

مقارنة بعض طرائق تقدير أنموذج دالة الانحدار اللامعلمي باستعمال المحاكاة

م.د. دريد حسين بدر*

م.د. هيفاء طه عبد**

المستخلص :

في هذا البحث تم تناول مقدرات أنموذج دالة الانحدار اللامعلمي من خلال تناول المقدر الخطي الموضوعي اللامعلمي (Local Linear Nonparametric Estimator) ، والمقدر التربيعي الموضوعي اللامعلمي (Local Quadratic Nonparametric Estimator) ، باستعمال معيار العبور الشرعي (Leave – One -Out Cross Validation) (CV) - Validation) ، وتناول إيجاد مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS) باستعمال المعيار الحصين (RCV) ، وكان لا بد من استعمال المحاكاة كأداة مثلى لدراسة حجوم عينات ولتوزيعات مختلفة للخطأ العشوائي منها الحالة الطبيعية وحالة تلوث البيانات بقيم شاذة أو ملوثة ولمستويات تباين مختلفة ولنماذج مختلفة ، للمقارنة بين طرائق التقدير المدروسة المعتمدة ومعرفة أفضل الطرائق ، وقد أثبتت النتائج أفضلية مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS) باستعمال المعيار الحصين (RCV) من خلال حصوله على أقل قيمة لمعيار متوسط المعدل النسبي للقيم المطلقة للأخطاء (AMAPE) .

الكلمات المفتاحية : أنموذج الأنحدار اللامعلمي ، المقدر الخطي الموضوعي اللامعلمي (LLNE) ، المقدر التربيعي الموضوعي اللامعلمي (LQNE) ، مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS) ، معيار العبور الشرعي (CV) ، المعيار الحصين (RCV) ، عرض الحزمة (h) ، معيار متوسط المعدل النسبي للقيم المطلقة للأخطاء (AMAPE) .

Comparison of some methods of estimating non-parameter regression model using simulation

Abstract

In this paper, the abilities of the non-parametric regression function model were studied by taking the Local Linear Nonparametric Estimator and the Local Quadratic Nonparametric Estimator using the Leave - One-Out Cross - Validation criterion. The robust Smoothing Spline Estimator (RSS) was developed using the RCV, and simulation had to be used as an optimal tool to study sample sizes and different random error distributions, including normal and contaminated data with abnormal or contaminated values and varying levels. Different and for different models, to compare Av DOCUMENTS assessment studied the approved and knowledge of the best methods, have demonstrated results destined preference primaries slides fortified (Robust Smoothing Spline) (RSS) using standard hippocampus (RCV) by receiving a lower value to the standard average relative rate of absolute values of errors (AMAPE)).

1.1 المقدمة : (Introduction)

تكتسب الطرق الإحصائية اللامعلمية جاذبية خاصة لدى العاملين في مجال البحث العلمي، وهي من الطرائق الإحصائية الاستدلالية التي يمكن استعمالها للتوصل إلى استنتاجات بشأن المجتمع في ضوء العينة بغض النظر عن نوع التوزيع النظري لذلك المجتمع وتكون أكثر مرونة في التعامل مع البيانات وتتعامل مع طرائق لا تحتاج إلى فرضيات أو على أقل تقدير تحتاج إلى فرضيات أقل من الأسلوب المعلمي خاصة فيما يتعلق بعدم الحصول على معلمات أو أن بيانات العينة تكون وصفية أو رتبوية أو نسبية وهي موجودة في الكثير من حالات التطبيق ، وعلى الرغم من أن نماذج الانحدار اللامعلمي هي اضعف وصفاً من نماذج الانحدار المعلمي إلا أنها في الوقت نفسه تحتاج إلى قيود أو شروط أقل على شكل الدالة التي تجعل أنموذج الانحدار خطي أو لاخطي ، إحصائي أو لوجستي ... الخ ، وان التطبيقات اللامعلمية لها القابلية على كشف تركيبة البيانات خصوصاً عندما يصغر حجم العينة من النماذج المعلمية للاستدلال على البيانات ، وهذا الأمر تحديداً هو الذي جعل من نماذج الانحدار اللامعلمية أداة مرغوبة جداً لدى الباحثين لكون إن البيانات الفعلية ليست دوماً لديها مواصفات ملائمة ، ولكن في حالة وجود بيانات غير ملائمة فيكون الاهتمام عادة بالاعتماد على إعطاء وصف عام للعلاقة وليس دراسة تفاصيل العلاقة الدقيقة ، مما دعا إلى طريقة التفكير بنماذج الانحدار اللامعلمي (3) ، (4).

2.1 هدف البحث:

يتمثل الهدف من هذا البحث بالمقارنة بين الطرائق في الأنموذج اللامعلمي ، للحصول على أفضل تقدير لاختبار الطريقة الأفضل والأكثر ملائمة للانموذج ، وذلك من خلال استعمال المحاكاة كأداة مثلى لدراسة حجوم عينات وتوزيعات مختلفة للخطأ العشوائي منها الحالة الطبيعية وحالة تلوث البيانات بقيم شاذة أو ملوثة والمستويات تباين مختلفة ولنماذج مختلفة ثم تعميم الاستنتاجات حول مدى فعالية طريقة التقدير المستعملة وذلك باستعمال معيار (AMAPE) .

2. الجانب النظري:**2.1 نماذج الانحدار اللامعلمية (Non-Parametric Regression Models)**

والشكل العام لهذه النماذج هو (2) ، (3) :

$$2-1) \quad \dots (Y_i = m(x_i) + \varepsilon_i), \quad i = 1, 2, \dots, n$$

إذ أن :

$m(\cdot)$: دالة التمهيد غير المعروفة .

Y_i : تمثل قيمة متغير الاستجابة.

x_i : تمثل قيمة المشاهدات للمتغير التوضيحي (التفسيري) وهو متجه من الدرجة $1 \times P$.

ε_i : قيمة الأخطاء العشوائية ذوات التوزيع الطبيعي مستقلة ومتماثلة التوزيع $(0, \sigma^2) \sim N$.

