توظيف تقنية الفروق في بعض مقدرات LIU لأنموذج الانحدار شبه المعلمي

سجى محمد حسين و ارشد حميد حسن

قسم الاحصاء - كلية الادارة والاقتصاد - جامعة بغداد

الخلاصة

يلقى موضوع الطرائق شبه المعلميه والذي يدمج الطرائق المعلميه والطرائق اللامعلميه اهتماماً واضحا في معظم الدراسات والتي تأخذ طابعا أكثر تقدما في عملية التحليل الإحصائي الدقيق الذي يهدف إلى الحصول على مقدرات ذات مستوى عالٍ من الكفاءة ،اذ يعد انموذج الانحدار الخطي الجزئي من اشهر انواع النماذج شبه المعلمية حيث يتكون من مركبة معلمية واخرى اللامعلمية ،ولغرض تقدير المركبة المعلمية التي تتمتع بخصائص معينة تعتمد على الافتراضات التي تتعلق بالمركبة المعلمية ،حيث ان عدم تحقق الافتراضات فان المركبة المعلمية سوف تعاني عدة مشكلات ومنها مشكلة التعدد الخطي اي اننا بصدد عدم تحقق فرض (ان المتغيرات التوضيحية غير مترابطة بعضها ببعض)،ولمعالجة هذه المشكلة نستخدم تقنية الفروق من خلال استخدام المقدرات المتحيزة ،ولغرض الحصول على مقدرات اقل تحيز واقل تباين نستخدم مقدر (Difference based liu estimator) ومقدر (Difference based almost unbiased liu estiomator) ومن خلال دراسة المحاكاة وبالاعتماد على معيار متوسط مربعات الخطأ استنتجنا بان مقدر (Difference based almost unbiased liu estiomator)افضل مقدر حيث يمتلك اقل متوسط مربعات الخطأ ونشير الى انه عندما يتم تقدير المركبة اللامعلمية يتم ازالة المركبة المعلمية من الانموذج شبه المعلمي ويتم تقدير المركبة اللامعلمية بأستخدام ممهد التجاور القريب (k-nearest neighbor smoother).

الكلمات المفتاحية : مقدر ليو المسند على الفروق (DBL) ، مقدر ليو غير المتحيز على الاغلب المسند على الفروق (DBAUL) ، ممهد التجاور القريب (K-NN)

Diyala Journal For Pure Sciences

توظيف تقنية الفروق في بعض مقدرات LIU لأنموذج الانحدار شبه المعلمي

سجى محمد حسين و ارشد حميد حسن

Employing difference technique in some Liu estimators to semiparametric regression model

Saja Mohammad Hussein[1] and Arshad Hameed Hassan[2]

[1,2]Department of statistics - college of Administration and Economics - Baghdad University

[1]sajamh@yahoo.com
[2]Arshad.hameed96@yahoo.com

Received: 18 January 2017 **Accepted: 19 February 2017**

Abstract

Semiparametric methods combined parametric methods and nonparametric methods ,it is important in most of studies which take in it's nature more progress in the procedure of accurate statistical analysis which aim getting estimators efficient, the partial linear regression model is considered the most popular type of semiparametric models, which consisted of parametric component and nonparametric component in order to estimate the parametric component that have certain properties depend on the assumptions concerning the parametric component, where the absence of assumptions, parametric component will have several problems for example multicollinearity means (explanatory variables are interrelated to each other) , To treat this problem we use a difference based through the use of biased estimators, in order to get less biased and variance estimators therefor we used difference based estimator liu and difference based almost unbiased liu estiomator. throughout studying simulation based upon mean square error, we concluded that difference based almost unbiased liu estiomator is better than difference based estimator liu since it has the smallest mean square error after that we estimate nonparametric component so removing parametric component and estimated Nonparametric using k-nearest neighbor smoother.

Keywords: Difference based liu estimator (DBL), Difference based almost unbiased liu estimator (DBAUL) , K-nearest neighbor smoother .

المقدمة

تحليل الانحدار هو اداة احصائية تقوم ببناء انموذج احصائي وذلك لتقدير العلاقة بين المتغير التابع وعدة متغيرات توضيحية وهي المتغيرات المستقلة ، بحيث ينتج معادلة احصائية توضح العلاقة بين المتغيرات ويمكن استخدام هذه المعادلة في التقدير والتنبؤ . إذ ان هذا الأسلوب يكون ذا فائدة عندما يكون متوسط الاستجابة خطياً اي ان

$$\boldsymbol{E(Y/X) = \beta' X}$$

ان هذا الافتراض لا يتحقق في اغلب التطبيقات العملية ، إذ أنه لا يأخذ بنظر الاعتبار التأثير اللاخطي للمتغيرات التوضيحية ، لذا كان من المناسب ايجاد اسلوب أخر يأخذ بنظر الاعتبار التأثير اللاخطي للمتغيرات وهو الانحدار اللامعلمي الذي تم اقتراحه من الباحث (Jacob Wolfowitz) عام (1942) إذ يكون متوسط الاستجابة بالشكل الاتي:

$$\boldsymbol{E(Y/X) = m(X)}$$

اذ تشير $m(\cdot)$ الى دالة الانحدار اللامعلمي P من الإبعاد (P-dimensional) . و قد لاحظ اغلب الباحثون ان الانحدار اللامعلمي يعاني أيضاً من مشكلة إلا وهي مشكلة الإبعاد (The Curse of Dimensionality) التي تحصل عند زيادة عدد المتغيرات. لذا فان الظروف و الأسباب الأنفة الذكر . ادت الى ظهور بعض الاساليب الحديثة في نماذج الانحدار ومنها نماذج الانحدار شبه المعلمية (semiparametric regression) لتحليل البيانات والتي تدمج المركبات المعلمية مع المركبات اللامعلمية التي تساهم في الحصول على قدرة عملية لتلافي مشكلة الابعاد . وعند وجود مشكلة التعدد الخطي (وتعني يرتبط اثنان او اكثر من المتغيرات التوضيحية بعلاقة خطية قوية جدا بحيث من الصعب فصل اثر كل متغير على المتغير الاستجابة) وعندما تكون قيمة احد المتغيرات التوضيحية متساوية لكافة المشاهدات او عندما تعتمد قيمة واحد او اكثر من المتغيرات التوضيحية في النموذج المدروس[4] ، لذلك نلجأ إلى استعمال احد أساليب معالجة هذه المشكلة وهو اسلوب انحدار الحرف عندما يكون هناك تعدد خطي شبه تام.حيث يتم توظيف تقنية الفروق لمعالجة مشكلة التعدد الخطي شبه التام في انموذج الانحدار شبه المعلمي ،ان تقنية الفروق تستعمل في تقدير المركبة المعلمية وازالة المركبة اللامعلمية في انموذح الانحدار شبه المعلمي.

1- هدف البحث

تهدف الدراسة الحالية الى تسليط الضوء على طرائق معالجة مشكلة التعدد الخطي في النماذج شبه المعلمية (Semiparametric Models) معتمدين بذلك انموذج انحدار خطي جزئي Partial linear model اذ يتم معالجة هذه المشكلة باستعمال تقنية الفروق لتقدير المركبة المعلمية وباستعمال ممهد التجاور القريب KNN في تقدير المركبة اللامعلمية ،ولعل اهم ما يركز عليه الباحث هو ابراز المقارنة بين المقدرات المعلمية التي تعالج مشكلة التعدد الخطي شبه التام وبيان افضل مقدر معتمدين بذلك على معيار متوسط مربعات الخطأ (MSE).

2- ادبيات البحث

سوف نستعرض اهم البحوث والدراسات الخاصة بموضوع تقدير انموذج الانحدار الخطي الجزئي شبه المعلمي والذي يحتوي على مشكلة التعدد الخطي . ففي عام (2008) قام الباحثان [Akdeniz and Esra][9] بدراسة مقدر liu في

الانحدار شبه المعلمي باستخدام المحاكاة واستنتجوا انه افضل من مقدر الفروق. وفي عام (2011) قام الباحثان [Esra and Akdeniz][10] بدراسة مقدر الفروق ومقدر liu في الانحدار شبه المعلمي . معتمدين بذلك على بيانات توزيع الكهرباء حيث اخذت البيانات من 81 محطة توزيع كهرباء في كندا وفي العام نفسه قام الباحثون [Fikri,Esra Hongchang][8] and بدراسة كفاءة مقدر liu في الانحدار شبه المعلمي ومقارنته مع اسلوب ذات المرحلتين في معالجة مشكلة التعدد الخطي واثبت كفاءة المقدر باستخدام اسلوب المحاكاة وفي عام (2012)قام كل من [Esra Akdeniz and Wolfgang karrl , Maria Osipenko] بدراسة كل من مقدر الحرف المسند الى الفروق ومقدر ليو المسند الى الفروق في انموذج الانحدار شبه المعلمي وتم تطبقها على البيانات الحقيقة وهي بيانات توزيع الكهرباء حيث اخذت البيانات من 81 محطة توزيع كهرباء في كندا وتوصلوا الى مقدر ليو المسند الى الفروق افضل من مقدر الحرف ،وفي عام (2016) قام الباحث [24] [jibo wu] بدراسة اداء مقدر almost unbiased liu المسند الى الفروق في انموذج الانحدار شبه المعلمي حيث تمت مقارنته مع مقدر الفروق ومقدر liu وباستعمال المحاكاة وبيانات حقيقية وتوصل الى ان اداء مقدر almost unbiased liu المسند الى الفروق كان افضل من المقدرين الاخرين .

