المجلة العراقية للعلوم الاحصائية (10) 2006
ص ص [61–84]

تحليل الإنحدار المضبب

حسن محمد الياس*   هبـــة علي طه الصـــباغ**

الملخص

يتناول هذا البحث الانحدار الخطي المضبب في حالة كون البيانـــات مضـــببة وغير مضببة وقد تم التحليل بواسطة عـــدة طرائـــق وهـــي الانحـــدار الخطـــي المضبب(FLR) والمعدل له، الانحـــدار الخطـــي المضـــبب بمربعـــات صـــغرى (FLSLR) اذ وظفت البرمجة الخطية (LP) في التحليل .

تم تقدير المعلمات المضببة ببيانات مضببة وغير مضـــببة وحـــدد نمـــوذج الانحدار باستخدام مفاهيم نظرية المجموعات المضببة، وطبقت هذه الطرائق فـــي المجال الطبي ببيانات حقيقية عن مرض هشاشة العظام وذلك من خلال قياس كثافة العظم من فحص لجهاز الـــ (DEXA) لثلاثين مريضا (10 ذكور ، 20 انثى).

أثبتت النتائج ان استخدام نموذج (Tanaka) المعدل افضـــل مـــن نمـــوذج (FLR) لتفادي ظهور المعلمات المقدرة غير المضببة. اما طريقة (FLSLR) فهي افضل من (FLR) وذلك من خلال نتائج درجة انتماء النموذج.

Fuzzy Regression Analysis with an Application

ABSTRACT

This work investigates the analysis of fuzzy regression in case of the data being fuzzy or non-fuzzy. The analysis is performed by several methods , such as fuzzy linear regression (FLR) and modified one, and fuzzy linear regression with least

---

* أستاذ مساعد /قسم الاحصاء/كلية علوم الحاسبات والرياضيات /جامعة الموصل.

** مدرس مساعد/ مركز الحاسبة الالكترونية / جامعة الموصل.

تاريخ التسلم : 12/ 7 /2005 ـــــــــــــــــــ تاريخ القبول : 19/ 12 /2005

squares (FLSLR), in which linear programming (LP) is used in the analysis.

Estimation of fuzzy parameters is carried out in case of fuzzy and non-fuzzy (crisp) data where the regression model is detected by fuzzy set theory.

These methods are applied on real data of Osteoporosis which are obtained by measuring the bone density of 30 patients (10 male and 20 female) by DEXA.

The results of the analysis show that, Tanaka model is better than fuzzy linear regression model (FLR) so that to avoid the appearance of non-fuzzy estimated parameters ,and . (FLSLR) is better than (FLR) in the sense of degree of membership of the model.

المقدمة

نتناول في هذه المقدمة توضيح عبارة الانحدار المضبب (Fuzzy Regression) إذ كل كلمة تفسر على حدة فان كلمة الانحدار تشتمل على اشكال وطرائق إحصائية واسعة الاستخدام في جميع العلوم المختلفة فهي توضح العلاقة بين متغير يسمى متغير الاستجابة (Response variable) ومتغير واحد اواكثر من متغيرات تسمى المتغيرات المفسرة (متغيرات توضيحية) (Explanatory variables) ، إذ ان متغير الاستجابة يكون بهيئة دالة للمتغيرات المفسرة وبذلك توضح قوة واتجاه العلاقة من خلال تقدير المعلمات.

اما كلمة المضبب فتعني المنطق المضبب وهو احد اشكال المنطق يستخدم في الأنظمة الخبيرة وتطبيقات الذكاء الاصطناعي ، نشأ هذا المنطق عام 1965 على يد العالم لطفي زاده من جامعة كاليفورنيا إذ طوره ليستخدم كطريقة افضل لمعالجة البيانات إذ يسمى هذا المنطق أحيانا بمنطق الغموض ليعالج التعابير الاكثر تعقيدا وغموضا (Kandel,1986).

ان مبدأ المنطق المضبب يقوم على وجود تابع قيمته عند عنصر معين هي قيمة حقيقية تقع بين (1 و0) تعبر عن انتماء هذا العنصر الى مجموعة ما ، فاذا كانت قيمة هذا التابع (1) فهذا العنصر ينتمي لها تماما ، وإذا كانت قيمته (0)

فالعنصر لاينتمي اليها ابدا، اما اذا كانت قيمته بين (1 و0) فتشير الى مدى انتماء العنصر الى هذه المجموعة (ويكيبديا الموسوعة الحرة، 2004).

ان المنطق المضبب قد تجاوز منطق الارسطو المعتمد في صياغته على مبدأ الحتمية العلمية وهذا المنطق يساعد العلم المعتمد على مبدأ اللايقين على الانطلاق والتنبؤ بالمستقبل فهو منطق يقبل التعدد لامجرد الثنائيات إذ انه يتعامل مع الشك والتعقيد الموجود في الواقع.

تربط هاتان الكلمتان مع بعضهما لتصبح العبارة (الانحدار المضبب) . اذ ان الانحدار المضبب هو وسيلة لايجاد العلاقة الدالية بين متغير الاستجابة ومتغير واحد اواكثر من متغيرات المفسرة في ظاهرة غامضة وعدم التأكدية إذ يتناول البيانات المضببة (Fuzzy Data) وبذلك سيعتمد الانحدار المضبب على مفاهيم نظرية المجموعات المضببة (Fuzzy Sets Theory) بينما اعتمد الانحدار التقليدي على نظرية الاحتمالات (Probability Theory) (Chang & Ayyub,2001).

"ان نظرية الاحتمال ونظرية المجموعات المضببة تميزان نوعين من اللاتاكدية (Uncertainty). فنظرية الاحتمال تتعامل مع مسألة توقع حدوث حوادث معينة بالمستقبل استنادا الى معلومات متوفرة في الماضي والحاضر. لذا فان نظرتنا الى اللاتأكدية من خلال نظرية الاحتمال سوف تكون باتجاه التنبؤ عن الحوادث، اما اذا نظرنا الى اللاتأكدية من خلال منظار نظرية المجموعات المضببة فنجد انها لاتتعلق بتوقع حدوث شئ معين، بل انها لاتأكدية ناجمة من عدم دقة المعنى لبعض المفاهيم والمصطلحات اللغوية. وتجدر الملاحظة ان هناك العديد من الموضوعات التي يظهر فيها النوعان من اللاتأكدية فعلى سبيل المثال ان التكهنات الجوية يمكن ان تشير الى انه "محتمل جدا " ان يكون الجو "غائما" فنجد مفهوم "غائم" هو مفهوم مضبب كما ان "محتمل جدا " هو ايضا مفهوم يتضمن العشوائية (Randomness) والتضبيبية (Fuzziness)" (الخياط، 2004).

