

دراسة تجريبية لاستعمال طرائق حصينة لتقدير الانموذجين

$SARFIMA(0, D, 0)_S$ و $SARFIMA(0, d, 0)(0, D, 0)_S$

م.د. نازك جعفر صادق
كلية الادارة والاقتصاد / جامعة بغداد

صلاح مهدي عزيز
باحث

P: ISSN : 1813-6729

<http://doi.org/10.31272/JAE.43.2020.125.14>

E : ISSN : 2707-1359

مقبول للنشر بتاريخ : 2020/9/1

تاريخ استلام البحث : 2020/7/21

المستخلص

ان وجود القيم الشاذة في اي بيانات يؤثر على تقدير معالم هذه البيانات لذلك وجد العلماء طرائق حصينة للتقدير تستبعد اثر القيم الشاذة على البيانات ومن هذه الطرائق طريقة المربعات الصغرى الموزونة المكررة واستعمال دوال اوزان مختلفة معها ان ما قمنا به في هذه البحث استعمال هذه الطريقة في تقدير معالم الفرق الاعتيادية والموسمية للانموذجين $SARFIMA(0, D, 0)_S$ و $SARFIMA(0, d, 0)(0, D, 0)_S$ مستعملين طريقة (Geweke and Porter-Hudak) المعتمدة على المخطط الدوري (periodogram) وتحليل دالة الطيف ومقاربتها لشكل دالة الانحدار وكذلك استعملنا (periodogram) المكيفة مع نوافذ طيفية مختلفة .

1- المقدمة

في نهاية القرن العشرين درس العديد من الباحثين السلاسل الزمنية بتغيير طويلة المدى (غالبًا ما تسمى خاصية الذاكرة الطويلة) حيث اقترح كل من و (Granger و Joyeux) في عام 1980 و (Hosking) في عامي 1984, 1981 الانموذج المختلط المتكامل كسريا (ARFIMA) لوصف ميزة الذاكرة الطويلة لسلسلة زمنية. تعتبر نماذج (Autoregressive fractionally integrated moving average ARFIMA) امتدادًا لنماذج (Autoregressive integrated moving average ARIMA)، لكن قيمة التكامل فيها تكون كسرية (1).

لقد تم النظر في نموذج المختلط المتكامل كسريا (ARFIMA) مع المكون الموسمي. (Seasonal Autoregressive fractionally integrated moving average SARFIMA) في العديد من الأعمال واستعمل لوصف عدداً كبيراً من الظواهر الدورية العالمية الحقيقية التي تظهر اعتماداً بعيد المدى حيث استعملت من قبل ، بورتر-هوداك (1990)، هاسلير (1994)، ارتيشي (2002)، ارتيشي & روبينسون (2000)، على سبيل المثال (2).

في بعض الحالات تحتوي السلاسل الزمنية على ملاحظات غير عادية تتأثر بالأحداث الخارجية التي يمكن أن تسبب تغييرات في ديناميكيتها. تُعرف هذه المشاهدات المؤقتة أو التغييرات الدائمة بأنها القيم الشاذة (يشار إليها أحياناً المشاهدات المتطرفة أو غير التقليدية)، ووفقاً لطبيعتها، يمكن أن تكون تأثيراتها على الاستدلال كبيرة. على سبيل المثال، يمكن أن يؤدي وجود القيم الشاذة إلى زيادة التباين المقدر للمتغيرات العشوائية، مما يعني انخفاضاً في الارتباطات التقديرية المقدرة وما يترتب على ذلك من فقد للمعلومات حول بنية الارتباط الذاتي (3).



مجلة الادارة والاقتصاد
العدد 125 / ايلول / 2020
الصفحات : 206-216

● مستل من رسالة ماجستير

2- مشكلة البحث

تعد نماذج $SARFIMA$ شائعة الاستعمال في التنبؤ عندما تكون السلسلة الزمنية لها ذاكرة طويلة ودورات موسمية وخاصة بيانات علوم الهيدرولوجيا لكن كثير من السلاسل الزمنية تحتوي على قيم شاذة لأسباب مختلفة تؤثر في دقة تقدير قيم معالم هذه السلسلة وبالتالي تنبؤات الانموذج لذلك يجب استعمال الطرائق الحسنة في التقدير لتجنب تأثير تلك القيم والوصول الى افضل نموذج للتنبؤ.