الجانب النظري

1- أنموذج الانحدار الخطي الجزئي (Partial Liner Regression Model)

يمكن تمثيل أنموذج الانحدار شبه المعلمي الذي اقترحه كلا من Engle , Granger, Rice and *weiss* في عام 1986[15]ويرمزله بـ (PLM) بالصيغة الاتية:

$$\underline{\boldsymbol{Y}} = \boldsymbol{X}\underline{\boldsymbol{\beta}} + \boldsymbol{m}(\boldsymbol{t}) + \underline{\boldsymbol{\varepsilon}} \qquad \textbf{(1.2)}$$

اذ يشير $\underline{Y}$ الى متجة متغير المعتمد أو متغير الاستجابة من الدرجة(nx1)

و X يمثل مصفوفة المتغيرات التوضيحية (المعلمية) من الدرجة (nxp)

$\underline{\beta}$ متجة المعالم المجهولة من درجة(px1)

اذ يمثل Xβ الجزء المعلمي للأنموذج المدروس.

t هو متغير مستمر ويمثل عادة المتغير اللامعلمي في البيانات من درجة (nx1) .

m(t) تمثل دالة تمهيدية غير معرفة من الدرجة(nx1).

$\underline{\varepsilon}$ متجة الاخطاء العشوائية من درجة (nx1) وتكون مستقلة بمتوسط صفر وتباين σ^2.

في البداية يتم تقدير الجزء المعلمي والذي يعاني من مشكلة التعدد الخطي باستعمال مقدرات المعتمدة على الفروق المقترحة من قبل yatchew(2003) ويتم تقدير الجزء الاخر اللامعلمي باستخدام ممهد التجاور القريب KNN

2- مقدر (DBL) (Difference Based liu estimator)

في عام (2012) اقترح الباحث (Akdeniz et al) واخرون مقدر (Difference based liu) حيث تم توظف تقنية الفروق (Difference based) في مقدر liu المقترح من قبل (Akdeniz etal) في عام (1993). ان تقنية الفروق تستعمل لتقدير المركبة المعلمية في انموذج الانحدار الخطي الجزئي شبه المعلمي .وهذه الطريقة المقترحة من قبل yatchew(2003) تستعمل لازالة المركبات اللامعلمية في انموذج الانحدار شبه المعلمي ومعالجة مشكلة التعدد الخطي شبه التام في المركبة المعلمية في انموذج الانحدار الخطي الجزئي . [20][17] [7] [5] وعلية فان النموذج يكون بالصيغة الاتية

$$\boldsymbol{Dy = Dx + D\varepsilon} \qquad (2.2)$$

$$\boldsymbol{\tilde{y} = \tilde{X}\beta + \tilde{\varepsilon}} \qquad (3.2)$$

حيث ان: $\boldsymbol{\tilde{y} = Dy}$ $\boldsymbol{\tilde{\varepsilon} = D\varepsilon}$, $\boldsymbol{\tilde{x} = DX}$,

$\boldsymbol{\tilde{y}}$: يمثل متجه مشاهدات متغير الاستجابة من الدرجة$(n - m \times 1)$

$\boldsymbol{\tilde{X}}$: يمثل مصفوفة مشاهدات المتغيرات التوضيحية من الدرجة $(n - m \times p)$

β: يمثل متجه المعلمات المجهولة من الدرجة $(p \times 1)$

D :مصفوفة الفروق من الدرجة $(n - m \times n)$

n : عدد المشاهدات. m : رتبة الفروق

$\boldsymbol{\tilde{\varepsilon}}$: يمثل موجه الأخطاء العشوائية المستقلة ومتماثلة التوزيع بمتوسط صفر من الدرجة $(n - m \times 1)$.

وتتمثل طريقة الحصول على مقدر ليو المستند على الفروق (DBL) كما يأتي:-

وجود مصفوفة P من درجة $(p \times p)$ تمثل المتجهات المميزة المقابلة للجذور المميزة لمصفوفة المعلومات $\tilde{x}'\tilde{x}$ وان $P'P = I$ ويمكن اعادة كتابة النموذج $(\mathbf{3.2})$ باستخدام الصيغة القانونية بالشكل الاتي :-

$$\tilde{y} = \tilde{z}\alpha + \tilde{\varepsilon} \qquad (4.2)$$

$$\tilde{z} = \tilde{X}P, \qquad \alpha = P'\beta$$

$\boldsymbol{\tilde{z}}$: يمثل مصفوفة مشاهدات المتغيرات التوضيحية من الدرجة $(n - m \times p)$

α: يمثل متجه المعلمات المجهولة من الدرجة $(p \times 1)$

ان صيغة مقدر **Difference Based Liu هي:**

$$\hat{\alpha}_{Ldiff} = (\Lambda + I)^{-1}(\tilde{z}'\tilde{y} + d\hat{\alpha})$$

$$\hat{\alpha} = (\tilde{z}'\tilde{z})^{-1}\tilde{z}'\tilde{y}$$

$$\hat{\alpha} = \Lambda^{-1}\tilde{z}'\tilde{y}$$

$$\hat{\beta}_{Ldiff} = p(\hat{\alpha}_{Ldiff})$$

$$\therefore \hat{\beta}_{Ldiff} = p(\Lambda + I)^{-1}(\tilde{z}'\tilde{y} + d\hat{\alpha}) \qquad (5.2)$$

