مقدرات R-في نماذج ARMA

محمد حبيب الشاروط •

أميرة جابر محيسن ••

1-المقدمة /

مقدرات -R هي احدى طرق التقدير الإحصائية التي تعتمد على الرتب ، وقد استخدمت هذه الطريقة اول مره عام (1963) من قبل (Hodges and lehman)[3] اللذين اقترحا مقدر –R الاول للموقع ، وقد طورت طريقتهما لاحقاً لتشمل كافة النماذج الخطية بوجود اخطاء مستعملة ، ولوحظ ان استخدام مقدرات –R في تحليل السلاسل الزمنية قليل جداً ، حيث ان اول محاولة ناجحة فبي هذا المجال كانت من قبل (ALLaL) في عام (1991)[2] الذي استخدم مقدرات – R لتقدير معالم نموذج (1) AR بالاعتماد على سلسلة خطية من الرتب الاحصائية وضعت من قبل (Hallin) عام (1985) و (Hallin and Puri)[5] عام (1988) وقد تم تطويرها فيما بعد لتشمل نماذج ذات مراتب أعلى (P) AR من قبل (Koul and Saleh) عام (1993)[6] والاحصاءات التي أستندت عليها طريقتهما لم تكن رتب فقط وانما استخدمت ايضاً الرتب للبواقي والمشاهدات الاصلية نفسها معاً ، وفي هذه الدراسة تم استخدام مقدرات – R في نماذج (p,q) ARMA وفق الصيغة :-

$$X_t - A_1X_{t-1} - \ldots - A_pX_{t-p} = E_t + B_1E_{t-1} + \ldots + B_qE_{t-q}$$

حيث ان :-

$\{E_t , t\in z\}$هو (white noise) مع دالة كثافة غير محدودة هي (g) .

• استاذ مساعد قسم الاحصاء جامعة القادسية

•• مدرس قسم الرياضيات كلية العلوم جامعة القادسية

المحور الإحصائي ـــــــ

فاذا افترضنا ان المتجه $(X_0, X-p+1, \ldots, \varepsilon_0, \varepsilon-q+1, \ldots)$ موجود فأنه تحت الفـرض $H_g^{(n)}(\vartheta_0)$ فان $\{Z_1^{(n)}(\vartheta_0), \ldots, Z_n^{(n)}(\vartheta_0)\}$ تسمى (white noise) مع دالة كثافـة احتماليـة (g) ، وتعتبر سلسلة الرتب الاحصائية من الرتبة (K) حيث (k=1,...,n-1) معرفة كرتـب الارتبـاط الذاتي من الرتبة k [3] :-

$$r_k^{(n)}(\vartheta_0) = \frac{\left\{(n-k)^{-1}\sum_{t=k+1}^{n} J_1\frac{(R_t^{(n)}(\vartheta_0))}{n+1} J_2\frac{(R_{t-k}^{(n)}(\vartheta_0))}{n+1} - m^{(n)}\right\}}{\sigma_k^{(n)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

حيث ان :-

J_2, J_1 هي نقاط دوال و $\sigma_k^{(n)}, m^{(n)}$ الثوابت الطبيعية

بحيث ان $(n-k)^{\frac{1}{2}} r_k^{(n)}(\vartheta_0)$ تكون قياسية ونرمز بالرمز $H^{(n)}(\vartheta_0)$ للفرضية التي بموجبها تبقـى دالة كثافة الاحتمال (g) غير محددة ، كما نرمز بالرمز $r_{k,g}^{(n)}$ لـ (g) من رتب الارتباط الذاتي من الرتبة (k) ، بعبارة أخرى فأن [3] :-

$$J_1 = \varphi g^0 G^{-1}$$
$$J_2 = G^{-1}$$

بحيث ان

$$\varphi g^{(.)} = \frac{-g'(.)}{g}$$

G^{-1} :- بحيث ان :- ة التجميعية (G) لدالة الكثافة (g) .