3- هدف الدراسة

تهدف الدراسة الى تقدير معالم الانموذجين $SARFIMA(0, d, 0)(0, D, 0)_S$ و $SARFIMA(0, D, 0)_S$ وهي جزء من نماذج $SARFIMA$ عن طريق توليد بيانات بالمحاكاة حسب الانموذجين المحددين حيث استعمل اولا طريقة (GPH) وهي اختصار (Geweke and Porter-Hudak) المعتمدة على المخطط الدوري (periodogram) وثانيا تم استعمال المخطط الدوري (periodogram) المعدل بنوافذ طيفية مختلفة وتطبيق طريقة المربعات الصغرى الموزونة المكررة (iteratively reweighted least-squares IRLS) مع استعمال اوزان مختلفة على نماذج الانحدار المختلفة الناتجة تعويض المخطط الدوري والمخطط الدوري المعدل

4- نماذج الدراسة:

1-4 الانموذج $SARFIMA(p, d, q) \times (P, D, Q)_S$ (1)

في حال اظهرت بيانات السلسلة انها تحتوي على ذاكرة طويلة في ارتباطاتها وتعاد دوراتها للفترة الموسمية S . يتحول الانموذج الى موسمي الى متكامل كسريا ويعرف بالصيغة ادناه

$$\phi(B)\Phi(B^S)\nabla^d\nabla_S^D(x_t - \mu) = \theta(B)\Theta(B^S)\varepsilon_t \quad t \in \mathbb{Z} \quad (1)$$

اذ ان

μ - هو متوسط السلسلة

$\{\varepsilon_t\}_{t \in \mathbb{Z}}$ هي الضجة البيضاء (white noise)

S - الفترة الموسمية

$D, d \in (-0.5, 0.5)$ - معالم الفرق

$\mathcal{B}$ - هي الارتداد الخلفي الذي هو $\mathcal{B}^k(X_t) = X_{t-k}$

$\mathcal{B}^{sk}(X_t) = X_{t-sk}$ الفرق الموسمي

وتعطى الفروق الموسمية وفق الصيغ ادناه ∇_S^D

$$\nabla_S^D \equiv (1 - \mathcal{B}^S)^D = \sum_{k \geq 0} \binom{D}{k} (-\mathcal{B}^S)^k \dots \quad (2)$$

$$\binom{D}{K} = \frac{\Gamma(1+D)}{[\Gamma(1+K)\Gamma(1+D-K)]} \dots \quad (3)$$

ومتعددة الحدود $\phi(\cdot), \theta(\cdot), \Phi(\cdot), \Theta(\cdot)$ لديها درجات p, q, P, Q على التوالي، وتعرف بـ

$$\phi(\beta) = \sum_{i=0}^p (-\phi_i) \beta^i, \quad (4)$$

$$\theta(\beta) = \sum_{j=0}^q (-\theta_j) \beta^j \quad (5)$$

$$\Phi(\beta) = \sum_{k=0}^P (-\Phi_k) \beta^k, \quad (6)$$

$$\Theta(\beta) = \sum_{m=0}^Q (-\Theta_m) \beta^m \quad (7)$$

اذ ان $1 \leq i \leq p, \phi_i$

$1 \leq j \leq q, \theta_j$

$1 \leq k \leq P, \Phi_k$

$1 \leq m \leq Q, \Theta_m$ و

و $\phi_0 = \theta_0 = -1 = \Phi_0 = \Theta_0$ ثوابت

- في حالة $P = p = 0 = q = Q$ فإن الانموذج المعطى في الصيغة (1) يصبح $SARFIMA(0, d, 0)(0, D, 0)_S$ ويكون وفق الصيغة

$$(1 - B)^d (1 - B^S)^D (X_t - \mu) = \varepsilon_t \quad (8)$$

- في حالة $d = 0$ و $P = p = 0 = q = Q$ في هذه الحالة تسمى العملية الموسمية الخاصة المتكاملة كسريا ويرمز لها ب $SARFIMA(0, D, 0)_S$ وتكون وفق الصيغة ادناه

$$\nabla_S^D (X_t - \mu) = (1 - \beta^S)^D (X_t - \mu) = \varepsilon_t \quad t \in \mathbb{Z} \dots (9)$$

- عندما $P = 0 = Q, D = 0, S = 1$ فإن عملية $SARFIMA(p, d, q)(P, D, Q)_S$

تصبح مجرد عملية $ARFIMA(p, d, q)$

- عندما $p = 0 = q, D = 0 = d$ فإن عملية $SARFIMA(p, d, q)(P, D, Q)_S$

تخضع الى عملية $SARIMA(p, q)$

2-4 الانموذج $SARFIMA(0, d, 0)(0, D, 0)_S$ (1)

اذ كانت $\{X_t\}_{t \in \mathbb{Z}}$ عملية وفق الانموذج $SARFIMA(0, d, 0)(0, D, 0)_S$ وان الفترة الموسمية $S \in \mathbb{N}$

فإن صيغة الانموذج تكون على النحو الأتي :

$$(1 - B)^d (1 - B^S)^D X_t = \varepsilon_t \quad (10)$$

وان خصائص هذا الانموذج تكون :

