# استخدام المحاكاة لتقدير أنموذج الانحدار اللامعلمي الضبابي

أ.م.د محمد جاسم محمد حسين* مؤيد سلمان عباس**

**المستخلص :**

في هذا البحث تم استخدام ثلاثة أساليب للتمهيد (أسلوب التمهيد الخطي الموضعي وأسلوب تمهيد النواة وأسلوب الجوار الأقرب-k ) لتقدير أنموذج الانحدار اللامعلمي الضبابي . أذ مُثلّ المتغير المعتمد باستخدام الأعداد الضبابية المثلثية (Fuzzy-Triangular-Numbers) وتم الاعتماد على مقياس المسافة للأعداد الضبابية والمُقترح من قبل (Diamond) ويتم أستخدام طريقة التمهيد الأمثل الطبيعي (Normal Optimal Smoothing) لاختيار القيمة المثلى لمعلمة التمهيد (h) حيث تم تضبيبها لملائمة الانموذج اللامعلمي الضبابي .

ولبيان افضل أسلوب للتمهيد يتم استخدام المحاكاة للمقارنة مابين تلك الأساليب باستخدام معيار التحيز لمعرفة أفضل مُقدر لتمثيل البيانات المُولّدة عن طريق أسلوب المحاكاة ، وعند دوال النواة المعتمدة والمقترحة ، وتبينّ من خلال معيار المقارنة بأن أقل قيمة لهذا المعيار كانت عند استعمال أسلوب الجوار الأقرب- k وعند دالة النواة (Gaussian) .

# Estimation Nonparametric Fuzzy Regression model using simulation

**Dr.Mohammed Jasim Mohammed Muayed Salman Abbas**

## Abstract

**In this research, three techniques of smoothing were used, local linear smoothing, kernel smoothing, k-Nearest Neighbor to estimate the model of the Fuzzy nonparametric regression. A dependent variable was represented by using Fuzzy Triangular Numbers . Based on the distance measure for the Fuzzy numbers which proposed by (Diamond) and we used the normal optimal smoothing method to select the optimal value of the smoothing parameter (h) where it was Fuzzified to fit the Fuzzy nonparametric model .**

**In order to demonstrate the best technique of smoothing, the simulation was used to compare these techniques using the (bias) criterion to determine the best estimate of the data generated by the simulation method and at the approved and proposed kernel functions. The comparison criterion showed that the lowest value of this criterion was when using the k- nearest neighbor technique at kernel functions (Gaussian).**

---

* جامعة بغداد / كلية الادارة والاقتصاد .

** باحث .

مستل من اطروحة دكتوراة

مقبول للنشر بتأريخ 2017/9/6

## 1. مقدمة

يهدف الانحدار الخطي الضبابي الى نمذجة ظاهرة غير دقيقة أو غامضة وذلك باستخدام معلمات الأنموذج الضبابي وفي حال تحقق الفرضيات الخاصة بطريقة المربعات الصغرى الاعتيادية فأن الانحدار الضبابي يكون اكثر فاعلية واكثر مرونة للمشاكل المختلفة كبديل للانحدار الكلاسيكي ، لقد تمت دراسة الانحدار الضبابي المعلمي من قبل الكثير من الباحثين ألا انه كان أنموذجاً ضعيفاً، ولتحسين هذا الأنموذج فقد قام الباحثون باقتراح أنموذج الانحدار اللامعلمي الضبابي بمتغير مستقل اعتيادي (crisp) ومتغير معتمد ضبابي (Fuzzy) وأن السبب وراء تسميته باللامعلمي هو عدم افتراضه لأي صيغة دالية لتوزيع الخطأ ، حيث يكون اعتماده على البيانات بصورة كبيرة ، في كافة المجالات مثل المال والاقتصاد والهندسة والعلوم الطبية والعلوم الاجتماعية وغيرها من المجالات عادة مايحتاج الباحثون الى تقدير مشاهداتهم ، ويتم ذلك عن طريق استخدام أنموذج الانحدار المناسب، لذلك سيتم دراسة أنموذج الانحدار اللامعلمي الضبابي أحادي المتغير بمتغير مستقل اعتيادي (crisp input) ومتغير معتمد ضبابي (Fuzzy output) من النوع المثلثي المتماثل $(\mathbf{LR})$ .