هدف الدراسة معرفة معنى الانحدار الخطي المضبب وتوضيح مفاهيمه واستخداماته وتناول بعض الطرائق المتبعة في الاستخدام وكذلك ميزاته مع المقارنة

بالانحدار التقليدي. اذ ساعدت نظرية المجموعات المضببة في تحليــل الانحــدار المضبب. وتم تطبيقه في المجال الطبي اذ تعد عملية التشخيص الطبي عملية معقدة وهي تشتمل على عدم التأكدية الناتجة عن اشـــتراك مجموعــة مـــن الاعــراض المرضية مع مجموعة من الامراض وبذلك تم تطبيق تحليل الانحدار المضبب على مرض ترقق العظام (هشاشة العظام).

كان اول من قدم تحليـــل الانحـــدار المضـــبب (Tanaka et al.,1982) إذ استخدموا نظامـــا خطيـــا مضـــببا مـــع مســـألة البرمجـــة الخطيـــة (Linear programming problem) باعتبار ان علاقة المتغيرات في نموذج الانحدار قد وضعت بضبابية، أي النموذج يكون بمدخلات غير مضببة ومعلمات مضببة.

تناولت دراسة الباحثين (Heshmaty & Kandel,1985) تطبيـــق نمـــوذج تكهن البيع في محيط لاتاكدي (Uncertain).

واوضح (Tanaka,1987) ان البيانات المضببة يمكن ان تعد توزيعا" امكانيا" وتحليله بواسطة نماذج خطية امكانية ومن ثم تقدير المعلمة المضببة إذ نوقشت في انظمة خطية امكانية مع مدخلات غير مضببة ومخرجات مضببة عرفت بواسطة الاعداد المضببة . إذ بين نوعين من التحليل وهما الاول يسمى مسألة الادنى (Min problem) والثاني مسألة الاعظم(Max problem) .

بين (Savic & Pedrycz,1991) اسلوبا لتخمـــين نمـــاذج الانحـــدار المضبب وذلك باجراء مرحلتين متتاليتين الاولـــى اســـتخدام طريقـــة المربعـــات الصغرى التقليدية ومن ثم استخدام النتائج من المرحلة الاولى لتوظيفها في مـــدخل (Tanaka et al.,1982) في المرحلة الثانية.

استخدم الباحثان (Chang & Ayyub,2001) مقارنة بين طرائق الانحـــدار المضبب مع الامثلة العددية إذ قدما مدخلا" وكان الاعتماد على نمـــوذج Tanaka هو الاساس اذ استخدم كقياس ذي اقل ضبابية ومن ثـــم قـــدما مـــدخل المربعـــات الصغرى المضببة واجريت مقارنة مع اسلوب الانحدار التقليدي .

وبين كل من (Nasrabadi & Nasrabadi,2004) مقترحا" اساسه البرمجة الرياضية لنماذج الانحدار الخطي المضبب مع مدخلات (غير مضببة / مضـــببة)

ومخرجات (غير مضببة / مضببة) إذ كانت من مميزات هذا المقترح انه بسيط في البرمجة والحسابات واقل اختلافا في مجموع الانتشار بين قيم المشاهدات والقيم المركزية.

ومن اللافت للنظر ان مجمل الدراسات السابقة في هذا المجال قد ركزت على استخدام البيانات المضببة (Fuzzy Data) أي الاعداد المضببة مع دوال الانتماء المثلثية المتماثلة (Symmetric Triangular Membership Functions).

الانحدار الخطي المضبب (FLR) Fuzzy Linear Regression

يستخدم الانحدار المضبب لتقدير العلاقة الدالة بين متغير الاستجابة والمتغيرات المفسرة (التوضيحية) في محيط مضبب مع دالة خطية وبذلك سمي بالانحدار الخطي المضبب (FLR).

إذ وظفت مسألة البرمجة الخطية (Linear Programming Problem) (LP) كطريقة اساسية لتقدير المعلمات المضببة اذ يستخدم نموذج خطي مضبب مع معلمات مضببة مثلثية متماثلة .

مسألة البرمجة الخطية

هي طريقة لايجاد احسن استخدام للموارد المحدودة والصيغة المستخدمة هي خطية لوصف العلاقة بين متغيرين اواكثر، وهذه العلاقة هي علاقة مباشرة وتتغير بالنسبة نفسها. مثال اذ تغيرت ساعات الانتاج بنسبة % 10 تغير الانتاج نسبة % 10 (المعزاوي ،1977).

المستلزمات الاساسية للبرمجة الخطية

1-ان يكون هناك هدف مطلوب تحقيقه مثل تحقيق أقصى أرباح ممكنة أو تخفيض التكاليف لأدنى حد ممكن، ويوضح هذا الهدف على شكل دالة تسمى بدالة الهدف.

2-تحديد مجموعة من القيود وهي التعبير الرياضي عن الحدود الموضوعة عن الموارد. يمكن التعبير عن دالة الهدف والقيود بمعادلات أو متباينات خطية.

فالهدف من تحليل الانحدار المضبب فضلا" عن تقدير المعلمات هـو ايجـاد نموذج الانحدار الذي يطابق كل البيانات المشاهدة المعطاة من خلال مقياس مطابق (Chang & Ayyub,2001).

اما صيغة نموذج (FLR) فتكتب كما ياتي:

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + \ldots + \tilde{A}_n x_n \qquad \ldots\ldots\ldots\ldots(1)$$
$$= \tilde{\mathbf{A}}' \mathbf{X}$$

إذ

مثلثي متماثل الشكل ويرمز لعناصر هذا المتجه بالشكل الآتي:

$\tilde{\mathbf{A}} = [\tilde{A}_0, \tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_n]'$ هو متجه المعلمات المضببة بشـكل عـدد مضبب

$\mathbf{X} = [1, x1, x2, \ldots, xn]$ هو متجه المتغيرات المفسرة

$$\tilde{A}_i = (\alpha_i, c_i)_L$$

إذ

α_i هي قيمة المركز (center value) للمعلمة المضببة .

c_i هي قيمة الانتشار (spread value) للمعلمة المضببة كما يطلق عليهـا (width) .

$i = 0,1,2, \ldots,n$

L تمثل دالة الانتماء للاتجاه الأيسر.

