$$I_K - D \text{ est une matrice idempotente de rang } K - \ell,\ \ell = p + q. \tag{3.3}$$

Donc

$$\sqrt{N}\, r \overset{a}{\sim} N_K(0, I_K) \tag{3.4}$$

$$\sqrt{N}\, \hat{r} \simeq (I_K - D)\sqrt{N}\, r \overset{a}{\sim} N_K(0, I_K - D) \tag{3.5}$$

$$Q(\hat{r}) = N \sum_{k=1}^{K} r_k(\hat{a})^2 \sim \chi^2(K - \ell)$$

$p + q$ n'inclut pas la constante.

## 4. Statistique de Ljung-Box

Pour des séries courtes et moyennes, l'approximation de la loi de $Q$ par une distribution $\chi^2(K - \ell)$ est très mauvaise [voir Davies, Triggs et Newbold (1977)]. En particulier,

$$E(Q) < E\left[\chi^2(K - \ell)\right] \ . \tag{4.1}$$

Ljung et Box (1978) ont proposé une modification qui améliore l'approximation. Considérons d'abord le cas d'un bruit blanc :

$$Var\left[r_k\left(a\right)\right] = \frac{N-k}{N\left(N+2\right)}\,, \quad k = 1, 2, \ldots, K\,,$$

$$\frac{r_k\left(a\right)}{\sqrt{\frac{N-k}{N(N+2)}}} \sim N\left(0,\, 1\right) \tag{4.2}$$

où

$$\frac{N-k}{N\left(N+2\right)} < \frac{1}{N}\,. \tag{4.3}$$

$$\tilde{Q}\left(r\right) = \sum_{k=1}^{K}\left[\frac{r_k\left(a\right)}{\sqrt{\frac{N-k}{N(N+2)}}}\right]^2 = N\left(N+2\right)\sum_{k=1}^{K}\frac{r_k\left(a\right)^2}{N-k}\,.$$

Cas des résidus estimés :

$$\tilde{Q}\left(\hat{r}\right) = N\left(N+2\right)\sum_{k=1}^{K}\left(N-k\right)^{-1} r_k\left(\hat{a}\right)^2 \sim \chi^2\left(K-\ell\right).$$

On appelle cette statistique la statistique de Ljung-Box.

## Références


Box, G. E. P. et Pierce, D. A. (1970), 'Distribution of residual autocorrelations in autoregressive-integrated moving average time series models', *Journal of the American Statistical Association* **65**, 1509–1526.

Davies, N., Triggs, C. M. et Newbold, P. (1977), 'Significance levels of the Box-Pierce portmanteau statistic in finite samples', *Biometrika* **64**, 517–522.

Ljung, G. M. et Box, G. E. P. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**, 297–303.