## 2. الجانب النظري

### 1.2 المشكلة واهمية البحث

هناك الكثير من المشاكل التي تواجهنا في الحياة اليومية ولكن المعلومات والبيانات الخاصة والمتعلقة بها تكون من الصعوبة تسجيلها أو جمعها بصورة دقيقة أو أنها تكون غامضة أوتكون متغيرات موصوفة عن طريق المصطلحات اللغوية، وعليه فأن نظرية المجاميع الضبابية تكون الوسيلة المناسبة والفاعلة في صياغة النماذج الإحصائية حيث تكون الأداة البارزة لمعالجة عدم الدقة أو الغموض عندما تكون المشاهدات ضبابية . فضلاً عن انه في العديد من المشاكل يكون من غير المنطقي أن يتم حساب علاقة الانحدار المعلمي الضبابي مسبقاً لذلك تم تطوير بعض الأساليب لمعالجة مشاكل الانحدار الضبابي في حالة كون الصيغة المحددة لعلاقة الانحدار غير مُعرّفة مسبقاً اي حالة الانحدار اللأمعلمي الضبابي موضوع بحثنا .

### 2.2 هدف البحث

يهدف البحث الى أستخدام المحاكاة لتقدير دالة الانحدار اللامعلمي الضبابي بمتغير مستقل اعتيادي أحادي المتغير (Univariate) ومتغير معتمد ببيانات ضبابية مثلثية (Triangular Fuzzy) من النوع (LR) وبالأعتماد على أساليب التمهيد ومنها أسلوب التمهيد الخطي الموضعي لمعالجة وتقدير دالة الأنحدار قيد البحث وبالأعتماد على مسافة (Diamond) وكذلك أستخدام أسلوب تمهيد النواة فضلاً عن أسلوب تمهيد الجوار الأقرب- K .

### 3.2 أنموذج الانحدار اللامعلمي الضبابي أحادي المتغير

"Univariate Fuzzy Nonparametric Regression Model"

يعد الأنحدار منهجية شائعة للتعبير عن العلاقة الدالية مابين المتغيرات ، وبما أن الصيغة الرياضية معلومة فأن القيمة لأحد المتغيرات يمكن التنبؤ بها من خلال قيم المتغيرات الأخرى ولكن وفي الكثير من الأحيان وفي بعض البحوث لايمكن الحصول بصورة دقيقة على البيانات العددية للمعلومات حول بعض الظواهر المدروسة ، وعليه فأن انحدار المربعات الصغرى التقليدية لا يمكن تطبيقها . لذلك وللتعامل مع هذا النوع من المشاكل يتم أستخدام مفهوم الأنحدار الضبابي لتقدير العلاقة الدالية مابين متغير الاستجابة والمتغيرات المستقلة في محيط ضبابي مع دالة خطية معلومة لذلك فقد سمي بالانحدار الخطي المعلمي الضبابي ، وبما أنه وفي الكثير من الحالات التطبيقية يكون من غير المنطقي التحديد المسبق لطبيعة وشكل العلاقة الدالية لعلاقة الانحدار المعلمي الضبابي أي (عدم الأعتماد على صيغة محددة) لذلك يتم الأعتماد على أنمُوذج الانحدار اللامعلمي الضبابي أحادي المتغير الأتي :

$$\mathbf{Y} = \boldsymbol{g}(\boldsymbol{x})\{+\}\boldsymbol{\varepsilon} = (\boldsymbol{l}(\boldsymbol{x}), \boldsymbol{m}(\boldsymbol{x}), \boldsymbol{u}(\boldsymbol{x}))\{+\}\boldsymbol{\varepsilon} \qquad \dots \quad (1)$$

إذ أن $(x)$ يُمثل المتغير المستقل الاعتيادي ومجاله الفترة بجميع قيم الأعداد الحقيقية ونرمز له (D) ، في حين أن(Y) يمثل المتغير المعتمد الضبابي(متغير الاستجابة) (Fuzzy dependent variable) ، وهو متغير ضبابي مثلثي متماثل (LR)[15].