اما دالة الانتماء للمعلمة المضببة فهي كما يأتي:

$$\mu_{\tilde{A}_i}(a_i) = \begin{cases} 1 - \dfrac{|\alpha_i - a_i|}{c_i} & \alpha_i - c_i \le a_i \le \alpha_i + c_i \\ 0 & \text{otherwise} \end{cases} \qquad \ldots\ldots(2)$$

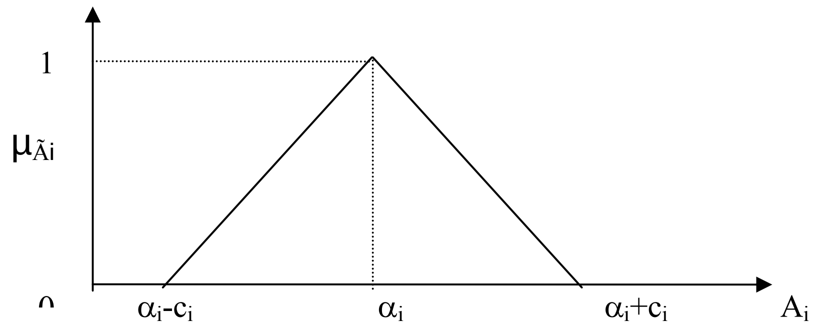
إذ

$a_i > 0$ و $i=1,2,\ldots,n$

وبذلك يمكن اعادة كتابة المعادلة (1) كما يأتي:

$$\widetilde{Y}_i = (\alpha_0, c_0) + (\alpha_1, c_1)x_1 + (\alpha_2, c_2)x_2 + \ldots + (\alpha_n, c_n)x_n$$

(Wang & Tsaur,2000) ، (McCauley & Wang,1999) . ويمكـــن تمثيل دالة الانتماء كما في الشكل (1).



الشكل (1) يوضح معلمة نموذج الانحدار المثلثية

طرائق الانحدار المضبب

يتناول هذا المبحث طرائق عدة للانحدار الخطي المضبب وهي:

1- نموذج Tanaka

اذ تعد هذه الطريقة النواة الاولى لـــ (FLR) التـــي قـــدمها (Tanaka et al.,1982) وسميت هذه الطريقة بعدئذ بنموذج (Tanaka) .

إذ طبق الباحثون في هذه الطريقة دالة خطية مضببة لتحديد الانحدار لظـــاهرة غامضة. فان الانحرافات بشكل عام بين القيم المشاهدة والقيم المقدرة في الانحدار التقليدي ناشئة عن أخطاء المشاهدات، اما هنا فهذه الانحرافات تعتمد علـــى عـــدم تحديد هيكلية النظام اذ تعد هذه الانحرافات ضبابية لمعلمـــات النظـــام (System parameters). طبق هذا المفهوم للحصول على نموذج لميكانيكية تســـعيرة بنـــاء دور يابانية، إذ كانت المعلمات المضببة قد درست بشكل دوال انتماء مثلثية .

وقد اقترح لتحديد علاقة خطية مضببة كما في الصيغة الاتية:

$$Y = A_0 + A_1 x_1 + \ldots + A_n x_n$$
$$= \mathbf{A}' \mathbf{X}$$

إذ

$\mathbf{X} = [1, x_1, x_2, \ldots, x_n]$ هو متجه لمتغيرات الموضحة

$\mathbf{A} = (A_0, A_1, \ldots\ldots, A_n)'$ هو متجه معلمات النموذج

باستخدام المعلمات المضببة A_i كما في الصيغة (2) كاعداد مضـــببة مثلثيـــة وتطبيق العمليات الحسابية لنظرية المجموعات المضببة تصبح دالة الانتماء لـــ Y

$$\mu_Y(y) = \begin{Bmatrix} 1 - \dfrac{|y - x'\alpha|}{c'|x|} & \text{for } x \neq 0 \\ 1 & x = 0, y = 0 \\ 0 & x = 0, y \neq 0 \end{Bmatrix} \quad \ldots\ldots\ldots (3)$$

من المعادلة (1) كما في الصيغة الاتية:

إذ

' يمثل عملية التبديل (Transposition)

α يرمز لمتجه القيم المركزية

c يرمز لمتجه قيم الانتشارات لكل معلمات النموذج

$$|x| = (|x_1|, \ldots, |x_n|)'$$

للبرهان انظر (Tanaka et al.,1982)

إذ استخدم الباحثون هذه الطريقة في حالة كون البيانات الخاصة بـــالمتغيرات الموضحة حادة (هشة) (crisp)بمعنى اخر غير مضببة. ومتغير الاستجابة فـــرض كعدد مضبب مع دالةانتماء مثلثية ويرمز له بـــ

$$Y = (y_i, e_i) \qquad (i = 1, 2, \ldots, N)$$

إذ

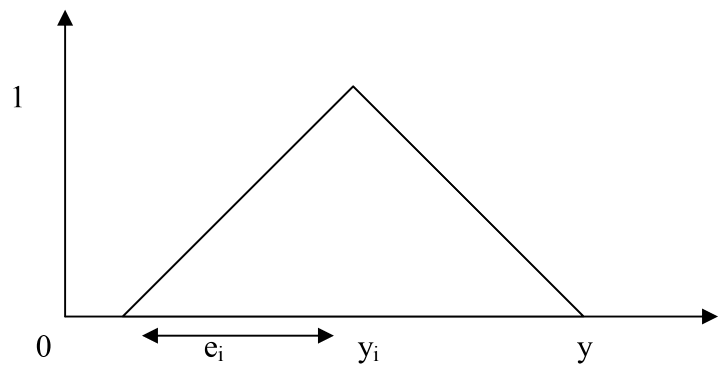
y_i قيمة المركز

e_i قيمة

اما دالة الانتماء لمتغير الاستجابة فهي كما في الصيغة الاتية:

$$\mu_{Y_i}(y_i) = \begin{cases} 1 - \dfrac{|y_i - y|}{e_i} & y_i - e_i < y < y_i + e_i \\ 0 & \text{otherwise} \end{cases} \quad ....(4)$$

ويمكن توضيح الدالة كما في الشكل (2) .



الشكل (2) متغير الاستجابة y كعدد مضبب بهيئة دالة انتماء مثلثية

هناك شروط فرضت لايجاد صيغة نموذج انحدار خطي مضبب والتي هي: (Savic and Pedrycz,1991).

