مقارنة بين الانحدار المضبب والانحدار المضبب الحصين

م.م. ذكاء يوسف عزيز
كلية الإدارة والاقتصاد
جامعة الموصل

المستخلص:

يهدف البحث الى المقارنة في استخدام الانحدار المضبب والانحدار المضبب الحصين. اذ ان هذه النماذج تستخدم في حالة كون أنموذج سببه الضبابية وليس العشوائية او كلاهما معا (الضبابية والعشوائية)، وقد تم استخدام الانحدار المضبب لحساب دالة الانتماء والتي تمثل مصفوفة الاوزان ومن خلالها يتم تحديد وزن كل مشاهدة يمثل مدى مساهمتها من المعلومات في عملية تقدير المعلمات أنموذج الانحدار المضبب الحصين. إذ تم اقتراح دالة انتماء لبناء مصفوفة الاوزان. وقد طبق الاسلوبين على متغيرات البيئية كمتغير استجابة (غاز ثاني اوكسيد الكاربون) والانشطة الاقتصادية كمتغيرات توضيحية (الناتج المحلي الاجمالي، عدد السكان، استهلاك الطاقة الاحفورية، الانفتاح الاقتصادي) ولغرض المقارنة بين الاسلوبين وللوقوف على مدى كفاءة التقدير فقد تم استخدام مقياس متوسط مربع الخطأ النسبي. وكانت النتائج تشير الى ان الانحدار المضبب الحصين كان أفضل في تقدير المعلمات من الانحدار المضبب.

A Comparative between Fuzzy Regression and Robust Fuzzy Regression

Asst. Lect: Thuka Y. Azez
College of Administration and Economics
Mosul University

Abstract:

The research aims at comparing between fuzzy regression and robust fuzzy regression. These models are used in case of the model suffer from fuzziness but not randomness, or both together (fuzziness and random). The fuzzy regression was used to calculate the membership function, which represents the weight matrix, by which the weight of each observation is determined and subsequently calculate the extent of each contribution to the information and then to be used in the process of determining the parameters. .For the purpose of analysis, environmental variables were adopted as response variable (carbon dioxide) and the economic activities as explanatory variables (GDP, population, energy consumption, economic openness). In order to compare the two methods and to determine the efficiency of the estimation, the mean square percentage error was used. The results indicated that robust fuzzy regression was better in estimating parameters than the fuzzy regression.

المقدمة

يعد تحليل الانحدار من الاساليب الإحصائية المستخدمة في معظم العلوم لبناء أنموذج بين متغير الاستجابة والمتغيرات التوضيحية وفي بعض الأحيان يكون في هذا الأنموذج عدم التأكد

الناتج عن المتغيرات العشوائية ولكن في كون عدم التأكد ناتج من الضبابية فان نظرية الاحتمالية لا يمكن استخدامها، اذ لابد من وجود أسلوب رياضي يعالج هذا النوع من عدم التأكد والذي يتمثل باستخدام نظرية المجموعات المضببة، وان هذه المتغيرات موجودة في الواقع العملي.

ويعد الباحث (Diamend) عام ١٩٨٧ اول من اقتراح طريقة المربعات الصغرى المضببة لإيجاد حسن المطابقة لأنموذج علاقة بين المتغيرات الاستجابة والمتغيرات التوضيحية ان موضوع الانحدار المضبب الحصين من المواضيع الحديثة اذ ان اول بحث نشر في هذا الموضوع والمقدم من قبل الباحثان (Watada and Yabuuchi) عام ١٩٩٤ حيث قدم الباحثان بناء أنموذج لبيانات تحتوي على أخطاء غير طبيعية.

تناول هذا البحث مواضيع تعريف المجموعة المضببة والدوال العضوية والارقام المضببة كونها تمثل جزء مهما في الاحصاء المضبب، كما تناول البحث استعراض الانحدار المضبب باستخدام طريقة مربعات الصغرى المضببة، وتم اقتراح دالة انتماء في مصفوفة الاوزان في اسلوب الانحدار المضبب الحصين وللمقارنة بين الاسلوبين تم استخدام متوسط مربع الخطأ النسبي. اما الجانب التطبيقي فقد تضمن اجراء التحليل من خلال البيانات التي تم الحصول عليها لدراسة العلاقة بين الانشطة الاقتصادية والبيئية.

