

مقارنة بين طريقتي انحدار المكونات الرئيسة وطريقة سبيكمان لإنموذج الخطي الجزئي شبه المعلمي

صالح عادل صالح / باحث / saleh000@uomustansiriyah.edu.iq
أم.د. امينة كريم عيسى/الجامعة المستنصرية/كلية الادارة والاقتصاد ameena@uomustansiriyah.edu.iq

P: ISSN : 1813-6729

E: ISSN : 2707-1359

<https://doi.org/10.31272/jae.i142.1030>

مقبول للنشر بتاريخ: 2022/12/14

تاريخ أستلام البحث : 2022/12/6

المستخلص

يهتم العديد من الباحثين بالطرق شبه المعلمية ، التي تجمع بين الأساليب المعلمية واللامعلمية ، لأنها توفر مستوى أكثر تقدماً من التحليل الإحصائي الذي يمكن استخدامه لتحقيق مقدرات ذات مستوى عالي ، تظهر أهمية البحث من خلال المقارنة بين طريقتي التقدير شبه المعلمية واستخدام الإنموذج الانحدار الخطي الجزئي شبه المعلمي حيث تم استخدام طريقتين من طرق للتقدير وهي طريقة سبيكمان (PLMS) وطريقة انحدار المكونات الرئيسية (PCR) وباستعمال اسلوب المحاكاة و بأحجام عينات مختلفة (128 , 256 , 512) وتباينات مختلفة وهي (0.5 , 1 , 2) ، ومن خلال نتائج المحاكاة توصل الباحث على ان افضل طريقة هي طريقة انحدار المكونات الرئيسية (PCR) هي حيث تغلبت على طريقة سبيكمان (PLMS) عندما تكون عدد المتغيرات ثمانية متغيرات وعند استخدام دوال التمهيد الخطية والتريبعية والاسية .

الكلمات المفتاحية : شبه معلمي ، إنموذج خطي جزئي ، طريقة سبيكمان ، انحدار المكونات الرئيسية .



مجلة الإدارة والاقتصاد
مجلد 49 العدد 142 / آذار / 2024
الصفحات : 79 - 91

* بحث مستل من رسالة ماجستير .

1- المبحث الأول

1-1 المقدمة [1]

تكمن فلسفة الاحصاء في محاولة محاكاة ظواهر مختلفة ونمذجتها بحيث تكون اقرب ما يكون الى الواقع الفعلي قدر الامكان وتقاس درجة قوة النماذج على حسب قربها من الخصائص الاحصائية وهي على انواع واشكال مختلفة وتعتمد صياغة النماذج على سبب ونتيجة السبب حيث تأتي في المقدمة نماذج الانحدار حيث ان هذه النماذج تستكشف العلاقة ما بين السبب والذي يدعى المتغيرات التفسيرية (التوضيحية) ونتيجة السبب وتدعى المتغيرات المعتمدة (الاستجابة)، حيث كانت التوزيعات المعلمية (Parametric distributions) موضوعا للنقاش من قبل العديد من الباحثين على مدار السنوات القليلة الماضية حيث ان الطرق المعلمية تفترض ان العينات من المجتمع له عائلة توزيعات معروفة مثل عائلة التوزيع الطبيعي وعائلة توزيع كاما وغيرها من عوائل التوزيعات ثم تقدير المعالم المجهولة لهذه العوائل باستخدام احدى طرق التقدير (OLS , مربع كاي , والامكان الاعظم) وغيرها الكثير من الطرائق او عمل اشتقاق لحدود الثقة او عمل اختبارات ولكن هذه الافتراضات تكون صارمة احيانا وذلك لان التوزيع المعلمي لا يكون التوزيع الفعلي للمسألة المراد حلها فاذا كان التوزيع المعلمي المفترض للمسألة غير صحيح يؤدي الى استنتاجات غير صحيحة عند استخدام الطرائق الاحصائية وبالتالي تكون التقديرات المستخرجة غير صحيحة وان اي تلوث طفيف بالبيانات يمكن ان يخلق استنتاجات غير صحيحة ولذلك فان بيانات من هذا النوع لا يكون لها عائلة معلمية مناسبة , وكما ان نقص المعلومات الكاملة وتعقد الظواهر كأن تكون ظاهرة حدثت لأول مرة حيث في مثل هذه الحالات يلجأ الباحث الى طريقة اخرى اكثر مرونة في التعامل مع هكذا بيانات وهي طرائق لا تتطلب عمل فرضيات او على الاكثر تتطلب افتراضات اقل من الاسلوب المعلمي وهذا ما يدعى بالنماذج اللامعلمية (Non parametric model) والتي تكون ملائمة لبيانات العينة , واما في حالة كانت الظاهرة مضطربة فان هذه الظاهرة تغير سلوكها بطريقة ما وفي وقت معين او فترة زمنية او مكان ما فسوف نطرح اسئلة حول جدوى استخدام النماذج المعلمية او اللامعلمية ومن هنا تأتي الحاجة الى استخدام النماذج شبه المعلمية (Sim parametric model) وهذه النماذج هي عبارة عن دمج المركبات المعلمية والمركبات اللامعلمية

2-1 المشكلة وهدف البحث

كثيرا ما يواجه الباحث مشكلة البيانات الحيوية , في ظل مشكلة اختلاف التباينات اذ ان هناك بيانات بين الخطية وغير خطية لذا من الصعوبة تقديرها بالطرق المتعارف عليها مثل النماذج المعلمية بسبب حلولها المعقدة وكذلك صعوبة تقديرها بالطرق اللامعلمية وذلك لانها تصف مشكلة الابعاد , لذا نلجأ الى استخدام النماذج شبه المعلمية التي تمتاز بالمرونة العالية عند التعامل مع البيانات ذات المتغيرات المعلمية واللامعلمية , وفي بحثنا هذا نلجأ الى دراسة طريقتين من طرائق تقدير الإنموذج الخطي الجزئي وهي طريقة سبيكمان Spekman وطريقة انحدار المكونات الرئيسية وعمل مقارنة بينهما من خلال استخدام الدوال اللبية وعرض الحزمة والمقارنة بينهما باستخدام متوسط مربعات الأخطاء (MSE) للتوصل الى افضل طريقة .

2- المبحث الثاني

1-2 النماذج شبه المعلمية [6][7][8]

تعتبر الكثير من النماذج الاقتصادية هي نماذج شبه المعلمية وكما هو معروف ان النماذج الشبه معلمية هي عبارة عن دمج نماذج الانحدار المعلمية مع نماذج الانحدار اللامعلمية في ان واحد لغرض بناء وملائمة صلاحية النماذج الاحصائية ويعتبر الجزء المعلمي هو دالة الانحدار الذي يفترض ان تكون خطية في مشاهدات متغيراتها التوضيحية والذي يوضح بعض الظواهر الاقتصادية المهمة مثل (الكلفة , الفائدة , الخ) بينما الجزء اللامعلمي يتضمن توزيع المجتمع وبسبب عشوائية البيانات فقد لجأ الكثير من الباحثين الى استخدام النماذج شبه المعلمية وذلك لانها تعتبر دمج الإنموذجين المعلمي واللامعلمي ويتكون الإنموذج شبه المعلمي الكثير من النماذج فمنها الإنموذج المؤشر المنفرد والإنموذج الخطي الجزئي وغيرها الكثير ولكن في بحثنا هذا سوف نركز على الإنموذج الخطي الجزئي ويدعى كذلك بالإنموذج شبه المعلمي البسيط وذلك بسبب اهميته الكبيرة ولسهولة حسابه فقد استخدم بكثرة في التجارب العلمية والعملية .