1-استخدم معيار لتصغير الابهام الكلي (S) اوالغموض ويعرف كمجموع الانتشارات المستقلة لمعلمات المضببة للنموذج.

$$\text{Min } S = c_1+c_2+\ldots+c_n \quad \ldots\ldots\ldots \ (5)$$

2- قيمة الانتماء لكل مشاهدة $Y_i = (y_i \ , \ e_i)$ هي اكبر وتساوي من حد الانتماء المفترض (h) (threshold) إذ $(h \in [0,1])$ أي ان

$$Y(y_i) \geq h \qquad \forall \ i=1,\ldots,N \quad \ldots\ldots\ldots \ (6)$$

$$Y_i^h \subseteq Y_i^{*h} \qquad \qquad \bar{h}_i$$

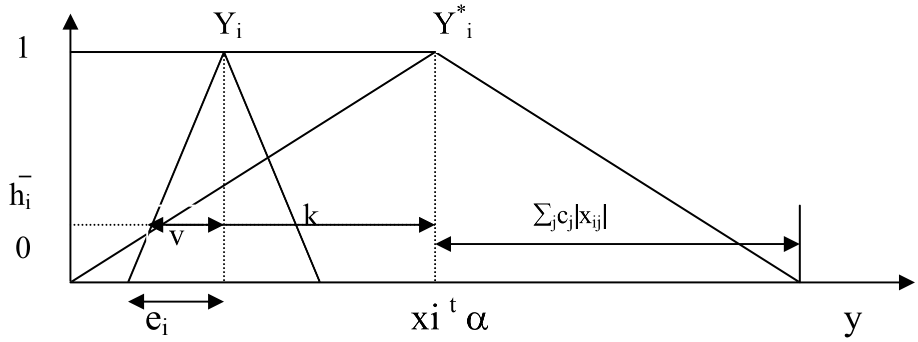
وتقاس بواسطة الدليل أي اعلى من قيمة (h) المختارة لـــ

3-درجة مطابقة تقدير نموذج خطي مضبب $Y_i^*=Ax_i^*$ لبيانات معطـــاة $Y_i=(y_i,e_i)$ إذ

$$Y_i^h = \{y \mid \mu_Y(y) \geq h\}$$
$$Y_i^{*h} = \{y \mid \mu_{Y^*}(y) \geq h\} \quad \ldots\ldots\ldots(7)$$

والتي هي مجموعات مستوى h (h-Level sets) .وان قيمة (h) تختار من قبل صانع القرار وتكون قيمتها محصورة بين الصفر والواحد (Kim et al.,1996).

اما الدليل $\bar{h}_i$ فيتوضح من خلال الشكل (3) (Tanaka et al.,1982) و (Heshmaty and Kandel,1985) .



الشكل (3) يوضح الدليل $\bar{h}_i$

اذ ان درجة مطابقة النموذج لكل البيانات المعطاة ($y_1,\ldots,y_N$) تعرف بواسطة $\min[\bar{h}_i]$

إذ $\bar{h}_i$ تكون درجة مطابقة نموذج الانحدار الخطي المضبب للعينة (i) ودرجة المطابقة للنموذج التي تكون فيه اقل قيمة $\bar{h}_i$ للمشاهدات.

ملاحظة:

عندما تكون قيمة الانتشار (e_i) مساوية للصفر أي ان بيانات المخرجات تكون غير مضببة فان قيمة $\bar{h}_i$ وقيمة دالة الانتماء للمتنبأ بها $\mu_Y*(y_i)$ تكونان متساويتين (Heshmaty and Kandel,1985).

لذا فان الدالة الموضحة في (3) والمتباينة (6) يمكن اعادة كتابتهما بالصيغة الاتية:

$$(1-h)c^t\,|x| - |y - x^t\alpha| \geq 0 \qquad x \neq 0 \qquad \ldots\ldots\ldots\ldots (8)$$

فان المتباينات المشروطة التي تحقق قيمة حد الانتماء ومقياس الغموض تكون خطية مع علاقة المعلمات غير المعلومة (المركز، الانتشار) ومن ثم توظف مسألة (LP) لتقدير معلمات الانحدار المضبب، التي تبنى على اساس دالة الهدف كما في المعادلة (5) والقيد (8) ويتمثل بشكل زوج من القيود المتباينة، وفيما ياتي الصيغة العامة لمسألة (LP) :

$$\left.\begin{array}{l}
\text{Min } S = \sum_{j=1}^{m} c_j \\
\text{s.t.} \\
\qquad \alpha^t x_i + (1-h)\sum_{j=1}^{m} c_j\,|x_{ij}| \geq y_i + (1-h)e_i \\
\qquad -\alpha^t x_i + (1-h)\sum_{j=1}^{m} c_j\,|x_{ij}| \geq -y_i + (1-h)e_i \\
c_j \geq 0\ ,\ x_i \geq 0 \\
\forall i = 1,2,\ldots,N
\end{array}\right] \ldots\ldots\ldots\ldots (9)$$

إذ α و c هما متجهان و S هو المجموع الكلي للغموض ، أي مجموع الانتشارات المستقلة للمعلمات المضببة للنموذج (Savic & Pedrycz,1991) .

اما عدد القيود المتمثلة في مسألة (LP) تكون (2xN) فهي عادة اكثر من عدد المتغيرات (m).

وقد اجري تعديل لنموذج (Tanaka) عام 1987 من قبل (Tanaka,1987) . إذ ظهرت انتقادات عديدة حول نموذج (Tanaka) الموضح اعلاه ، ومن هذه الانتقادات هي ان العديد من قيم الانتشارات (c_j) قد تصبح قيمتها مساوية للصفر عند اجراء الحل لـ (LP). وبذلك فان العلاقة بين متغير الاستجابة والمتغيرات التوضيحية تكون حادة (هشة) ، ولتفادي هذه المشكلة فقد اجري التعديل على دالة الهدف لنموذج (Tanaka) (5) ، التي كانت ذات اقل مجموع كلي لانتشارات المعلمات المضببة A_i في حين أصبحت ذات اقل مجموع كلي لانتشار قيمة التنبؤ لـ $(\tilde{Y}_i)$ لان هذه القيمة هي ايضا تكون مضببة.

عندها تصبح مسألة (LP) كما يأتي:

$$\left.\begin{array}{l}
\text{Min } S = \sum_{i=0}^{n}\sum_{j=1}^{m} c_j x_{ij} \\
\text{s.t.} \\
\quad \alpha^t x_i + (1-h)\sum_{j=1}^{m} c_j \,|\, x_{ij} \,|\, \ge y_i + (1-h)e_i \\
\quad -\alpha^t x_i + (1-h)\sum_{j=1}^{m} c_j \,|\, x_{ij} \,|\, \ge -y_i + (1-h)e_i \\
c_j \ge 0,\ x_i \ge 0 \\
\forall i = 1,2,\dots,N
\end{array}\right] \quad \dots\dots\dots\dots(10)$$

(Tanaka,1987) ، (Redden and Woodall,1994) و (Hojati et al.,2005) .