أما الفرضيات التي أعتمدتها هذه الدراسة والتـي تخـص دالـة الكثافـة للخطـأ (g) ونقـاط الدوال J_2, J_1 فهي كالاتي :-

* الفرضية الاولى (A) // تنقسم هذه الفرضية الى مايلي :-

$$\int Xg(X)dX = 0 \ , \ 0\langle\sigma^2 \ , \ \int X^2 g(X)dX\langle\infty \qquad \text{// A-1}$$

A-2 // دالة الكثافة للخطأ (g) هي دالة مستمرة مطلقاً ولها مشتقة (g^1) .مع الاشارة هنا الـــى افتراض ان معلومات فيشر (انظر (ALLal & Paindaveine(1999) :-

$$I(g) = \int \frac{\{g(X)\}^2}{g(X)} g(X)dX$$

تكون محدودة

A- 3// المتجه $(\xi\text{-}q+1 \ \ldots\ldots, \xi\ o, X\text{-}p+1, \ldots, X_o)$

$X\varepsilon R \ \forall\ g(x) \geq O$ يمتلك دالة كثافة مشتركة مخفية $(\delta\ ،\ .\ )$ g^o بحيث ان :-

$Hg(\theta_o)$ وبتحقق الفرضية $g^o(.\ ،\ \theta)^n - g^o\ (.\ ،\ \theta) = O\ p\ (1)$

فأن $\theta^{(n)} \rightarrow \theta_o$

الفرضية الثانية (B) //

لكل من J_2, J_1 فأن $\int_0^1 J_i(u)du = 0, i = 1,2$

3-2// ايجاد المقدرات المعتمدة على الرتب based on ranks Estimators

في هذا الجزء تقدم مجموعة من مقدرات R– التي تعتمد على قيمة اولية للمعلمة المـــراد تقـــديرها (θ) في نماذج ARMA ونرمز لها بالرمز (θ_o) وكالآتي [3] ،[6] :-

3 -2 -1 // تعريـــف مقـــدرات R– of The R - estimators Definition

يعرف متجه الرتب الإحصائية كالآتي :-

$$\sqrt{n}T^{n}_{J_2,J_1}(\vartheta_0) = \left\{\sum_{k=1}^{n-1}(n-k)^{\frac{1}{2}}\psi_k k^{(1)}(\vartheta_0)r_k^{(n)}(\vartheta_0),\ldots,\sum_{k=1}^{n-1}(n-k)^{\frac{1}{2}}\psi_k^{\frac{p}{q}}(\vartheta_0)r_k^{(n)}(\vartheta_0)\right\}$$

وباستخدام العلاقة التالية :-

$$\Delta^{(n)}_{J_1,J_2}(\vartheta_0) = n^{\frac{1}{2}}M(\vartheta_0)G\psi^{-1}(\vartheta_0)T^{(n)}_{J_1,J_2}(\vartheta_0)$$

وعندما تكون قيم النقاط قيم النقاط (.) J_1، (.) J_2 بالتتابع هي (.) G^{-1} (.) φg_0، (.) G^{-1} فاننا نحصل على متتابعة مركزية لعامل موجب وبدقة اكثر (3)، (6):-

$$\Delta^{(n)}g(\vartheta_0) = \sigma_1^{\frac{1}{2}}(g)\Delta^{(n)}\varphi g^0 G^{-1}, G^{-1}(\vartheta_0)$$

وبتحقق الفرضية $Hg^{(n)}(\vartheta_0)$ فأنه يمكن بسهولة ان نثبت بأن المتجه $\Delta^{(N)}_{J_1,J_2}(\vartheta_0)$ يتوزع تقريباً توزيع طبيعي بمتوسط صفر ومصفوفة تباين $\psi(\vartheta_0)$ ويمكن تقدير المعلمة (ϑ_0) من الصيغة الاتية :-

$$\vartheta^{-n} = \arg\ \min\ \left\|\Delta^{(n)}_{J_1,J_2}(\vartheta)\right\| \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