١. **المجموعة المضببة:** تعرف المجموعة المضببة على انها مجموعة من العناصر التي تمتلك درجة انتماء أو عضوية مداها ضمن الفترة [0, 1]، أي أن عندما يمتلك العنصر درجة انتماء (صفر) فهذا يعني أن العنصر ينتمي بدرجة صفر إلى المجموعة المضببة وإذا كانت درجة انتماء العنصر مساوية إلى الواحد فهذا يعني أن العنصر ينتمي بالتمام الى المجموعة وهنالك درجات تقع بين الصفر والواحد أي أذا كان العنصر يمتلك درجة انتماء 0.5 فهذا يعني أن العنصر ينتمي إلى المجموعة بنسبة 50% وبنفس النسبة لا ينتمي إلى المجموعة ويدعى هذا العنصر بنقطة التوازن (Equilibrium). (Klir, 1997, 10)

٢. **دالة الانتماء او العضوية (Membership Function):** يقال ان العضوية هي تعبير عن درجة الانتماء، وهي الدالة التي عن طريقها يتم حساب درجة عضوية عنصر ما إلى المجموعة المضببة ويرمز لها:

$\mu_A(x)$ التي تمثل درجة انتماء المتغير x إلى المجموعة المضببة A

$$\mu_A(x) = x \rightarrow [0, 1]$$

$$\mu_A(x) = 0$$ عدم الانتماء

$$\mu_A(x) = 1$$ تام الانتماء

١-٢. أنواع الدوال العضوية: هنالك عدة أنواع لدوال العضوية منها ما يأتي:

أ. دالة العضوية القياسية (Standard Membership Fucnction)

وهي دالة متزايدة وكما موضح في المعادلة الآتية:

$$\mu_A(x) = \begin{cases} 0 & \text{for} \quad \chi \le \alpha \\ 2\left[\dfrac{\chi - \alpha}{c - \alpha}\right]^2 & \text{for}\ \alpha \le \chi \le \beta \\ 1 - 2\left[\dfrac{\chi - c}{c - \alpha}\right]^2 & \text{for}\ \beta \le \chi \le c \\ 1 & \text{for} \quad \chi \le c \end{cases} \qquad ....1$$

اذ ان:

χ : تمثل عناصر المجموعة المضببة

α : تمثل عنصر من المجموعة قيمة الدالة له (0.0)

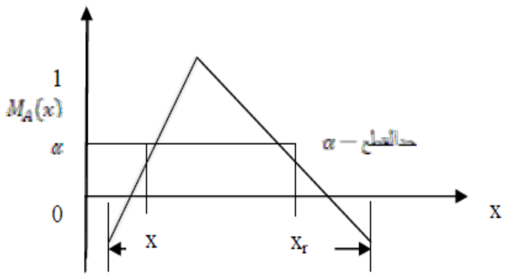
β : تمثل عنصر من المجموعة قيمة الدالة له (0.5)

c : تمثل عنصر من المجموعة قيمة الدالة له (1.0)

ب. الدالة العضوية المثلثية (Triangular Membership Function)

وهي دالة خطية تعطي شكل المثلث عند رسمها قاعدتها الفترة المحددة في الصيغة ورأسه مركز العدد المضبب وكما في الشكل رقم (١) ويمكن كتابتها كما يأتي:

$$\mu_A(x) = \left\{ 1 - \frac{|x-a|}{s} \quad \text{when} \quad a - s \le x \le a + s \right\} \qquad ---(2)$$



الشكل (1): الدالة العضوية المثلثية

ج. دالة العضوية الجرسية (Bell-Shaped Membership Function)

وهي دالة آسية غير خطية تأخذ شكل المنحنى الطبيعي وتكون بالصيغة الآتية:

$$\mu_A(x) = \left\{ c e^{\frac{(x-a)^2}{b}} \qquad -\infty < x < \infty \right\} \qquad ---(3)$$

٣. الأرقام المضببة (Fuzzy Numbers)

يقال ان المجموعة المضببة A في R بأنها أرقام مضببة إذا تحققت الشروط الآتية:

- يجب ان تكون المجموعة (A) محددة في الفترة [0,1] ∋ α
- دالة الانتماء للمجموعة المضببة (A) تكون مستمرة وجزئية (piecewise).
- يجب ان تكون المجموعة (A) مجموعة مضببة طبيعية ومحدبة (gearge, 1995, 68).