ومن الأنموذج (2-1) فإنه ليس هنالك حاجة لتقدير معالم هي عبارة عن نقاط ، وبالتالي نحتاج إلى تمهيد (Smoothing) وهو أداة تستخدم لوصف اتجاه وتأثير المتغير التوضيحي في متغير الاستجابة ، وهذه الأداة يمكن تمثيلها كدالة بدلالة واحد أو أكثر من المتغيرات التوضيحية لمنحني الأنموذج $m(x)$ ، وهو أسلوب بياني لامعلمي يمكن استخدامه لتقدير أي دالة إحصائية ، إذ أن هدف أساليب التمهيد هو تقريب دالة الانحدار اللامعلمي التقريبية إلى دالة الانحدار اللامعلمية الحقيقية ، ثم تستغل البيانات بعد ذلك في تحديد ما يمثل دالة الانحدار الحقيقية ، إذ إن تحليل الانحدار اللامعلمي هو الأفضل في حالة غياب أي معلومات عن الدالة الحقيقية ، أي أن تقنيات الانحدار اللامعلمي تعول كثيراً على البيانات في تحديد تلك الدالة أكثر من اعتماده على الافتراضات ، للحصول على معلومات

عن m بدلا من الصيغة الدالية وفي بعض الحالات ينتهي تحليل الانحدار اللامعلمي بتوفيق منحني للبيانات دون تحديد شكل معين للدالة ، إذا ما قورنت بنظيراتها المعلمية ، إذ إن الانحدار اللامعلمي معروف جيدا بأنه يجعل البيانات أكثر تكيفا ، وأقل تقييدا من المنهج المعلمي ، لأنه يتعامل مع طرائق لاتحتاج إلى فرضيات أو في أقل تقدير تحتاج إلى فرضيات أقل وخاصة فيما يتعلق بعدم الحصول على المعلمات أو أن بيانات العينة تكون رتبية أو نسبية أو وصفية ، لذلك برز الانحدار اللامعلمي كبديل لمعالجة البيانات. (14) ، (15) .

2.2 المقدرات اللامعلمية لأنموذج دالة الانحدار اللامعلمي .

(Nonparametric Estimation of Nonparametric regression function)

تحدث بعض المشاكل الأساسية جداً بالانحدار اللامعلمي ، وتوجد أساليب للتعامل مع هذه المشكلة ولكنها جميعها تعتمد بشكل حاسم على تناول طرائق تقدير للدالة بالاعتماد على إيجاد عرض الحزمة الأمثل للحصول على عرض الحزمة المصحح للارتباط باستعمال طرائق اختيار عرض الحزمة ، وتبين بان شكل دالة الوزن الدالة اللبية (Kernel) مهم جداً ، وقبل تناول المقدرات اللامعلمية لأنموذج دالة الانحدار اللامعلمية لابد من تناول أهم الدوال (Kernel) اللبية المستعملة وكما في الجدول الآتي (1) ، (5) :

جدول (1)

أهم دوال (Kernel) اللبية المستعملة

رقم الدالة	Kernel function	K (x)	حدود الدالة	رسم الدالة
1	Epanechnikov (or quadratic)	$(1-x^2) \frac{3}{4}$	[-1 , 1]	
2	Gaussian (or normal)	$(2\pi)^{-1/2} \exp(-x^2/2)$	$(-\infty, \infty) I$	

Sores : Blanchet. J and Wadsworth. J. (2012)

وان دالة (Kernel) اللبية يجب إن تحقق الشروط الآتية (11) ، (15) :

1. $K(x)$ دالة مستمرة وذات قيم حقيقية موجبة.

$$\int_{-\infty}^{\infty} K(x) dx = 1 \quad 2.$$

3. $K(x)$ دالة متماثلة حول الصفر، قيمة العزم الأول لدالة (Kernel) اللبية .

$$(\mu_1(K) = \int x K(x) dx = 0) \text{ ، وقيمة العزم الثاني } (\mu_2(K) = \int x^2 K(x) dx < \infty) .$$

$$\int K^2(x) dx < \infty \quad 4.$$

3.2 مقدر متعدد الحدود الموضوعي اللامعلمي (Local polynomial regression)

أن أحد الأساليب الأحصائية المستعملة لتقدير منحني أنموذج دالة الأنحدار اللامعلمي (2-1) للوصول إلى منحني تقدير يملك كمية تمهيد مثلى فضلا عن بروزه كمقدر يراد به مطابقة البيانات ، وخاصة في حالة البيانات غير الخطية (8) ، وسيتم التطرق للمقدر لأن مقدر ناداريا واتسون (Nadarya – Watson) الذي يعتبر من دوال (Kernel)

وهو حالة خاصة من مقدر (Local polynomial regression) ، وبفرض $(p+1)^{th}$ هي مشتقة دالة الانحدار (التمهيد) $m(x)$ عند النقطة x_0 موجودة ، وبتقريب دالة الانحدار (التمهيد) $m(x)$ باستخدام متعدد الحدود من الدرجة p ، وباستخدام توسيع تايلور نحصل على :

$$m(x) \approx m(x_0) + m'(x_0)(x-x_0) + \frac{m''(x_0)}{2!}(x-x_0)^2 + \dots + \frac{m^{(p)}(x_0)}{p!}(x-x_0)^p \quad \dots(2-2)$$

وباستخدام متعدد الحدود الموضعي للـ x_0 بتقليل انحدار المربعات الصغرى الموزونة نحصل على : (13)

$$\min_B \sum_{k=1}^n \left\{ y_i - B_0 - B_1(X_k - x_0) - \dots - B_p(X_k - x_0)^p \right\}^2 K_h(X_k - x_0) \quad \dots(2-3)$$

إذ إن :

h : عرض الحزمة .

K : الدالة اللبية (Kernel) .

$K_h(\cdot) = K(\cdot/h) / h$: تحديد أوزان إلى كل نقطة بيانات .

وبالتالي نحصل على مقدرات المربعات الصغرى الموزونة وبالصيغة التالية :

$$\dots(2-4) \hat{B}(x) = (X'WX)^{-1} X'WY$$

إذ إن :

$$X = \begin{pmatrix} 1 & (X_1 - x_0) & \dots & (X_1 - x_0)^p \\ 1 & (X_2 - x_0) & \dots & (X_2 - x_0)^p \\ \vdots & \vdots & \dots & \vdots \\ 1 & (X_n - x_0) & \dots & (X_n - x_0)^p \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$W = \begin{pmatrix} K_h(X_0 - x_1) & 0 & \dots & 0 \\ 0 & K_h(X_0 - x_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K_h(X_0 - x_n) \end{pmatrix}$$

وبالتالي فإن مقدر متعدد الحدود الموضعي (Local polynomial regression) لدالة الانحدار (التمهيد) m تكون كالاتي :

$$\dots(2-5) \hat{m}_{p,h}(X) = \hat{B}_0(X)$$

وان $m(X) \approx B_0(X)$ ، فعندما تكون درجة متعدد الحدود $(p = 0)$ نحصل على مقدر الثابت الموضعي (ناداريا واتسون) (Local constant estimator) (Nadaraya-Watson) ، وعندما تكون درجة متعدد الحدود $(p = 1)$ فنحصل على المقدر الخطي الموضعي (ناداريا واتسون) (Local Linear) (Nadaraya-Watson estimator) ، وعندما تكون درجة متعدد الحدود $(p = 2)$ فنحصل على المقدر التربيعي الموضعي (ناداريا واتسون) (Local Quadratic estimator) (Nadaraya-Watson) ، وعندما تكون درجة متعدد الحدود $(p = 3)$ فنحصل على مقدر التكعيبي الموضعي (ناداريا واتسون) (Local Cubic estimator) (Nadaraya-Watson) ، الخ..... وعندما تكون درجة متعدد الحدود (p) فنحصل على مقدر متعدد الحدود الموضعي

(Local polynomial regression) ، وسيتم في هذا البحث تناول كل من ، المقدر الخطي الموضوعي اللامعلمي (ناداريا واتسون) (Local linear nonparametric estimator) ، و المقدر التربيعي الموضوعي اللامعلمي (ناداريا واتسون) (Local Quadratic estimator) .