- عندما $|D| < 0.5, |d + D| < 0.5$ فإن العملية $\{X_t\}_{t \in \mathbb{Z}}$ تمتلك دالة كثافة الطيف الاتية :

$$f_x(w) = \frac{\sigma_\varepsilon^2}{2\pi} \left[2 \sin\left(\frac{w}{2}\right) \right]^{-2d} \left[2 \sin\left(\frac{Sw}{2}\right) \right]^{-2D}, \quad 0 < w \leq \pi \quad (11)$$

5 عندما $d + D < 0.5, D < 0.5$ فإن العملية $\{X_t\}_{t \in \mathbb{Z}}$ مستقرة

6 عندما $0 < d + D < 0.5$ و $0 < D < 0.5$ فإن العملية $\{X_t\}_{t \in \mathbb{Z}}$ تمتلك خاصية ذاكرة طويلة

7 عندما $-0.5 < d + D < 0$ و $-0.5 < D < 0$ فإن العملية $\{X_t\}_{t \in \mathbb{Z}}$ تمتلك خاصية ذاكرة متوسطة

8 تمتلك العملية (x_t) دالة تباين ذاتي من الرتبة $h, h \in \mathbb{Z}$ تعطى وفق الصيغة

$$\gamma x_h = \begin{cases} \sigma^2 \sum_{k \in \mathbb{Z}} \gamma_Z(sk) \gamma_Y(h-sk), & h=sl, \quad l \in \mathbb{Z} \\ 0 & h = sl + \zeta, \quad \zeta \in A \end{cases} \quad (12)$$

إذ ان $A = \{1, \dots, S-1\}$

وان $Z_{t \dots t \in \mathbb{Z}}$ هي $sarfima(0, D, 0)_S$ التي سيمر ذكرها لاحقا والتي تمتلك دالة تباين ذاتي من الرتبة

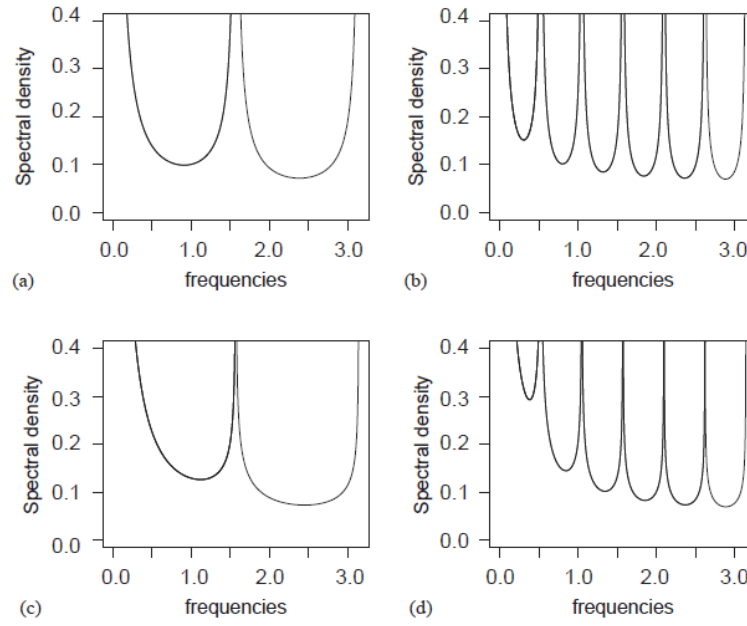
$k, k \in \mathbb{Z}$ وتعطى وفق الصيغة ادناه

$$\gamma_Z(sk - \zeta) = \begin{cases} \frac{(-1)^k \Gamma(1 - 2D)}{\Gamma(K - D + 1) \Gamma(1 - K - D)} = \gamma_X(k), & \zeta = 0 \\ 0 & \zeta \in A \end{cases} \quad (13)$$

وان y_t تتبع الانموذج $ARFIMA(0, d, 0)$ مع دالة تباين ذاتي تعطى وفق الصيغة ادناه

$$\gamma_Y(k) = \frac{(-1)^k \Gamma(1 - 2d)}{\Gamma(K - d + 1) \Gamma(1 - K - d)} \quad (14)$$

ويمكن تمثيل دالة كثافة الطيف للانموذج $SARFIMA(0, d, 0)(0, D, 0)_S$ كما في الشكل



شكل يبين رسم دالة كثافة الطيف للأنموذج $SARFIMA(0, d, 0)(0, D, 0)_S$

وان تفاصيل الرسومات اعلاه كانت كما يأتي :

a : عندما تكون الـ $d = 0.2, D = 0.4, S = 4$

b : عندما تكون الـ $d = 0.2, D = 0.4, S = 12$

c : عندما تكون الـ $d = 0.4, D = 0.2, S = 4$

d : عندما تكون الـ $d = 0.4, D = 0.2, S = 12$

3-4 الأنموذج $SARFIMA(0, D, 0)_S$ (4)