ودرس الباحثون المعالجات الممكنة وقدموا جهودا كبيرة لتطوير الاساليب التي تقلل من مشكلة الابعاد في نماذج الانحدار التي تتكون نتيجة تزايد عدد التغيرات التوضيحية الداخلة في التحليل مما يجعل دقة التقدير منخفضة , اذ ان أنموذج الانحدار الشبه معلمي يحقق الخصائص العامة للانحدار المعلمي واللامعلمي ويتفق معها في نفس الغاية وهي الحصول على افضل منحى للبيانات يقترب او يطابق مع منحى متغير الاستجابة (Y) وذلك بالدمج بين اساليب التقدير المعلمية واللامعلمية وبسبب هذا تلقت النماذج شبه المعلمية اهتماماً كبيراً في مجال الاحصاء ولديها مجموعة كبيرة و واسعة من التطبيقات ومنها تطبيقات طبية حيوية وتطبيقات اقتصادية وتعتبر النماذج شبه المعلمية هي الاكثر شعبية في وقتنا الحاضر من النماذج الاخرى لانها تحافظ على

مرونة النماذج اللامعلمية مع الحفاظ على القوة التفسيرية للنماذج المعلمية وفي بحثنا هذا سوف نتناول إحدى النماذج شبه المعلمية وهو الانموذج الخطي الجزئي وذلك بسبب أهمية هذا الانموذج وكثرة استخدامه في العديد من المجالات والتطبيقات

2-2 تقديرات كيرنل [2][5][4][3]

ينصب تركيزنا في هذا الموضوع على تقدير دوال الانحدار اللامعلمي باستعمال إحدى الطرق اللامعلمية المهمة ، وهي طريقة مقدرات النواة (اللبية) ، حيث تعتمد اساس هذه الطريقة على اخذ الوسط الحسابي الموزون لقيم المتغير المعتمد Y_i من قبل الدالة اللبية $K(u)$ والتي تعد القاعدة الاولى للتقديرات اللبية اذ ان غالبية الابحاث تؤكد على ان اختيار الدالة اللبية المناسبة يجب تحقيق فيه الشروط الاربعة وان هذه الشروط هي (ان تكون الدالة مستمرة ، ان تكون دالة متماثلة حول نقطة الصفر ، ان تكون الدالة موجبة ، يجب ان تحتوي الدالة على مشتقة ثانية) ولتقدير دالة الانحدار اللامعلمي حسب الصيغة سابقة الذكر هي

$$Y = m(X_i) + e_i \dots (1) \quad i = 1, 2, 3, \dots, n \quad E(e_i) = 0 \quad E(e_i)^2 = \sigma^2 < \infty$$

وبشكل اكثر شمولية نعبر عن إنموذج الانحدار العام بالشكل التالي :-

$$\hat{Y} = \hat{m}(x) \dots (2)$$

ان المقدر لهذه الدالة $m(\bullet)$ يدعى بالمهمد ، ولايجاد معادلة التقدير هناك عدة طرق وسوف نذكر بعضها منها :

- 1- مهمد ناداريا واتسون (N.W) Nadaraya – Watson Kernel
- 2- مهمد (G.M) Gasser- muller kernel
- 3- مهمد الانحدار الخطي الموضعي (L.L.S) Local linear smoother

2-3 طرائق اختيار معلمة التمهيد [8]

ان اختيار معلمة التمهيد ، يعتبر جزءاً اساسياً من تقارب دالة الانحدار اللامعلمية مع الدالة الاصلية ، ونحصل على هذا التقارب لغرض الموازنة بين التباين والتحيز بحيث يكون الخطا اقل ما يمكن ، كما ان حجم العينة للمعلمة التمهيدية تحقق الشرطين

$$1- \lim_{n \rightarrow \infty} t \quad h = 0$$

$$2- \lim_{n \rightarrow \infty} nt \quad nh = \infty$$

توجد طرق عدة لمعلمة التمهيد وسوف نذكر بعضها منها وهي [طريقة العبور الشرعي (Gross-val) ، طريقة الدوال ، طريقة التقريب القياسية ، طريقة قاعدة الابهام ، طريقة الملى (plug-in) ، طريقة Local (risk estimation)]

2-4 طريقة التقدير سبيكمان (Speckman Method) [13][2]

في عام (1988) اقترح العالم (Speckman) طريقة لتقدير إنموذج الانحدار الخطي الجزئي حيث تتلخص هذه الطريقة بتقدير الجزء المعلمي حيث تم اخذ التوقع الشرطي للإنموذج الخطي الجزئي

$$Y = X\beta + m(t_i) + e \dots (3) \quad i = 1, 2, \dots, n$$

$$E(Y/X) = X\beta + m(t_i) \dots (4) \quad i = 1, 2, \dots, n$$

حيث وصف المقدر W ليمثل عناصر الدالة اللبية وان ال W مصفوفة كاملة الرتبة وتشير الى اوزان دالة (Kernel) ويشير ال γ الى متجه المعالم المضافة وبامكاننا كتابة الإنموذج بالشكل التالي :-

$$Y = X\beta + W\gamma + e \dots (5)$$

ويمكن تقدير ال β و γ بطريقة المربعات الصغرى وبالشكل التالي :-

$$e = Y - X\beta - W\gamma$$

$$e' = (Y - X\beta - W\gamma)'$$

$$ee' = (Y - X\beta - W\gamma)(Y - X\beta - W\gamma)'$$

يتم اشتقاق بالنسبة لل β و γ ومساواتها بالصفر نحصل على

$$\frac{\partial e'e}{\partial \beta'} = -X'(Y - X\hat{\beta} - W\hat{\gamma}) = 0 \quad , \quad \frac{\partial e'e}{\partial \gamma'} = -W'(Y - X\hat{\beta} - W\hat{\gamma}) = 0$$

$$X'X\hat{\beta} = X'(Y - W\hat{\gamma}) \dots (6)$$

$$W'W\hat{\gamma} = W'(Y - X\hat{\beta}) \dots (7)$$

نضرب المعادلة (6) في $(W'W)^{-1}$ ليصبح لدينا :-

$$\hat{\gamma} = (W'W)^{-1}W'(Y - X\hat{\beta}) \dots (8)$$

وبضرب المعادلة في (W) ينتج :-

$$W\hat{\gamma} = W(W'W)^{-1}W'(Y - X\hat{\beta}) \dots (9)$$

ولنفرض ان

$$P_{\omega} = W(W'W)^{-1}W'$$

هي مصفوفة متماثلة وصماء لذلك تصبح المعادلة (9) بالشكل التالي :-

$$W\hat{Y} = P_{\omega}(Y - X\hat{\beta}) \quad \dots\dots\dots(10)$$

نعوض المعادلة (10) في (6) حيث سينتج التالي :-

$$X'X\hat{\beta} = X'(Y - P_{\omega}(Y - X\hat{\beta}))$$

$$X'(I - P_{\omega})X\hat{\beta} = X'(I - P_{\omega})Y$$

بفتح الأقواس واخذ عامل مشترك للطرفين ينتج
بقسمة طرفي المعادلة على معامل $\hat{\beta}$ ينتج

$$\hat{\beta} = (X'(I - P_{\omega})X)^{-1}X'(I - P_{\omega})Y \quad \dots\dots\dots(11)$$