علما ان المصفوفة Λ من درجة $(p \times p)$

وبعد معرفة الصيغة التقديرية المقدر **Difference Based Liu** نوضح اهم خصائصه :-

اولا:- التحيز

$$Bias(\hat{\beta}_{Ldiff}) = E(\hat{\beta}_{Ldiff}) - \beta$$

$$Bias(\hat{\beta}_{Ldiff}) = \beta - (1-d)(\Lambda + I)^{-1}\beta - \beta$$

$$Bias(\hat{\beta}_{Ldiff}) = -(1-d)(\tilde{x}'\tilde{x} + I)^{-1}\beta \qquad (6.2)$$

ثانيا:- مصفوفة التباين والتباين المشترك

$$cov(\hat{\beta}_{Ldiff}) = \sigma^2(\Lambda + I)^{-1}(\Lambda + dI)\Lambda^{-1}(\Lambda + I)^{-1} \qquad (7.2)$$

ثالثا :-مصفوفة متوسط مربعات الخطأ(MSEM)

$$MSEM(\hat{\beta}_{Ldiff}) = cov(\hat{\beta}_{Ldiff}) + Bias(\hat{\beta}_{Ldiff})Bias(\hat{\beta}_{Ldiff})'$$

$$MSEM(\hat{\beta}_{Ldiff}) = \sigma^2(\Lambda + I)^{-1}(\Lambda + dI)\Lambda^{-1}(\Lambda + I)^{-1} + (1-d)^2(\Lambda + I)^{-1}\beta\beta'(\Lambda + I)^{-1}$$

4- مقدر (Difference based almost unbiased liu)(DBAUL)

في عام 2016 م اقترح الباحث (Jibo wu) مقدر Difference based almost unbiased liu الذي يستند على توظيف تقنية الفروق في مقدر(almost unbiased liu(AULE حيث تتمثل هذه الطريقة كما يأتي:-

يفترض وجود مصفوفة P من درجة $(p \times p)$ تمثل المتجهات المميزة لمصفوفة $\tilde{x}'\tilde{x}$ وان $P'P = I$ ويمكن اعادة كتابة النموذج (3.2) باستخدام الصيغة القانونية بالشكل الاتي :-

$$\tilde{y} = \tilde{z}\alpha + \tilde{\varepsilon} \qquad (8.2)$$

$$\tilde{z} = \tilde{X}P, \qquad \alpha = P'\beta$$

فان صيغة مقدر (DBAUL) تكون بالشكل التالي:

$$\begin{aligned}
\hat{\alpha}_{AULEdiff} &= [I-(1-d)^2(\Lambda+I)^{-2}]\hat{\alpha} \\
\hat{\beta}_{AULEdiff} &= P(\hat{\alpha}_{AULEdiff}) \\
\therefore \hat{\beta}_{AULEdiff} &= p[I-(1-d)^2(\Lambda+I)^{-2}]\hat{\alpha} \\
\hat{\alpha} &= (\tilde{z}'\tilde{z})^{-1}\tilde{z}'\tilde{y} \\
\hat{\alpha} &= \Lambda^{-1}\tilde{z}'\tilde{y}
\end{aligned} \tag{9.2}$$

علما ان المصفوفة Λ من درجة $(p \times p)$

اما اهم خصائصه:-

اولا:- التحيز

$$Bias(\hat{\beta}_{AULEdiff}) = -(1-d)^2(\Lambda+I)^{-2}\beta \tag{10.2}$$

ثانياً:- مصفوفة التباين والتباين المشترك

$$COV(\hat{\beta}_{AULEdiff}) = \sigma^2[I-(1-d)^2(\Lambda+I)^{-2}]\Lambda^{-1}[I-(1-d)^2(\Lambda+I)^{-2}]' \tag{11.2}$$

ثالثاً:- مصفوفة متوسط مربعات الخطأ

$$MSEM(\hat{\beta}_{AULEdiff}) = cov(\hat{\beta}_{AULEdiff}) + Bias(\hat{\beta}_{AULEdiff})Bias(\hat{\beta}_{AULEdiff})'$$

$$MSEM(\hat{\beta}_{AULEdiff}) = \sigma^2 U\Lambda^{-1}U' + (1-d)^4\beta'(\Lambda+I)^4\beta \tag{12.2}$$

$$U = [I-(1-d)^2(\Lambda+I)^{-2}]$$

5- حساب معلمة ليو(Liu)