2- انحدار خطي مضبب بمربعات صغرى

Fuzzy Least Squares Linear Regression (FLSLR)

إن هذه الطريقة قــدمها (Savic and Pedrycz,1991) كصــيغة اخــرى لطريقة الانحدار الخطي المضبب ، فقد تم تنفيذ طريقة الحل بمرحلتين هما:

المرحلة الاولى:

ايجاد مطابقة خط انحدار بواسطة استخدام المعلومات المتاحــة حــول نقــاط المركز للمشاهدات ، أي ان بيانات المدخلات تكون غير مضببة (الحل يتم بواسطة طريقة المربعات الصغرى الاعتيادية) ، وبعد ايجاد المقدرات من هذه المرحلــة إذ تعد كمتجه (α^*) يستخدم كمجموعة بيانات لتنفيذ المرحلة الثانية.

وتستخدم طريقة المربعات الصغرى التقليدية لتقدير نموذج الانحــدار الخطــي العام إذ ان:

$$\alpha^* = (x'x)^{-1} x'y$$

المرحلة الثانية :

يتم تنفيذ مسألة (LP) أي استخدام طريقة معيار اقل مجموع كلــي للغمــوض (5) مع متباينات الشرطية لحد الإنتماء (6).

ان الاختلاف الواضح بين نموذج (Tanaka) وهذه الطريقة هــو اســتبدال المعلمة (α) بالمعلمة (α*) الناتجة من المرحلة الاولى وكما ياتي:

$$\left.\begin{array}{l} \text{Min } S = \sum_{j=0}^{m} c_j \\ \text{s.t.} \\ \quad \alpha^* + (1-h)\sum_{j=0}^{m} c_j \mid x_{ij} \mid \geq y_i + (1-h)e_i \\ \quad -\alpha^* + (1-h)\sum_{j=0}^{m} c_j \mid x_{ij} \mid \geq -y_i + (1-h)e_i \\ c_j \geq 0 \\ \forall i = 1,2,\ldots,N \end{array}\right] \quad \ldots\ldots\ldots\ldots\ldots(11)$$

إذ

c متجه للمتغيرات غير المعلومة.

h قيمة يحددها صانع القرار.

ومن الجدير بالذكر ان اختيار طريقة المربعات الصغرى الاعتيادية في المرحلة الاولى قد اكد على استخدامها (Redden & Woodall,1994) في دراسة لاحقة وذلك لسهولة الحساب.

مقارنة بين طريقتي الانحدار التقليدي والمضبب

ان كلا الانحدارين التقليدي والمضبب يمتلكان تطورا في مختلف المفاهيم النظرية والمنهجية، ومن ثم هناك اختلاف كبير بين هاتين الطريقتين من ناحية الفروض الاساسية وتقدير المعلمات، ومن خلال النقاط الاتية يمكن توضيح اوجه المقارنة :

1-الانحرافات في الانحدار التقليدي بين القيم المشاهدة والقيم التقديرية ، يفترض ان تكون مقياسا" للأخطاء . وهذه الانحرافات تاتي من المصدرين:

ا- اهمال بعض المتغيرات الموضحة في حين انها مهمة.

ب- اخطاء القياسات العشوائية في مشاهدات مسجلة .

اما الانحدار المضبب فان هذه الاخطاء تاتي من الافتراض بعدم تحديد هيكلية النظام والالتباس في معرفة النموذج (Tanaka et al.,1982)، فتكون الانحرافات

بين قيم المشاهدات والقيم المقدرة معتمدة على عدم تحديد وغموض المعلمات التي تغطي تركيبة النظام أي بمعنى اخر ناشئة من ضبابية النظام (Chang & Ayyub,2001).

2-الانحدار التقليدي يستخدم مجموعة البيانات غير المضـــببة (البيانـــات الحـــادة اوالهشة) . فتحليل الانحدار هو اداة احصائية واسعة الاستخدام لوصف العلاقة بين المتغيرات والتنبؤ لقيم ظاهرة ما. اما الانحدار المضبب فيستخدم مجموعة البيانات المضببة التي تكون بهيئة اعداد مضببة . والغــرض مــن الانحــدار المضبب هو لايجاد العلاقة بين المتغيرات في ظاهرة غامضة وعــدم الدقــة باستخدام دوال مضببة تعرف بواسطة نظرية المجموعات المضببة للعالم لطفي زاده.

3-الانحدار المضبب اساسه نظرية المجموعات المضببة ، إذ تعرف درجة انتماء مجموعة من البيانات بواسطة دالة الإنتماء بينما الانحدار التقليــدي اساســه نظرية المجموعات التقليدية.

4-نموذج الانحدار التقليدي عند توافر n من المتغيرات المفسرة فيكون النموذج بالشكل الاتي:

$$Y_i=\beta_0+\beta_1X_{i1}+\ldots+\beta_nX_{in}+\varepsilon_i$$

إذ Y_i قيمة متغير الاستجابة

X_{ij} قيمة لـــ j من المتغير المفسر في .i من المشاهدات $j=1,\ldots,n$.

β_j معلمات غير المعلومة للانحدار

ε_i حد الخطأ العشوائي او حد الاضطراب يتوزع توزيعا طبيعيا بوسط حسابي يساوي الصفر وتباين ثابت (Gunst & Mason,1980).

اما نموذج الانحدار المضبب فهو كما ياتي:

$$= \mathbf{A'X}$$
$$Y_i=A_0+A_1x_{i1}+\ldots+A_nx_{in}$$

إذ ان Y_i قيمة متغير الاستجابة

$\mathbf{A} = (A_0, A_1, ..., A_n)'$ معلمات غير معلومة ويمكن ان يرمز لها بصيغة متجه كما في الصيغة الاتية:

$$\alpha = \{\alpha_0, \alpha_1, ..., \alpha_n\}'$$
$$c = \{c_0, c_1, ..., c_n\}'$$
$$A = \{\alpha , c\}$$

α_j قيمة الوسط أي المركز للمعلمة A_j

c_j قيمة الانتشار للمعلمة A_j ، $j=0,1,....,n$

$\mathbf{X} = [1, x_{i1}, \ldots, x_{in}]$ المتغيرات المفسرة

المعلمات المقدرة في الانحدار التقليدي تكون اعدادا" حادة اوهشة بينما في الانحدار المضبب تكون اعدادا" مضببة أي تشتمل على قيمة المركز وقيمة الانتشار، ويمكن تمثيل القيمة بشكل فترة أي درجة انتماء لأي عدد حقيقي في فترة هي (قيمة المركز – قيمة الانتشار ، قيمة المركز + قيمة الانتشار)، وتعرف بواسطة دالة الانتماء . وبما ان المعلمات المقدرة هي اعداد مضببة فيصبح تقدير متغير الاستجابة هو عددا مضببا ايضا.