$\hat{\beta}$ بما ان P_{ω} عبارة عن مصفوفة تماثلية وصماء لذلك بإمكاننا كتابة المعادلة اعلاه بالشكل التالي:-

$$(X'(I - P_{\omega})'(I - P_{\omega})X)^{-1}X'(I - P_{\omega})'(I - P_{\omega})Y$$

وعلى فرض ان $\tilde{Y} = (I - P_{\omega})Y$ و $\tilde{X} = (I - P_{\omega})X$ فان المقدر المعلمي كالاتي:-

$$\hat{\beta}_{LS} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y} \quad \dots\dots\dots(12)$$

ولتقدير الجزء اللامعلمي نستخدم ممد Local polynomial smoother (ممد متعدد الحدود الموضوعي)
(ومتسلسلة تايلور (Taylor expansion) , وباستخدام طريقة المربعات الصغرى الموزونة نحصل على :-

$$\hat{\beta}_{LP} = (X'\mathcal{R}X)^{-1}X'\mathcal{R}Y \quad \dots\dots\dots(13)$$

2-5-2 طريقة تقدير انحدار المكونات الرئيسية (مقدر PCR) [10][11][12]

Principal components regression estimator method

في تحليل الأنحدار غالبا ما يواجه الباحثون مشكلة العلاقة الخطية المتعددة ، بالنسبة لأنموذج الانحدار الخطي القياسي يؤدي مقدر المربعات الصغرى اداء ضعيفا في وجود علاقة خطية متعددة ، وقد يؤدي وجود هذه العلاقة الى اعطاء تقديرات بعلامات خاطئة ، وقد يؤدي كذلك الى فترات ثقة واسعة للمعلمت الفردية (تقديرات غير مستقرة) وقد يؤثر ذلك على قرارنا بشكل مباشر .

حيث في اختبار الفرضية قد تجعل العلاقات الخطية المتعددة حادة او شديدة نوعا ما غير مستقرة لدرجة انها تصبح عديمة الفائدة عمليا وللتغلب على هذه المشكلة تم اقتراح العديد من المقدرات المتحيزة الجديدة كوسيلة لتحسين دقة تقدير المعلمة في الأنموذج عند وجود علاقة خطية متعددة ومن هذه المقدرات (مقدر stein , مقدر انحدار المكونات الرئيسية (PCR) , مقدر انحدار Ridge , مقدر Liu) ففي هذا البحث استخدم مقدر (PCR) الجديد للنماذج المعلمية β لذلك فان الهدف الرئيسي هو استخدام هذا المقدر الجديد ليشمل النماذج شبه المعلمية حيث سنقوم ببناء مقدر PCR للمركب المعلمي وعلى طريقة (Profile least-squares) حيث في البداية يجب علينا ان نقدم طريقة (Profile least-squares) وذلك لاي معلمة β

$$\text{Let } Y_i^* = y_i - X_i^T \beta$$

حيث يكون لدينا نموذج لا معلمي بالصيغة التالية :-

$$Y_i^* = m(t_i) + e_i \quad i = 1, 2, 3, \dots, n \quad \dots\dots\dots(14)$$

والان سوف نقوم بتقدير دالة الجزء اللامعلمي $m(o)$ لأي نقطة ولتكن (t_0) حيث يمكن تقريباها بواسطة الدالة الخطية

$$m(t) \approx m(t_0) + m'(t_0)(t - t_0) \quad \dots\dots\dots(15)$$

حيث ان t هو مقدر تقريبي للنقطة t_0 .

وليكن $k(o)$ هي دالة كيرنل و h هو عرض الحزمة (band width) فان الانحدار الخطي المقرب وجد المعلمت $m'(t_0)$ و $m(t_0)$ بحيث تكون اقل ما يمكن

$$\sum [y_i^* - m(t_0) - m'(t_0)(t_i - t_0)]^2 k_h(t_i - t_0) \quad \dots\dots\dots(16)$$

بما ان $K_h(o) = K(o/h)/h$ سوف تصبح المعادلة (16) بالشكل التالي :-

$$(\hat{m}(t_0), \hat{m}'(t_0))^T = (D_{t_0}^T W_{t_0} D_{t_0})^{-1} D_{t_0}^T W_{t_0} (Y - X\beta) \quad \dots\dots\dots(17)$$

W_{t_0} : تمثل مصفوفة يمكن التعبير عنها بالشكل التالي :-

$$W_{t_0} = \text{diag} \{k_h(t_1 - t_0), k_h(t_2 - t_0), \dots, k_h(t_n - t_0)\}$$

حيث ان باستبدال $m(t_i)$ في معادلة (14) بـ $\hat{m}(t_0)$ واجراء بعض العمليات الجبرية يمكننا الحصول على إنموذج انحدار خطي

$$\bar{Y} = \bar{X}\beta + e \quad \dots\dots\dots(18)$$

$$\bar{Y} = (\bar{y}_1, \bar{y}_2, \bar{y}_3, \dots, \bar{y}_n)^T = (I_n - S) Y$$

$$\bar{X} = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n)^T = (I_n - S) X$$

$$S = \begin{bmatrix} (1 & 0) & (D_{t_1}^T W_{t_1} D_{t_1})^{-1} D_{t_1}^T W_{t_1} \\ (1 & 0) & (D_{t_2}^T W_{t_2} D_{t_2})^{-1} D_{t_2}^T W_{t_2} \\ \vdots & & \vdots \\ (1 & 0) & (D_{t_n}^T W_{t_n} D_{t_n})^{-1} D_{t_n}^T W_{t_n} \end{bmatrix}$$

و بتطبيق نهج المربعات الصغرى على المعادلة رقم (18) يمكننا الحصول على مقدر (Profile least-squares) لل β

$$\hat{\beta} = (\bar{X}^T \bar{X})^{-1} \bar{X}^T \bar{Y} \dots \dots \dots (19)$$

وسوف نقوم بتطبيق مقدر (PCR) على المعادلة رقم (18) حيث سيظهر التحويل التالي على نموذج المعادلة رقم (20)

$$\bar{Y} = \bar{X} \beta + e = \bar{X} \Phi \Phi^T \beta + e = Z \alpha + e \dots \dots \dots (20)$$

حيث ان

$$\Phi = (\phi_1, \phi_2, \phi_3, \dots, \phi_p) \text{ orthogonal matrix (p \times p)}$$

$$\Phi^T \bar{X}^T \bar{X} \Phi = Z^T Z = \Lambda \text{ is diag}$$

$$\Lambda = (\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p) \quad , \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$$