$g(x)$ تمثل دالة انحدار ضبابي غير معلومة (دالة تمهيد) حدودها هي القيم العليا والمتوسطة والدنيا على التتابع $l(x)$ ،$m(x)$،$u(x)$ ،ومجالها من (D) إلى $\Re_{LR}$.

إذ أن

$$\Re_{LR} = \{k: k = (l_k, m_k, u_k)_{LR}\}$$

تمثل مجموعة تضم جميع الأعداد الضبابية (k) مـن النوع $(\mathbf{LR})$ وأن $L(\cdot)$ و $R(\cdot)$ تكون معلومة [15] .

وأن دالة الانتماء له مبينة في المعادلة الأتية [13]:

$$\mu_k(t)=\begin{cases} L\left(\frac{m-t}{m-l_k}\right) & \text{if } l_k \le t \le m \\ R\left(\frac{t-m}{u_k-m}\right) & \text{if } m \le t \le u_k \\ 0 & \text{otherwise} \end{cases} \quad \dots (2)$$

$\varepsilon$ يمثل حد الخطأ والذي يمثل الفرق مابين القيمة المشاهدة والقيمة المُقدرّة .

## 4.2 التمهيد وأساليب التمهيد

أن الفكرة الأساس للتمهيد تقوم على الافتراض بأننا نمتلك عينة عشوائية بحجم (n) من أزواج المشاهدات$(x_i, y_i)$، وان العلاقة مابين المتغيرين يمكن وصفها كما في الصيغة (1) . إذ أن التمهيد هو تقدير للدالة $g(x)$ [8].

أن التمهيد الموضعي يستند على توسيع تايلور (Taylor Expansion) حيث أنه لأي $y$ في جوار $x$

$$g(y) \approx g(x) + g'(x)(y-x)$$

أي أنه يمكن تقريب كل دالة تمهيد موضعياً عن طريق الدالة الخطية [4]

وعليه فأن التمهيد يستند على أسلوب المعدل الموضعي وحسب الصيغة الآتية [7] :

$$\hat{g}(x) = n^{-1}\sum_{i=1}^{n} W_{ni}(x)Y_i \qquad \cdots (3)$$

أذ أن :

$W_{ni}(x)$، $i = 1, 2, \dots, n$

تمثل سلسلة الأوزان التي من الممكن ان تعتمد على المتجه $\{X_i\}_{i=1}^{n}$ .

أن الفكرة المهمة والأساس لجميع أساليب التمهيد هو أن دالة الانحدار غير المعلومة $g$ تكون ممهدة الى حد ما والمشاهدات التي تكون قريبة من $(x)$ يجب أن تحتوي على المعلومات حول قيمة $(g)$ عند $(x)$ وعليه فمن الممكن أستخدام المعدل الموضعي للبيانات المشاهدة بالقرب من النقطة $(x)$ وذلك لإيجاد التقدير لـ $g(x)$ [3] . وبحسب الدراسات السابقة فقد أثبتت تلك الأساليب بأنها مفيدة ولاسيما في معالجة مشاكل الانحدار اللامعلمي [15] ، حيث أنها تمتلك معلمة للضبط تقوم بتحديد مقدار التمهيد الذي نقوم به .

سيتم دراسة أساليب التمهيد الآتية :-

1. التمهيد الخطي الموضعي (Local Linear Smoothing) .
2. تمهيد النواة (smoothing Kernel) .
3. الجوار الأقرب- K (K-Nearest Neighbor) .

## 5.2 تقديرات الانحدار اللامعلمي الضبابي

لأنموذج الانحدار اللامعلمي الضبابي في المعادلة رقم (1) وعلى أفتراض توفر عينة بحجم (n) من أزواج المشاهدات $(x_i , y_i)$ لقيم المتغير المستقل الأعتيادي والمتغير المعتمد الضبابي من النوع (LR) فأن هدف الأنحدار اللامعلمي الضبابي هو تقدير دالة التمهيد$g(x)$ عند أي $x \in D$ والمستندة على العينة المشاهدة $(x_i , y_i)$[15] .