هناك عدة أساليب يمكن من خلالها الوصول إلى أفضل قيمة لمعلمة التحيز ، حيث تعتمد قيمة معلمة انحدار Liu (d) كمية صغيرة موجبة تقع قيمتها بين$(0 < d < 1)$ بالأساس على مدى خبرة ومعرفة الإحصائي في كيفية اختيارها والتحكم بها ، وان هذه القيمة تعمل على تقليل متوسط مربعات الخطأ (MSE) ، وبالتالي سوف تكون المقدرات المحصل عليها هي الأفضل. قدمت هذه الطريقة من قبل الباحثين[12] (Alhrrty and Golam) ، وصيغتها هي :

$$d = \frac{\sum_{i=1}^{p} \frac{\lambda_i(\hat{\beta}_i^2 - \hat{\sigma}^2}{\lambda_i(\lambda_i+1)^2}}{\sum_{i=1}^{p} \frac{(\hat{\sigma}^2 + \lambda_i\hat{\beta}_i^2)}{\lambda_i(\lambda_i+1)^2}} \qquad (13.2)$$

$\hat{\beta}_i$: تمثل تقدير معلمات بطريقة المربعات الصغرى الاعتيادية

$$\hat{\sigma}^2 = \frac{Y'Y - \hat{\beta}' \;\; X'Y}{n-m-1} \qquad , \qquad i = 1, \dots, P$$

6- طرق تمهيد الجزء اللامعلمي

1-6 ممهد التجاور القريب (K-Nearest Nieghbor Smoother)

ان سلسلة الوزن لممهد(*K-NN*) قد قدمت لأول مرة من قبل *Loftsgaarden and Qaesenberry* في عام (*1965*) واستخدمه *Cover and Hart* في عام *1967*حيث كانت سلسلة الوزن بالشكل الاتي:-

$$\boldsymbol{W_{ki}} = \begin{cases} \boldsymbol{\frac{n}{k}} & \boldsymbol{if\ i \in j_x} \\ \boldsymbol{0} & \boldsymbol{otherwise} \end{cases} \qquad (14.2)$$

حيث ان سلسلة الوزن تعرف خلال المدى $\boldsymbol{j_x}$ والذي يمثل النقاط القريبة من *x* .

في عام (*1977*) اقترح الباحث *Stone* باضافة بعض دوال *kernel* كدوال وزنية لممهد التجاور الاقرب واضاف *Unform* و *Triangular* و *Quadratic* حيث ان سلسلة الوزن لهذا الممهد تكون بالشكل الاتي[15].

$$W_{ni}(t) = \frac{k\left(\frac{t-T_i}{k_n}\right)}{\sum_{i=1}^{n} k\left(\frac{t-T_i}{k_n}\right)}. \qquad (15.2)$$

لذلك يكون تقدير الدالة *m* بالشكل الاتي

$$\hat{m}_k(t) = n^{-1}\sum_{i=1}^{n} W_{ni}(t)\, y_i \qquad (16.2)$$

ان قياس التقارب بين المشاهدات يجب ان يحدد باستخدام

مقياس من بين العديد من المقاييس المختلفة ،لذلك يتم باستخدام دالة المسافة الاقليدية لحساب المسافة بين المشاهدات *ti* و *tj* في

مجموعة بيانات وفق الصيغة الاتية [1]

$$k_n = \sqrt{\sum_{k=1}^{n} (t_{ik} - t_{jk})^2} \qquad , i, j = 1,2\ldots\ldots\ldots\ldots n \qquad (17.2)$$

دراسة المحاكاة

من خلال دراسة المحاكاة يتم مقارنة مقدرات الفروق مع المقدرات المقترحة بالاعتماد على معايير المقارنة حيث ان دراسة المحاكاة تعتمد على دراسة McDonald and Galarneau [21] وفي هذه التجربة يتم توليد ثلاث متغيرات توضيحية ويتم توليد قيم المتغيرات التوضيحية X_{ij} ،والعمل على السيطرة على التذبذبات في قيم الارتباطات بين المتغيرات التوضيحية المرتبطة ، وعملية التوليد والسيطرة تتم من خلال المعادلة الآتية

$$X_{ij} = (1-\ \rho^2)^{1/2}\, u_{ij} +\ \rho\, u_{iP} \quad , i = 1,2, \dots , n \quad j = 1,2, \dots , P \qquad (1.3)$$

حيث u_{ij} الإعداد العشوائية المولدة والتي تتبع التوزيع الطبيعي القياسي وان ρ^2 تمثل الارتباط بين كل اثنين من المتغيرات التوضيحية وفي هذه الدراسة يتم استخدام الارتباطات وهي $\rho =$ 0.90,0.99,0.999 وتم استعمال حجوم عينات (200 ,100 ,60, 30)واستخدام سلسلة الفروق المثلى (1 ≤ m≤ 10) حيث نستخدم الفروق من الرتبة الخامسة ويتم تكرار هذه التجربة 1000 مرة حيث يتم توليد مشاهدات المتغير(Yi) وحسب الصيغة الاتية:-

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + x_{i3}\beta_3 + m(t_i) + \varepsilon_i \qquad (2.3)$$

اما بالنسبة للمتغير اللامعلمي t_i فيتم توليده بالاعتماد على المعادلة الاتية [10] .