اذ ان λ_p هي عبارة عن جذور مميزة (eigen values) لمصفوفة $(\bar{X}^T \bar{X})$ حيث ان :-

$$Z = \bar{X} \Phi = (\bar{X} \phi_1, \bar{X} \phi_2, \bar{X} \phi_3, \dots, \bar{X} \phi_p)$$

$$\alpha = \Phi \beta$$

وبافتراض ان

$$\lambda_{p-k+1}, \lambda_{p-k+2}, \lambda_{p-k+3}, \dots, \lambda_p$$

تكون قريبة من الصفر عندها يمكن تقسيم (Φ) وكما يلي :-

$$\Phi = (\Phi_1, \Phi_2) \quad \text{wher} \quad \Phi_1 = (\phi_1, \phi_2, \phi_3, \dots, \phi_k)$$

$$\Phi_2 = (\phi_1, \phi_2, \phi_3, \dots, \phi_p)$$

$$Z = (Z_1, Z_2) \quad \text{wher} \quad Z_1 = (\bar{X} \phi_1, \bar{X} \phi_2, \dots, \bar{X} \phi_k)$$

$$Z_2 = (\bar{X} \phi_{p-k+1}, \dots, \bar{X} \phi_p)$$

$$\Lambda = \begin{bmatrix} \Lambda_1 & 0_{k \times (p-k)} \\ 0_{(p-k) \times k} & \Lambda_2 \end{bmatrix} \quad , \quad \alpha = \Phi^T \beta = \begin{bmatrix} \Phi_1^T \beta \\ \Phi_2^T \beta \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

والان يمكن كتابة المعادلة رقم (20) بالصيغة التالية

$$\bar{y} = Z \alpha + e = Z_1 \alpha_1 + Z_2 \alpha_2 + e \dots \dots \dots (21)$$

والان يمكننا تعريف مقدر (Profile least-squares) ل α_1 مع α_2

$$\hat{\alpha}_1 = (Z_1^T Z_1)^{-1} Z_1^T \bar{Y} \\ = (\Phi_1^T \bar{X}^T \bar{X} \Phi_1)^{-1} \Phi_1^T \bar{X}^T \bar{Y} \\ = \Lambda_1^{-1} \Phi_1^T \bar{X}^T \bar{Y} \dots \dots \dots (22)$$

واخيرا نقوم بحساب المعادلة $\beta = \Phi_1 \alpha_1 + \Phi_2 \alpha_2$ من خلال مقدر (PCR) للمعلمة β ويمكن كتابته :-

$$\hat{\beta}_{PCR} = \Phi_1 \hat{\alpha}_1 = (\Phi_1 (\Phi_1^T \bar{X}^T \bar{X} \Phi_1)^{-1} \Phi_1^T \bar{X}^T \bar{Y}) \\ = \Phi_1 \Lambda_1^{-1} \Phi_1^T \bar{X}^T \bar{Y} \dots \dots \dots (23)$$

وكذلك يمكن كتابة التوقع و التباين لمقدر $\hat{\beta}_{PCR}$ بالشكل التالي :-

$$E(\hat{\beta}_{PCR} - \beta) = -\Phi_2 \Phi_2^T \beta + \Phi_1 \Lambda_1^{-1} \Phi_1^T \bar{X}^T (I_n - S) f \\ \text{Var}(\hat{\beta}_{PCR}) = \sigma^2 \Phi_1 \Lambda_1^{-1} \Phi_1^T \bar{X}^T (I_n - S) (I_n - S)^T \bar{X} \Phi_1 \Lambda_1^{-1} \Phi_1^T$$

3- المبحث الثالث الجانب التجريبي

1-3 وصف التجربة

تضمنت تجارب المحاكاة توليد المتغيرات العشوائية بأحجام عينات وقيم معلمات افتراضية مختلفة وذلك بالاعتماد على البيانات الحقيقية، كما تم تقدير نموذج الانحدار الجزئي وفق الطرائق المذكورة في الجانب النظري، ومن ثم تم اجراء مقارنة بين طرائق التقدير بالاعتماد على معيار متوسط مربعات الخطأ، اذ تم توليد متغيرات نماذج الانحدار الجزئي المدروسة وفق الخطوات الاتية :-

1. توليد المتغيرات التوضيحية (x_m) بحيث تتبع التوزيع المنتظم، اذ ان:

$$X^T = (x_{1i} \ x_{2i} \dots \ x_{mi}), \ X \sim U(a, b)$$

ويتم توليد المتغير التوضيحي اللامعلمي (t) بحيث يتبع التوزيع المنتظم، أي ان:

$$t \sim U(a, b)$$

2. يتم توليد الاخطاء العشوائية e_i بحيث يتبع التوزيع الطبيعي بمتوسط صفر وتباين σ_e^2 وفق طريقة (Box-Muller)، أي ان:

$$e_i \sim N(0, \sigma_e^2), \ i = 1, 2, \dots, n$$

ولقد تم افتراض ثلاث قيم لتباين الخطأ العشوائي، $\sigma_e^2 = 0.5, 1, 2$.

3. فيما يخص دوال التمهيد اللامعلمية $m(t_i)$ فهي عديدة ومتنوعة، وقد تم اختيار الدوال (الخطية، التربيعية، جيبية اسية) وبالصغى الاتية وعلى الترتيب:

$$m_1(t_i) = i/n$$

$$m_2(t_i) = 3.2t^2 - 1$$

$$m_3(t_i) = \sin(2t) + 2\exp(-16t^2)$$

4. يتم الحصول على المتغير المعتمد y_i من خلال تجارب المحاكاة المختلفة وذلك من خلال افتراض قيم معلمات مختلفة، والمتغيرات التوضيحية التي تم توليدها في الفقرة (1) فضلا عن إضافة الأخطاء العشوائية والتي تم توليدها في الفقرة (2)، ودوال التمهيد المشار إليها في الفقرة (3)، ولكل نموذج من النماذج المدروسة، أي ان:

$$y_i = X\beta + m_j(t_i) + e_i, \ i = 1, 2, \dots, n, \ j = 1, 2, 3$$

5. كما تم تحديد قيم المعلمات لانموذج الانحدار الجزئي من خلال تقديرها بطريقة المربعات الصغرى وبشكل يتلائم مع طبيعة البيانات المدروسة في الجانب التطبيقي وكما ياتي:

- $\beta_1 = -0.5, \beta_2 = 1, \beta_3 = -0.05, \beta_4 = 0.25,$
 $\beta_5 = -1.2, \beta_6 = 1.3, \beta_7 = -0.07, \beta_8 = 0.25$

6. فيما يخص طرائق التقدير المدروسة فقد تم تطبيقها على التجارب المذكورة أعلاه وذلك بالاعتماد على التفاصيل والخطوات التي تم تناولها في الجانب النظري، اذ تم توظيف طريقة التمهيد الخطي الموضعي (Local Linear Smooth) بالدالة اللببية (Gaussian) وطريقة العبور الشرعي (Cross-Validation) في تقدير عرض الحزمة (h)، وتم الاعتماد على معيار (Kaiser) لتحديد عدد العوامل المؤثرة على المتغير المعتمد، اما بالنسبة لطريقة التقدير المويجي فقد تم توظيف مرشحات بالمستوى الرابع (Filters Level=4) والعتبة الشاملة الناعمة (Universal Soft Threshold)

7. لتنفيذ تجارب المحاكاة المختلفة فقد تم استعمال ثلاثة احجام مختلفة $n = 128, 256, 512$ ، وتكرار الخطوات أعلاه هو ($L=5000$)، ومن ثم اخذ المعدل لإيجاد المقدرات والحصول على القيم النهائية للمقدرات ومتوسط مربعات الخطأ.