1.3.2 معيار تقاطع الشرعية (العبور الشرعي) (CV) (Leave – One-Out Cross - Validation)

بعد أن يتم وصف طريقة التقدير للأنموذج الانحدار اللامعلمية للبيانات ، هناك مشكلة خطيرة هي في اختيار عرض المعلمة التمهيدية (h) ، وبالتالي نحن بحاجة لايجاد طريقة سريعة لتحديد دقيق لمعلمة التمهيد ، ومعالجة لمشكلة اختيار عرض الحزمة ، إذ سيتم تناول معيار تقاطع الشرعية (العبور الشرعي) في حالة Leave – one-out cross validation) ويسمى ايضاً (C V) ، وهو من أشهر الطرق التي تقيس جودة الأداء لمقدرات دالة الانحدار اللامعلمية ، ولكي نتجنب اختيار قيم المعالم بواسطة التجزئة والخطأ ويوصف بالصيغة التالية (13) :

$$LCV(h) = \frac{1}{n} \sum_{i=1}^n (\hat{m}_{(-i)}(x_i) - y_i)^2 \quad \dots(2-6)$$

إذ أن :

$\hat{m}_{(-i)}(x_i)$: دالة التقدير التمهيدية لمقدي الخطي والتربيعي الموضوعي باستبعاد (2L+1) من مشاهدات المتغير التوضيحي ، (L=0) ، وبالتالي فان قيمة (h) هي التي تقابل اقل (C V) .

2. 4 مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS)

إذ إن الشرائح التمهيدية الحصينة (Robust Smoothing Spline) تعرف من إيجاد دالة التمهيد \hat{m} ، ويكون عن طريق استعمال دالة وزن بطريقة (M) بالصيغة (2-8) وكالاتي (13) :

$$\frac{1}{n} \sum_{i=1}^n \psi \left| y_i - \hat{m}(x_i) \right| + \lambda \int (\hat{m}'')^2 \quad \dots (2-7)$$

إذ إن :

$$\psi(x) = \begin{cases} \frac{x^5}{3c} & , \text{if } |x| \leq c \\ \left| x - \frac{c}{3} \right| & , \text{if } |x| > c \end{cases} \quad \dots (2-8)$$

$$c = 10^{-16} sd(x) ,$$

$\psi(x)$: دالة الوزن بطريقة (M) والتي تمتاز بخاصية (Scale invariant) المستعملة لإيجاد مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) وتتطلب تحقق شرط التماثل .

$$\hat{m} , \hat{m}'' : \text{ دالة تمهيد } , \int |\hat{m}''(x)| < \infty .$$

λ : معلمة التمهيد > 0 ، وان أكبر قيمة لها تمثل تقديرات التمهيد .

وان الحد الأول للصيغة (2-7) هو الحد المطلق لمجموع مربعات البواقي الجزائية بوجود دالة الوزن بطريقة (M) ، بينما الحد الثاني لنفس الصيغ يسمى حد الجزاء وهو يعمل على إيجاد التوازن بين تقدير المنحنى ومطابقتة للأنموذج ومن ثم التخلص من التقدير المتذبذب (5) .

لنحصل على قيم \hat{y}_i الملائمة لتقدير نموذج دالة الانحدار اللامعلمية (2-1) لإيجاد مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) وجعله أكثر حصانة ومقاومة للقيم الشاذة (الملوثة) (8).

1.4.2 استعمال معيار (RCV) لمقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline):

سيتم تناول طريقة حصينة لاختيار معلمة التمهيد (λ) ، لإيجاد مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) فتم استعمال معيار حصين لاختيار معلمة التمهيد (Wang and Scott,1994) ، (Robust Cross - Validation) (RCV) والذي يوصف بالصيغة الأتية (13) :

$$RCV(\lambda) = \frac{1}{n} \sum_{i=1}^n \psi(y_i - \hat{m}_{\lambda}^{(i)}(x_i))^2 \quad \dots (2-9)$$

إذ أن :

$\psi(x)$: دالة الوزن المذكورة في الفقرة (4.2) بالصيغة (2-8) لإيجاد مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) .

$\hat{m}_{\lambda}^{(i)}(x_i)$: دالة التمهيد لمقدر الشرائح التمهيدية الحصينة باستبعاد مشاهدة من مشاهدات المتغير التوضيحي ، ومن ثم فإن قيمة (λ) هي التي تقابل اقل (RCV) .

5.2 معيار المقارنة

لغرض إيجاد أفضل طريقة لتقدير لأنموذج دالة الانحدار اللامعلمي بالصيغة (2-1) سيتم استعمال المعيار المذكور لاحقاً والتي تناولتها أغلب البحوث المنشورة والتي تضمنت موضوع الدراسة ، إذ أن تناقص قيم هذا المعيار يشير بالتأكيد إلى جودة التقدير وهي كما يأتي:

2.5.1 معيار متوسط المعدل النسبي للقيم المطلقة للأخطاء AMAPE

هو عبارة عن معدل متوسط المجموع الكلي للقيم المطلقة لأنحرافات القيم التقديرية عن القيم الحقيقية مقسوماً على القيم الحقيقية وتمثل قيمته بما يلي : (6) ، (7)

$$AMAPE = \frac{1}{R} \sum_{i=1}^R MAPE(i) \quad (2-10)$$

: يمثل عدد مرات تكرار التجربة i . (R) إذ أن :

: تمثل متوسط الخطأ النسبي المطلق في التجربة i وتكون صيغته كما يأتي : ($MAPE(i)$)

$$MAPE(i) = \frac{1}{n} \sum_{i=1}^n \left| \frac{\theta - \hat{\theta}}{\theta} \right|$$

n يمثل حجم العينة :

$\hat{\theta}$ تشير الى المعلمات المقدرة بواسطة المحاكاة :

θ تشير الى القيم الحقيقية المحتسبة :

3. الجانب التجريبي:

يتم اللجوء إلى أسلوب المحاكاة في الحالات التي يصعب فيها الحصول على بيانات واقعية ودقيقة وذلك بهدف محاكاة عدد كبير جداً من الحالات الافتراضية والتي يمكن أن تحصل في الواقع العملي ، وكذلك عندما يعجز البرهان الرياضي عن بيان أفضلية طرائق معينة ، إذ تم تنفيذ تجارب المحاكاة وبذلك سيكون لدينا ثلاث حجومات للعينات صغيرة ($N=50$) ومتوسطة ($N=100$) وأخيراً كبيراً ($N=200$) ، وبتكرارات ($\text{Replicates}=1000$) لكل تجربة محاكاة للحصول على نتائج متنسقة ولحالات متنوعة لتوزيع الخطأ العشوائي وعند الحالة الطبيعية ولحالات تلوث (10%) و (20%) وكما يأتي :