حيث ان :-

$\|\cdot\|$:- أي طبيعة قياسية في R^{p+q}

3- 2- 2 // التقارب الطبيعي لمقدرات R- **Asymptotic normality of R- estimates**

اذا افترضنا تحقق الفرضيات A-1، A-2 وافترضنا ان (ϑ^{-n}) تكون الحل $(\sqrt{n})$ المتسق للمعادلة (3) فان التوزيع التقاربي لـ $\sqrt{n}(\vartheta^{-(n)} - \vartheta_0)$ تحت الفرضية $Hg^{(n)}(\vartheta_0)$ هو $N(0, C^{-1}(J_1, J_2, g)\Gamma^{-1}(\vartheta_0))$، (3) .

3- 3 // تركيب مقدرات R – التقاربية **Construction of asymptotic R-estimates //**

في هذا الجزء سوف نعتمد طريقة Le cam- Hajek لتركيب مجموعة من مقدراتR التقاربية $\hat{\vartheta}^{(n)}$ التي تحقق مايلي (3)، (4) :-

$$\Delta^{(n)}_{J_1,J_2}(\hat{\vartheta}^{(n)}) = O_p\vartheta_0(1)$$

لهذا الغرض نفرض $(\hat{\vartheta}^{(n)})$ ترمز لـ $\sqrt{n}$ من التقديرات الأولية المتسقة المتقطعة للمعلمة ϑ_0 بعبارة أخرى متتابعة من التقديرات بحيث :-

$$(i)\sqrt{n}(\widetilde{\vartheta}^{(n)} - \vartheta_0) = O_p\vartheta_0(1), \overline{as} \quad n \to \infty$$

(ii) العدد من القيم الممكنة للمعلمة $(\widetilde{\vartheta}^{(n)})$ من الصيغة :-

$$\left\{\vartheta \in R^{p+q}, \sqrt{n}\|\vartheta - \vartheta_0\| \le c\right\}, c\rangle 0 \quad fixed \quad as \quad n \to \infty$$

فان شرط الاتساق $\overline{\sqrt{n}}$ يكون متحقق بواسطة كل التقديرات التي تكون عادة في سياق نماذج ARMA واذا افترضنا بان الفرضيات A-1 ، A-2 متحققة وعرفنا التقدير الاتي ،(3) :-

$$\hat{\vartheta}^{(n)} = \widetilde{\vartheta}^{(n)} + \frac{1}{\sqrt{n}}C^{-1}(J_1,J_2,g)\Gamma^{-1}(\widetilde{\vartheta}^{(n)})\Delta^{(n)}_{J_1,J_2}(\widetilde{\vartheta}^{(n)}) \quad \ldots\ldots(4)$$

فأن :-

$$\sqrt{n}(\hat{\vartheta}^{(n)} - \vartheta_0) = C^{-1}(J_1,J_2,g)\Gamma^{-1}(\vartheta_0)\Delta^{(n)}_{J_1,J_2}(\vartheta_0) + O_p\vartheta_0(1) \quad \ldots\ldots(5)$$

والتوزيع التقاربي لـ $\sqrt{n}(\hat{\vartheta}^{(n)} - \vartheta_0)$ تحت الفرضية $H^{(n)}_g(\vartheta_0)$ هو :-

$N(0, C^{-2}(J_1,J_2,g)\Gamma^{-1}(\vartheta_0))$ بالاضافة الى ، (3) :-

$$\Delta^{(n)}_{J_1,J_2}(\hat{\vartheta}^{(n)}) = O_p\vartheta_0(1)$$

اخيرا فان المتتابعة من المقدرات $(\hat{\vartheta}^{(n)})$ وفقا للنقاط $J_1 = \varphi g^0 G^{-1}, J_2 = G^{-1}$ تكون (locally and asymptotically minimax) تحت دالة الكثافة الجديدة g .

