

مقارنة بين المقدرات الاعتيادية والحسينة لنموذج السلسلة الزمنية المختلطة احادية

المتغيرات من الرتب الدنيا

أ. د. عبد المجيد حمزة الناصر*
د. صفاء يونس الصفاوي**

المدخل

ان الحصول على مقدرات ذات كفاءة تعد من اهم مراحل التحليل الاحصائي، وعليه يجب الاهتمام باختبار الطرائق المناسبة للوصول الى ذلك المستوى من التقدير وخصوصا عندما تكون بيانات السلسلة الزمنية تحتوي على بعض البيانات الملوثة التي تكون غير متسقة مع بقية المشاهدات.

وتهدف هذه الدراسة الى اجراء مقارنة بين المقدرات الاعتيادية والمقدرات الحسينية لنموذج المختلطة الاحادية من الرتب الدنيا ARMA(1,1).

وقد تبين بان مقدر الامكان الاعظم كان هو الافضل في حالة عدم وجود الشوارد. اما في حالة افهام الشوارد بنسبة 10%. كانت المقدرات الحسينية هي الافضل وقد تم الاعتماد على معيار متوسط الاخطاء النسبية المطلقة MAPE كمعيار احصائي للمقارنة.

Comparison Between Ordinary Method and Robust Method to estimate the Parameters of the Univariate Mixed Model with Low Order

Abstract

A condense study was done to compare between the ordinary estimators. In particular the maximum likelihood estimator and the robust estimator, to estimate the parameters of the mixed model of order one, namely ARMA(1,1) model.

Simulation study was done for a varieties the model. using: small, moderate and large sample sizes, were some new results were obtained. MAPE was used as a statistical criterion for comparison.

* استاذ/ جامعة بغداد

** استاذ مساعد/ كلية علوم الحاسوب والرياضيات/ جامعة الموصل

1-1 المبحث الأول: المقدمة

حظيت معظم الدراسات والبحوث التي اهتمت بالسلسل الزمنية والخاصة بالنمذج المختلطة من الدرجة الأولى باهتمام واسع. كما ان فقدان المقدرات لخصائصها الجيدة، دفع الكثير من الباحثين لدراسة الشوارد فضلاً عن دراسة طائق التقدير الحصينة التي تعد وسيلة لحل مشكلة الشوارد. حيث اقترح (Fox, 1972) اول نموذج للشوارد اذ حدد نوعين من الشوارد وفي عام 1973 قام بدراسة تأثير الشوارد في مقدرات المربعات الصغرى وقدم طريقته ودرس خصائصها النظرية والتجريبية. وفي عام 1979 قدم Demby and Martin دراستهما المتميزة والتي تعتبر نقطة انطلاق ومرجعاً للعديد من الباحثين الذين درسوا الشوارد في السلسل الزمنية. وتوالت الدراسات مثل Abraham and Chang عام 1989 و Li and Hui عام 1996 و Marnna and Yahai سنة 1996 و Raftery and Martin عام 1997 و Martinez and Mentz عام 2002 والشرجي عام 2003.

يهدف هذا البحث الى اجراء مقارنة كفاءة المقدرات بين الطائق الاعتيادي مثل (مقدرات المربعات الصغرى ومقدرات الامكان الاعظم) والمقدرات الحصينة للنمذج المختلطة الاحادية المتغيرات من الرتب الدنيا ARMA(1,1) في حالة عدم وجود الشوارد وفي حالة وجود الشوارد.

وقد قسم هذا البحث الى ثلاثة مباحث تضمن المبحث الاول المقدمة والهدف من البحث في حين شمل المبحث الثاني على الجانب النظري اما المبحث الثالث فقد انصب على الجانب التجاري.

المبحث الثاني: الجانب النظري

2-1 قيم الشوارد Outliers - Values

لقيت مشكلة الشوارد في البيانات اهتماماً كبيراً في السنوات الأخيرة بسبب ادراك الكثير من الباحثين لخطورة الاساليب التقليدية في تقدير المعلومات عند ظهور هذه المشكلة فضلاً عن ان البيانات لا يمكن ان تخضع بشكل كامل للافتراءات الاعتيادية (كالتوزيع الطبيعي وغيره).

لذا كان لزاماً البحث عن طائق بديلة تكون حصينة أو مقاومة لظهور القيم الشاردة في البيانات او خرق الشروط المطلوبة . وبما أن موضع السلسل الزمنية من المواضيع المهمة في الدراسات الاحصائية، وان استخدام الطائق التقليدية سيكون تأثيره كبيراً في تقدير المعلومات وذلك لأن وجود الشوارد في بعض متغيرات السلسل الزمنية سوف يؤثر في دقة تلك المقدرات وأن كشف الشوارد في البيانات ومعالجتها مثبت مساحة واسعة في البحث الاحصائي بسبب كون الاستدلال الاحصائي المبني على اساس التوزيع الطبيعي حساساً تجاه الشوارد Lye, Jenny N.& Vance, L. Martin, (1993).

وقد عُرف الشوارد العديد من الباحثين منها تعريف Barnett و Lewis ، الا ان جميع هذه التعريفات تتصل في مفهوم واحد كونها المشاهدة غير المتسبة مع بقية المشاهدات، لأن ظهور القيم الشاردة ناشئ عن احد الاسباب الآتية:

1. أخطاء القياس الناتجة من أخطاء التسجيل.
2. أخطاء المعاينة الناتجة من سوء اختيار العينة وعدم تمثيلها للمجتمع تمثيلاً جيداً، أو بسبب التحيز في اختيار العينة.
3. أو تحدث لسبب طبيعي تفرضه حالة معينة.

يوجد اسلوبان لمعالجة مسألة وجود الشوارد في البيانات، يمكن الاسلوب الاول في اكتشاف الشوارد وتصحيفها (تعديلها) ومن ثم اجراء التحليل الاحصائي التقليدي المعتمد. اما الاسلوب الثاني فيجري فيه التعامل مع مجموعة البيانات كما هي دون أي تعديل ولكن يتم التحليل بأجراءات او طرائق حصينة تكون حساسيتها أقل اتجاه الشوارد من الطرائق التقليدية اذ تتم بعض الطرائق فيها من خلال اعطاء اوزان صغيره (غالبا صفر) للمشاهدات الشاردة والبعض الاخر يتم اجراء الحصانة من خلال وضع قيود معينة على معلمتي التموضع والقياس يتم بعدها اكتشاف الشوارد من خلال مسافات مهلونبس الحصينة لبيانات متعدد المتغيرات.