1-3 توليد المتغيرات والعينة المستخدمة:

توليد المتغير التوضيحي X بتوزيع منتظم $X \sim U(0,1)$ بالاستناد إلى طريقة *Box-Muller* والتي تعتمد على أسلوب توليد متغيرين عشوائيين U_1, U_2 يتبعان التوزيع المنتظم $U(0,1)$ ثم يتم تحويل المتغيرين إلى متغيرين عشوائيين مستقلين Z_1, Z_2 يتبعان التوزيع الطبيعي القياسي وفقاً للصيغة :

$$Z_1 = -2(\ln U_1)^{1/2} \cos(2\pi U_2)$$

$$Z_2 = -2(\ln U_1)^{1/2} \sin(2\pi U_2)$$

ولتحويل المتغيرات من التوزيع الطبيعي القياسي إلى التوزيع الطبيعي بمتوسط صفر وتباين σ^2 يستخدم التحويل $U=0+\sigma^2 Z$

الخطأ العشوائي ويتم توليده ليتوزع توزيعاً طبيعياً بمتوسط صفر وتباين σ^2 أي $\epsilon \sim N(0, \sigma^2)$ وقد تم تناول ثلاثة مستويات من التباين لكل دالة من لدوال التي ستذكر لاحقاً وهي تباين عال (*High Noise*) حيث:

$$\sigma = 1/2 \times \text{Function Range}$$

حيث (Medium Noise) تباين متوسط

$$\sigma = 1/4 \times \text{Function Range}$$

حيث (Low Noise) وتباين واطئ

$$\sigma = 1/8 \times \text{Function Range}$$

أما التوزيع الأخر للخطأ العشوائي فهو التوزيع الملوث ويستخدم في حالة تلوث البيانات بقيم شاذة إذ يتم توليد المتغيرات التي تتبع التوزيع الملوث بتوليد متغيرات تتبع التوزيع الطبيعي بنسبة $(1-\epsilon)$ من البيانات بالأسلوب السابق نفسه وكذلك النسبة (ϵ) من البيانات يتم توليدها بأسلوب توليد متغيرات التوزيع الطبيعي نفسه لكن بتباين $\sigma^2 \lambda^2$ أي

$$\epsilon \sim [(1-\epsilon) N(0, \sigma^2) , \epsilon N(0, \lambda^2 \sigma^2)] \quad \lambda^2 > 1$$

يتم احتساب متغير الاستجابة y مباشرة من خلال النماذج المستخدمة في تجارب المحاكاة وذلك باستخدام المتغيرات التوضيحية التي تم توليدها في الفقرة I، مضافاً إليها الأخطاء العشوائية المولدة في الفقرة II، ولكل أنموذج من النماذج المدروسة.

2-3 النماذج المستخدمة في تجارب المحاكاة :

هناك عدد كبير من النماذج التي لا يمكن تناولها جميعها في آن واحد في تجارب المحاكاة نظراً لما تتطلبه من الجهد والمساحة الواسعة وقد حاولنا قدر المستطاع أن تكون النماذج المتناولة متنوعة لتلائم جميع الحالات منها الخطية وغير الخطية إضافة لدالة الانحدار غير المتجانسة (*Heterogeneous Regression Function*) والأنموذج الخاص (التغير الحيزي الواطئ في التذبذب) (*Low Spatial Variability of Oscillations*) والأنموذج الخطي من الدرجة الرابعة وقد أخذت النماذج من بحوث منشورة وهي كالاتي :

1 . دالة التمهيد الخطية بوجود حد ثابت وصيغتها (12)

$$m(x) = x - 3$$

2 . دالة التمهيد غير الخطية وصيغتها (10)

$$m(x) = \cos(2x)$$

3 . دالة التمهيد الأسية غير المتجانسة وصيغتها :

(9) (Heterogeneous Smoothing Function)

$$m(x) = \exp(-18(x-0.5)^4)$$

وبالاعتماد على دوال التمهيد المذكورة أعلاه يمكن كتابة صيغة لأنموذج التأثيرات العشوائية اللامعلمية للبيانات المزدوجة والتي ستجري محاكاتها وكالاتي :

1 . الأنموذج الخطي بوجود حد ثابت وصيغته :

$$y = (x-3) + \varepsilon$$

2 . الأنموذج غير الخطي وصيغته :

$$y = \cos(2x) + \varepsilon$$

3 . الأنموذج غير الخطي الخاص بدالة التمهيد الأسية غير المتجانسة وصيغته :

$$y = \exp(-18(x-0.5)^4) + \varepsilon$$

أما دوال Kernel فهي كما معطاة في الجدول (1) .

3.3 تنفيذ تجارب المحاكاة لطرائق الانحدار اللامعلمي لأنموذج الانحدار اللامعلمي:

لكل أنموذج من نماذج الأنحدار اللامعلمي والتي تم التطرق لها في الفقرة (3.2) تم القيام بما يأتي :

أولاً : توليد المتغيرات :

يتم توليد المتغير التوضيحي (التفسيري) X ليتوزع توزيعاً منتظماً بمتوسط (0) وتباين (1) ، مع توليد أخطاء عشوائية تتبع توزيعاً طبيعياً قياسياً ، فضلاً عن توزيع ملوث للخطأ العشوائي وهو التوزيع الملوث وبنسبة (0.20 ، 0.10 ، 0) .

ثانياً : طرائق التقدير (Estimation Methods) :

بعد أن يتم توليد البيانات ولكل حالة وللمشكلة قيد الدراسة ، يتم إيجاد المقدرات اللامعلمية لأنموذج الأنحدار اللامعلمي ، باستعمال المقدرين الخطي والتريبيعي الموضوعي اللامعلمي (ناداريا واتسون) في الفقرة (2-3) ، وتم تطبيق طريقة مقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) عن طريق توظيف دالة وزن بطريقة (M) المشار له بالفقرة (4.2) .

ثالثاً : استعمال معايير اختيار معلمة التمهيد (عرض الحزمة) :

إما بالنسبة لتحديد عرض الحزمة لطرائق التقدير ، فتم استعمال معيار العبور الشرعي (CV) لإيجاد معلمة التمهيد (h) للمقدين الخطي والتريبيعي الموضوعي اللامعلمي المشار له بالفقرة (1.3.2) ، وفيما يخص طريقة مقدر الشرائح التمهيدية الحصينة فتم تناول استعمال معيار (RCV) المشار له بالفقرة (1.4.2) لإيجاد معلمة التمهيد (λ) .

3.3 تقدير المعاملات وتحليل النتائج:

نفذت جميع التجارب باستعمال برنامج الحزم الجاهزة (R.Version.3.5.1 (2/7/2018)) وتم تحليل نتائج المحاكاة لكل أنموذج ، وقد تم تفسير النماذج الثلاث بالاعتماد على استعمال معيار (AMAPE)، لبيان الطريقة الأكثر ملائمة للأنموذج .