5-يخصص صانع القرار مستوى الثقة (Confidence level) في الانحدار التقليدي كقيمة (95%) ، هذه القيمة تكون مناظرة لقيمة المعلمة h التي يختارها صانع القرار في الانحدار المضبب كما يسميها بعض الباحثين حد الانتماء.

6-هدف الانحدار المضبب هو اقل ضبابية في قيم التنبؤ لمتغير الاستجابة وهذا يتم بواسطة مجموع اقل الانتشارات للمعلمات المضببة أي القيمة الدنيا لـ $(c_0+c_1+.....+c_n)$ ، بينما في تحليل الانحدار التقليدي ما يناظره هو مقياس المربعات الصغرى.

وفي النهاية فان الانحدار التقليدي والمضبب يعدان جزءا" من اللاتاكدية أي العشوائية (Randomness) والضبابية (Fuzziness) فهما نوعان من عدم الوضوحية في تحليل الانحدار (Chang & Ayyub,2001).

استخدامات الانحدار المضبب

1-يعالج مشاكل الانحدار ناقصة المعنوية حول البيانات لتحديد نماذج الانحـــدار علاقـــة غامضـــة بـــين متغيـــرات الاســـتجابة والمفســـرة (Savic & Pedrycz,1991).

2-يستخدم في حالة كون حجم العينات قيد الدراسة تكون صغيرة (McCauley & Wang,1999)، اذ يعد الانحدار المضبب بديلا" متاحا" للانحدار التقليـــدي في تقدير المعلمات عندما تكون مجموعة البيانات غير كافية لاســـناد تحليـــل الانحدار التقليدي، والملائمة لنمـــوذج الانحـــدار تكـــون ضـــعيفة (Kim et al.,1996).

3-يستخدم في حالة مجموعة بيانات المدخلات غير المضببة ولكن المخرجـــات مضببة اوكلتاهما مضببتان. او البيانات غير مضببة لكـــن العلاقـــة تكـــون مضببة وغامضة. فعندما تكون نمذجة نظام مضبب مع دوال خطية مضـــببة فان الغموض لبيانات المخرجات ربما سببها هو اللاتاكدية لمعلمات النمـــوذج والغموض في بيانات المدخلات (Lee & Chen,2001).

الجانب التطبيقي

هشاشة العظام (ترقق العظام) Osteoporosis

التعريف بمرض هشاشة العظام

هو نقص في الكتلة العظمية مما يؤدي الى نقص في قوة مخاطر حدوث الكسور وزيادتها ، ويعد من اكثر امراض العظام شيوعا خصوصا لجهة ارتفاع نسبة الاصابة بكسور خاصة في العمود الفقري ، الورك والساعد، وتؤدي الى نسبة عالية من هذا المرض الى الوفيات فضلا عن إعاقات جسدية متقدمة(Haslett et al., 1999) و (بلوط،2001) .

وهو مرض يصيب المسنين وخصوصا النساء فوق سن الخمسين ، سببه نقص شديد في كمية الكالسيوم في العظام ، العظام تتكون من طبقة خارجية قاسية تحوي

في داخلها نسيجا يشبه في بنيانه خلايا النحل هذا العظم يهدم بشكل مستمر ولكن يتم تعويضه مباشرة بنسيج جديد مما يبقي تركيب العظم قويا ومدعما" ، بعد تجاوز سن (35) سنة سواء عند الرجال او النساء فان عملية فقدان العظم تبدأ بالتغلب على عملية التعويض العظمي مما يؤدي الى فقد الكلس بشكل بطئ ولكن مستمر. يدعى مرض ترقق العظام بالوباء الصامت لانه غالبا مايكون الكسر هو اول دليل على وجوده. ان عملية فقدان العظم تبدأ بالازدياد بشكل حاد عند النساء بعد سن اليأس بسبب التناقص في مقادير هرمون الاستروجين في الجسم (Osteoporosis,2002).

تطبيق طرائق الانحدار الخطي المضبب

اعتمد هذا المبحث على نوعين من البيانات لتطبيقها في طرائق الانحدار المضبب، النوع الاول تكون فيها بيانات المدخلات والمخرجات غير مضببة (هشة) ، وذلك من خلال قياس كثافة العظم من فحص لجهاز الـ (DEXA) لثلاثين مريضا (10 ذكور ، 20 انثى).

(x_i,y_i) $\quad i=1,2,\ldots,30$

إذ

y_i متغير الاستجابة الذي بدوره يمثل كثافة العظم (Bone Mineral Density) (BMD).

x_i المتغير المفسر والذي يمثل العمر.

اما النوع الثاني من البيانات فتكون فيها المدخلات غير مضببة والمخرجات مضببة.

1- بيانات المدخلات والمخرجات غير المضببة

نموذج Tanaka (FLR)

لتطبيق البيانات على النموذج تستخدم مسألة (LP) من (9) التي تمثل دالة الهدف مع متباينات القيود، سيتولد 60 قيدا باعتبار (2xN=60) وبذلك تكون مسألة

$$\left.\begin{array}{l} \text{Min } S=\sum_{j=1}^{m} c_j \\ \text{s.t.} \\ \quad \alpha^t x_i + (1-h)\sum_{j=1}^{m} c_j \,|\, x_{ij} \,| \geq y_i \\ \quad -\alpha^t x_i + (1-h)\sum_{j=1}^{m} c_j \,|\, x_{ij} \,| \geq -y_i \\ c_j \geq 0,\quad x_i \geq 0 \qquad\qquad \forall i=1,2,\ldots,N \end{array}\right] \qquad \ldots\ldots\ldots\ldots\ldots\ldots(12)$$

(LP) كما في المنظومة (12) .