$$t_i = \frac{i-0.5}{n}, \quad i = 1,2,3, \dots \dots , n \qquad (3.3)$$

وان الأخطاء العشوائية وتتوزع توزيعاً طبيعياً بمتوسط (0) وتباين (0.05,0.5,0.9) $= \sigma^2$ اي ان

$$\varepsilon_i \sim \mathrm{N}(0, \sigma^2), i = 1,2, \dots , n$$

النماذج المستخدمة في المحاكاة

تختلف النماذج باختلاف الظواهر التي تمثلها، ولذا لا يمكن عرض جميع انماط النماذج في الوقت نفسه، إلا انه تم استيفاء هذه النماذج من بحوث منشورة فعلا والتي تتماشى مع التطبيق العملي لهذا البحث، ويتكون أنموذج الانحدار شبه المعلمي من دالة الانحدار المعلمية $X_i'\beta$ ودالة الانحدار اللامعلمية $m(t_i)$ بالاضافة الى حد الخطأ العشوائي ε_i ،حيث تم توليد المتغيرات التوضيحية X_1, X_2, X_3 حسب الصيغة رقم (1.3). مع توليد أخطاء عشوائية تتبع التوزيع الطبيعي بمتوسط (0) وتباين σ^2 اما المتغير اللامعلمي فقد تم توليده حسب الصيغة (3.3) وسيتم مقارنة المقدرات باستعمال معيار معدل متوسط مربعات الخطا (Mean of Means squares Error) لتقييم اداء طرائق التقدير، وباستعمال البرنامج (MATLAB 2014a) تم تطبيق الأفكار النظرية الواردة في الفصل الثاني وتم الحصول على نتائج عديدةً ولكل أنموذج من النماذج المدروسة وكما في ادناه.

$$Y_i = X_i'\beta + m(t_i) + \varepsilon_i \qquad (4.3)$$

وبذا كانت المركبة اللامعلمية للنموذج اعلاه كما يأتي:

1. الأنموذج الأول:[22] وتكون المركبة اللامعلمية فيه وفق الصيغة الاتية :-

$$m_1(t_i) = 4.26e^{-3.25t_i} - 4e^{-6.5t_i} + 3e^{-9.75t_i} \tag{5.3}$$

2. الأنموذج الثاني [10]

$$m_2(t_i) = \sqrt{t_i(1-t_i)}\,\sin(2.1\prod/(t_i+0.05)) \tag{6.3}$$

جدول(3-1) يوضح متوسط مربعات الخطأ للمقدرات المركبة المعلمية

Model	n	σ^2	ρ	Estimators	
1				DBL	DBAUL
	30	0.05	0.90	6.6355e-04	4.7632e-04
			0.99	5.4287e-04	3.7251e-04
			0.999	3.7938e-04	3.0974e-04
		0.5	0.90	3.1146e-04	1.9094e-04
			0.99	3.9998e-04	2.4417e-04
			0.999	1.6668e-04	1.3919e-04
		0.9	0.90	1.2490e-04	4.5285e-05
			0.99	3.2453e-04	1.5368e-04
			0.999	5.7484e-05	4.4187e-05
	60	0.05	0.90	5.1287e-04	3.7440e-04
			0.99	4.2455e-04	3.3690e-04
			0.999	3.4494e-04	2.5979e-04
		0.5	0.90	0.0058	6.9206e-04
			0.99	1.0343e-05	2.8894e-04
			0.999	0.0436	0.0341
		0.9	0.90	0.2265	0.9903
			0.99	2.6039e-04	0.0022
			0.999	1.7863e-05	0.0010
	100	0.05	0.90	0.0014	7.5290e-04
			0.99	9.0800e-04	4.4563e-04
			0.999	6.6255e-04	3.0473e-04
		0.5	0.90	0.0058	0.0030
			0.99	4.7049e-05	5.9780e-06
			0.999	4.9768e-05	2.7136e-04
		0.9	0.90	0.0139	0.0069
			0.99	9.5369e-05	5.3191e-04
			0.999	1.7430e-04	0.0014
	200	0.05	0.90	1.1012e-04	3.3680e-05
			0.99	0.0455	0.0418
			0.999	0.0333	0.0256
		0.5	0.90	0.0431	0.0628
			0.99	2.2180	3.0569
			0.999	0.6860	1.1691
		0.9	0.90	0.1637	0.2194
			0.99	6.7902	9.6932
			0.999	1.8524	3.4811