ان دالة الانتماء (Membership Function) للمتغير المعتمد الضبابي المُقدر يجب ان تكون اقرب ما يمكن الى العدد الضبابي المشاهد المقابل ، لذلك يجب تقدير $u(x), m(x), l(x)$ لكل $x \in D$ أي بمعنى المطابقة الأفضل وهنالك بعض المسافات التي يمكن ان تقيس القرب (Closeness) مابين دوال الانتماء للمُخرج الضبابي المُقدر والمشاهد المقابل[9] .

وسنستخدم في بحثنا هذا المسافة المقترحة من قبل(Diamond) كمقياس للمطابقة وكما يأتي:

نفترض بأن

$$C = (l_c, m_c, u_c)_{LR} \ , \ \ D = (l_d, m_d, u_d)_{LR} \ \text{،} \ \ m_c, u_c, m_d, u_d \ge 0$$

يمثلان أي عددين ضبابيين من النوع (LR) في $\Re_{LR}$ .

حيث عرّف (Diamond) المسافة مابين C وD كما يأتي:

$$d^2(C,D) = (l_c - l_d)^2 + (m_c - m_d)^2 + (u_c - u_d)^2 \qquad \cdots \quad (4)$$

أن$L(\cdot)$ و$R(\cdot)$ في المعادلة (2) تُبين بأن دالة الانتماء للعدد الضبابي من النوع $(LR)$ يمكن ايجادها عن طريق المتوسط والحدود الدنيا والعليا للعدد الضبابي[15] .

## 1.5.2 أسلوب التمهيد الخطي الموضعي

لتوسيع اسلوب التمهيد الخطي الموضعي المستخدم في الإحصاء التقليدي سنعتمد على مسافة (Diamond) وذلك لمطابقة الأنموذج اللامعلمي الضبابي (1) وكما يأتي :

نفترض بأن $l(x)$ ،$m(x)$،$u(x)$ في دالة الأنحدار $g(x)$ تمتلك مشتقات مستمرة في المجال (D)، عندئذ وللقيمة المفترضة $(x_0)$ والتي تنتمي لـ (D) وباستخدام توسيع تايلور فأن$l(x)$ ،$m(x)$،$u(x)$ يمكن ان تُقرّب موضعياً (Locally) في جوار القيمة $(x_0)$على التوالي عن طريق الدوال الخطية الأتية[15]:

$$l(x) \approx \tilde{l}(x) = l(x_0) + l\,'(x_0)(x - x_0) \qquad \dots (5)$$

$$m(x) \approx \widetilde{m}(x) = m(x_0) + m\,'(x_0)(x - x_0) \qquad \dots (6)$$

$$u(x) \approx \widetilde{u}(x) = u(x_0) + u\,'(x_0)(x - x_0) \qquad \dots (7)$$

أذ ان $l'(x_0)$ , $m'(x_0)$, $u'(x_0)$ تمثل مشتقات لـــ $l(x_0)$ و $m(x_0)$ و$u(x_0)$ على التتابع عند النقطة $(x_0)$.

وبما أن المتغير المعتمد يكون ضبابياً وبالشكل الآتي :

$$\mathrm{Y}_i = \left(l_{yi}\,, m_{yi}, u_{yi}\right)_{LR} \quad , \quad i = 1, 2, \dots, n$$

فأن أزواج مشاهدات العينة قيد البحث ستكون كالأتي :

$$(\mathrm{X}_i, \mathrm{Y}_i) = \left(x_1, x_2, \dots, x_n\,, \left(l_{yi}, m_{yi}, u_{yi}\right)_{LR}\right) \quad , i = 1, 2, \dots, n$$

وعليه فأن تقدير دالة الانحدار $\widehat{g}(x_o)$ وبالاعتماد على مسافة (Diamond) وكما في المعادلة (4) وأسلوب التمهيد الخطي الموضعي سيقودنا الى مسألة المربعات الصغرى الموزونة موضعياً( Locally Weighted Least squares) أي تصغير المعيار الاتي:

Minimize

$$\sum_{i=1}^{n} d^2\left(\left(l_{yi}, m_{yi}, u_{yi}\right)_{LR} ، \left(\tilde{l}(x_i), \widetilde{m}(x_i), \widetilde{u}(x_i)\right)_{LR}\right) K_h(|x_i - x_0|) \qquad \cdots (8)$$