وكالاتي :

جدول (2)

المعيار *AMAPE* المعتمد في طريقة التقدير الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي عند الحالة الطبيعية للأنموذج الأول ولجميع أحجام العينات ($N=50, N=100, N=200$) ولجميع مستويات التباين

The Method	Sample Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
		AMAPE	AMAPE	AMAPE
LLNE . h by using CV	50	1.036612	1.036759	1.048422
	100	1.019645	1.019701	1.028433
	200	1.009105	1.009108	1.009111
LQNE . h by using CV	50	1.027237	1.034027	1.035483
	100	1.016847	1.017163	1.017890
	200	1.007407	1.007813	1.007930
Robust smoothing λ by Spline using RCV	50	0.7284275	0.7882291	0.7888791
	100	0.5358899	0.5368923	0.5436758
	200	0.2534092	0.2590671	0.2592563

جدول (3)

المعيار *AMAPE* المعتمد في طريقة التقدير الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.10) للأنموذج الأول ولجميع أحجام العينات ($N=50, N=100, N=200$) لجميع مستويات التباين

%	The Method	Sample Size (N)	$2\sigma=1$	$\sigma=1/4$	$8\sigma=1$
			AMAPE	AMAPE	AMAPE
10%	LENE . h by using CV	50	1.016199 1.035776	1.017960 1.043227	1.043307
		100	1.007603	1.009517	1.018300
		200			1.009563
	LQNE . h by using CV	50	1.012065 1.024919	1.013459 1.029553	1.033264
		100	1.005466	1.008070	1.015375
		200			1.008491
λ by using RCV-Robust smoothing Spline	50	0.5382813 0.7849568	0.5406585 0.7894426	0.5483791 0.7902373	
	100				
	200	0.2533728	0.2549455	0.2580516	

جدول (4)

المعيار $AMAPE$ المعتمد في طريقة التقدير الأفضل لإيجاد تقدير نموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.20) للأنموذج الأول ولجميع أحجام العينات ($N=50, N=100, N=200$) لجميع مستويات التباين

ξ	The Method	Sample. Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
			AMAPE	AMAPE	AMAPE
20 %	LLNE . h by using CV	50	1.017630 1.032114	1.045712	1.049169
		100		1.021949	1.022949
		200	1.009727	1.010254	1.010801
	LQNE . h by using CV	50	1.012175 1.025274	1.036970	1.042545
		100		1.018680	1.019642
		200	1.008152	1.008603	1.009214
	Robust smoothing λ by 'Spline using RCV	50	0.7836203	0.7863497	0.787618
		100	0.5392421	0.5399122	0.5417697
		200	0.2545051	0.2569546	0.2588889

الأنموذج الأول :

يلاحظ من نتائج الجداول (2)،(3)،(4) على التوالي للأنموذج الأول المتضمنة نتائج طرائق التقدير اللامعلمية الاعتيادية عند الحالة الطبيعية وحالة التلوث للبيانات بنسبة ($\xi = 0.10, 0.20$) قيم معيار (AMAPE) سجل انخفاضاً ترافق مع زيادة حجم العينة بين مستويات الانحراف المعياري لحد خطأ للأنموذج مما يشير إلى مدى تأثير المقدرات بقلّة حجم العينة وازدياد قدرة ودقة تلك المقدرات عند تزايد حجم العينة , وكذلك سجل زيادةً عند مستوى التباين العالي ($1 \setminus 2$) و $\sigma = (1 \setminus 4)$ والمتوسط مقارنة مع مستوى التباين الواطئ ($1 \setminus 8$) وجميع أحجام العينات مما يشير إلى تأثير تلك المقدرات بتشتت البيانات وقلّة قدرتها عند تزايد تشتت البيانات ، في حين أشارت النتائج للجداول أفضلية واضحة وبصورة مطلقة لمقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS) باستعمال المعيار الحصين (RCV) على طرائق التقدير الأخرى المستعملة في تقدير أنموذج دالة الانحدار اللامعلمي ولجميع أحجام العينات ولجميع الحالات ومستويات الانحراف المعياري لحد الخطأ للأنموذج الأول .

جدول (5)

المعيار $AMAPE$ المعتمد في طريقة التقدير الأفضل لإيجاد تقدير نموذج دالة الانحدار اللامعلمي عند الحالة الطبيعية للأنموذج الثاني ولجميع أحجام العينات ($N=50, N=100, N=200$) ولجميع مستويات التباين

The Method	Sample. Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
		AMAPE	AMAPE	AMAPE
LLNE . h by using CV	50	1.029921	1.033944	1.034818
	100		1.018164	1.018336
	200	1.008377 1.017589	1.009180	1.008236
LQNE . h by using CV	50	1.014571 1.030112	1.029660	1.030248
	100		1.016261	1.017030
	200	1.007944	1.008715	1.007880

Robust smoothing λ by \cdot Spline using RCV	50	0.7854248	0.7764756	0.7878597
	100	0.5337045	0.5405122	0.5434832
	200	0.258834	0.26282	0.2533721

جدول (6)

المعيار AMAPE المعتمد في طريقة التقدير الأفضل لإيجاد تقدير نموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.10) للأنموذج الثاني ولجميع أحجام العينات ($N=50, N=100, N=200$) لجميع مستويات التباين

ξ	The Method	Sample Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
			AMAPE	AMAPE	AMAPE
10 %	LLNE . h by using CV	50	1.021410 1.034998	1.035881	1.033964
		100		1.021532	1.020300
		200	1.008739	1.009186	1.009737
	LQNE . h by using CV	50	1.013050 1.027289	1.028598	1.026431
		100		1.016273	1.016992
		200	1.008311	1.008682	1.008770
	Robust smoothing λ by \cdot Spline using RCV	50	0.791833	0.792824	0.7829028
		100	0.5420644	0.547551	0.5387019
		200	0.2553106	0.2567802	0.2576431

جدول (7)

المعيار AMAPE المعتمد في طريقة التقدير الأفضل لإيجاد تقدير نموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.20) للأنموذج الثاني ولجميع أحجام العينات ($N=50, N=100, N=200$) لجميع مستويات التباين

ξ	The Method	Sample Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
			AMAPE	AMAPE	AMAPE
20 %	LLNE . h by using CV	50	1.020745 1.048723	1.020932 1.049591	1.048117
		100			1.019917
		200	1.008254	1.008440	1.011085
	LQNE . h by using CV	50	1.020091 1.043982	1.020273 1.044592	1.042334
		100			1.019490
		200	1.007831	1.007911	1.010584
	*Robust smoothing Spline λ by using RCV	50	0.5394408 0.7897926	0.547551 0.795186	0.5386761 0.7860522
		100		0.2583813	
		200	0.252381		0.2588818

الأنموذج الثاني :