وتكون قيمة (h) المختارة مساوية لــ (0.5) اما التحليل فقد تم تنفيذه في برنامج الجاهز(LINDO) (Winston,1994) للحصول على المعلمات المضببة والمجموع الاقل لانتشارات المعلمات الذي يتمثل بدالة الهدف. وكان النموذج الخطي كما ياتي:

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_i$$

$$\tilde{Y}_i = (\alpha_0 + c_0) + (\alpha_1 + c_1) x_i$$

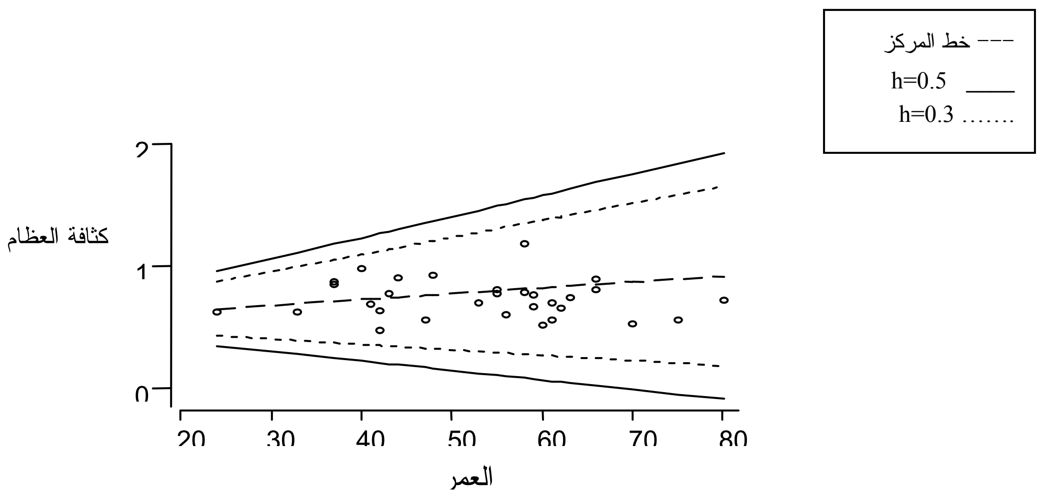
$$\tilde{Y}_i^{0.5} = (0.53455,0) + (0.0048,0.012601)\, x_i \quad \ldots\ldots\ldots\ldots (13)$$

إذ ان (0.53455,0) هوالمقطع المضبب، القيمة المركزية (0.53455) والانتشار (0).

وعند اختيار قيمة (h) مساوية لــ (0.3) فكان النموذج الخطي المضبب كما ياتي:

$$\tilde{Y}_i^{0.3} = (0.53455,0) + (0.0048,0.009) x_i \quad \ldots\ldots\ldots\ldots\ldots\ldots (14)$$

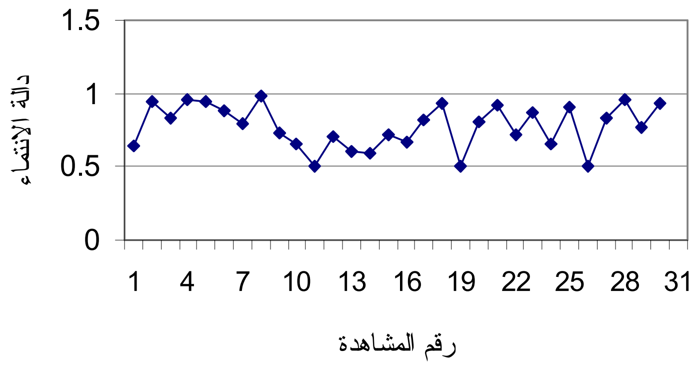
ومن خلال المعادلتين (13) و(14) يمكن توضيحها بالشكل (4) كحدين اعلى



وادنى لكل من قيم (h) المختارة والتي تكون قيمتها محصورة بين (0,1).

الشكل (4) نموذج انحدار مضبب عند قيم h=0.3 و h=0.5 لبيانات (x,y) غير مضببة

اما درجة مطابقة النموذج الخطي المضبب المقدر($\tilde{Y} = \tilde{A}\, x_i$) للبيانات المعطاة (Y_i) فيتم بواسطة قياس الدليل $\bar{h}_i$ الذي يحقق ($\tilde{Y}_i^h \supseteq Y_i^h$) عند قيمة h=0.5 وكما موضح في الشكل (5).



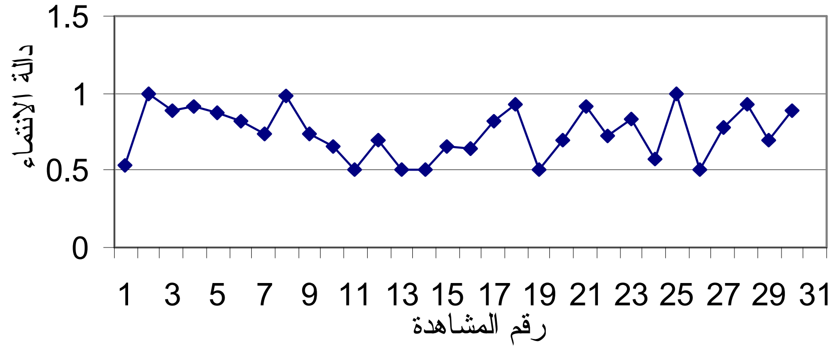
الشكل (5) درجة مطابقة نموذج (FLR)

اما في حالة تطبيق البيانات على نموذج (Tanaka)المعدل باستخدام المنظومة (10) أي فقط تغيير دالة الهدف.

كان النموذج الخطي المضبب، عند قيمة (h=0.5) هو:

$$\tilde{Y}_i^{0.5} = (0.462, 0.145) + (0.0065, 0.0091)x_i \quad \ldots\ldots\ldots\ldots \ (15)$$

كما كانت درجة مطابقة النموذج المقدر($\tilde{Y}_i^h$) للبيانات المعطاة (Y_i) تحقق الشرط ($\tilde{Y}_i^h \supseteq Y_i^h$) كما هو موضح في الشكل (6) .