اذ ان هذه المشاهدات يكون لها أعلى المسافات (المسافة التي تزيد عن قيمة القطع Cut-off value تعتبر شاردة).

ففي حالة المتغير الواحد يمكن ملاحظة القيمة الشاردة من خلال ترتيب البيانات ترتيبا تصاعديا او تنازليا حيث تقع في ذيل الترتيب.
اما في حالة متعدد المتغيرات فإن عملية الكشف عن الشوارد تصبح من الصعوبة بشكل يجعل من الضروري اتخاذ صيغة رياضية حصينة للكشف عنها. اذ اننا في هذه الحالة يجب ان تتجاوز اسلوب الفحص البصري والتشخيص الشخصي

2- الكشف عن الشوارد في السلسلة الزمنية:

ان عملية الكشف عن الشوارد وتحديد طبيعتها لها اهمية في عملية معالجة تأثيرات الشوارد السلبية في تقديرات معلمات النماذج المختلفة، وبعد مقياس الامكان من المقاييس المهمة والفعالة في هذا المجال اذ يتم استخدامه في الحالتين الآتتين:

الحالة الاولى: عندما يكون التباين ومعلمات النموذج معروفة.

ان نموذج السلسلة الزمنية في حالة وجود الشوارد النمطية (IO Innovations Outlier) يتم تقدير التأثير (w) وفق الصيغة:

$$Y_t = y_t + (\theta(B)/\phi(B)a(B))w\varepsilon_t^T \quad \dots \quad (1a)$$

حيث

$$a(B) = (1-B)^d$$

وبالاعتماد على حدود (a_t) يمكن اعادة كتابة المعادلة (1-1a) كما ياتي:

$$Y_t = (\theta(B)/\phi(B)a(B)) \{ a_t + w\varepsilon_t^T \} \dots \quad (1b)$$

اما النموذج في حالة وجود الشوارد المضافة (A0 Additive Effects Outliers) كما في المعادلة الآتية :

$$\phi(B) y_t = \theta(B) a_t$$

ويمكن كتابته على حدود (a_t) كما يلي:

$$Y_t = y_t + w\varepsilon_t^T$$

$$y_t = \frac{\theta(B)}{\phi(B)a(B)} a_t + w_t^T$$

كما ان الشكل العام لنموذج الشوارد هو

$$Y_t = \sum W_j V_j(B) a_t^T + y_t$$

اذ ان

$$y_t = \frac{\theta(B)}{\phi(B)a(B)} a_t$$

$$V_j(B) = \theta(B)/\phi(B)a(B)$$

في حالة وجود الشوارد من النوع (IO) عند الزمن ($t=T$)

$V_{j=1}$

في حالة وجود الشوارد من النوع (AO) عند الولمن ($t=T$)

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad \dots \dots \quad (2)$$

حیث ان:

ϕ هي متعدد الحدود في B لمعلمات نموذج الانحدار الذاتي (ϕ_1, \dots, ϕ_p) .

$\theta(B)$ هي متعدد الحدود في B لمعلمات نموذج الاوساط المتحركة $(\theta_1, \dots, \theta_q)$

وأن B هو معامل الارتداد (الازاحة) الخلفي.

وأن a_t هي سلسلة التشويش الابيض بوسط حسابي صفر وتبين ثابت σ_a^2

و عند تعريف y_t بانها عملية عشوائية خطية مستقرة بعدد محدد (t) من المركبات فانه يمكن التعبير عن y_t باتجاه واحد من الاوساط المتحركة (MA) وكما يأتي:

$$y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

$$\Psi_0 = 1$$

$$y_t = \psi(B)a_t \dots \quad (3)$$

وان (B) تتحقق الشرط الآتي :

$$\psi_j \rightarrow 0 \quad \text{as } j \rightarrow \infty$$

و هذا الشرط يعكس مدى تقارب متعدد الحدود (B) ψ الا انه من الصعب بمكان الوصول الى تقدير النموذج بسبب ان (B) ψ تحتوي على عدد لا نهائي من المعلومات ولذلك يمكن استخدام نسب تقريبية وهي (Kaiser , R. & Maravall , A., 2001).

$$\psi(\mathbf{B}) = \theta(\mathbf{B})/\phi(\mathbf{B}) \quad \dots \quad (4)$$

حیث ان

پائی: (B) ϕ متعدد الحدود في B من درجة p و q على التوالي وعليه يمكن كتابة y_t كما

$$\phi(\mathbf{B})\mathbf{y}_t = \theta(\mathbf{B})\mathbf{a}_t \quad , \quad \mathbf{y}_t = [\theta(\mathbf{B})/\phi(\mathbf{B})]\mathbf{a}_t$$

۱۰

$$[I + \phi_1 B + \dots + \phi_p B^p] y_t = [1 + \theta B + \dots + \theta_q B^q] a_t \quad \dots \dots \quad (5)$$

وإذا ضربنا طرفي المعادلة بالمقدار y_{t-k} حيث ان $k > q$ ومن ثم نأخذ التوقع سنحصل على المعادلة الفرقية (Difference Equation) الآتية:

$$\gamma_k + \phi_1\gamma_{k-1} + \dots + \phi_p\gamma_{k-p} = 0 \quad \dots \dots \dots \quad (6)$$

۱۰۷

$$\phi(\mathbf{B})\gamma_k = 0$$

وإن $k > q$ هي دالة لـ k

اما ايجاد قيمة معامل الارتباط الذاتي بمعادلة الفروق المتجانسة (1-6) التي ستكون

$$r^p + \phi_1 r^{p-1} + \dots + \phi_{p-1} r + \phi_p = 0 \quad \dots \dots \dots \quad (7)$$

وإذا كانت (r_1, r_2, \dots, r_p) هي جذور المعادلة $\phi(B)\gamma_k = 0$ فيمكن كتابة γ_k على وفق الصيغة الآتية:

$$\gamma_k = \sum_{i=1}^p r_i^k$$

وهي التي تقارب الى الصفر كلما كانت k كبيرة جدا، أي :

$$\gamma_k \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

او ان

$$|r_i| < 1, i = 1, 2, \dots, p$$

وبالعودة الى المعادلة (4-1) تكون الدالة المولدة للتباين المشترك AGF

$$\gamma(B, F) = \theta(B)\theta(F)/\phi(B)\phi(F) \quad \dots \dots \dots \quad (8)$$

وهي الصيغة التي تحقق شروط الاستقرارية من تماثل وتقارب .