يلاحظ من نتائج الجداول (5)،(6)،(7) على التوالي للأنموذج الثاني المتضمنة نتائج طرائق التقدير اللامعلمية الاعتيادية عند الحالة الطبيعية وحالة التلوث للبيانات بنسبة ($\xi = 0.10, 0.20$) قيم معيار (AMAPE) سجل انخفاضاً ترافق مع زيادة حجم العينة بين مستويات الانحراف المعياري لحد خطأ للأنموذج مما يشير إلى مدى تأثير المقدرات بقلّة حجم العينة وازدياد قدرة ودقة تلك المقدرات عند تزايد حجم العينة , وكذلك سجل زيادةً عند مستوى التباين العالي ($\sigma = (1 \backslash 2)$) و المتوسط ($\sigma = (1 \backslash 4)$) مقارنة مع مستوى التباين الواسع ($\sigma = (1 \backslash 8)$) ولجميع أحجام

العينات مما يشير إلى تأثير تلك المقدرات بتشتت البيانات وقلة قدرتها عند تزايد تشتت البيانات , في حين أشارت النتائج للجداول أفضلية واضحة وبصورة مطلقة لمقدر الشرائح التمهيدية الحصينة (Robust Smoothing) (RSS) (Spline) باستعمال المعيار الحصين (RCV) على طرائق التقدير الأخرى المستعملة في تقدير أنموذج دالة الانحدار اللامعلمي ولجميع حجوم العينات ولجميع الحالات ومستويات الانحراف المعياري لحد الخطأ للأنموذج الثاني .

جدول (8)

المعيار AMAPE المعتمد في طريقة التقدير الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي عند الحالة الطبيعية للأنموذج الثالث ولجميع حجوم العينات (N=50, N=100, N=200) ولجميع مستويات التباين

The Method	Sample. Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
		AMAPE	AMAPE	AMAPE
LLNE . h by using CV	50		1.468067	1.497056
	100	1.225881 1.472534 1.090082	1.226565	1.217088
	200		1.096590	1.096764
LQNE . h by using CV	50		1.385476	1.396203
	100	1.193660 1.392142	1.194367	1.192718
	200	1.071477	1.080169	1.080877
Robust smoothing λ by Spline using RCV	50	0.8842555		0.8984113
	100	0.6373237	0.8794424 0.6382509	0.6371031
	200	0.3087866	0.3125316	0.3136811

جدول (9)

المعيار AMAPE المعتمد في طريقة التقدير الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.10) للأنموذج الثالث ولجميع حجوم العينات (N=50, N=100, N=200) لجميع مستويات التباين

ξ	The Method	Sample. Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
			AMAPE	AMAPE	AMAPE
10 %	LLNE . h by using CV	50	1.467448	1.454424	1.484542
		100	1.210498	1.241338	1.242092
		200	1.103615	1.097581	1.107783
	LQNE . h by using CV	50	1.405050	1.392822	1.417117
		100	1.186711	1.190195	1.193365
		200	1.086024	1.085293	1.086460
	Robust smoothing λ by Spline using RCV	50	0.889126	0.8887806	0.8990171
		100	0.6269548	0.6385882	0.6399534
		200	0.2143571	0.2123657	0.2198074

جدول (10)

المعيار AMAPE المعتمد في طريقة التقدير الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.20) للأنموذج الثالث ولجميع حجوم العينات (N=50, N=100, N=200) لجميع مستويات التباين

ξ	The Method	Sample. Size (N)	$2\sigma=1\backslash$	$\sigma=1/4$	$8\sigma=1/$
			AMAPE	AMAPE	AMAPE
20 %	LLNE . h by using CV	50			1.501646
		100	1.207225 1.497531	1.494000 1.213226	1.204057
		200	1.090581	1.092828	1.097790
	LQNE . h by using CV	50			1.493018
		100	1.178562 1.450880	1.437168 1.185744	1.175908
		200	1.078097	1.078603	1.085371
	Robust smoothing λ by using -Spline RCV	50			0.635512 0.8960128
		100	0.638799 0.8927073	0.8854227 0.6399649	0.2138529
		200	0.2086792	0.2097985	

الأنموذج الثالث :

يلاحظ من نتائج الجداول (8)،(9)،(10) على التوالي للأنموذج الثالث المتضمنة نتائج طرائق التقدير اللامعلمية الاعتيادية عند الحالة الطبيعية وحالة التلوث للبيانات بنسبة ($\xi = 0.10, 0.20$) قيم معيار (AMAPE) سجل انخفاضاً ترافق مع زيادة حجم العينة بين مستويات الانحراف المعياري لحد خطأ للأنموذج مما يشير إلى مدى تأثير المقدرات بقلّة حجم العينة وازدياد قدرة ودقة تلك المقدرات عند تزايد حجم العينة , وكذلك سجل زيادة عند مستوى التباين العالي ($2 \backslash 1$) $\sigma =$ والمتوسط ($4 \backslash 1$) $\sigma =$ مقارنة مع مستوى التباين الواطئ ($8 \backslash 1$) $\sigma =$ ولجميع حجوم العينات مما يشير إلى تأثير تلك المقدرات بتشتت البيانات وقلّة قدرتها عند تزايد تشتت البيانات , في حين أشارت النتائج للجداول أفضلية واضحة وبصورة مطلقة لمقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS) باستعمال المعيار الحصين (RCV) على طرائق التقدير الأخرى المستعملة في تقدير أنموذج دالة الأندحار اللامعلمي ولجميع حجوم العينات ولجميع الحالات ومستويات الانحراف المعياري لحد الخطأ للأنموذج الثالث .

4. الاستنتاجات:

من خلال تنفيذ تجارب المحاكاة وما تم عرضه من نتائج في الجانب التجريبي تم التوصل إلى الاستنتاجات الآتية :

1- لوحظ أن قيم معيار المفاضلة (AMAPE) لمعدل متوسط مربعات الخطأ لكافة تكرارات تجارب المحاكاة لكل حالة من حجوم العينات الثلاثة ومستويات الانحراف المعياري الثلاثة وكذلك للنماذج الثلاث المعتمدة في المحاكاة المتضمنة نتائج طرائق التقدير اللامعلمية الاعتيادية وعند الحالة الطبيعية وحالة التلوث للبيانات بنسبة ($\xi = 0.10, 0.20$) ، كانت صغيرة ومقاربة تقريبا مما يشير إلى تجانس قيم متوسط مربعات الخطأ لكل حالة محاكاة .

2- تبدأ قيم معيار المفاضلة (AMAPE) بالازدياد في اغلب الأحيان عند مستوى التباين العالي ($2 \backslash 1$) $\sigma =$

والمتوسط ($4 \backslash 1$) $\sigma =$ مقارنة مع مستوى التباين الواطئ ($8 \backslash 1$) $\sigma =$ للنماذج الثلاث المعتمدة في المحاكاة المتضمنة نتائج طرائق التقدير اللامعلمية الاعتيادية لكل حالة من حجوم العينات الثلاثة وعند الحالة الطبيعية وحالة التلوث للبيانات بنسبة ($\xi = 0.10, 0.20$) مما يشير إلى تأثير تلك المقدرات بتشتت البيانات وقلّة قدرتها عند تزايد تشتت البيانات .