3-3 عرض وتحليل نتائج المحاكاة

تم استخراج نتائج تجارب المحاكاة والتي تتضمن طرائق تقدير نموذج الانحدار الخطي الجزئي في الجداول من (1-3) الى (10-3) ادناه، كذلك تضمنت نتائج المقارنة وفقا للمعيار (MSE)، فضلا عن الاشكال والرسومات لمنحنيات التمهيد المقدر، وكما في ادناه:

الجدول رقم (3-1)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.54818	-0.50986		$\hat{\beta}_5$	-1.21312	-1.21202
	$\hat{\beta}_2$	0.94745	1.01504		$\hat{\beta}_6$	1.26308	1.28727
	$\hat{\beta}_3$	0.01740	-0.05610		$\hat{\beta}_7$	-0.06585	-0.06705
	$\hat{\beta}_4$	0.21286	0.23224		$\hat{\beta}_8$	0.23030	0.22717

مقارنة بين طريقتي اندجار المكونات الرئيسية وطريقة سبيكمان لانموذج الخطي
الجزئي شبه المعلمي

256	$\hat{\beta}_5$	-1.29003	-1.25617	512	$\hat{\beta}_1$	-0.47933	-0.48645
	$\hat{\beta}_6$	1.29540	1.30163		$\hat{\beta}_2$	0.97180	0.99874
	$\hat{\beta}_7$	-0.10035	-0.06879		$\hat{\beta}_3$	-0.05712	-0.04876
	$\hat{\beta}_8$	0.32106	0.26989		$\hat{\beta}_4$	0.25685	0.25146
	$\hat{\beta}_1$	-0.51448	-0.49769		$\hat{\beta}_5$	-1.17160	-1.20339
	$\hat{\beta}_2$	1.03279	0.98871		$\hat{\beta}_6$	1.27023	1.29934
	$\hat{\beta}_3$	-0.08435	-0.05875		$\hat{\beta}_7$	-0.05567	-0.06926
	$\hat{\beta}_4$	0.25156	0.24227		$\hat{\beta}_8$	0.25190	0.24551

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 0.5$ عند اجراء التجربة لدالة التمهيد (M1) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (2-3)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.56886	-0.51438	512	$\hat{\beta}_5$	-1.21875	-1.21247
	$\hat{\beta}_2$	0.92479	1.01239		$\hat{\beta}_6$	1.24722	1.28404
	$\hat{\beta}_3$	0.04634	-0.05274		$\hat{\beta}_7$	-0.06406	-0.06639
	$\hat{\beta}_4$	0.19719	0.23151		$\hat{\beta}_8$	0.22188	0.22797
	$\hat{\beta}_5$	-1.32875	-1.26492		$\hat{\beta}_1$	-0.47047	-0.48452
	$\hat{\beta}_6$	1.29339	1.30311		$\hat{\beta}_2$	0.95972	0.99760
	$\hat{\beta}_7$	-0.11335	-0.07261		$\hat{\beta}_3$	-0.06016	-0.04834
	$\hat{\beta}_8$	0.35150	0.27780		$\hat{\beta}_4$	0.25978	0.25291
256	$\hat{\beta}_1$	-0.52064	-0.49950	$\hat{\beta}_5$	-1.15943	-1.20236	
	$\hat{\beta}_2$	1.04686	0.99403	$\hat{\beta}_6$	1.25746	1.29824	
	$\hat{\beta}_3$	-0.09903	-0.06114	$\hat{\beta}_7$	-0.04955	-0.06782	
	$\hat{\beta}_4$	0.25226	0.24272	$\hat{\beta}_8$	0.25271	0.24590	

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 1$ عند اجراء التجربة لدالة التمهيد (M1) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (3-3)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.59642	-0.52040	512	$\hat{\beta}_5$	-1.22625	-1.21130
	$\hat{\beta}_2$	0.89457	1.00886		$\hat{\beta}_6$	1.22608	1.27888
	$\hat{\beta}_3$	0.08492	-0.04826		$\hat{\beta}_7$	-0.06168	-0.06571
	$\hat{\beta}_4$	0.17629	0.23055		$\hat{\beta}_8$	0.21065	0.22819
	$\hat{\beta}_5$	-1.38037	-1.27660		$\hat{\beta}_1$	-0.45865	-0.48192
	$\hat{\beta}_6$	1.29071	1.30508		$\hat{\beta}_2$	0.94361	0.99594
	$\hat{\beta}_7$	-0.13068	-0.07771		$\hat{\beta}_3$	-0.06423	-0.04777
	$\hat{\beta}_8$	0.39209	0.28833		$\hat{\beta}_4$	0.26370	0.25468
256	$\hat{\beta}_1$	-0.52887	-0.50073	$\hat{\beta}_5$	-1.14321	-1.20073	
	$\hat{\beta}_2$	1.06562	1.00010	$\hat{\beta}_6$	1.24043	1.29641	
	$\hat{\beta}_3$	-0.11860	-0.06431	$\hat{\beta}_7$	-0.04138	-0.06565	
	$\hat{\beta}_4$	0.25319	0.24305	$\hat{\beta}_8$	0.25379	0.24640	

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 2$ عند اجراء التجربة لدالة التمهيد (M1) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (4-3)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.57183	-0.60442	512	$\hat{\beta}_5$	-1.19478	-1.07694
	$\hat{\beta}_2$	0.93154	0.84341		$\hat{\beta}_6$	1.22637	1.15912
	$\hat{\beta}_3$	-0.02928	0.00614		$\hat{\beta}_7$	-0.08033	-0.04786
	$\hat{\beta}_4$	0.28983	0.35787		$\hat{\beta}_8$	0.26068	0.29985
	$\hat{\beta}_5$	-1.33137	-1.11491		$\hat{\beta}_1$	-0.46616	-0.47863
	$\hat{\beta}_6$	1.27849	1.17329		$\hat{\beta}_2$	0.97443	0.90424
	$\hat{\beta}_7$	-0.11600	-0.14228		$\hat{\beta}_3$	-0.05506	-0.03558

مقارنة بين طريقتي اندجار المكونات الرئيسية وطريقة سبيكمان للنموذج الخطي
الجزئي شبه المعلمي

256	$\hat{\beta}_8$	0.35514	0.30952		$\hat{\beta}_4$	0.24652	0.23762
	$\hat{\beta}_1$	-0.54071	-0.52727		$\hat{\beta}_5$	-1.18959	-1.03144
	$\hat{\beta}_2$	1.03506	1.04148		$\hat{\beta}_6$	1.28103	1.14553
	$\hat{\beta}_3$	-0.10802	-0.04549		$\hat{\beta}_7$	-0.06657	0.00805
	$\hat{\beta}_4$	0.24282	0.23687		$\hat{\beta}_8$	0.24851	0.26543