وعن طريق تعويض المعادلات (5) و(6) و(7) بالمعيار (8) نحصل على المعيار الآتي:

$$\text{Minimize} \sum_{i=1}^{n}\left(l_{yi} - l(x_0) - l'(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n}\left(m_{yi} - m(x_0) - m'(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|) \qquad \cdots (9)$$

$$+ \sum_{i=1}^{n}\left(u_{yi} - u(x_0) - u'(x_0)(x_i - x_0)\right)^2 K_h(|x_i - x_0|)$$

أي تصغير المعيار (9) بالنسبة الى كل من $l(x_0)$،$m(x_0)$، $u(x_0)$ و$l'(x_0)$, $m'(x_0)$ ،$u'(x_0)$ ولــدالة (Kernel) المُعطاة $k_{\mathrm{h}}(.)$ ومعلمة التمهيد (h) إذ أن:

$$\mathrm{k_h}(|x_\mathrm{i} - x_0|) = \frac{\mathrm{K}\left(\frac{|x_\mathrm{i} - x_0|}{h}\right)}{\mathrm{h}} ، \qquad i = 1, 2, \cdots, n \qquad \cdots (10)$$

تمثل سلسلة الأوزان عند النقطة $(x_0)$ التي تجعل البيانات القريبة من$(x_0)$ تساهم بشكل أكبر في عملية تقدير المعلمات عند$(x_0)$ من تلك التي تكون بعيدة مع تعديل في قيمة (h) .

وبحل مسألة المربعات الصغرى الموزونة سنحصل على تقديرات كلاً من $\hat{l}(x_0)$ و$\widehat{m}(x_0)$ و$\hat{u}(x_0)$ عند $(x_0)$ وكذلك الحصول على مشتقاتها $l'(x_0)$ ،$m'(x_0)$, $u'(x_0)$ وحيث أن الهدف من البحث تقدير دالة الانحدار اللامعلمي الضبابي$g(x)$ عند النقطة $(x_0)$ ، لذلك سنقوم بأخذ قيم $l(x_0)$ ،$m(x_0)$،$u(x_0)$ في المعادلة (9) والتي مُثلت عن طريق تقديراتها ، أي أن تقدير $g(x)$ عند$(x_0)$ سيكون :

$$\widehat{g}(x_o) = \left(\hat{l}(x_0), \widehat{m}(x_0), \widehat{u}(x_0)\right)_{LR} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (11)$$

$$= \left(\widehat{m}(x_0) - \widehat{\alpha}(x_0), \widehat{m}(x_0), \widehat{m}(x_0) + \widehat{\beta}(x_0)\right)_{LR}$$

نلاحظ بأن الدالة في المعادلة (9) متكونة من حاصل جمع ثلاثة أجزاء إذ ان كل جزء يتضمن بشكل منفصل على مجموعة مختلفة من المعلمات غير المعلومة، وعليه فسنقوم بأخذ المشتقات الجزئية للدالة بالنسبة الى المعلمات المجهولة وجعلها صيغ صفرية وبذلك سنحصل على ثلاثة مجاميع من المعادلات الخطية والتي تحتوي بصورة منفصلة على المعلمات الآتية [15] :

$$\left(l(x_0), l'(x_0)\right), \left(m(x_0), m'(x_0)\right)\left(u(x_0), u'(x_0)\right)$$

وبحل المعادلات الخطية وللمجاميع الصفرية كلاً على حدة نحصل على تقديرات تلك المعلمات .