3- أظهرت النتائج أفضلية واضحة وبصورة مطلقة وبلا اعتماد على قيم معيار المفاضلة (AMAPE) وعند الحالة الطبيعية وحالة التلوث للبيانات بنسبة ($\xi = 0.10, 0.20$) أفضلية واضحة وبصورة مطلقة لمقدر الشرائح التمهيدية الحصينة (Robust Smoothing Spline) (RSS) باستعمال معيار الحصين (RCV) على طرائق

التقدير الأخرى المستعملة في تقدير أنموذج الانحدار اللامعلمي ولجميع حجوم العينات ولجميع الحالات ومستويات الانحراف المعياري الثلاثة وكذلك للنماذج الثلاث المعتمدة في المحاكاة .
4- من الجداول المحاكاة (2) الى (10) نلاحظ إن أكبر القيم لمعدل متوسط مربعات الخطأ (AMAPE) كانت مع مقدر الانحدار التربيعي الموضوعي اللامعلمي (LQNE) باستعمال معيار العبور الشرعي (CV) ، لتأثرها بنقاط الحد مما سبب تلك الزيادة ، وبالتالي عدم كفاءة المقدر .

5 . التوصيات:

- 1- نوصي باستعمال طرائق التقدير في تقدير نماذج الانحدار الشبه المعلمية .
- 2- نوصي باستعمال طرق تقدير لامعلمية لمعالجة مشكلتي عدم التجانس والتعدد الخطي ووجود القيم الشاذة أو الملوثة.
- 3- أجراء دراسات مستقبلية تتناول مقارنة طرائق لامعلمية أخرى مثل طريقة تمهيد الشريحة التكعيبية Cubic (Smoothing Spline) وطريقة الموجة الصغيرة (Wavelet) مع الطرائق التي تم اعتمادها في البحث .
- 4- إمكانية استعمال أكثر من معيار ملائم .

المصادر

1. البكري ، حسام عبد الرزاق ، (2014) ، " الممهيات الأعملمية لأنموذج المعاملات المتغيرة والمتغيرة جزئياً " اطروحة دكتوراه في الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد .
2. يوسف، خلود يوسف خمو، (2004)، "مقارنة اساليب بيز مع طرائق اخرى لتقدير منحى الانحدار اللامعلمي"، اطروحة دكتوراه في الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد .
3. Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica* 77, 1229-1279 .
4. Baltagi, B.H. and Q. Li. (1996). On instrumental variable estimation of parametric dynamic panel data models. *Economics Letters* 76: 1- 9 .
5. Blanchet . J and Wadsworth. J . (2012) . " Applied Statistics " , Institute of Mathematics, Analysis, and Applications EPF Lausanne An MSc Course for Applied Mathematicians , working paper .
6. Chad Schafer and Larry Wasserman, (2014) . " Tutorial on Nonparametric Inference With R " Carnegie Mellon University working paper .
7. Fox , J. (2005) . " Introduction to Nonparametric Regression " , McMaster University Canada , Journal of Economic ans Social Research Council .
8. Henderson, D.J. and Carroll , Qi. and Li. (2008). " Nonparametric estimation and testing of fixed effects panel data models". *J Econom.* 2008 ; 144(1): 257–275. doi:10.1016/j.jeconom.2008.01.005 .
9. Huansha Wang , (2014) " . Essays on Semiparametric Ridge-Type Shrinkage Estimation, Model Averaging and Nonparametric Panel Data Model Estimation" A Dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Economics , Uuiversity of California .
10. Green, C., W. Long and Hsiao , C. (2014) . " Testing Error Serial Correlation in Fixed Effects Nonparametric Panel Data Models" , working paper .
11. Kris De Brabanter , Johan A.K. Suykens , Bart De Moor (2012) . " Nonparametric Regression via Stat LSSVM " *Journal of Statistical Software Published by the American Statistical Association Volume VV, Issue II* .
12. Kromann , B . (1995) . " Nonparametric Kernel Density Estimation" , *Statistical Science* , working paper .

- Lee, J., and Cox. D. (2008). " Robust Smoothing: Smoothing Parameter Selection and Applications to Fluorescence Spectroscopy "Carnegie Mellon University Research Showcase @ CMU. .13
- Linton, O. Xiao, Z. (2001). " A Nonparametric Regression Estimator That Adapts to Error Distribution of Unknown Form " Working Paper . o.linton@lse.ac.uk. .14
- Mundra, K. (2005). Nonparametric slope estimation for fixed effects panel data. Working Paper , Department of Economics, San Diego State University. .15

الملحق

جدول (1)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي عند الحالة الطبيعية للأنموذج الأول ولجميع حجوم العينات

(N=50, N=100, N=200) ولجميع مستويات التباين .

The Method	Sample Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.5641898	h = 0.33307	h = 0.2899646
	100	h = 0.5315521	h = 0.46851	h = 0.5783734
	200	h = 0.2860971	h = 0.42878	h = 0.3705311
LQNE . h by using CV	50	h = 0.543222	h = 0.21209	h = 0.2431098
	100	h = 0.4321876	h = 0.45321	h = 0.4360923
	200	h = 0.2765431	h = 0.24532	h = 0.3465891
Robust smoothing Spline λ by using RCV	50	$\lambda = 48.69286$	$\lambda = 0.00024734$	$\lambda = 0.01345294$
	100	$\lambda = 0.170094$	$\lambda = 16.29788$	$\lambda = 0.07624071$
	200	$\lambda = 0.04207592$	$\lambda = 3.078194$	$\lambda = 0.05503163$

جدول (2)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.10) للأنموذج الأول ولجميع حجوم العينات

(N=50, N=100, N=200) ولجميع مستويات التباين

The Method	Sample. Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.5705509	h = 0.6471394	h = 0.561132
	100	h = 0.5315521	h = 0.4683259	h = 0.5783734
	200	h = 0.2860971	h = 0.2569487	h = 0.3705311
LQNE . h by using CV	50	h = 0.5431298	h = 0.6540987	h = 0.4094532
	100	h = 0.5478901	h = 0.4360987	h = 0.3769087
	200	h = 0.27813459	h = 0.23912645	h = 0.38903214
Robust smoothing Spline . λ by using RCV	50	$\lambda = 1.246228$	$\lambda = 0.000192139$	$\lambda = 5.097745$
	100	$\lambda = 48.70513$	$\lambda = 16.29261$	$\lambda = 0.05337038$
	200	$\lambda = 0.0007122555$	$\lambda = 8.195854$	$\lambda = 11.1303$

جدول (3)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة توليوت (0.20) للأنموذج الأول ولجميع أحجام العينات