DIYALA JOURNAL FOR PURE SCIENCES

جدول(2-3) يوضح متوسط مربعات الخطأ للمقدرات المركبة المعلمية

Model	n	σ^2	ρ	Estimators	
2				DBL	DBAUL
	30	0.05	0.90	e-052.4345	2.7502e-05
			0.99	1.3945e-05	1.6404e-05
			0.999	1.0721e-06	1.5095e-06
		0.5	0.90	1.5235e-04	1.7545e-04
			0.99	4.4019e-05	5.9766e-05
			0.999	4.4885e-05	4.9295e-05
		0.9	0.90	3.3256e-04	4.1389e-04
			0.99	6.7590e-05	1.2062e-04
			0.999	1.3568e-04	1.4810e-04
	60	0.05	0.90	7.6094e-07	1.2270e-07
			0.99	2.2919e-05	3.3533e-05
			0.999	3.1497e-06	1.9550e-05
		0.5	0.90	0.0230	0.0094
			0.99	1.9670e-04	0.0016
			0.999	3.2847e-06	8.3623e-04
		0.9	0.90	2.3190	90.6711
			0.99	2.4933e-04	0.0042
			0.999	0.0012	0.0033
	100	0.05	0.90	2.2826e-05	1.6893e-05
			0.99	6.6241e-06	6.5606e-07
			0.999	4.3277e-05	7.9755e-06
		0.5	0.90	0.0019	9.3187e-04
			0.99	1.3283e-04	4.9805e-04
			0.999	0.0018	2.0554e-04
		0.9	0.90	0.0071	0.0032
			0.99	6.8323e-04	0.0018
			0.999	0.0158	0.0019
	200	0.05	0.90	3.7107e-05	4.3625e-04
			0.99	0.0353	0.0322
			0.999	0.0206	0.0162
		0.5	0.90	0.0494	0.0704
			0.99	2.2112	3.0004
			0.999	0.7334	1.2064
		0.9	0.90	0.1755	0.2336
			0.99	6.7902	9.6001
			0.999	1.9544	3.5822

نتائج المحاكاة

1. انموذج الانحدار شبه المعلمي الاول

اظهرت نتائج الجدول رقم (3-1) عندما (n=30)ولجميع حالات التباين وقيم الارتباط فان افضل مقدر هو DBAUL كونه يمتلك اقل متوسط مربعات خطا ثم يليه مقدر DBL اما عندما حجم العينة (n=60) عند قيمة$\sigma^2 = 0.05$ فان مقدر DBAUL افضل في جميع مستويات معامل الارتباط وعند قيمة $\sigma^2 = 0.5$ وعند قيمة معامل الارتباط 0.90 فان مقدر DBAUL افضل من مقدر DBL وعند قيمة معامل الارتباط 0.99 يكون مقدر DBL افضل من مقدر DBAUL وعند قيمة معامل الارتباط 0.999 يكون مقدر DBAUL افضل من مقدرDBL وعندما تكون قيمة $\sigma^2 = 0.9$ فان مقدر DBL افضل من مقدر DBAUL في جميع مستويات الارتباط، اما عندما (n=100) فان مقدر DBUAL يكون افضل من مقدر DBL الا في حالتين الحالة الاولى عندما $\sigma^2 = 0.5$ وقيمة معامل الارتباط 0.999 والحالة الثانية عندما $\sigma^2 = 0.9$ وقيمة معامل الارتباط 0.999 اما عندما حجم العينة (n=200) عندما قيمة التباين $\sigma^2 = 0.05$. ولجميع قيم معامل الارتباط فيكون مقدر DBAULافضل من مقدر DBL اما عندما قيمة ($\sigma^2 = 0.5$ ،0.9) ولجميع قيم الارتباط فيكون مقدر DBL افضل من مقدر DBAUL.

2. انموذج الانحدار شبه المعلمي الثاني

اظهرت نتائج جدول رقم (3-2)عندما (n=30)ولجميع حالات التباين وقيم الارتباط فان افضل مقدر هو DBL كونه يمتلك اقل متوسط مربعات خطا في جميع حالات ثم يليه مقدر DBAUL اما عندما حجم العينة (n=60) وعند قيمة التباين $\sigma^2 = 0.05$ عند قيمة معامل الارتباط 0.90 فان مقدر DBAUL افضل من مقدر DBL اما عندما تكون قيم معامل الارتباط 0.99,0.999 فان مقدر DBL افضل من مقدر DBAUL وعندما تكون قيم التباين ($\sigma^2 = 0.5$ ،0.9) فان مقدر DBL افضل من مقدر DBAUL ولجميع مستويات الارتباط ،اما عندما (n=100) فان مقدر DBUAL يكون افضل من مقدر DBL الا في حالتين الحالة الاولى عندما $\sigma 2 = 0.5$ وقيمة معامل الارتباط 0.99 والحالة الثانية عندما $\sigma 2 = 0.9$ وقيمة معامل الارتباط 0.99 اما عندما حجم العينة (n=200)يكون مقدر DBAUL افضل من مقدر DBL في حالتين فقط وهي عندما قيمة التباين $\sigma 2 = 0.05$. وقيم معامل الارتباط (0.99,0.999) اما بقية الحالات فيكون مقدر DBL افضل كونه يمتلك اقل متوسط مربعات خطا .