الشكل (6) درجة مطابقة نموذج Tanaka المعدل

وقد تم تنفيذ مسألة (LP) لقيم عدة من (h) (اختيار صانع القرار) والقيم هي (0,0.3,0.5,0.7,0.9) ، وتبين من خلال المقارنة بين نتائج هذه القيم انه كلما كانت قيمة(h) اصغر كانت المسافة بين الحد الاعلى والادنى اقصر، كما موضحة في الشكل (4). وفي اغلب الدراسات كان يتم اختيار القيم الوسطى لــ (h) كقيمة (0.5) او (0.7) (Chang& Ayyub,2001). وكذلك عند تنفيذ مسألة (LP) لقيم عدة لنموذج (Tanaka) والمعدل ظهر كيف ان قيم المراكز لم تتغير عند تغير قيم (h)، اما قيم الانتشار فيجرى عليها التغيير (Hojati et al.,2005) بمعنى اخر كلما زادت قيم (h) زادت قيم الانتشار (Moskowitz & Kim,1993)، كما في الجدولين (1) و(2) على التوالي، وكذلك تم توضيح قيم الانتشار عند نموذج (Tanaka) لعام 1982 ربما تظهر قيمتها صفر في حين نموذج المعدل لم تظهر. كما يبين الجدولان عند زيادة قيم (h) يزداد مجموع الانتشار الذي يمثل دالة الهدف من مسألة (LP) (Chang & Lee,1994).

الجدول (1) قيم المعلمات المضببة لنموذج (Tanaka et al.,1982).

Min S	A_1		A_0		**h**
	α_1	c_1	α_0	c_0	
0.0065	0.0048	0.0065	0.534	0	**0**
0.0092	0.0048	0.0092	0.534	0	**0.3**
0.0126	0.0048	0.0126	0.534	0	**0.5**
0.0212	0.0048	0.0212	0.534	0	**0.7**
0.0632	0.0048	0.0632	0.534	0	**0.9**

الجدول (2) قيم المعلمات المضببة لنموذج Tanaka المعدل (1987).

Min S	A_1		A_0		**h**
	α_1	c_1	α_0	c_0	
9.470	0.006	0.004	0.462	0.072	**0**
13.529	0.006	0.006	0.462	0.013	**0.3**
18.941	0.006	0.009	0.462	0.145	**0.5**
31.568	0.006	0.015	0.462	0.241	**0.7**
94.704	0.006	0.045	0.462	0.725	**0.9**

الاســـتنتاجات:

من خلال دراسة الجانب التطبيقي للبحث ثم التوصل الى الاستنتاجات الآتية:

1- اذا كانت البيانات مضببة فيجب استخدام الانحدار المضبب في التحليل وكذلك في حالة ايجاد العلاقة بين المتغيرات في ظاهرة غامضة.

2-نموذج Tanaka هو عموما يستخدم لتقدير المعلمـــات المضـــببة فـــي نمـــاذج (FLR) وحتى في حالة كون البيانات غير مضببة. ويفضل اســـتخدام نمـــوذج Tanaka المعدل بدلا من النموذج الاولي لتفادي ظهور المعلمات المقدرة غير المضببة.

المصادر

الخياط، باسل يونس ذنون الخياط ، 2004. اللاتأكدية من خلال نظرية الاحتمال ونظرية المجموعات المضببة، المجلة العراقية للعلوم الاحصائية، العدد 6، المجلد 4، 18-31.

المعزاوي ،علي عبد السلام،1977. بحوث العمليات ،دار العلوم الحديثة بيروت- لبنان.

بلوط، هاجر احمد، 2001. الصحة والحياة..مرض ترقق العظام ، متاح على الموقع
http://www.baqiatollah.org/archeive/2001/114/osrah/osra/osra003.html.

ويكيبديا الموسوعة الحرة ،2004. فروع الرياضيات متاح على الموقع http://ar.wikipedia.org.

Chang, P.-T., Lee, E.S., 1994. Fuzzy Linear Regression with Spreads Unrestricted in Sign. Comput. Math. Appl.28, 61–70.

Chang,Y.-H. and Ayyub, B.M., 2001. Fuzzy Regression Methods—A Comparative Assessment. Fuzzy Sets and Systems 119, 187–203.

Gunst, R.F. and Mason,R.L.,1980. Regression Analysis and its Application, Marcel Dekker ,INC. New York.

Haslett, C., Chilver, E.R., Hunter, J.A.A. and Boon, N.A. , 1999. Davidson's Principles and Practice of Medicine, Harcourt Brace and Company U.K.

Heshmaty, B. and Kandel, A., 1985. Fuzzy Linear Regression and Its Applications to Forecasting in Uncertain Environment. Fuzzy Sets Systems 15, 159-191.

Hojati,M. ,Bector,C.R. and Smimou, K.,2005. A Simple Method for Computation of Fuzzy Linear Regression. European Journal of Operational Research (In Press).

Kandel, A.1986. Fuzzy Mathematical Techniques with Applications, Addison-Wesley Publishing Company, England.

Kim, K.J., Moskowitz, H., and Koksalan, M., 1996. Fuzzy Versus Statistical Linear Regression. Eur. J. Oper. Res. 92, 417–434.

Lee,H.T. and Chen, S. H. ,2001. Fuzzy Regression Model with Fuzzy Input and Output Data for Manpower Forecasting. Fuzzy Sets and Systems, 119, 205-213.

McCauley P. B. and Wang, H.,1999. Measurement of Cumulative Trauma Disorders Risks of Tasks Using Fuzzy linear regression , IEEE Transactions on Systems , Man , and Cybernetics, Vol.29,1,pp.1-14.

Moskowitz, H. and Kim, K., 1993. On assessing the H Value in Fuzzy Linear Regression. Fuzzy Sets and Systems 58, 303–327.

Nasrabadi, M.M. and Nasrabadi, E. , 2004. A Mathematical – Programming Approach to Fuzzy Linear Regression Analysis, Applied Mathematics and Computation, Vol.155, 3, 873-881.

Osteoporosis, 2002. Available at: http://www.chamclinic.com/edu/ osteoporosis _pre.doc. Accessed June 20, 2004.

Redden, D.T. and Woodall, W.H., 1994. Properties of Certain Fuzzy Linear Regression Methods. Fuzzy Sets and Systems 64, 361–375.

Savic, D.A. and Pedrycz, W. ,1991. Evaluation of Fuzzy Linear Regression Models. Fuzzy Sets and Systems, 39, 51-63.

Tanaka, H., Uejima, S.and Asai, K., 1982. Linear Regression Analysis with Fuzzy Model. IEEE Trans. Systems, Man, Cybernet. 12, 903–907.

Tanaka, H.,1987. Fuzzy Data Analysis by Possibilistic Linear Models, Fuzzy Sets and Systems 24 ,363 - 375.

Wang, H.-F. and Tsaur, R.-C., 2000. Insight of a Fuzzy Regression Model . Fuzzy Sets and Systems 112 , 355-369

Winston,W.L.,1994. Operations Research Applications of Algorithms. Duxbury Press, USA.