اذ كانت $\{X_t\}_{t \in \mathbb{Z}}$ عملية وفق الأنموذج $SARFIMA(0, D, 0)_S$ وان الفترة الموسمية $S \in \mathbb{N}$ فإن صيغة الأنموذج تكون على النحو الآتي :

$$(1 - B^S)^D X_t = \varepsilon_t \quad (15)$$

وان دالة التباين الذاتي والارتباط الذاتي تعطى وفق الصيغ

$$\gamma(sk) = \frac{(-1)^k \Gamma(1 - 2D)}{\Gamma(K - D + 1) \Gamma(1 - K - D)} \sigma_e^2, k = 1, 2, \dots \quad (16)$$

$$\rho(sk) = \frac{\Gamma(1 - D) \Gamma(k + D)}{\Gamma(D) \Gamma(K - D + 1)}, k = 1, 2, \dots, \quad (17)$$

وعندما

$$k \rightarrow \infty \quad \rho(sk) \sim \frac{\Gamma(1 - D)}{\Gamma(D)} k^{2D-1}$$

وان دالة كثافة الطيف للأنموذج تعطى وفق للصيغة

$$f_x(w) = \frac{\sigma_e}{2\pi} \left[2 \sin\left(\frac{Sw}{2}\right) \right]^{-2D}, 0 \leq w \leq \pi \quad (18)$$

وأن دالة الكثافة الطيفية غير محدودة عند الترددات

$$\frac{2\pi v}{s}, v = 1, 2, \dots, \left\lfloor \frac{s}{2} \right\rfloor$$

5- طرائق التقدير 1-5 مقدر GPH (2)

في هذا القسم ، نقدم المقدر الذي اقترحه (Geweke and Porter-Hudak(1983))، والذي يعتمد على طريقة الانحدار باستعمال دالة المخطط الدوري . افترض $\{X_t\} t \in \mathbb{Z}$ عبارة عن انموذج $SARFIMA(0, d, 0) \times (0, D, 0)_S$ و $D, d \in (0.5, -0.5)$ ؛ ودالة الكثافة الطيفية لها يعبر عنها بـ

$$f_x(w) = \frac{\sigma_\varepsilon^2}{2\pi} \left[2 \sin\left(\frac{w}{2}\right) \right]^{-2d} \left[2 \sin\left(\frac{Sw}{2}\right) \right]^{-2D}, 0 < w \leq \pi \quad (19)$$

يمكننا إعادة كتابة الصيغة اعلاه على النحو التالي

$$f_x(w) = f_v(w) \left[2 \sin\left(\frac{w}{2}\right) \right]^{-2d} \left[2 \sin\left(\frac{Sw}{2}\right) \right]^{-2D} \quad (20)$$

و

$$f_v(w) = \frac{\sigma_\varepsilon^2}{2\pi}$$

وان w تقع ضمن الفترة $0 < w \leq \pi$ وباخذ اللوغاريتم على طرفي المعادلة (2)

$$\ln|f_x(w)| = \ln|f_v(w)| - d \ln \left[2 \sin\left(\frac{w}{2}\right) \right]^2 - D \ln \left[2 \sin\left(\frac{Sw}{2}\right) \right]^2 \quad (21)$$

وإضافة $\ln|f_v(0)|$ و $\ln|I(w)|$ لكلا طرفي المعادلة (3) ، حيث $I(w)$ هي دالة الرسم البياني يعبر عنها بـ

$$I(w) = \frac{1}{2\pi n} \left| \sum_{t=0}^{n-1} X_t e^{-itw} \right|^2 \quad (22)$$

والاستفادة من خصائص اللوغاريتم ، نحصل عليها

$$\begin{aligned} \ln|I(w)| &= \ln|f_v(0)| \\ &\quad - d \ln \left| 2 \sin\left(\frac{w}{2}\right) \right|^2 \\ &\quad - D \ln \left| 2 \sin\left(\frac{Sw}{2}\right) \right|^2 + \ln \left[\frac{f_v(w)}{f_v(0)} \right] + \ln \left[\frac{I(w)}{f_x(w_j)} \right] \dots \end{aligned} \quad (23)$$

استبدال w في المعادلة (23) ، بترددات فوريير $w_j = \frac{2\pi j}{n}$ ، و $j \in \{0, 1, \dots, g(n)\}$ وان $g(n)$ يمثل عرض الحزمة

$$\begin{aligned} \ln|I(w_j)| &= \ln|f_v(0)| \\ &\quad - d \ln \left| 2 \sin\left(\frac{w_j}{2}\right) \right|^2 \\ &\quad - D \ln \left| 2 \sin\left(\frac{Sw_j}{2}\right) \right|^2 + \ln \left[\frac{f_v(w_j)}{f_v(0)} \right] + \ln \left[\frac{I(w_j)}{f_x(w_j)} \right] \dots \end{aligned} \quad (24)$$