وهناك ارقام مضببة ذات انتشار من اليمين واليسا(LR) وكما في الشكل رقم (٢)، اذ كانت المجموعة المضببة ارقامها مضببة من نوع (LR) وحيث ان دالة الانتماء تكون كما في الصيغة الاتية:

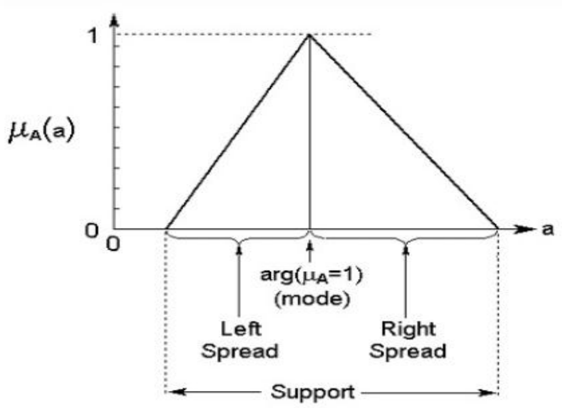
$$\mu_A(X)=\begin{cases} \mathrm{L}\left(\dfrac{m-\chi}{\alpha}\right) & \chi \le m \\ \mathrm{R}\left(\dfrac{\chi-m}{\beta}\right) & \chi \ge m \end{cases} \qquad \alpha > 0 \;\; , \beta > 0 \qquad ---(4)$$

اذ ان:

m : تمثل المتوسط او المنوال.

α : تمثل المدى من اليمين.

β : تمثل المدى من اليسار. (Jann & Chwan, 2008, 2)



الشكل (2): الارقام القياسية

٤. الانحدار المضبب (Fzzy Regression)

يعد الانحدار المضبب من المواضيع الحديثة حيث أن أول من كتب في هذا الموضوع هو العالم (Tanaka) واخرون عام ١٩٨٠، وتم تطويره في الاعوام (١٩٨٢-١٩٨٧)، ويعد الباحث (Diamond) عام ١٩٨٧ اول من اقتراح طريقة المربعات الصغرى المضببة لإيجاد حسن المطابقة لنموذج تكون البيانات الداخلة قطعية والعلاقة مضببة وفي اغلب الاحيان يستخدم الانحدار التقليدي لوضع نموذج علاقة بين المتغير التابع والمتغيرات التوضيحية ويكون في هذا النموذج عدم التأكد ناتج من العشوائية ولكن في حالة كون عدم التأكد ناتج من الضبابية (Fuzziness) فان النظرية الاحتمالية لا يمكن استخدامها، وإنما سيتم استخدام نظرية المجموعات المضببة وان عدم التأكد في الانحدار المضبب ناتج من:

- ان تكون البيانات الداخلة اي المتغير التابع والمتغيرات التوضيحية قطعية والعلاقة الناتجة اي بين المتغير التابع والمتغيرات التوضيحية مضببة.
- ان تكون بيانات المتغير التابع والمتغيرات التوضيحية مضببة والعلاقة بينهم ايضا مضببة.
- او في حالة البيانات مضببة العلاقة مضببة. (Kamile & Aysen, 2004, 74)

يمكن تقدير معلمات نموذج الانحدار باستخدام طريقة المربعات الصغرى المضببة (Fuzzy Laest Square) التي تعتمد على تقليل المسافة بين النتائج المضببة المتنبئ بها للأنموذج يمكن التعبير عن الانحدار المضبب الخطي المتعدد كما يأتي:

$$Y_i = A_0 + A_1 X_{11} + A_2 X_{21} + \text{---} + \mathrm{A_P} X_{P1} + ei \qquad \text{---}(5)$$
$$\mathrm{i} = 1,2,3,\text{---}\mathrm{n}$$