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 0.5$ عند اجراء التجربة لدالة التمهيد (M2) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (5-3)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.60430	-0.62156	512	$\hat{\beta}_5$	-1.19253	-1.08334
	$\hat{\beta}_2$	0.90767	0.83997		$\hat{\beta}_6$	1.19636	1.13960
	$\hat{\beta}_3$	0.00126	0.02929		$\hat{\beta}_7$	-0.07733	-0.04473
	$\hat{\beta}_4$	0.27588	0.36394		$\hat{\beta}_8$	0.25664	0.31194
	$\hat{\beta}_5$	-1.40006	-1.23757		$\hat{\beta}_1$	-0.44636	-0.45897
	$\hat{\beta}_6$	1.28895	1.22577		$\hat{\beta}_2$	0.95781	0.89061
	$\hat{\beta}_7$	-0.14826	-0.17068		$\hat{\beta}_3$	-0.05442	-0.03080
	$\hat{\beta}_8$	0.41565	0.37539		$\hat{\beta}_4$	0.25123	0.23477
256	$\hat{\beta}_1$	-0.54504	-0.53302	$\hat{\beta}_5$	-1.17360	-1.01938	
	$\hat{\beta}_2$	1.06770	1.07894	$\hat{\beta}_6$	1.26132	1.12619	
	$\hat{\beta}_3$	-0.13033	-0.05826	$\hat{\beta}_7$	-0.05349	0.02281	
	$\hat{\beta}_4$	0.24355	0.24711	$\hat{\beta}_8$	0.24966	0.26800	

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 1$ عند اجراء التجربة لدالة التمهيد (M2) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (6-3)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.63027	-0.63218	512	$\hat{\beta}_5$	-1.19074	-1.08710
	$\hat{\beta}_2$	0.88857	0.82486		$\hat{\beta}_6$	1.17234	1.12355
	$\hat{\beta}_3$	0.02570	0.05039		$\hat{\beta}_7$	-0.07492	-0.04171
	$\hat{\beta}_4$	0.26471	0.37573		$\hat{\beta}_8$	0.25340	0.32046
	$\hat{\beta}_5$	-1.45502	-1.30491		$\hat{\beta}_1$	-0.43052	-0.45132
	$\hat{\beta}_6$	1.29732	1.23764		$\hat{\beta}_2$	0.94452	0.88677
	$\hat{\beta}_7$	-0.17406	-0.19624		$\hat{\beta}_3$	-0.05390	-0.02856
	$\hat{\beta}_8$	0.46405	0.41488		$\hat{\beta}_4$	0.25501	0.23451
256	$\hat{\beta}_1$	-0.54850	-0.53656	$\hat{\beta}_5$	-1.16081	-1.01826	
	$\hat{\beta}_2$	1.09382	1.10690	$\hat{\beta}_6$	1.24556	1.12062	
	$\hat{\beta}_3$	-0.14819	-0.06829	$\hat{\beta}_7$	-0.04303	0.03416	
	$\hat{\beta}_4$	0.24414	0.25461	$\hat{\beta}_8$	0.25058	0.27078	

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 2$ عند اجراء التجربة لدالة التمهيد (M2) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (7-3)

n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.53443	-0.53694	512	$\hat{\beta}_5$	-1.19638	-1.20707
	$\hat{\beta}_2$	0.94463	0.88322		$\hat{\beta}_6$	1.24900	1.25275
	$\hat{\beta}_3$	0.02642	-0.02518		$\hat{\beta}_7$	-0.06515	-0.08250
	$\hat{\beta}_4$	0.20655	0.28230		$\hat{\beta}_8$	0.22923	0.28697
	$\hat{\beta}_5$	-1.33110	-1.38417		$\hat{\beta}_1$	-0.45515	-0.47242
	$\hat{\beta}_6$	1.33913	1.31735		$\hat{\beta}_2$	0.96603	0.99707
	$\hat{\beta}_7$	-0.14508	-0.12346		$\hat{\beta}_3$	-0.05084	-0.03302
	$\hat{\beta}_8$	0.33830	0.42903		$\hat{\beta}_4$	0.26114	0.26159
256	$\hat{\beta}_1$	-0.51531	-0.52407	$\hat{\beta}_5$	-1.16493	-1.13480	
	$\hat{\beta}_2$	1.06322	1.08034	$\hat{\beta}_6$	1.25547	1.22664	

مقارنة بين طريقتي اندجار المكونات الرئيسية وطريقة سبيكمان للنموذج الخطي
الجزئي شبه المعلمي

$\hat{\beta}_3$	-0.09152	-0.06533	$\hat{\beta}_7$	-0.03683	-0.02875
$\hat{\beta}_4$	0.24426	0.22588	$\hat{\beta}_8$	0.25069	0.24429

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 0.5$ عند اجراء التجربة لدالة التمهيد (M3) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (3- 8)

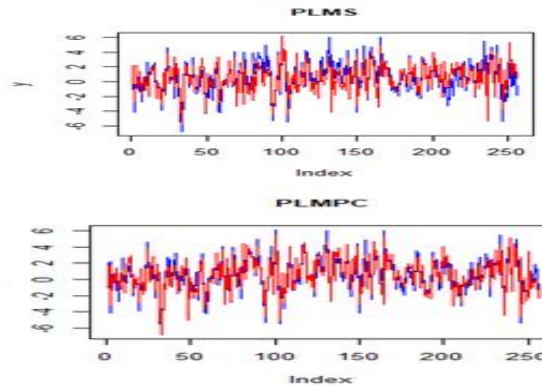
n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.54742	-0.53026	512	$\hat{\beta}_5$	-1.19548	-1.20541
	$\hat{\beta}_2$	0.93508	0.87049		$\hat{\beta}_6$	1.23699	1.24275
	$\hat{\beta}_3$	0.03863	0.00003		$\hat{\beta}_7$	-0.06394	-0.08107
	$\hat{\beta}_4$	0.20097	0.29778		$\hat{\beta}_8$	0.22761	0.28773
	$\hat{\beta}_5$	-1.35858	-1.36859		$\hat{\beta}_1$	-0.44723	-0.46708
	$\hat{\beta}_6$	1.34332	1.30402		$\hat{\beta}_2$	0.95938	0.99205
	$\hat{\beta}_7$	-0.15798	-0.14894		$\hat{\beta}_3$	-0.05058	-0.03256
	$\hat{\beta}_8$	0.36250	0.42684		$\hat{\beta}_4$	0.26303	0.26326
256	$\hat{\beta}_1$	-0.51704	-0.52698	$\hat{\beta}_5$	-1.15854	-1.12869	
	$\hat{\beta}_2$	1.07628	1.09287	$\hat{\beta}_6$	1.24759	1.21935	
	$\hat{\beta}_3$	-0.10045	-0.07168	$\hat{\beta}_7$	-0.03159	-0.02210	
	$\hat{\beta}_4$	0.24455	0.22664	$\hat{\beta}_8$	0.25115	0.24544	

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 1$ عند اجراء التجربة لدالة التمهيد (M3) والتي تحتوي على (ثمانية متغيرات)

الجدول رقم (3- 9)