والان اذا افترضنا ان A-1، A-2 متحققة وافترضنا ان $(\vartheta^{-(n)})$ هي الحل المتسق $\sqrt{n}$ للمعادلة (3) فان :-

$$\sqrt{n}(\hat{\vartheta}^{(n)} - \vartheta^{-(n)}) = O_p(1)$$

تحت الفرضية $H^{(n)}_g(\vartheta_0)$

4- الكفاءة النسبية التقريبية Asymptotic relative efficiencies

تحسب الكفاءة النسية التقريبية (ARE) لمقدرات R– المقترحة بالنسبة الى مقدرات Gaussian الاعتيادية المسماة مقدرات (Ls) أي الكفاءة النسيية التقاربية (ARE) للمعلمة $(\hat{\vartheta}_n)$ المقدرة حسب طرق R– بالنسبة للمعلمة $(\hat{\vartheta}^{(n)}LS)$ بالشكل الاتي ، (3)،(4)،(6) :–

$$ARE(\hat{\vartheta}^{(n)} / \hat{\vartheta}^{(n)}LS) = C^2(J_1, J_2, g) \quad \ldots\ldots(6)$$

والجدول رقم (1) يلخص القيم العددية للكفاءات النسيية التقريبية (AREs) لمقدرات R– التي اقترحت بالنسبة الى مقدرات Ls لنقاط دوال مختلفة ولتوزيعات مختلفة وهذا الجدول يوضح فائدة مقدرات R– من خلال قيم (ARE) ونلاحظ انه عندما نستخدم نقاط (Van der wearden) فان (AREs) تكون أكبر من أو مساوية للواحد بأنتظام ((أنظر Chernoff & Savage 1958))

جدول رقم (1) يوضح القيم العددية للكفاءات النسبية التقريبية لمقدرات R– المقترحة بالنسبة لمقدرات Ls–

Score functions J1 (u) ,J2(V)	Density types		
	Normal	Logistic	Double- exp
Van der waerden $\Phi^{-1}(u)$ $\Phi^{-1}(v)$	1	1.048	1.226
Wilcoxon(2u-1) Ln(v/1-v)	0.948	1.098	1.482
Laplace sign (2u-1)$F^{-1}e^{(v)}$	0.612	0.812	2

5 – الدراسة التطبيقية

في هذه الفقرة نعرض نتائج التطبيق العملي للدراسة النظرية اعلاه اذ تم استخدام دراسة المحاكاة باستخدام (250) سلسلة زمنية اضافة الى ذلك تم اجراء تحليل بيانات واقعية وفي كلتا الحالتين تبين ان الطرق المستخدمة كانت كفوءة وكالاتي :–

أ – دراسة المحاكاة //

double-exponential dist) على التوالي ولجميع السلاسل الزمنية أمكن احتساب سبعة مقدرات مختلفة اعتمادا على التقدير الاولي لمعلمة الانحدار الذاتي $\phi_{\hat{a}} = \rho_{\hat{a}} = 0.6$ وهذه المقدرات هي كالاتي :-

مقدرات Koul and Saleh's لكل من (Laplace ، Wilcoxon ، Van der Waerden) وكذلك المقدرات one – step etimator المقترحة في الجانب النظري لكل من (Laplace ، Wilcoxon ، Van der waerden) إضافة إلى مقدرات LS الاعتيادية وقد تم احتساب معيارين للمفاضلة بين التقديرات أعلاه هما متوسط الانحراف :-

$$\frac{1}{M}\sum_{m=1}^{M}\hat{\rho}_m^{(n)} - \rho_0$$

ومتوسط مربعات الأخطاء:-

$$\frac{1}{M}\sum_{m=1}^{M}(\hat{\rho}_m^{(n)} - \rho_0)^2$$

ويمكن توضيح نتائج التقدير كما في الجدول رقم (2) ومنه نلاحظ جودة وكفاءة التقديرات المقترحة في الجانب النظري ، اذ تبين ان افضل المقدرات هو المقدر – R المقترح لـ Laplace عندما يكون التوزيع طبيعي قياسي حيث كانت MSE له هي (0.00000121) وهي اقل MSE .