من خلال تصغير الدالة اعلاه سوف نحصل على تقدير المعلمات:

$$\min(\mathrm{S}) = \min\ (A_0, \mathrm{A_1}, \mathrm{A_2}, \mathrm{A_3}, \text{---}, \mathrm{A_P}) \qquad \text{---}(6)$$
$$\min(\mathrm{S}) = d(A_0 + A_1 X_{1i} + \mathrm{A_2 Xi} + \mathrm{A_3} X_{3i} + \text{---} + \mathrm{A_P\ X_{Pi}}) \qquad \text{---}(7)$$

اذ ان d يمكن كتابتها بالشكل الاتي: (Srabani & Madumangal, 2009, 60)

$$= [A_0 + A_1 X_{1i} + \mathrm{A_2}\ X_{2i} + \text{----} + A_p X_{pi} - y_i - (A_1 I x_1 + A_2 I x_2 + \text{---} A_p x_{pi} - Iy)]^2$$
$$+ [A_0 + A_1 x_{1i} + \text{----} + A_p x_{pi} - y_i + A_0 + A_1 r_{x1} + A_2 r_{2i} + \text{----} + A_P r_{pi} - r_y)]^2$$
$$+ (A_0 X_{1i} +_{1i} + \mathrm{A_2}\ X_{2i} + \text{----} + A_p X_{pi} - y_i)^2 \qquad \text{---}(8)$$

وبأخذ التفاضل الجزئي للمعادلة أعلاه نسبة إلى المعلمات التي يراد تقديرها ($A_0, A_2, \ldots\ldots., A_p$) سوف نحصل على p+1 من المعادلات وبمساواتها بالصفر وحلها أنيا سنحصل على تقديرات المربعات الصغرى، وباستخدام المصفوفات للتعبير عن نموذج الانحدار المتعدد للحصول على A وكما يأتي: (Kamile & Aysen, 2004, 77)

$$A = (CC' + XX' + DD')^{-1}(X'Y + C'E + D'F) \qquad \text{---}(9)$$

$$X = \begin{bmatrix} 1 & \chi_{11} & \cdots & \chi_{p1} \\ 1 & \chi_{12} & \cdots & \chi_{p2} \\ \vdots & \vdots & \ddots & \cdots \\ 1 & \chi_{np} & \cdots & \chi_{np} \end{bmatrix} \quad \mathrm{Y} = \begin{bmatrix} \mathrm{y_1} \\ \vdots \\ \mathrm{y_n} \end{bmatrix} \quad ، \quad \mathrm{E} = \begin{bmatrix} \mathrm{y_1} - \mathrm{I_{y_1}} \\ \vdots \\ \mathrm{y_n} - \mathrm{I_{y_n}} \end{bmatrix} \quad ، \quad \mathrm{F} = \begin{bmatrix} \mathrm{y_1} + \mathrm{r_{y_1}} \\ \vdots \\ \mathrm{y_n} + \mathrm{r_{y_n}} \end{bmatrix}$$

---(10)

$$\mathrm{C} = \begin{bmatrix} 1 & (\chi_{11} - \mathrm{I}_{\chi_{11}}) & \cdots & (\chi_{1p} - \mathrm{I}_{\chi_{1p}}) \\ 1 & (\chi_{12} - \mathrm{I}_{\chi_{12}}) & \cdots & (\chi_{2p} - \mathrm{I}_{\chi_{2p}}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (\chi_{n1} - \mathrm{I}_{\chi_{n1}}) & \cdots & (\chi_{np} - \mathrm{I}_{\chi np}) \end{bmatrix} \quad ، \quad \mathrm{D} = \begin{bmatrix} 1 & (\chi_{11} - \mathrm{r}_{\chi_{11}}) & \cdots & (\chi_{1p} - \mathrm{r}_{\chi_{1p}}) \\ 1 & (\chi_{12} - \mathrm{r}_{\chi_{12}}) & \cdots & (\chi_{2p} - \mathrm{r}_{\chi_{2p}}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (\chi_{n1} - \mathrm{r}_{\chi_{n1}}) & \cdots & (\chi_{np} - \mathrm{r}_{\chi np}) \end{bmatrix}$$