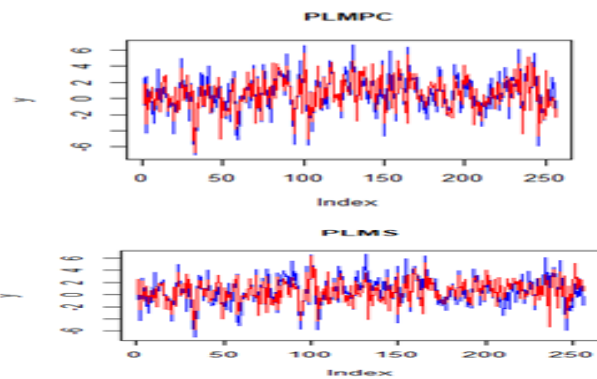
n	$\hat{\beta}$	PCR	PLMS	n	$\hat{\beta}$	PCR	PLMS
128	$\hat{\beta}_1$	-0.56041	-0.54148	512	$\hat{\beta}_5$	-1.19458	-1.20739
	$\hat{\beta}_2$	0.92553	0.86320		$\hat{\beta}_6$	1.22499	1.23452
	$\hat{\beta}_3$	0.05085	0.00967		$\hat{\beta}_7$	-0.06274	-0.07924
	$\hat{\beta}_4$	0.19539	0.29657		$\hat{\beta}_8$	0.22599	0.29023
	$\hat{\beta}_5$	-1.38606	-1.39028		$\hat{\beta}_1$	-0.43931	-0.46174
	$\hat{\beta}_6$	1.34750	1.30728		$\hat{\beta}_2$	0.95273	0.98702
	$\hat{\beta}_7$	-0.17088	-0.15977		$\hat{\beta}_3$	-0.05033	-0.03212
	$\hat{\beta}_8$	0.38671	0.44692		$\hat{\beta}_4$	0.26492	0.26492
256	$\hat{\beta}_1$	-0.51877	-0.53234	$\hat{\beta}_5$	-1.15215	-1.12256	
	$\hat{\beta}_2$	1.08934	1.10751	$\hat{\beta}_6$	1.23971	1.21206	
	$\hat{\beta}_3$	-0.10938	-0.07806	$\hat{\beta}_7$	-0.02636	-0.01545	
	$\hat{\beta}_4$	0.24485	0.22795	$\hat{\beta}_8$	0.25161	0.24660	

الجدول اعلاه يمثل تقدير المعلمات عندما $\sigma_e = 2$ عند اجراء التجربة لدالة التمهيد (M3) والتي تحتوي على (ثمانية متغيرات)

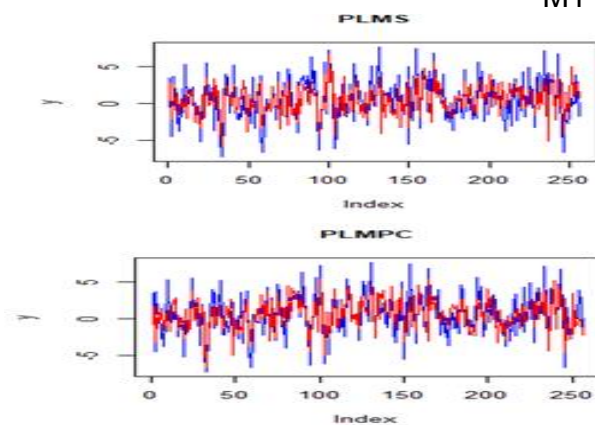


الشكل رقم (1-3) ادناه يمثل قيم عندما $\sigma_e = 0.5$ و حجم العينة ($n=256$) وعدد المتغيرات تساوي (ثمانية متغيرات) باستخدام المعادلة M1

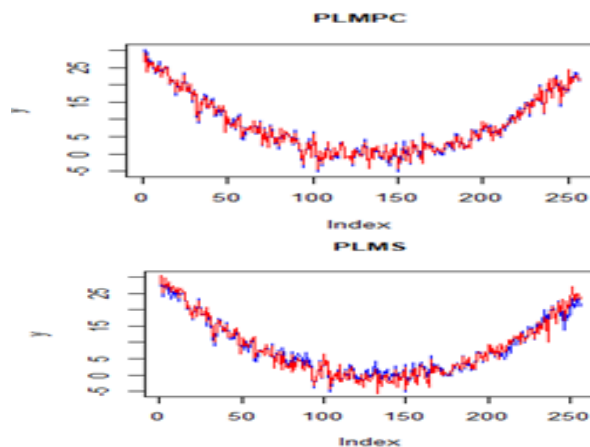
مقارنة بين طريقتي انحدار المكونات الرئيسية وطريقة سيكمان لإنموذج الخطي
الجزئي تشبه المعلمي



الشكل رقم (2-3) اعلاه يمثل قيم $\sigma_e = 1$ و حجم العينة ($n=256$) وعدد المتغيرات تساوي (ثمانية متغيرات) باستخدام المعادلة M1

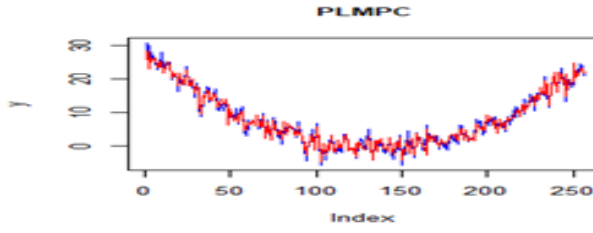


الشكل رقم (3-3) اعلاه يمثل قيم $\sigma_e = 2$ و حجم العينة ($n=256$) وعدد المتغيرات تساوي (ثمانية متغيرات) باستخدام المعادلة M1

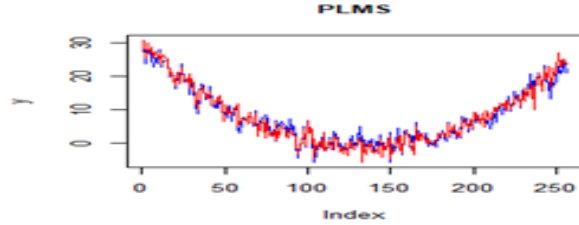


مقارنة بين طريقتي انحدار المكونات الرئيسية وطريقة سبيكمان للنموذج الخطي
الجزئي شبه المعلمي

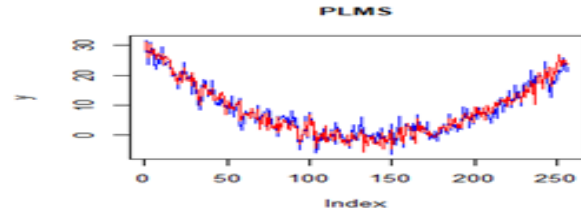
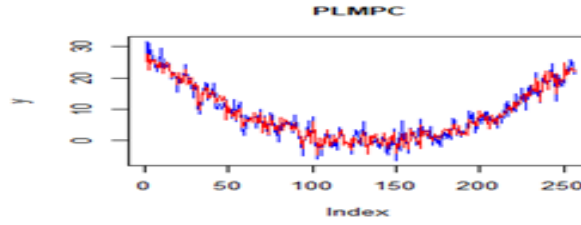
الشكل رقم (3-4) اعلاه يمثل قيم عندما $\sigma_e = 0.5$ و حجم العينة ($n=256$) وعدد المتغيرات تساوي (



ثمانية متغيرات) باستخدام المعادلة M2



الشكل رقم (3-5) اعلاه يمثل قيم عندما $\sigma_e = 1$ و حجم العينة ($n=256$) وعدد المتغيرات تساوي (ثمانية متغيرات) باستخدام المعادلة M2



الشكل رقم (3-6) اعلاه يمثل قيم عندما $\sigma_e = 2$ و حجم العينة ($n=256$) وعدد المتغيرات تساوي (ثمانية متغيرات) باستخدام المعادلة M2

الجدول رقم (3-10)