وطبقاً لمبدأ المربعات الصغرى الموزونة وباستخدام رموز المصفوفات نستطيع الحصول على:

$$\left(\hat{l}(x_0), \hat{l}'(x_0)\right)^T = \left(X^T(x_0)W(x_0; h)X(x_0)\right)^{-1} X^T(x_0)W(x_0; h)L_y \,,$$

$$(\hat{m}(x_0),\hat{m}'(x_0))^T=\left(X^T(x_0)W(x_0;h)X(x_0)\right)^{-1}X^T(x_0)W(x_0;h)M_y\,, \qquad \cdots(12)$$
$$(\hat{u}(x_0),\hat{u}'(x_0))^T=\left(X^T(x_0)W(x_0;h)X(x_0)\right)^{-1}X^T(x_0)W(x_0;h)U_y\,,$$

فبدلاً من العدد الاعتيادي للمتغير المعتمد (Y) تم استخدام العدد الضبابي (Fuzzy number) .
إذ أن:

$$X(x_0)=\begin{pmatrix}1 & x_1-x_0\\ 1 & x_2-x_0\\ \vdots & \vdots\\ 1 & x_n-x_0\end{pmatrix}, L_y=\begin{pmatrix}l_{y_1}\\ l_{y_2}\\ \vdots\\ l_{y_n}\end{pmatrix}, M_y=\begin{pmatrix}m_{y_1}\\ m_{y_2}\\ \vdots\\ m_{y_n}\end{pmatrix}, U_y=\begin{pmatrix}u_{y_1}\\ u_{y_2}\\ \vdots\\ u_{y_n}\end{pmatrix}$$

وأن

$$W(x_0,h)=Diag\left(K_h(|x_1-x_0|)،K_h(|x_2-x_0|),\cdots,K_h(|x_n-x_0|)\right) \quad \cdots(13)$$

تمثل مصفوفة قطرية من الدرجة $n\times n$ وعناصرها القطرية مساوية الى:

$K_h(|x_i-x_0|)(i=1,2,\dots,n)$

يمكن كتابة التقدير لـ $g(x)$ في حالة كون المتغير المعتمد ضبابي مثلثي كما يأتي[15]:

$$\hat{g}(x_o)=\left(\hat{l}(x_0),\hat{m}(x_0),\hat{u}(x_0)\right)_{LR}$$
$$\hat{g}(x_o)=\left(\frac{\sum_{i=1}^n W_i l_{y_i}}{\sum_{i=1}^n W_i},\frac{\sum_{i=1}^n W_i m_{y_i}}{\sum_{i=1}^n W_i},\frac{\sum_{i=2}^n W_i u_{y_i}}{\sum_{i=1}^n W_i}\right) \qquad \dots(14)$$

وفي التطبيق العملي وللتبسيط يمكن تعريف الوزن[4]:

$$W_i=K\left(\frac{x-x_i}{h}\right)\left[Z_{n،2}-(x-x_i)Z_{n،1}\right] \qquad \cdots(15)$$

أذ أن :

$$Z_{n،j}=\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)(x-x_i)^j,\ j=1،2 \qquad \cdots(16)$$

## 2.5.2 أسلوب تمهيد النواة (smoothing Kernel)

وبنفس الطريقة أعلاه يمكن إيجاد التقديرات باستعمال تمهيد النواة (K-S) ولكن بسلسلة الاوزان عند $(x_0)$ بحسب الصيغة الاتية [15] :

$$W_i(x_0)=W_i(x_0;h)=\frac{K_h(|x_i-x_0|)}{\sum_{i=1}^n K_h(x_j-x_0)}\quad ,i=1,2,\dots,n \qquad \cdots(17)$$

لنحصل على

$$\hat{g}(x_0)=\left(\frac{\sum_{i=1}^n K_h(|x_i-x_0|)l_{y_i}}{\sum_{i=1}^n K_h(|x_i-x_0|)},\frac{\sum_{i=1}^n K_h(|x_i-x_0|)m_{y_i}}{\sum_{i=1}^n K_h(|x_i-x_0|)},\frac{\sum_{i=1}^n K_h(|x_i-x_0|)u_{y_i}}{\sum_{i=1}^n K_h(|x_i-x_0|)}\right)_{LR} \qquad \cdots(18)$$

إذ أن $\mathrm{K_h}(\cdot)=\frac{\mathrm{K}\left(\frac{\cdot}{\mathrm{h}}\right)}{\mathrm{h}}$ يمثل نفس الوزن في طريقة التمهيد الخطي الموضعي (L-L-S).