الحالة الثانية : عندما تكون معلمات النموذج والتباين مجهولة:

وهي الحالة التي تلاحظ دائما في التطبيقات العملية، ولكن يمكن الحصول على التباين وكذلك المعلمات من خلال تقديرها سوية مع (W) بالنسبة الى كل نموذج سواء كان (IO) ام Maximum Likelihood (ML) او ذلك باستخدام اسلوب الامكان الأعظم (AO) ومن ثم ايجاد كل من $\hat{\lambda}_{1T}$, $\hat{\lambda}_{2T}$ وكما يأتي (الشرجي، 2003) :

$$\begin{aligned} \hat{\lambda}_{1T} &= \hat{W}_{IO} / \hat{\sigma}_a \\ \hat{\lambda}_{2T} &= \hat{W}_{AO} / \hat{\rho} \hat{\sigma}_a \end{aligned} \quad \dots \dots \dots \quad (9)$$

وان

$$\begin{aligned} \hat{W}_{IO} &= \hat{e}_T \\ \hat{W}_{AO}^2 &= \hat{\rho}^2 (1 - \hat{\Psi}_1 F - \hat{\Psi}_2 F^2 - \dots - \hat{\Psi}_{n-T} F^{n-T}) \hat{e}_T \\ \hat{\rho}^2 &= (1 + \hat{\Psi}_1 + \hat{\Psi}_2 + \dots + \hat{\Psi}_{n-T})^{-1} \end{aligned} \quad \dots \dots \dots \quad (10)$$

وهذه في الحقيقة تكافئ تقييم مقياس نسبة الامكان عندما تكون معلمات النموذج والتباين معلومة.

وللكشف عن (IO) والـ (AO) ذات الواقع المجهولة فإنه يمكن اختبار ذلك عن طريق

$$\hat{\eta}_{IO} = \max |\hat{\lambda}_{1t}| > c, \quad t=1,2,\dots,n$$

$$\hat{\eta}_{AO} = \max |\hat{\lambda}_{2t}| > c, \quad t=1,2,\dots,n \quad \dots \dots \dots \quad (11)$$

حيث ان c عدد ثابت موجب اختياري وللتمييز بين انواع الشوارد (AO) و (IO) يتم استخدام المتباينات الآتية:

اذا كانت $|\hat{\lambda}_{1T}| > |\hat{\lambda}_{2T}|$ فهذا يعني وجود شوارد من النوع (IO).

وهذه كانت $|\hat{\lambda}_{1T}| \leq |\hat{\lambda}_{2T}|$ فهذا يعني وجود شوارد من النوع (AO).

3-2 مقدرات معلمات نموذج الائتمان الذاتي بالطريق العينادي:

2-3-1 طریقة المربعات الصغری Least Squares Method

ستعرض في هذا المبحث الى توضيح الاجراءات الخاصة بتقدير معلمات نموذج الانحدار الذاتي ARMA(p,0) عندما $1 \geq p$ باستخدام طريقة المربعات الصغرى. وهي احدى طرائق التقدير الاعتيادية في الحصول على المقدرات الاولية Initial Estimates اللازمة للبدء باجراءات الطرائق الحصينة. (Harba 1981).

بافتراض السلسلة الزمنية $\{y_t\}$, ($t = 0, 1, \dots, T$) والتي تعطي على وفق نموذج الانحدار الذاتي ARMA (p,0)

$$y_t = \Phi_j y_{t-j} + \varepsilon_t \quad \dots \quad \dots \dots \quad (12) \sum_{i=1}^p$$

حيث ان Φ تمثل معلمات الانحدار الذاتي، ϵ تمثل حد الخطأ العشوائي بوسط صفر وتبالين σ^2 ، وبجعل Φ و ϵ معرفة بصيغة المصفوفات

$$\Phi = (\phi_1, \phi_2, \dots, \phi_p)'$$

لذلك يمكن كتابة نموذج الانحدار الذاتي ARMA(p,0) بالصيغة الآتية:

$$\mathbf{y}_t = \Phi^T \mathbf{Y}_t + \varepsilon_t \quad \dots \quad (13)$$

وبافتراض ان فترة الاشارة Signal Interval هي (T) ، تحتوي على (I) من العينات وهذا يعني بان $I \Delta T = T$ ، حيث ΔT تمثل مجال العينة .
وعليه فان اول عينة في هذه الفترة تكون y_t ، في حين تكون العينة الاخيرة y_{t+r} . وخلال هذه الفترة يتطلب تدبير J_r بحيث ان تكون دالة الكلفة (Cost Function) J_r اقل ما يمكن ، حيث ان

$$J_r = \sum (y_m - \hat{\phi}_r^T Y_m)^2 \quad \dots \quad (14)$$

وعلیه فان $\hat{\Phi}$ يجب ان تتحقق

$$(\partial J_r / \partial \hat{\phi}_r) = 0$$

بیویٹ ان

$$\left[\sum_{m=t}^{t+r} Y_m Y_m^T \right] \hat{\phi}_r = \sum_{m=t}^{t+r} y_m Y_m \quad \dots \quad (15)$$

وِتَّجْدِيدٌ

$$\mathbf{P}^{-1}\mathbf{r} = \sum_{m=t}^{t+r} \mathbf{Y}_m \mathbf{Y}_m^T \quad \dots \quad (16)$$