يمكننا أن نرى أن المعادلة (24) هي شكل تقريبي لمعادلة الانحدار المتعدد كما في الصيغة ادناه

$$y_j = B_0 + B_1 X_{j1} + B_2 X_{j2} + e_j \quad (25)$$

لكل $j = 1, 2, \dots, g(n)$ حيث

$$y_j = \ln|I(w_j)| \quad (26)$$

$$X_{j1} = \ln \left[2 \sin\left(\frac{w_j}{2}\right) \right]^2 \quad (27)$$

$$X_{j2} = \ln \left[2 \sin\left(\frac{Sw_j}{2}\right) \right]^2 \quad (28)$$

$$e_j = \ln \left[\frac{f_v(w_j)}{f_v(0)} \right] + \ln \left[\frac{I(w_j)}{f_x(w_j)} \right] \quad (29)$$

$$\begin{aligned} B_0 &= \ln|f_v(0)| \\ B_1 &= -d, \quad B_2 = -D \end{aligned} \quad (30)$$

2-5 المخطط الدوري المكيف Modified Periodogram

وفي الطرائق الاخر التي استعملت في المصدر (2) استعمل دالة Modified Periodogram مضروبة بنوافذ طيفية والدالة المعدلة تكون وفق الصيغة ادناه

$$I(w) = \frac{1}{2\pi n} \left| \sum_{t=0}^{n-1} h_t X_t e^{-itw} \right|^2 \quad (31)$$

$$-\pi < w \leq \pi$$

وان h_t تمثل نافذة طيفية تستعمل لتقليل التباين الحاصل. وان هناك عدة نوافذ طيفية وسوف نستعمل اربع نوافذ منها

1-2-5 النوافذ الطيفية

a- نافذة PARZEN⁽⁶⁾ ويعطى وفق الصيغة ادناه

$$\Lambda(U) = \begin{cases} 1 - 6U^2 + 6|U|^3 & |U| \leq 1/2 \\ 2(1 - |U|)^3 & -1/2 < U \leq 1 \\ 0 & |U| > 1 \end{cases} \quad (32)$$

اذ ان $U = \frac{k}{m}$ وان $0 < \beta < 1$ $m = n^\beta$

b- نافذة Bartlett⁽⁶⁾

$$w_{bartlett} = \begin{cases} \frac{2n}{N} & 0 \leq n \leq \frac{N}{2} \\ 2 - \frac{2n}{N} & \frac{N}{2} \leq n \leq N \end{cases} \quad (33)$$

c- نافذة Gaussian⁽⁶⁾

$$w_{Gaussian} = \exp\left[-\frac{1}{2} \left(\alpha \frac{n}{N}\right)^2\right] \quad \alpha = (2.5, 3, 3.5) \quad (34)$$

d- نافذة Hann⁽⁶⁾

$$w_{Hann} = \begin{cases} \frac{1}{2} \left[1 + \cos\left(\frac{2\pi\tau}{N}\right)\right] & -\frac{N}{2} \leq \tau \leq \frac{N}{2} \\ \frac{1}{2} \left[1 - \cos\left(\frac{2\pi\tau}{N}\right)\right] & 0 \leq \tau \leq N \end{cases} \quad (35)$$

3-5 طريقة المربعات الصغرى الموزونة المكررة (iteratively reweighted least-squares IRLS)

بعدها نطبق طريقة المربعات الصغرى الموزونة المكررة⁽⁸⁾ (iteratively reweighted least-squares IRLS) وهي طريقة حسنة لتقدير معالم الانحدار نستعملها لتقدير معالم النماذج الناتجة من الفرات 5-1 و 5-2

وتفاصيل الطريقة كالآتي

الصيغة العامة لطريقة التقدير M تعتمد على تقليل دالة الهدف

$$\sum_{i=1}^n a(e_i) = \sum_{i=1}^n a(y_i - x_i \hat{b}) \quad (36)$$

حيث تمثل الدالة ρ مساهمة كل المتبقي في الوظيفة الهدف. و تمتلك الخصائص التالية :