الاستنتاجات

1. ان مقدر DBAUL افضل من مقدر DBL لأغلب حجوم العينات وقيم التباين وقيم معامل الارتباط في
2. الانموذج شبه المعلمي الاول.
3. اظهر استعمال مقدر DBL تقدما واضحا على مقدر DBAUL وذلك في اغلب حجوم العينات وقيم التباين وقيم معامل الارتباط في انموذج الانحدار شبه المعلمي الثاني .
4. ان مقدر DBL افضل من مقدر DBAUL لأغلب حجوم العينات وقيم التباين وقيم معامل الارتباط في
5. الانموذج شبه المعلمي الاول.

DOI : http://dx.doi.org/10.24237/djps.1304.295B
P-ISSN: 2222-8373
E-ISSN: 2518-9255

6. ان قيم معيار متوسط مربعات الخطأ تزداد عند زيادة قيم التباين(العلاقة طرديه) وتقل عند زيادة حجم
7. العينة(العلاقه عكسيه).
8. عند زيادة قيم التباين فان قيم معيار متوسط مربعات الخطأ تزداد وتقل عند زيادة حجم العينة.
9. نستنتج ومن كلا النموذجين ولجميع حجوم العينات وقيم التباين ومعامل الارتباط ان افضل مقدر هو DBAUL كونه يمتلك اقل متوسط مربعات خطا .

المصادر

1. الصفاوي ،صفاء يونس ومتى،نور صباح.(2011)، " تقدير دوال الانحدار باستخدام بعض اساليب التمهيد " المجلة العراقية للعلوم الإحصائية العدد 20 .
2. حسين ، سجى محمد ويوسف، حنين مراد(2014) "مقارنة الطريقة المقترحة (AUGRR) مع الطرائق المتحيزة لتقدير انحدار الحرف العام بوجود التعدد الخطي"،مجلة كلية الرافدين الجامعة ،العدد 37 .
3. حسين،سجى محمد وحسن ،ارشد حميد (2016)" مقارنة الطريقتين المقترحتين(DBMJRR) و(DBJRR) في انموذج الانحدار شبه المعلمي باستعمال المحاكاة" ،وقائع مؤتمر العلمي الدولي الرابع عشر للجمعية العراقية للعلوم الاحصائية .
4. كاظم ، أموري هادي و مسلم ، باسم شليبة (2002) ، " القياس الاقتصادي المتقدم النظرية والتطبيق"، مكتبة دنيا الامل ، بغداد.
5. A.Yatchew ,Semiparametric regression for the applied econometrican. Cambridge University press, Cambridge,2003.
6. A.Yatchew A. An elementary estimator of the partial linear model. Econom Lett. 1997;57:135–143.
7. A.yatchew,differencing method in nonparametric regression ,simple techniques for the applied econometricician.(1999).
8. Akdeniz, Esra. Akdeniz, Fikri, - Hu, Hongchang, Efficiency of a Liu-type estimator in semiparametric regression models, Journal of Statistical Computation and Simulation.(2008).
9. Akdeniz, Esra. Akdeniz, Fikri. Liu-type estimator in semiparametric regression models regression models , Journal of Statistical Computation and Simulation , (2010).
10. Akdeniz, Esra. Akdeniz, Fikri. New difference-based estimator of parameters in semiparametric.(2011).

11. Akdeniz, F., Kacıranlar, S." On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE". Commun. Statist. Theor. Meth. 24:1789–1797. (1995).

12. Alheety, M. I. and B. M. G. Kibria." On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors". Surveys in Mathematics and its Applications, 4, 155-167. (2009).

13. G. Tabakan, F. Akdeniz, Difference-based ridge estimator of parameters in partial linear model, Statistical Papers 51 ,357–368.(2010)

14. Hardle. W. "Applied nonparametric regression" Gambridg- MA:cambirg university press. (1990).

15. Härdle.W, H. Liang, J. Gao, Partially Linear Models, Physika Verlag, Heidelberg, (2000).

16. Hu, Hongchang, Ridge estimation of a semiparametric regression model, Journal of Computational and Applied Mathematics,(2005),pp(215-222).

17. J. Fan andY.Wu, Semiparametric estimation of covariance matrices for longitudinal data, J. Amer.(1997).

18. Liu, Kejian " A new class of biased estimate in linear regression". Communications in Statistics–Theory and Methods 22, 393–402. (1993).

19. L. Brown, M. Levine, Variance estimation in nonparametric regression via the difference sequence method, Annals of Statistics 35 ,(2007), 2219–2232

20. M. Lokshin, Semi-parametric difference-based estimation of partial linear regression models, Stata J. 6 (3) ,(2006), 377–383.

21. McDonald GC, Galarneau DI. AMonte Carlo evaluation of some Ridge-type estimators. J Amer Statist Assoc. (1975);70:407–416

22. M. Arashi · T. Valizadeh." Performance of Kibria's methods in partial linear ridge regression model". Springer-Verlag Berlin Heidelberg (2014).

23. Roozbeh, Mahdi, Shrinkage ridge estimators in semiparametric regression models, Journal of Multivariate Analysis 136 ,(2015), 56–74.

24. wu.jibo."Performance of the difference-based almost unbiased Liu estimator in partial linear model" Journal of Statistical Computation and Simulation. .(2016)