فإن $r^{-1}P$ تكون قابلة للعكس Invertible فقط إذا كانت $p > r$ والمتمثلة بدرجة النموذج وعليه فإن المعادلة (16) ستكون

$$P_r^{-1} \hat{\phi}_r = \sum_{m=t}^{t+r} y_m Y_m^T \quad \dots \quad (17)$$

$$\hat{\phi}_r = P_r \sum_{m=t}^{t+r} y_m Y_m^T \quad \dots \quad (18)$$

ويلاحظ انه بالرغم من ان (Y_m^T) تكون مصفوفة غير احادية non-singular فان المصفوفة P_r^{-1} تكون مصفوفة غير احادية non-singular وذلك لكون المجموع هو على مدى (m) . هذا ويمكن اعادة كتابة المعادلة (18) بالصيغة الآتية:

$$P_r^{-1} \hat{\phi}_r = \sum_{m=t}^{t+r-1} y_m Y_m + y_r Y_r \quad \dots \quad (19)$$

ومن الصيغة (16) نلاحظ ان:

$$\sum_{m=t}^{t+r-1} y_m Y_m = \left[\sum_{m=t}^{t+r-1} Y_m Y_m^T \right] \hat{\phi}_{r-1} \quad \dots \quad (20)$$

لذلك فان المعادلة (19) ستكون:

$$P_r^{-1} \hat{\phi}_r = \left[\sum_{m=t}^{t+r-1} Y_m Y_m^T \right] + y_r Y_r \quad \dots \quad (21) \quad \hat{\phi}_{r-1}$$

باضافة وطرح $y_r Y_r^T$ من الطرف اليمين من المعادلة (21) فان:

$$\begin{aligned} P_r^{-1} \hat{\phi}_r &= \left[\sum_{m=t}^{t+r-1} Y_m Y_m^T \right] + Y_r (y_r - Y_r^T \hat{\phi}_{r-1}) + Y_r Y_r^T \hat{\phi}_{r-1} \hat{\phi}_{r-1} \\ &= \left[\sum_{m=t}^{t+r} Y_m Y_m^T \right] + Y_r (y_r - Y_r^T \hat{\phi}_{r-1}) \quad \dots \quad (22) \quad \hat{\phi}_{r-1} \end{aligned}$$

وبتعويض قيمة p_r^{-1} واجراء بعض التبسيطات نحصل على مقدار $\hat{\Phi}_r$ بالصيغة الآتية:

$$(23) \quad \hat{\Phi}_r = \hat{\phi}_{r-1} + P_r Y_r (y_r - Y_r^T \hat{\phi}_{r-1})$$

اذا $\hat{\Phi}_r$ يتم الحصول عليها بالتتابع من خلال المقدار السابق $(\hat{\phi}_{r-1})$ ، ومن قياسات (y_r) و (Y_r) التي تشرط بان P_r يتم الحصول عليها بالتتابع ايضاً . وإن المقدار الابتدائي لـ (p) يتم اختياره بطريقة اعتباطية ولكن لمقتضيات إنجاز عملية التقارب Convergence ، بشكل متتابع فإنه من الملائم استخدام المقدار الابتدائي المتمثل بالصورة الآتية :

$$P_0 = 1/\varepsilon \quad , \quad \varepsilon > 0$$

حيث أن (ε) عدد صغير اعتباطي، بحيث أن $(\varepsilon/1)$ تكون بأي موقع بين العدد (10) وأعلى قيمة يمكن أن يحتوي عليها الحاسوب.

طريقة بديلة لايجاد مقدار الامكان الاعظم الناتم لنموذج ARMA(1,0) من الواضح ان نموذج الانحدار الذاتي من الدرجة الاولى ARMA(1,0) يمكن كتابته بالصيغة الآتية:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \quad \dots \quad (24)$$

حيث يتبع الخطأ العشوائي توزيعاً طبيعياً معنى ذلك ان $\{\varepsilon_t\}$ يمكن كتابته بدلالة الاخطاء العشوائية على وفق الصيغة الآتية:

$$y_t = \sum_{t=0}^{\infty} \phi_1^t \varepsilon_{t-i} \quad \dots \quad (25)$$

وعليه يمكن البرهنة على ان التوزيع الحدي للسلسلة الزمنية y_t يتبع توزيعاً طبيعياً على وفق الصيغة الآتية:

$$y_t \sim N(0, \frac{\sigma_\varepsilon^2}{1 - \phi_1^2})$$

وواضح ان $\{y_t\}$ مترابطة فيما بينها ولذلك يشترط اجراء تحويل (Transformation) مناسب لايجاد دالة الامكان وعلى النحو الآتي:
من الواضح ان :

$$\varepsilon_2 = y_2 - \phi_1 y_1$$

$$\varepsilon_3 = y_3 - \phi_1 y_2$$

.

$$\varepsilon_n = y_n - \phi_1 y_{n-1}$$

مستقلة بعضها عن البعض الآخر وتوزع كل منها طبيعياً بوسط حسابي صفر وتبالين ثابت σ^2
دالة الامكان للعينة هي:

$$\ell(y_1, \varepsilon_2, \dots, \varepsilon_n) = \left[\frac{(1-\phi_1^2)}{2\pi\sigma_\varepsilon^2} \right]^{\frac{1}{2}} e^{-\frac{y_1^2}{2\sigma_\varepsilon^2(1-\phi_1^2)}} \left[\frac{1}{2\pi\sigma_\varepsilon^2} \right]^{\frac{n-1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^n \varepsilon_t^2} \quad (26)$$

ومن الواضح ان دالة التحويل العيقوبي (Jacobian) اللازمة للتحويل المحدد لها يساوي واحداً أي
ان $|J|=1$ وهذا يقود الى دالة الامكان التامة (Exact):

$$\ell(y_1, y_2, \dots, y_n) = \left[\frac{1-\phi_1^2}{2\pi\sigma_\varepsilon^2} \right]^{\frac{1}{2}} e^{-\frac{y_1^2}{2\sigma_\varepsilon^2(1-\phi_1^2)}} \left(\frac{1}{2\pi\sigma_\varepsilon^2} \right)^{\frac{n-1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^n (y_t - \phi_1 y_{t-1})^2} \quad (27)$$