- $a(e) \geq 0$
- $a(0) = 0$
- $a(e) = \rho(-e)$

لكل $a(e_i) \geq a(e'_i) \dots \dots |a(e_i)| > |a(e'_i)|$ ■

افترض ان $\hat{a} = \psi$ لتمييز دالة الهدف فيما يتعلق بالمعاملات b ، وتحديد المشتقات الجزئية إلى 0 ، ينتج نظام من المعادلات التقديرية ل $k + 1$ من المعاملات:

$$\sum_{i=1}^n \psi(y_i - x_i \hat{b}) \hat{x}_i = 0 \quad (37)$$

وتعرف دالة الوزن $w(e)$ كالآتي

$$w(e) = \frac{\psi(e)}{e} \quad (38)$$

وافترض $w_i = w(e_i)$ ويمكن اعادة كتابة دالة الهدف بالصيغة

$$\sum_{i=1}^n w_i (y_i - x_i \hat{b}) \hat{x}_i = 0 \quad (39)$$

هدف حل المعادلات المقدره بطريقة المربعات الصغرى المرجحة ، هو تقليل $\sum w_i^2 e_i^2$ ومع ذلك ، تعتمد الأوزان على لبواقي (residuals) ، وتعتمد البواقي على المعاملات المقدره ، والمعاملات المقدره تعتمد على الأوزان. والحل تكراري تدعى المربعات الصغرى المعاد وزنها ، لذلك مطلوب:

- 1- تحديد التقديرات الأولية b_0 ، مثل تقديرات المربعات الصغرى.
- 2- لكل تكرار t يحسب البواقي $e_i^{(t-1)}$ والاوزان المرتبطة $w_i^{(t-1)} = w[e_i^{(t-1)}]$ من التكرار السابق
- 3- حل معادلة المربعات الصغرى الموزونة

$$b^{(t)} = [X'W^{(t-1)}X]^{-1}X'W^{(t-1)}y \quad (40)$$

حيث تمثل $W^{(t-1)} = \{w_i^{(t-1)}\}$ مصفوفة قطرية للاوزان الحالية

1-3-5 الاوزان المستعملة في طريقة التقدير الحصينة (5)

اسم الدالة	دالة الهدف	الوزن	ثابت الضبط K
huber	$\rho_H(e) = \begin{cases} \frac{1}{2}e^2 & e \leq k \\ k e - \frac{1}{2}e^2 & e > k \end{cases}$	$w_H(e) = \begin{cases} 1 & e \leq k \\ \frac{k}{ e } & e > k \end{cases}$	1.345
bisquare	$\rho_b(e) = \begin{cases} \left(\frac{1}{2}e^2\right)\{1 - [1 - (\frac{e}{2})^2]^3\} & e \leq k \\ \frac{1}{6}k^2 & e > k \end{cases}$	$w_b(e) = \begin{cases} \left[1 - \left(\frac{e}{2}\right)^2\right]^2 & e \leq k \\ 0 & e > k \end{cases}$	4.685
cauchy	$\left(\frac{k^2}{2}\right) \log[1 + \left(\frac{e}{k}\right)^2]$	$[1 + \left(\frac{e}{k}\right)^2]^{-1}$	2.385
fair	$k^2 \left[\frac{ e }{k - \log\left(1 + \frac{ e }{k}\right)} \right]$	$(1 + e /k)^{-1}$	1.400
talwar	$\rho_T(e) = \begin{cases} \frac{1}{2}e^2 & e \leq k \\ \frac{k^2}{2} & e > k \end{cases}$	$w_T(e) = \begin{cases} 1 & e \leq k \\ 0 & e > k \end{cases}$	2.795
welsch	$\left(\frac{k^2}{2}\right) [1 - \exp[1 - \left(\frac{e}{k}\right)^2]]$	$\exp[-\left(\frac{e}{k}\right)^2]$	2.985

6- تجارب المحاكاة

تم توليد بيانات عينات باحجام (150,300,450) لغرض المقارنة بين الطرائق المستعملة للتقدير وكررت التجارب (1000) مرة لكل عينة وقد استعمل المعيار جذر متوسط مربع الخطأ (root-mean-square error RMSE) للمقارنة بين مخرجات طرائق التقدير

1-6 التجربة الاولى

اعتمد فيها على انموذج التكامل الكسري الموسمي $SARFIMA(0, d, 0)(0, D, 0)_S$ والمعرف في الصيغة (10) حيث تم توليد البيانات متجه الاخطاء العشوائية (ϵ_t) والذي يتبع التوزيع الطبيعي القياسي $N(0, I_n)$ والذي يسمى ايضا (white noise) عن طريق دالة NORMRND في برنامج ماتلاب مع تلويت للمتجه بنسب مختلفة (10%، 20%، 30%) وتم تعويضه في الانموذج اعلاه بعد فك الانموذج باستعمال fractional binomial theorem وتعويض قيم اولية لمعالم الفرق (d, D) تقع ضمن منطقة الاستقرار حيث تم تعويض القيم التالية (0.2,0,1) ، (0.25,0,2) ، (0.15,0,3) وتم تكرار التجربة (1000) مرة مع كل نسبة تلويت ولكل حجم من الاحجام وكانت النتائج كما في الجدول رقم (1)

جدول (1)