٥. الانحدار المضبب الحصين: (Robust Fuzzy Regression)

كان اول اقتراح لموضوع الانحدار الضبابي الحصين في عام ١٩٩٤ من قبل الباحثان (Yabuuchi and Watada)، حيث قدم الباحثان نموذج انحدار لتقليل الخطأ الكلي بين نموذج للبيانات التي تحتوي على أخطاء لا تتبع التوزيع الطبيعية، وللحصول على المعلمات الحصينة سوف نستخدم طريقة المربعات الصغرى الموزونة الضبابية وكما مبين بصيغة الآتيــة:(Yang & Chen, 2013, 229)

$$A = (CWC' \ + XWX' + DWD')^{-1}(X'WY + C'WE + D'F) \qquad ---(11)$$

حيث ان W تمثل مصفوفة الاوزان وعن طريقها يتم تحديد وزن كل مشاهد التي تساهم في عملية تقدير المعلمات، ولبناء مصفوفة الاوزان فان ذلك يتم من خلال الخطوات الاتية:

- ايجاد قيمة الخطأ (r_i) باستخدام المعادلة $\left(y_i - y_i^{\wedge}\right)$ حيث y_i تمثل قيمة متغير الاستجابة $y_i^{\wedge}$ يمثل القيمة التقديرية لمتغير الاستجابة
- يتم ايجاد دالة المسافة D كما في المعادلة الاتية: (Sanli & Fatih, 2012, 229)

$$D_i = \|\text{abs}(r_i) - \ \text{median}(\text{abs}(r_i)\| \qquad i = 1,2,3......n$$

حيث أن (‖.‖) تمثل مسافة اقليدس (Euclidean Distance): وهي المسافة المستقيمة بين نقطتين في فضاء الاقليدس.

- وفقا لقيمة المسافة (d) يتم تعريف دالة الانتماء وكما يأتي:

حيث أن:

$$b = \max(D_I) + \text{d} \qquad , \qquad \text{a} = median\,(\text{D}_\text{i})$$

$$d = median \mid r_i \text{ - } median\,(r_i)\mid / \, 0.6745$$

- بالاعتماد على دالة الانتماء يتم بناء مصفوفة الاوزان التي تمثل قيم دالة الانتماء فيها قيم القطر الرئيس لمصفوفة الاوزان.

$$\mu(x) = \begin{cases} 1 & r \le a \\ \dfrac{b - |r|}{b - a} & a \le r \le b \\ 0 & ow \end{cases} \qquad ---(12)$$

أي:

$$W = diag(\mu_1, \mu_2, \mu_3 ....\mu_n \ )$$

- باستخدام معادلة (١١) يتم الحصول على التقديرات للمعلمات الانحدار المضبب الحصين.

ولغرض المقارنة بين الانحدار المضبب والمضبب الحصين وللوقف على مدى كفاءة التقدير تم استخدام مقياس متوسط مربع الخطأ النسبي $MSPE$ وفق الصيغة الاتية:

$$MSPE = \frac{\sum_{\text{i=1}}^{\text{n}} (PE)^2}{n} \qquad ---13$$

PE : يمثل الخطأ النسبي

٦. الجانب التطبيقي

لغرض المقارنة بين تقدير الانحدار المضبب والمضبب الحصين تم استخدام البيانات التي تم الحصول عليها من اعداد متسلسلة للمجموعة السنوية للأمم المتحدة (Stastistical Year book) وموقع البنك الدولي (http:// Data worldbank .ostalog/work)، ويبحث الجانب التطبيقي شكل العلاقة بين المتغيرات البيئة كمتغير استجابة هو ثاني اوكسيد الكاربون، أما المتغيرات التوضيحية فهي (الناتج المحلي، ،عدد السكان، الاستهلاك الطاقة الاحفورية، الانفتاح الاقتصادي) من سنة (١٩٩٠-٢٠١٥).