ب– تحليل بيانات واقعية

سوف نتاول هنا تحليل بيانات واقعية ميدانية تمثل سلسلة عدد حوادث المرور الشهرية في محافظة القادسية للسنوات (1995 _1999) بواقع (60) مشاهدة و لغرض التعرف على خصائص هذه السلسلة فقد تم رسمها ضد الزمن كما في الشكل رقم (1) اذ تبين ان السلسلة مستقرة الى حد ما في المتوسط والتاثيرات الموسمية فيها ليست واضحة ، وقد تم أختبار مدى مطابقة السلسلة للتوزيع الطبيع القياسي باستخدام اختبار كولمكروف – سميرنوف وتبين ان قيمة احصاءة كولمكروف –سميرنوف المحتسبة (0.12031) وهي اقل من القيمة النظرية بمستوى معنوية (0.05) وبمستوى معنوية (0.01) إذ كانت على التوالي (0.1723) و (0.2067) مما يدعم الفرضية القائلة بان سلسلة حوادث المرور الشهرية تتبع للتوزيع الطبيعي ضمن الفترة الزمنية أعلاه .

ولغرض تشخيص وتحديد درجة النموذج تم اولاً استخراج ورسم دالة الارتباط الذاتي كما في الشكل رقم (2) واستخراج ورسم دالة الارتباط الذاتي الجزئي كما في الشكل رقم (3) ويتضح من خلال دالتي الارتباط الذاتي للسلسلة والارتباط الذاتي الجزئي لها بانه ليس هناك نموذج اوساط متحركة

الفرضية القائلة بان سلسلة حوادث المرور الشهرية تتبع للتوزيع الطبيعي ضمن الفترة الزمنية أعلاه .

ولغرض تشخيص وتحديد درجة النموذج تم اولاً استخراج ورسم دالة الارتباط الذاتي كما في الشكل رقم (2) واستخراج ورسم دالة الارتباط الذاتي الجزئي كما في الشكل رقم (3) ويتضح من خـــلال دالتي الارتباط الذاتي للسلسلة والارتباط الذاتي الجزئي لها بانه ليس هناك نموذج اوساط متحركـــة تخضع له البيانات لكون دالة الارتباط الذاتي تتناقص تدريجياً بينما تخضع البيانات لنموذج الانحدار الذاتي في الدرجة الاولى (1) AR لكون دالة الارتباط الذاتي الجزئي تتلاشى بعد اول ازاحـــة ، وعليه فيمكن اخذ نموذج الانحدار الذاتي من الدرجة الاولى كنموذج مشخص لبيانات سلسلة حوادث المرور الشهرية للفترة 1995 – 1999 .

بعدما تم تشخيص النموذج امكن تقدير معلمة النموذج بتطبيق طرق التقدير المعروضة في الجزء النظري على بيانات السلسلة الزمنية أعلاه ، وقد تم اعتبار التقدير الأولى للمعلمة هي (Ø0 =0.6) والذي استخدم كتقدير اولي للطرق المدروسة ، وثم استخدام معيار (MSE) للمفاضلة بين الطـــرق المستخدمة في التقدير ، ولكون كافة الطرق المستخدمة قد أعطت كفاءة عالية لـــذا يمكـــن اعتبـــار النموذج الافضل هو النموذج الذي يعطينا أقل (MSE) ويمكن توضيح نتائج التقدير التي تم التوصل اليها في الجدول رقم (3) .

جدول رقم (2) يوضح نتائج التقدير للمحاكاة

Estimators	Density types		
	Normal	Logistic	Double-exponential
LS estimate			
MD	-0.0032	-0.00089	-0.00116
MSE	0.00000684	0.00000756	0.00000716
Van der waerden R-estimate			
MD	-0.00362	-0.00113	-0.00118
MSE	0.00000781	0.00000711	0.00000591
ARE	0.9556	1.0633	1.2115
Wilcoxon R-estimate			
MD	-0.00362	-0.00102	-0.00114
MSE	0.00000798	0.00000676	0.00000478
ARE	0.8634	1.1183	1.4979
Laplace R-estimate			
MD	-0.0041	-0.00123	-0.00112
MSE	0.00000121	0.00000894	0.00000361