يبين اقل قيمة (RMSE) حسب نسبة التلويت للاحجام والقيم الابتدائية المختلفة

للانموذج $SARFIMA(0, d, 0)(0, D, 0)_S$

اوزان طريقة التقدير الحسنة	نماذج الانحدار مع النوافذ الطيفية					
		PERIODOGRAM	Gausswin	hann	parzen	bartlett
	تلويث بنسبة 10%					
	bisquare					
	cauchy					
	faire					
	huber		2	1		1
	talwar		2	1		2
	welsch					
	تلويث بنسبة 20%					
	bisquare					
	cauchy					
	faire					
	huber				5	
	talwar				4	
	welsch					
	تلويث بنسبة 30%					
	bisquare					
	cauchy					
	faire					
	huber				5	
	talwar				4	
	welsch					

حيث يتبين حسب نسبة التلويت الاتي

- نسبة التلويت (10%) اظهرت اقل قيمة لـ RMSE بتكرار (2) عند استعمال (PERIODOGRAM) المعدلة ب النافذه (Gausswin) في طريقة المربعات الصغرى الموزونة المكررة عند الوزن (Huber) والوزن (talwar) ومع النافذه (Bartlett) عند الوزن (talwar) وتكرار (1) في المضروب بالنافذه (hann) عند الاوزان (Huber) و (talwar) وتكرار (1) عند النافذه (Bartlett) مع الوزن (Huber).

- نسبة تلويت (20%) اظهرت اقل قيمة لـ RMSE بتكرار (5) عند استعمال (PERIODOGRAM) المعدلة ب النافذه (Parzen) في طريقة المربعات الصغرى الموزونة المكررة عند الوزن (Huber) وب تكرار (4) عند الوزن (talwar).

- نسبة تلوث (30%) اظهرت اقل قيمة RMSE بتكرار (5) عند استعمال (PERIODOGRAM) المعدلة ب النافذة (Parzen) في طريقة المربعات الصغر الموزونة المكررة عند الوزن (Huber) وبتكرار (4) عند الوزن (talwar).

2-6 التجربة الثانية

اعتمد فيها على انموذج التكامل الكسري الموسمي $SARFIMA(0, D, 0)_S$ وفق الصيغة (15) وايضا تم توليد متجه الخطأ العشوائي كما في التجربة الاول وتعويضه في الانموذج المعرف بالصيغة (15) وتعويض قيم اولية مختلفة لمعلمة الفرق الموسمية (D) (0.2, 0.1, 0.3) وتم تكرار التجربة (1000) مره وكانت النتائج وفق الجدول رقم (2)

جدول (2)

يبين اقل قيمة (RMSE) حسب نسبة التلوث للاحجام والقيم الابتدائية المختلفة للانموذج $SARFIMA(0, D, 0)_S$

اوزان طريقة التقدير الحسنة	نماذج الانحدار مع النوافذ الطيفية					
		PERIODOGRAM	Gausswin	hann	parzen	bartlett
		D	D	D	D	D
	تلويث بنسبة 10%					
	bisquare					
	cauchy					
	faire	7			1	1
	huber					
	talwar					
	welsch					
	تلويث بنسبة 20%					
	bisquare					
	cauchy			1		
	faire	6				2
	huber					
	talwar					
	welsch					
	تلويث بنسبة 30%					
	bisquare					
	cauchy					
	faire	1		1	3	4
	huber					
	talwar					
	welsch					

حيث يتبين حسب نسبة التلوث الاتي

- نسبة التلوث (10%) اظهرت اقل قيمة ل RMSE بتكرار (7) عند استعمال (PERIODOGRAM) في طريقة المربعات الصغر الموزونة المكررة عند الوزن (faire) وبتكرار (1) مع النافذة (Bartlett) والنافذة (Parzen) عند الوزن (faire) ايضا
- نسبة تلوث (20%) اظهرت اقل قيمة RMSE بتكرار (6) عند استعمال (PERIODOGRAM) في طريقة المربعات الصغر الموزونة المكررة عند الوزن (faire) وبتكرار (2) عند التعديل بالنافذة الطيفية (bartlett) وبتكرار (1) عند التعديل بالنافذة (hann) واستعمال الوزن (cauchy)
- نسبة تلوث (30%) اظهرت اقل قيمة RMSE بتكرار (1) عند استعمال (PERIODOGRAM) وعند التعديل بالنافذة (hann) وبتكرار (3) عند التعديل بالنافذة (parzen) وبتكرار (4) عند التعديل بالنافذة (Bartlett) في طريقة المربعات الصغر الموزونة المكررة عند الوزن (faire)

7- الاستنتاجات والتوصيات

من خلال التحليل الاحصائي لنتائج المحاكاة هذه الدراسة ظهرت لنا مجموعة من الاستنتاجات والتوصيات فيما يخص نسبة التلوث وتأثيرها على طرائق التقدير في كلا الانموذجين المستعملين وكما مبين ادناه

7-1 الاستنتاجات :

لوحظ بشكل عام ان نسبة التلوث في البيانات لها تأثير على الطريقة المستعملة في التقدير مقرونة بالنافذة الطيفية المستعملة في التصحيح .