σ_u	Method	n	PCR	PLMS	σ_u	Method	n	PCR	PLMS
0.5	M1	128	0.11741	0.20426	2	M3	512	0.04409	0.07593
		256	0.09757	0.16460			128	0.08101	0.17334
		512	0.07029	0.10764			256	0.07103	0.11321
		128	0.02967	0.10618			512	0.04585	0.08300
		256	0.02171	0.06951			128	0.20928	0.29922
		512	0.01505	0.03860			256	0.17640	0.24735
	M2	M1	128	0.06025		0.14421	512	0.13071	0.16907
			256	0.04837		0.08951	128	0.14247	0.23327
			512	0.03267		0.06412	256	0.11516	0.17215
			128	0.15041		0.23826	512	0.07782	0.12013
			256	0.12587		0.19442	128	0.10557	0.20694
			512	0.09194		0.12967	256	0.09119	0.14091
1	M3	128	0.07043	0.16562	M2	128	0.10557	0.20694	
		256	0.06348	0.11651		256	0.09119	0.14091	
		512	0.06249	0.10506		128	0.10557	0.20694	
		512	0.06249	0.10506		256	0.09119	0.14091	

الجدول اعلاه يمثل قيم متوسط مربعات الخطأ المقدرة MSE (Mean Squared error) عند اجراء التجربة لثمانية متغيرات

تبين من خلال الجدول اعلاه ان متوسط المربعات الصغرى الاعتيادية عند اجراء التجربة التي تحتوي على (ثمانية متغيرات) لأحجام عينات مختلفة (128 , 256 , 512) على التوالي وتباينات مختلفة

(0.5,1,2) وباستخدام دوال التمهيد الخطية والاسية والتربيعية اثبتت من خلال المحاكاة ان طريقة المكونات الرئيسية (PCR) هي افضل من طريقة سبيكمان (PLMS) عند استخدام جميع التباينات وجميع دوال التمهيد

4- المبحث الرابع

الاستنتاجات والتوصيات

- 1- ان قيم MSE لطريقتي التقدير جاءت مطابقة للنظرية الاحصائية , اذ تقل MSE عندما يزداد حجم العينة وتزداد قيم MSE كلما زادت قيمة الخطأ العشوائي .
- 2- كلما زاد عدد المتغيرات التوضيحية الداخلة في الإنموذج و لطريقتي التقدير فان ذلك يؤدي الى زيادة قيم متوسط مربعات الخطأ لذلك الإنموذج .
- 3- من خلال هذا البحث تظهر طريقة المكونات الرئيسية (PCR) افضل من طريقة تقدير سبيكمان بجميع دوال التمهيد ومختلف التباينات واحجام العينات
- 4- وجد انه كلما زاد حجم العينة قلت قيم متوسط مربعات الخطأ مما يعني افضلية في النتائج
- 5- لكافة طرائق التقدير كانت دالة التمهيد التربيعية قدمت اداءً افضل من اداء دالتي التمهيد الجيبية-الاسية و الخطية , وان دالة التمهيد الجيبية-الاسية قدمت اداءً افضل من اداء دالة التمهيد الخطية .
- 6- نوصي باستخدام طريقة المكونات الرئيسية (PCR) باستخدام عدد اكبر من المتغيرات ودوال تمهيد مختلفة

المصادر العربية

- 1- حسن, ارشد حميد , (2017م), " استعمال تقنية الفروق في تقدير انموذج الانحدار شبه المعلمي في ظل مشكلة التعدد الخطي شبه التام مع تطبيق عملي " رسالة ماجستير في الإحصاء كلية الادارة والاقتصاد, جامعة بغداد.
- 2- حمزة , سعد كاظم , (2009), "مقارنة بعض الطرائق أليليه في تقدير نماذج الانحدار اللامعلميه بوجود بيانات تامة و غير تامة", رسالة ماجستير في الإحصاء, كلية الإدارة والاقتصاد , جامعة بغداد .
- 3- حمزة , سعد كاظم , (2015) , " التحليل المويجي لانموذج الانحدار في ظل بيانات مفقودة" اطروحة دكتوراة في الاحصاء , كلية الادارة والاقتصاد , جامعة بغداد .
- 4- حمود , مناف يوسف , (2000), "مقارنة مقدرات kernel اللامعلميه لتقدير دوال الانحدار " , رسالة ماجستير في الإحصاء , كلية الإدارة والاقتصاد , جامعة بغداد .
- 5- طراد , علاء جابر , (2013) , " نماذج الانحدار المعلمي وشبه المعلمي (دراسة مقارنة) " , رسالة ماجستير في الاحصاء , كلية الادارة والاقتصاد , جامعة المستنصرية .
- 6- عيسى, اسيل مسلم, (2011م), "مقارنة بعض المقدرات شبه المعلمية لتقدير دالة استهلاك الطاقة الكهربائية لمدينة بغداد" رسالة ماجستير في الإحصاء كلية الادارة والاقتصاد, جامعة بغداد.
- 7- كامل , ريم طلال , (2021) م , " استعمال بعض طرائق تقدير نماذج البيانات الطولية مع تطبيق عملي " رسالة ماجستير في الاحصاء , كلية الادارة والاقتصاد , جامعة بغداد
- 8- محمد , مياسة كاطع , (2014) , " مقارنة النماذج اللامعلمية وشبه المعلمية بوجود قيم مفقودة مع تطبيق عملي للنتائج المحلي الإجمالي العراقي للمدة (1971-2010)م " , رسالة ماجستير في الإحصاء , كلية الإدارة والاقتصاد , جامعة بغداد .

المصادر الأجنبية

- 9- Akdeniz Duran E, Akdeniz F, Hu H. Efficiency of a Liu-type estimator in semiparametric regression models J Comput Appl Math. 2011;235(5):1418–1428.
- 10- Chunling Liu, Shuang Guo and Chuanhua Wei , (2016) , " Principal components regression estimator of the parameters in partially linear models "
- 11 - Donald SG, Newey WK. Series estimation of semilinear models. J Multivariate Anal. 1994;50:30–40.
- 12 - Engle RF, Granger CWJ, Rice J, Weiss A. Semi parametric estimates of the relation between weather and electricity sales. J Amer Statist Assoc. 1986;81:310–320
- 13 - Paul Speckman , (1988) , " Kernel Smoothing in Partial Linear Models " , Journal of the Royal Statistical Society. Series B (Methodological), Vol. 50, No. 3 , pp. 413-436 .

A comparison between the two principal components regression methods and the Spekman method for the semi-parametric partial linear model

Saleh Adel Saleh / Researcher. saleh000@uomustansirivah.edu.iq .
Professor Amina Karim Issa ameena@uomustansirivah.edu.iq

Abstract

Many researchers are interested in semi-parametric methods , which combine parametric and non-parametric methods because they provide a more advanced level of statistical analysis that can be used to achieve high-level estimators . The importance of the research appears through the comparison between the two semi-parametric estimation methods and the use of the semi-parametric partial linear regression model . Two estimation methods were used , namely the Spykman method (PLMS) and the principal components regression method (PCR) , using the simulation method and with different sample sizes (128, 256, 512) and different variances , which are (0.5, 1, 2) , and through the simulation results it was concluded The researcher concluded that the best method is the principal components regression (PCR) method , which outperforms the Spekman method (PLMS) when the number of variables is eight variables and when linear , quadratic , and exponential smoothing functions are used.

Keywords : semi-parametric , partial linear model , Spekman method , principal components regression .