Wilcoxon ks-estimate MD MSE ARE	 -0.0064 0.00000803 0.8580	 -0.00449 0.00000679 1.1134	 -0.00399 0.000004691 1.5267
Laplace ks-estimate MD MSE ARE	 -0.0084 0.000001231 0.5597	 -0.006898 0.00000894 0.8456	 -0.00629 0.00000357 2.0056

جدول رقم (3) يوضح نتائج التقدير لبيانات سلسلة حوادث المرور

Estimators	Ô	MSE	ARE
Ls estimate	-0.349	0.00451	
Van der waerden R-estimate	-0.213	0.00448	1.0067
Wilcoxon R-estimate	-0.212	0.00479	0.09415
Lapiace R-estimate	-0.331	0.001202	0.4501
Van der waerden Ks-estimate	-0.415	0.00465	0.9699
Wilcoxon Ks-estimate	-0.564	0.00496	0.9093
Laplace Ks-estimate	0.645	0.001012	0.4457

6- الاستنتاجات

نستنتج مما تقدم ما يلي :-

1- ان كافة مقدرات -R أعطت كفاءة عالية في تقدير معالم نماذج السلاسل الزمنية وان افضل تقدير تم الحصول عليه باستخدام المحاكاة هو مقدر -R المقترح لـ Laplace عندما يكون التوزيع طبيعي قياسي حيث اعطى اقل قيمة لـ MSE والتي بلغت (0.00000121) .

2- ان بيانات سلسلة حوادث المرور الشهرية في محافظة القادسية للسنوات (1995-1999) تتبع التوزيع الطبيعي القياسي وكما وضح ذلك اختبار كولمكروف- سميرنوف اذ كانت قيمة

نستنتج مما تقدم ما يلي :-

1- ان كافة مقدرات -R أعطت كفاءة عالية في تقدير معالم نماذج السلاسل الزمنية وان افضل تقدير تم الحصول عليه باستخدام المحاكاة هو مقدر -R المقترح لـ Laplace عندما يكون التوزيع طبيعي قياسي حيث اعطى اقل قيمة لـ MSE والتي بلغت (0.00000121) .

2- ان بيانات سلسلة حوادث المرور الشهرية في محافظة القادسية للسنوات (1995-1999) تتبع التوزيع الطبيعي القياسي وكما وضح ذلك اختبار كولمكروف- سميرنوف اذ كانت قيمة احصاءة كولمكروف-سميرنوف المحتسبة (0.12031) وهي أقل من القيمة النظرية بمستوى معنوية (0.05) وبمستوى معنوية (0.01) اذ كانت على التوالي (0.1723) و (0.2067)

3-ان بيانات سلسلة حوادث المرور الشهرية تخضع لنموذج الانحدار الذاتي من الدرجة الاولى AR(1) لكون دالة الارتباط الذاتي الجزئي تتلاشى بعد اول ازاحة .

4-ان افضل نموذج مقدر لسلسة حوادث المرور الشهرية في محافظة القادسية هو مقدر -R المقترح لـ Laplace الذي اعطى اقل قيمة لـ MSE والتي بلغت (0.001012) .

7- التوصيات

1- استخدام مقدرات -R لتقدير كافة نماذج السلاسل الزمنية وذلك لكونها تعطي مقدرات كفوءة جداً .

2- اجراء دراسات لمقدرات -R في نماذج الانحدار الاخرى الخطية وغير الخطية بالاستفادة من هذه الدراسة التي اجريت في مجال السلاسل الزمنية .

3-اجراء دراسات مقارنة بين مقدرات -R والطرف الآخر للتقدير بقية الوقوف على الافضل .

4-توفير برامج احصائية جاهزة لايجاد مقدرات -R وفي كافة المواضيع الاحصائية لتسهيل مهمة البحث العلمي .

8- المصادر

1. Adichie,J.N. (1967). "Estimation of regression parameters based on rank tests" , Ann.Math.Statist. 38,894-904.