ولبناء مصفوفة الاوزان (W) اقترحنا دالة وزن والتي يمكن ايجادها من خلال الخطوات الاتية:

- حساب قيم الخطأ العشوائي من معادلة ٩ عند تطبيق الانحدار المضبب.
- ايجاد قيمة الانحراف المعياري والوسط الحسابي لقيم الخطأ العشوائي.
- ايجاد دالة الانتماء للأوزان باستخدام الدالة الاتية.

$$\mu(x)=\left[0.5+(1+\left(\frac{ri-\mathrm{X}'_{\mathrm{ri}}}{\mathrm{S}_{\mathrm{ri}}}\right)\right]^{-2} \qquad \ldots 14$$

حيث ان:

ri : تمثل قيم الخطأ العشوائي

$\mathrm{X}'_{\mathrm{ri}}$: تمثل الوسط الحسابي لقيم الخطأ العشوائي

S_{ri} : تمثل الانحراف المعياري لقيم الخطأ العشوائي

الجدول (1)

تقديرات معلمات لأنموذج الانحدار المضبب والمضبب الحصين لمتغير co2

المعلمات	$b0$	$b1$	$b2$	$b3$	$b4$
FR	0. 442	0.3184	0.4872	0.6846	0.6728
RFR	0.3006	0.5346	0.5307	0.461	0.446

يوضح الجدول (1) تقديرات معلمات نموذج الانحدار المضبب (FR) باستخدام معادلة رقم (9) وتقديرات معلمات نموذج الانحدار والمضبب الحصين (RFR) باستخدام معادلة (11) بعد احتساب مصفوفة الاوزان (w) وفق المعادلة رقم (14) لمتغير الاستجابة (co_2) متغيرات الانشطة الاقتصادية متغيرات التوضيحية.

ولغرض المقارنة تم احتساب متوسط مربع الخطأ النسبي بموجب المعادلة رقم (14) وكانت النتائج كما في الجدول الاتي:

الجدول (2)

قيم متوسط مربع الخطأ النسبي

الانحدار المضبب الحصين (RFR)	الانحدار المضبب (FR)	
0.7347	0.8490	متوسط مربع الخطأ النسبي

إذ يظهر الجدول (2) ان الانحدار المضبب الحصين أفضل من الانحدار المضبب، لان قيمة (MSPE) في الانحدار المضبب الحصين اقل من قيمته في الانحدار المضبب.

الاستنتاجات

١. الانحدار المضبب الحصين (RFR) أفضل من الانحدار المضبب (FR)، لأن قيمة (MSPE) في الانحدار المضبب الحصين أفضل من الانحدار المضبب.
٢. كلا الاسلوبين اعطى نتائج متقاربة للمعلمات الانموذج.
٣. وجود علاقة طردية بين متغير البيئة (co_2) والانشطة الاقتصادية.

المصادر:

1. George, J. k. and BoYuan, (1995), "**Fuzzy Sets and Fuzzy Logic Theory and applications**", publ. By Prentice Hall PTR. New Jersey, 07458.
2. Jann, H and Chwan, C., (2008), "**Astudy of Fuzzy Linear Regresision"**.
3. Kamile, S. and Aysen, (2004), "**The Fuzzy Robrst Regression Analysis the Case Fuzzy Data Set HAS"**, G. U. Journal of Science Vol. 17 (13), pp71-84.
4. Klir, G. J., (1997), **"Fuzzy Set Theoury"** publ. By Prentice Hall PTR.
5. Sanli, K and Fatih, Ta. and Erbay, T., (2012), "**Astudy on Fuzzy robust Regression and Its applications to Lnsurance,** Mathematcial **and** Computational Applieation, Vol. 17. No 3, pp 223-234.
6. Srabani, s. and Modhumangal, p., (2009), "**Multiple Regression of Fuzzy Valued Variable"**, Journal of Physical Science,Vol. 13, pp57-66 ISSNKli, G.J, (1997).
7. Watada J. and Yabuuchi, Y., (1994), **Fuzzy Robust Regression Analysis**, Fuzzy Systems, 1994, IEEE World Congress on Computational Intelligence International Conference; Proceedings of the Third IEEE International Conference on. Vol. 2, PP: 1370-1376.
8. Yang, Z, and Chen, Y., (2013), "**Robust Fuzzy Coeffcient Regression Analysis with Crisp Inputs and Gaussian Zuzzy Output,**Journal Computing Science and **Engineering,** Vol. 7, No. 40, pp. 263-271.