(N=50, N=100, N=200) ولجميع مستويات التباين

The Method	Sample. Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.5705509	h = 0.6471394	h = 0.561132
	100	h = 0.4384112	h = 0.4785956	h = 0.4801367
	200	h = 0.3154232	h = 0.2300997	h = 0.3946966
LQNE . h by using CV	50	h = 0.5670987	h = 0.650871	h = 0.4390876
	100	h = 0.4321225	h = 0.480123	h = 0.4820981
	200	h = 0.35679087	h = 0.976109	h = 0.39087124
Robust smoothing Spline . λ by using RCV	50	$\lambda = 0.07895061$	$\lambda = 0.000228587$	$\lambda = 5.292204$
	100	$\lambda = 0.0002146717$	$\lambda = 8.557234$	$\lambda = 1.12987$
	200	$\lambda = 0.002526227$	$\lambda = 0.03321856$	$\lambda = 0.0001927655$

جدول (4)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي عند الحالة الطبيعية للأنموذج الثاني ولجميع أحجام العينات

(N=50, N=100, N=200) ولجميع مستويات التباين

The Method	Sample. Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.4249547	h = 33.05023	h = 0.5658414
	100	h = 0.5315521	h = 0.53454	h = 0.5658414
	200	h = 0.368281	h = 0.428548	h = 0.3431401
LQNE . h by using CV	50	h = 0.43217651	h = 33.123876	h = 0.57890123
	100	h = 0.54321890	h = 0.571579	h = 0.57634218
	200	h = 0.37654312	h = 0.4362105	h = 0.32587912
Robust smoothing Spline $\cdot \lambda$ by using RCV	50	$\lambda = 0.6662982$	$\lambda = 0.05772954$	$\lambda = 8.377106$
	100	$\lambda = 0.170094$	$\lambda = 0.009327003$	$\lambda = 6.377106$
	200	$\lambda = 0.04207592$	$\lambda = 0.001136574$	$\lambda = 5.263681$

جدول (5)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.10) للأنموذج الثاني ولجميع حجوم العينات

ولجميع مستويات التباين (N=50, N=100, N=200)

The Method	Sample Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.6825674	h = 0.765945	h = 0.5658414
	100	h = 0.5356618	h = 0.423522	h = 0.4579105
	200	h = 0.368281	h = 0.565414	h = 0.3431401
LQNE . h by using CV	50	h = 0.6512980	h = 0.713265	h = 0.5429870
	100	h = 0.5321873	h = 0.439675	h = 0.5901765
	200	h = 0.3713298	h = 0.518690	h = 0.3298456
Robust smoothing Spline . λ by using RCV	50	λ = 3.176484	λ = 3.95358	λ = 0.0001232355
	100	λ = 0.1060042	λ = 0.2803762	λ = 8.377106
	200	λ = 0.04235228	λ = 0.00217281	λ = 5.265928

جدول (6)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة توليث (0.20) للأنموذج الثاني ولجميع حجوم العينات

ولجميع مستويات التباين (N=50, N=100, N=200)

The Method	Sample Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.6662982	h = 0.22508	h = 0.6103625
	100	h = 0.127132	h = 0.56361	h = 0.4694965
	200	h = 0.4148104	h = 0.37256	h = 0.3973959
LQNE . h by using CV	50	h = 0.62334190	h = 0.218765	h = 0.6329081
	100	h = 0.18765432	h = 0.543212	h = 0.4321879
	200	h = 0.42131980	h = 0.364321	h = 0.4390812
Robust smoothing Spline . λ by using RCV	50	λ = 33.05091	λ = 3.164631	λ = 0.0002913873
	100	λ = 35.51859	λ = 0.00642884	λ = 9.894623
	200	λ = 0.06546021	λ = 0.07111227	λ = 8.684983

جدول (7)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي عند الحالة الطبيعية للأنموذج الثالث ولجميع حجوم العينات

ولجميع مستويات التباين (N=50, N=100, N=200)

The Method	Sample. Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.6829128	h = 0.51985	h = 0.5833623
	100	h = 0.3038602	h = 0.43670	h = 0.4840952
	200	h = 0.3823103	h = 0.48271	h = 0.4216681
LQNE . h by using CV	50	h = 0.69876512	h = 0.52189	h = 0.5439801
	100	h = 0.3217894	h = 0.43109	h = 0.4389076
	200	h = 0.3458901	h = 0.45310	h = 0.4218901
Robust smoothing Spline - λ by using RCV	50	$\lambda = 7.487638$	$\lambda = 2.978291$	$\lambda = 8.611511$
	100	$\lambda = 1.680803$	$\lambda = 8.15411$	$\lambda = 4.916628$
	200	$\lambda = 4.434855$	$\lambda = 3.734134$	$\lambda = 3.523602$

جدول (8)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.10) للأنموذج الثالث ولجميع حجوم العينات

(N=50 ,N=100, N=200) ولجميع مستويات التباين

The Method	Sample. Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.6829128	h = 0.51985	h = 0.5833623
	100	h = 0.3038602	h = 0.43746	h = 0.4840952
	200	h = 0.3823103	h = 0.48478	h = 0.3488754
LQNE . h by using CV	50	h = 0.6751298	h = 0.45480	h = 0.52318901
	100	h = 0.3129812	h = 0.42188	h = 0.41888723
	200	h = 0.34219011	h = 0.21958	h = 0.34308871
Robust smoothing Spline - λ by using RCV	50	$\lambda = 7.487638$	$\lambda = 2.978291$	$\lambda = 8.611511$
	100	$\lambda = 1.680803$	$\lambda = 8.02098$	$\lambda = 4.916628$
	200	$\lambda = 4.434855$	$\lambda = 3.734134$	$\lambda = 3.523602$

جدول (9)

قيم معالم التمهيد المستعملة لإيجاد معيار المقارنة (AMAPE) المعتمد في طريقة التقدير اللامعلمية الأفضل لإيجاد تقدير أنموذج دالة الانحدار اللامعلمي وبنسبة تلويث (0.20) للأنموذج الثالث ولجميع حجوم العينات

(N=50 ,N=100, N=200) ولجميع مستويات التباين

The Method	Sample. Size (N)	$\sigma=1/2$	$\sigma=1/4$	$\sigma=1/8$
		Bandwidth	Bandwidth	Bandwidth
LLNE . h by using CV	50	h = 0.2332662	h = 0.69698	h = 0.649271
	100	h = 0.4394356	h = 0.55435	h = 0.4397255
	200	h = 0.4478086	h = 0.18732	h = 0.2784917

LQNE . h by using CV	50	h =0.24319876	h =0.51890	h = 0.6143298
	100	h = 0.42196543	h =0.16754	h = 0.4218901
	200	h =0.18356981	h =0.17542	h = 0.2653907
Robust smoothing Spline · λ by using RCV	50	λ =2.927901	λ =1.021657	λ =5.646306
	100	λ = 8.444337	λ =3.377268	λ =1.758618
	200	λ =4.038464	λ = 1.568002	λ =3.599213

.....

.....

.....