وبعد التبسيط نحصل على دالة الامكان

$$\ell(y_1, y_2, \dots, y_n) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} (1-\phi_1^2)^{\frac{1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2} [(1-\phi_1^2)y_1^2 + \sum_{t=2}^n (y_t - \phi_1 y_{t-1})^2]}. \quad (28)$$

وعند اخذ اللوغاريتم الطبيعي لدالة الامكان والتفاضل بالنسبة الى المعلمة ϕ_1 نحصل على المعادلة من الدرجة الثالثة وكما يأتي:

$$\hat{\phi}_1^3 - \frac{n-2}{n-1} \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=3}^n y_{t-1}^2} \hat{\phi}_1^2 - \left[1 + \frac{1}{n-1} \left(1 + \frac{\sum_{t=1}^n y_t^2}{\sum_{t=3}^n y_{t-1}^2} \right) \right] \hat{\phi}_1 + \frac{n}{n-1} \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=3}^n y_{t-1}^2} = 0 \quad \dots \quad (29)$$

وبعد حل المعادلة نحصل على مقدر الامكان الاعظم التام للمعلمة ϕ_1 والتي تحقق الاستقرارية:

$$\hat{\phi}_1 = \frac{n-2}{n-1} \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=3}^n y_{t-1}^2} \dots \quad (30)$$

4 تقديم معلمات النموذج بالطريقة الحصينة:

إن القيم الشاردة للسلسلة الزمنية يمكن أن تؤثر عكسياً في كل من مقدرات المربعات الصغرى (LS) ومعامل مقدرات M لمعلمات الانحدار الذاتي. إن الاهتمام إنصب هنا للحصول على مقدرات حصينة لمعامل الانحدار الذاتي من الدرجة الأولى (x_i). إذ أن المشاهدات هي $y_t = x_t + v_t$ ، ($v_t = 0$) مع Innovation Outliers (IO) Possibly non-Gaussian Additive Effects Outliers (AO) possibly quite large small fraction من الزمن و X_t طبيعية، والتصنيف العام لمقدرات M تكون مفترضة والتي لها خصائص حصينة لمتوسط مربعات الخطأ باتجاه كلا النموذجين (IO) و(AO) بطريقة Gaussian Method.

في هذا المبحث سندرس مشاكل الحصول على مقدرات حصينة للأنحدار الذاتي من الرتب الدنيا ، أي التحويل بالنسبة إلى السلسلة الزمنية للشوارد. فعند إقتراح إجراءات حصينة لتقيير معلمات السلسلة الزمنية، فإنه يقتضي تمييز السلسلة الزمنية الملوثة بالشوارد بنماذج احتمالية مناسبة.

وبسبب صعوبة صياغة النماذج الاحتمالية التامة (Martin 1979) فإنه يبدو إزاماً البدء بالنماذج المولدة لتعيين الشوارد البسيطة، والتي يمكنها تهيئة بيانات حقيقة تحتوي على الشوارد، وقد أثبتت عملياً، أن سلوك الشوارد غالباً ما يتبع أحد الأشكال الآتية:

- أ- السلوك المحتمل الأول لحدوث الشوارد هو أن تكون فرصة حدوثها مرتبطة عادة بالمتبقى من مفردات العينة، بأستثناء حالة Initial Jump التي تعرف بالتغيير الفجائي الابتدائي.
- ب- السلوك المحتمل الثاني، يعرف بشوارد الخطأ الكبير، الذي قد يعود لأسباب مختلفة، أمثل خطأ التسجيل.
- ج- السلوك المحتمل الثالث، يعرف بالشوارد مختلفة الأنواع ذات السلوك اللا متصل بسلوك بقية مفردات العينة. وهذا النوع قد يعود إلى القصور في استخدام وسيلة التسجيل.

ما تقدم، فإن أنواع السلوك المحتملة أعلاه يمكن أن تتحقق بنماذج مناسبة، فالنوع الأول من السلوك يمكن الحصول عليه مع نموذج الشوارد النمطية (Innovation Outlier. IO)، فإذا كانت قيم المشاهدات تساوي قيمة جوهر العملية (Zch, J.E., 1979)، أي

$$y_t = \hat{x}_t \dots \dots \dots \quad (31)$$

وإن توزيع أخطاء (IO) متناظر، ثقيل الأطراف ، عندئذ فإن نموذج الشوارد يسمى بالنموذج النمطي (IO) ، امثال توزيع t أو توزيع طبيعي ملوث آخر (G) حيث أن:

$$G(p, \sigma_1, \sigma_2) = (1-p)N(0, \sigma_1^2) + pN(0, \sigma_2^2)$$

بحيث ان $\sigma_1^2 > \sigma_2^2$ وان قيمة p تكون صغيرة عادة.

وبعبارة أخرى إذا كانت قيم (e_t) الضوضاء الأبيض تحقق شرط (iid) للمتغير العشوائي ذي التوزيع المتماثل (G) بمتوسط صفر ومعلمة القياس (σ) فإن قيم المتغير العشوائي تسمى Innovation. أما بالنسبة الى نوع السلوك الثاني و الثالث، فإن النموذج الملائم ، يعرف بنموذج الشوارد المضافة أو التجمعية (AO) Additive outlier حيث ان:

$$Y_t = X_t + V_t \dots \dots \dots \quad (32)$$

حيث V_t متغير عشوائي توزيعه مستقل عن X_t وتوزيعه الحدي (عندما تكون p صغيرة نسبيا) هو

$$P(V_t = 0) = 1 - \gamma$$

هذا وقد اثبتت التجربة في حقل السلسل الزمنية ، ان مدى γ يتحقق ما بين (0.01, 0.25) كما يمكن ان يكون توزيع V_t طبيعيا مختلطا (Stochinger & Duter, 1987)

$$CND(p, \sigma_3) = (1-\gamma)\delta_0 + \gamma N(p, \sigma_3^2) \dots \dots \dots \quad (33)$$

حيث تشير δ_0 إلى التوزيع المنحل Degenerated الذي تتركز كلته عند مركز الثقل .