فيما يخص الانموذج $SARFIMA(0,d,0)(0,D,0)_S$

يلاحظ عندما تكون نسبة التلوث (10%) هناك عدة طرائق تحمل اقل اخطاء يمكن استعمالها في تقدير في حالة نسبة التلوث (20%) و (30%) ان افضل طريقة في التقدير هي استعمال الوزن (Huber) ويليه بالدرجة الثانية الوزن (talwar) في حالة استعمال المخطط الدوري المعدل بالنافذة الطيفية (parzen)

فيما يخص الانموذج $SARFIMA(0,D,0)_S$

يلاحظ في حالة نسبة التلوث (10%) و (20%) ان افضل طريقة في التقدير هي استعمال الوزن (FAIRE) وتطبيقها في حالة استعمال المخطط الدوري (PERIODOGRAM)

في حالة نسبة التلوث (30%) ان افضل طريقة في التقدير هي استعمال الوزن (FAIRE) في حالة استعمال المخطط الدوري المعدل بالنافذة الطيفية (BARTLETT) في الدرجة الاولى ويليه المعدل بالنافذة الطيفية (PARZEN)

عدم وجود تأثير ملموس في حالة استعمال قيم اولية مختلفة لمعالم الفرق ضمن الفترة المحددة للمواصفات هناك تأثير لحجم العينة على طريقة التقدير مقرونا بنسب تلوث

7-2 التوصيات :

- التوسع في اجراء الدراسة لتطبيق الطرائق الحسينة للتقدير على نموذج موسمي مختلط متكامل كسريا اي هناك قيم p, q و P, Q
- اجراء دراسة في حالة وجود اكثر من فترة موسمية لنفس السلسلة
- اجراء دراسة في حال كان الانموذج مضاعف جمعي وليس ضربي

8- المصادر

- 1- Bisognin, C, and Lopes, S.R.C, Properties of seasonal long memory processes, Mathematical and Computer Modelling ,49, 1837_1851, 2009
- 2- Bisognin, Cleber ,ESTIMACAO E PREVISAO EM PROCESSOS SARFIMA (p, d, q) (P,D,Q)s NA PRESENCA DE OUTLIER UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL INSTITUTO DE MATEMATICA PROGRAMA DE POS-GRADUACAO EM MATEMATICA.2007
- 3- Mostafaei, Hamidreza, Using SARFIMA Model to Study and Predict the Iran's Oil Supply, International Journal of Energy Economics and Policy, Vol. 2, No. 1, , pp.41-49, 2012,.
- 4- Erol Egrioglu, et.al, The CSS and The Two-Staged Methods for Parameter Estimation in SARFIMA Models Journal of Probability and Statistics, Volume 2011, Article ID 691058, 11 pages
- 5- Holland Paul W. Robust regression using iteratively reweighted least-squares E. Welsch (1977) Robust regression using iteratively ,reweighted least-squares, Communications in Statistics - Theory and Methods, 6:9, 813-827,
- 6- FREDRIC J. HARRIS On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform PROCEEDINGS OF THE IEEE, VOL. 66, NO. 1, JANUARY 1978
- 7- AL.Astal, Jannat Abdallah, Comparison of Methods for Detecting Outliers in Medical Data, Al-Azhar University –Gaza,P.12-14 2018
- 8- Rasheed, Bello Abdulkadir et.al, Robust Weighted Least Squares Estimation of Regression Parameter in the Presence of Outliers and Heteroscedastic Errors, Jurnal Teknologi p 13-14 ,2014

An experimental study using invulnerable methods to estimate $SARFIMA(0, d, 0)(0, D, 0)_S$ and $SARFIMA(0, D, 0)_S$.

Salah Mahdi Aziz, a Researcher

L. Dr. Nazek Jafar Sadiq

Abstract

The presence of outliers in any data affects the estimation of the parameters of these data. Therefore, scientists have found invulnerable methods of estimation that exclude the effect of outliers on the data. Among these methods is the repeated weighted least squares method and the use of different weight functions with it. What we have done in this research is the use of this method in estimating the normal and seasonal difference parameters of the two models $SARFIMA(0, d, 0)(0, D, 0)_S$ and $SARFIMA(0, D, 0)_S$ using the (Geweke and Porter-Hudak) method based on the periodogram and spectrum function analysis. It approximates the shape of the regression function, as well as using a periodogram adapted with different spectral windows.

.....
.....
.....