ان هذا النوع من الشوارد يمكن حدوثه، إذا اسقطت فرضية الاستقلالية V_t وقد أشار الى هذا النوع من الشوارد لأول مرة (Fox 1972) ، إذ اقترح نوعين من الشوارد، تلك التي تؤثر في المشاهدة فقط عند حدوثها والذي عُرف بعدئذ بالشوارد المتتجدة أو النمطية (النوع الأول) وتلك التي تؤثر في المشاهدات عموما والذي عُرف بالشوارد المضافة أو التجمعية (النوع الثاني)، وعلاوة على ذلك فقد تقدم Fox في السنة نفسها أيضا باقتراحين لتحديد نوع الشوارد، الأول يبني على فكرة الفحص للنموذجين ومن ثم اختيار النموذج الذي تكون المشاهدة الشاردة فيه أكثر تطرفا، والثاني باختيار النموذج عندما تهيا فرصة الكشف عن مدى تأثير المشاهدة الشاردة في المشاهدات اللاحقة لها.

3- صياغة نموذج المحاكاة:

تعد المحاكاة عملية تشبيه او تقليد للواقع الحقيقي، أي ايجاد صورة طبق الاصل من أي نظام او نموذج دون اخذ ذلك النظام او النموذج ذاته ، وخصوصاً ان بعض هذه المشاكل والنظريات الاحصائية يصعب برها فيها رياضياً، مما دفع الباحثين الى ترجمتها على مجتمعات تجريبية ثم سحب عدد من العينات العشوائية منها ليتم التوصل الى الحلول المثلث لمثل هذه المشكلات. لهذا وسعياً لتحقيق الهدف الاساسي لهذا البحث فقد صيغ نموذج المحاكاة على وفق الخطوات التالية:

أولاً: المجموعة الاولى من التجارب لمقارنة الطائق الاعتيادية والحسينة لتقدير معلمات ARMA(1,1) وتتضمن هذه التجارب عدة مراحل لتوليد البيانات، التقدير ثم المقارنة، وقد تمت كتابة البرامج الخاصة بذلك بالاعتماد على برنامج (Minitab) ومن خلال امكاناته باستخدام الـ (Macros). وتوليد البيانات لغرض تقدير معلمات النموذج المختلط من الرتبة الاولى ARMA(1,1) وحسب الحالات الآتية:

1. تحديد عدد المتغيرات في النموذج $m=1$

$$y_t = \phi_1 y_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

حيث

$$\phi_1 = \frac{\rho_2}{\rho_1} = \frac{\sum (y_{t-2} - \bar{y})(y_t - \bar{y})}{\sum (y_{t-1} - \bar{y})(y_t - \bar{y})}$$

$$\theta_1 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

2. تحديد حجم العينة (n) حيث تم اختيار ثلاثة احجام للعينات هي (25 , 50 , 100).
3. تحديد القيم الافتراضية لمعلمات النموذج المختلط من الدرجة الاولى ARMA(1,1) حيث تم تحديد القيم الافتراضية :

$$(\phi_1, \theta_1) = (0.8, -0.8), (0.8, -0.6), (0.8, -0.2)$$

وجميعها تحقق الاستقرارية وقابلية العكس وقد تم اخذ توافق القيم الافتراضية وعددها 36 حالة وتم تقدير معلمات النموذج المختلط على وفق طرائق المربعات الصغرى (LS) والامكان الاعظم (ML) والحسينية (RO) في حالة عدم اقحام الشوارد وكذلك في حالة اقحام الشوارد حيث تم اقحام الشوارد بنسبة 10% وفي جميع الطرائق تم اعتماد متوسط الاخطاء النسبية المطلقة (MAPE) معايير احصائية حيث كانت طريقة الامكان الاعظم هي الافضل قبل اقحام الشوارد اما في حالة اقحام الشوارد فكانت الطريقة الحسينية هي الافضل كما في الجدول (1). كما لوحظ ان متوسط الاخطاء النسبية المطلقة يتناقص مع زيادة حجم العينة وهذا يتطابق تماماً والنظرية الاحصائية.

المبحث الثالث: الاستنتاجات

ما تقدم نستنتج ما ياتي:

- أ. في حالة كون حجم العينة صغيراً فان مقدر الامكان الاعظم باعتماد معيار متوسط الاخطاء النسبية (MAPE) هو المقدر المفضل بنسبة 86% في حالة عدم وجود الشوارد في حين تفوقت الطريقة الحسينية عند وجود الشوارد وبنسبة 92%.
- ب. في حالة كون حجم العينة متوسطاً فان مقدر الامكان الاعظم باعتماد معيار متوسط الاخطاء النسبية (MAPE) هو المقدر المفضل بنسبة 96% في حالة عدم وجود الشوارد في حين تفوقت الطريقة الحسينية عند وجود الشوارد وبنسبة 100%.
- ج. في حالة كون حجم العينة كبيراً فان مقدر الامكان الاعظم باعتماد معيار متوسط الاخطاء النسبية (MAPE) هو المقدر المفضل بنسبة 86% في حالة عدم اقحام الشوارد في حين تفوقت الطريقة الحسينية عند وجود الشوارد وبنسبة 98%.

الجدول (1)

يبين متوسط الاخطاء النسبية المطلقة لمعلمات النموذج المختلط المقدرة ARMA(1,1)
عندما يكون حجم العينة (25)

رقم التشكيلة	طريقة التقدير الافتراضية لمعلمات	قبل اقحام الشوارد				بعد اقحام الشوارد			
		LS	ML	Robust	Best	LS	ML	Robust	Best
1	$\phi = 0.8$ $\theta = -0.8$	0.00463029 0.00283523	0.00439570 0.00105452	0.009413040 0.008713000	ML ML	0.000996376 0.00467967	0.000996527 0.00105258	0.000104288 0.000941054	RO RO
2	$\phi = 0.8$ $\theta = -0.6$	0.00081820 8 0.00231701	0.000825783 0.00111706	0.000930120 0.008833290	LS ML	0.00117350 0.00529653	0.00116627 0.00105018	0.000857707 0.000913140	RO RO
3	$\phi = 0.8$ $\theta = -0.2$	0.00286717 0.00626127	0.00278937 0.00124991	0.009183490 0.005705290	ML ML	0.000913989 0.00772764	0.000917573 0.00197049	0.000117342 0.000731583	RO RO
4	$\phi = 0.8$ $\theta = 0.2$	0.0170527 0.00118536	0.0163838 0.000741829	0.00902436 0.00146248	ML ML	0.000934881 0.0119543	0.000937595 0.00252196	0.000106919 0.00127661	RO RO
5	$\phi = 0.8$ $\theta = 0.6$	0.00472037 0.0136528	0.00456535 0.000880842	0.00869391 0.00115520	ML ML	0.00119371 0.00312255	0.00118564 0.00109207	0.001134403 0.00109452	RO ML
6	$\phi = 0.8$ $\theta = 0.8$	0.00126354 0.00387362	0.00125256 0.000985539	0.00862717 0.00111930	ML ML	0.000640919 0.000146489	0.000655881 0.000411784	0.000140192 0.000106975	RO RO

قبل اقحام الشوارد						بعد اقحام الشوارد				
رقم التشكيلة	طريقة التقدير	القيمة الافتراضية للمعلمات	LS	ML	Robust	Best	LS	ML	Robust	Best
7	$\phi = 0.6$ $\theta = -0.8$	0.00275341 0.00282882	0.00268035 0.00102013	0.00900925 0.00859432	ML ML	0.00150306 0.00264026	0.00148210 0.00370954	0.000933047 0.000938648	RO RO	
8	$\phi = 0.6$ $\theta = -0.6$	0.00377653 0.00635461	0.00366084 0.00102683	0.00896936 0.00810186	ML ML	0.00225181 0.000690708	0.00219965 0.00115875	0.00107317 0.000614427	RO RO	
9	$\phi = 0.6$ $\theta = -0.2$	0.0137672 0.00568743	0.0131519 0.00115125	0.08949400 0.00372315	ML ML	0.00275948 0.00151989	0.00268616 0.00141587	0.000946449 0.000744000	RO RO	
10	$\phi = 0.6$ $\theta = 0.2$	0.00226954 0.00422176	0.00213331 0.00103025	0.00890763 0.00165588	ML ML	0.00322045 0.0133190	0.00312793 0.000827771	0.00124967 0.00126774	RO ML	
11	$\phi = 0.6$ $\theta = 0.6$	0.00196166 0.00081193 5	0.00192159 0.000929183	0.00882144 0.00121548	ML LS	0.000269031 0.000270593	0.000216155 0.000855581	0.000130883 0.000109491	RO RO	
12	$\phi = 0.6$ $\theta = 0.8$	0.00186700 0.00215007	0.00183087 0.000898923	0.00881413 0.00116984	ML ML	0.00118158 0.00527499	0.00117401 0.000953264	0.001160485 0.000106715	RO RO	
13	$\phi = 0.2$ $\theta = -0.8$	0.00430972 0.00196659	0.00417181 0.00103745	0.00712071 0.007095560	ML ML	0.00157419 0.00142212	0.00155026 0.00105878	0.000181885 0.000893671	RO RO	
14	$\phi = 0.2$ $\theta = -0.6$	0.00258095 0.00285102	0.00251508 0.00107400	0.00708650 0.00419368	ML ML	0.001030020 0.00348267	0.00102877 0.000982917	0.00101384 0.000926789	RO RO	
15	$\phi = 0.2$ $\theta = -0.2$	0.00106091 0.00613817	0.00105838 0.00125443	0.00700833 0.00970477	ML ML	0.000774569 0.00821737	0.000783962 0.00119015	0.000556272 0.001172180	RO RO	

قبل اقحام الشوارد							بعد اقحام الشوارد			
رقم التشكيلة	طريقة التقدير	القيم الافتراضية للمعلمات	LS	ML	Robust	Best	LS	ML	Robust	Best
16	$\phi = 0.2$ $\theta = 0.8$	0.00050905 8 0.00016719 4	0.000529514 0.000913850	0.000708238 0.00167160	LS LS		0.000227533 0.000492360	0.000176386 0.000968271	0.000968271 0.000106737	ML RO
17	$\phi = 0.2$ $\theta = 0.6$	0.00344673 0.00131983	0.00334478 0.000865931	0.00705764 0.00137121	ML ML		0.00119343 0.00179691	0.00118537 0.00111105	0.000251296 0.00109330	RO RO
18	$\phi = 0.2$ $\theta = 0.2$	0.00058570 5 0.00437912	0.000602967 0.000687902	0.000707054 0.00219561	LS ML		0.00267751 0.00359475	0.00260761 0.000204914	0.00139605 0.000128262	RO RO
19	$\phi = -0.2$ $\theta = -0.8$	0.0629432 0.00222156	0.0602789 0.00105406	0.12770400 0.00514471	ML ML		0.000758055 0.00309385	0.000768136 0.00113433	0.000131504 0.000969255	RO RO
20	$\phi = -0.2$ $\theta = -0.6$	0.0681697 0.00161950	0.0653710 0.00108685	0.001288211 0.00375883	RO ML		0.00195617 0.000377845 7	0.00191633 0.00107108	0.00290783 0.000999113	ML LS
21	$\phi = -0.2$ $\theta = -0.2$	0.00174646 0.00574273	0.00163202 0.00125361	0.00168048 0.0526373	ML ML		0.00131327 0.00266635	0.00130022 0.00116159	0.001187196 0.000956145	RO RO
22	$\phi = -0.2$ $\theta = 0.8$	0.00061483 1 0.00076137 8	0.000547546 0.00101969	0.00128544 0.000887049	ML LS		0.000619709 0.000583410	0.000552222 0.00111969	0.000522657 0.000102363	RO RO

قبل اقحام الشوارد						بعد اقحام الشوارد			
رقم التشكيلة	طريقة التقدير الافتراضية للمعلمات	LS	ML	Robust	Best	LS	ML	Robust	Best
23	$\phi = -0.2$ $\theta = 0.6$	0.00253578 0.00179237	0.00238846 0.000896017	0.00248506 0.00179682	ML ML	0.00107095 0.00466677	0.00106800 0.000991863	0.000157456 0.000103459	RO RO
24	$\phi = -0.2$ $\theta = 0.2$	0.00534864 0.00008374 1	0.00516744 0.00118440	0.00628463 0.00331487	ML LS	0.00202892 0.00503197	0.00198605 0.000791014	0.00249839 0.000108948	ML RO
25	$\phi = -0.6$ $\theta = -0.8$	0.00194223 0.00224099	0.00190297 0.00101969	0.00198480 0.00928447	ML ML	0.000823676 0.00140632	0.000831023 0.000993557	0.000815367 0.000102104	RO RO
26	$\phi = -0.6$ $\theta = -0.6$	0.00359719 0.00311809	0.00348898 0.00107486	0.01009053 0.300807136	ML ML	0.00103039 0.00171842	0.001025912 0.00108249	0.00012392 0.00103621	RO RO
27	$\phi = -0.6$ $\theta = -0.2$	0.00480416 0.00486864	0.00464566 0.00132751	0.01092360 0.00631003	ML ML	0.000449171 0.00470919	0.000472122 0.00109705	0.000106894 0.00109145	RO RO
28	$\phi = -0.6$ $\theta = 0.8$	0.00141914 0.00067432 1	0.00140168 0.000902693	0.0110455 0.00108514	ML LS	0.00119456 0.00127040	0.00118645 0.00122944	0.00114596 0.000978811	RO RO
29	$\phi = -0.6$ $\theta = 0.6$	0.00151499 0.00042593 1	0.00149353 0.000414401	0.01111102 0.00113350	ML ML	0.00296014 0.00169032	0.00287846 0.004295460	0.001236240 0.000972346	RO RO
30	$\phi = -0.6$ $\theta = 0.2$	0.00256511 0.00617285	0.00249989 0.00108744	0.01096900 0.00141917	ML ML	0.000025188 0.00366989	0.000017528 0.00217112	0.0000086514 0.000909985	RO RO

قبل اقحام الشوارد						بعد اقحام الشوارد			
رقم التشكيلة	طريقة التقدير القيمة الافتراضية للمعلمات	LS	ML	Robust	Best	LS	ML	Robust	Best
31	$\phi = -0.8$ $\theta = -0.8$	0.00463058 0.00361925	0.00447930 0.00127526	0.0103694 0.01009360	ML ML	0.00148161 0.000310973	0.00146154 0.00102984	0.000951735 0.000104207	RO RO
32	$\phi = -0.8$ $\theta = -0.6$	0.00162042 0.00328502	0.00159457 0.00107710	0.01057300 0.009918800	ML ML	0.00199782 0.00251583	0.00195625 0.000737236	0.00103815 0.000105117	RO RO
33	$\phi = -0.8$ $\theta = -0.2$	0.00253717 0.00576393	0.00238979 0.00135396	0.01067480 0.009271850	ML ML	0.0045748 0.0367147	0.00239675 0.00109408	0.000813360 0.00104672	RO RO
34	$\phi = -0.8$ $\theta = 0.8$	0.00218391 0.0043354	0.00213458 0.000942272	0.01102129 0.00102129	ML ML	0.000910076 0.00056345	0.000913823 0.000985332	0.000141907 0.0000967017	RO RO
35	$\phi = -0.8$ $\theta = 0.6$	0.00221704 0.00130092	0.00208300 0.000960749	0.01108811 0.00102727	ML ML	0.000201500 0.00331784	0.000151438 0.000967617	0.000137908 0.000953546	RO RO
36	$\phi = -0.8$ $\theta = 0.2$	0.00022669 1 0.00869704	0.000258913 0.000715501	0.00108140 0.00109839	LS ML	0.00186147 0.00395514	0.00182557 0.00116320	0.00115563 0.000856588	RO RO

في حالة كون حجم العينة صغيراً ($n=25$) فان طريقة الامكان الاعظم تفوقت بنسبة 86% على بقية الطرائق قبل اقحام الشوارد ولكل قيم المعلمات المختارة. في حين تفوقت الطريقة الحصينة في جميع الحالات عند اقحام الشوارد وبنسبة 96% معتمدين على معيار متوسط الاخطاء النسبية المطلقة.

المصادر

1. الشرجي، قاسم عبده علي (2003) "التقديرات الحصينة لنماذج ARMA مع تطبيق عملي" اطروحة دكتوراه في الأحصاء، كلية الادارة و الاقتصاد، الجامعة المستنصرية.
2. الصفاوي، صفاء يونس طليع. (2005). "مقارنة بين المقدرات الاعتيادية وال Hutchinson لنماذج السلاسل الزمنية المختلطة الثنائية من الرتب الدنيا"، اطروحة دكتوراه غيرمنشورة، كلية علوم الحاسوبات والرياضيات، جامعة الموصل.
3. Fox, A.J., (1973), "Outliers In Time Series", J.R. Statistics Soc. B. 34, 350-363.
4. Harba, M.I.A., (1981), "Signal Processing & Digital Computer Techniques Applied To Surface Electromyography Ph.D. Dissertation University Of Bristol.
5. Kaiser, R. & Maravall, A., (2001),"Notes Of Time Series Analysis ARIMA & Signal Extraction", Bonco De Espane – Servicio Estudios.
6. Lye, Jenny N. & Vance, L. Martin, (1993), "Robust Estimation Non-Normalities & Generalized Exponential Distribution", JASA 88, 261-267.
7. Martin , R.D., (1980), "Robust Estimation of Autoregressive Models", Indirection in Time Series, eds, D.R. Brillinger & G. C. Tiao, Hayward, C.A. : Institute of Mathematical Statistics.
8. Stochinger, N. & Duter, R., (1987),"Robust Time Series Analysis A survey", Supplement To The J. Kybernetika.
9. Zch, J.E., (1979),"Efficiency Robustness Of Generalized M-Estimates For Auto regression & Their Use In Determining Outlier", Ph.D. Dissertation Univ. Washington, seattle, U.S.A.