## 3.5.2 أسلوب تمهيد الجوار الأقرب –K (K-Nearest Neighbor )

وكما سبق يمكن إيجاد تقديرات الانحدار اللامعلمي الضبابي باستعمال تمهيد الجوار الاقرب-K ، أذ أن تقدير(K-N-N) يعتمد على الجوار والذي يتم تعريفه بالاعتماد على متغيرات (X) والتي تكون الجوارات الاقرب لـــ $(x)$ ويعتبر تقدير (K-N-N) معدل موزون في جوار متغير[7].
أن المُقدر(K-N-N) يكون بالصيغة الاتية :

$$\hat{g}(x)=\frac{\sum_{i=1}^n w\left(\frac{\|x-x_i\|}{R_x}\right)y_i}{\sum_{i=1}^n w\left(\frac{\|x-x_i\|}{R_x}\right)} \qquad \cdots(19)$$

والذي يمثل المعدل الموزون للجوارات الاقرب -k وأن $R_x$ تمثل المسافة الاقليدية [10].

نفترض بأن لدينا عينة من المشاهدات $\{X_1،X_2،\dots،X_n\}$ [6]وأن $X\in\mathbb{R}^q$ وعليه فأنه لأي نقطة ثابتة ولتكن

$x_0 \in \mathbb{R}^q$ نستطيع ان نقوم بحساب مدى قرب كل مشاهدة لـ $x_0$ وذلك بأستخدام المسافة الاقليدية $R_x$ (Euclidean Distance) [14] .
إذ أن المسافة الاقليدية يمكن تعريفها كما يأتي[1]:

$$R_x = \sqrt{\sum (x - \mathrm{x}_i)^2} \;=\left((x-\mathrm{x_i})^T(x-\mathrm{x}_i)\right)^{1/2} \qquad \cdots (20)$$

أي المسافة ما بين $x$ و(k) من الجوارات الاقرب من بين قيم $(\mathrm{x}_i)$ ، وعليه فالإحصاءات المرتبة( Order Statistics) للمسافات $R_i$ تكون :

$0 \le R_{(1)} \le R_{(2)} \le \cdots \le R_{(k)}$

أن المشاهدات المقابلة لتلك المسافات تمثل الجوارات الأقرب لـ $(x)$، إذ أن الجوار الأقرب الأول يمثل المشاهدة الأقرب لـ $(x_0)$ والجوار الأقرب الثاني يمثل المشاهدة الأقرب الثانية وهكذا .
نستطيع معاملة $R_x$ كعرض حزمة (Bandwidth) واستخدام دالة النواة (Kernel) المنتظمة .
أما ما يخص بحثنا ولأنموذج الانحدار اللامعلمي الضبابي المبين في المعادلة (1) فأن التقـــــدير باستخدام تمهيد(K-N-N) لدالة الانحدار الضبابي عند $x_0 \in D$ يمكن التعبير عنه بالصيغة في أدناه :

$$\hat{g}(x_0) = \left(\hat{l}(x_0), \hat{m}(x_0), \hat{u}(x_0)\right)_{LR}$$

$$\hat{g}(x_0) = \left(\sum_{i=1}^{n} W_i(x_0) l_{y_i}, \sum_{i=1}^{n} W_i(x_0) m_{y_i}, \sum_{i=1}^{n} W_i(x_0) u_{y_i}\right)_{LR} \qquad \cdots (21)$$

أذ أن $w_i(x_0)$ تمثل سلسلة الوزن عند $(x_0)$ وأن $(l_{yi})$ و $(m_{yi})$ و$(u_{yi})$ تمثل الحد الأدنى والمتوسط والحد الاعلى على التوالي للمتغير المعتمد الضبابي المشاهد من النوع (LR) [3] .
أن مُقدر الجوار الأقرب- K يتأثر بصورة كبيرة بمعلمة التمهيد (k) والتي تكون ثابتة بغض النظر عن$(x)$ وهذا هو الفرق عن مُقدر النواة (kernel) التقليدي والذي فيه العدد المؤثر للمشاهدات يتغير مع $(x)$ ، وتعمل معلمة (k) على ضبط درجة التمهيد للمنحني المُقدر، إذ أن دور معلمة التمهيد (k) يكون مشابهاً لدور معلمة عرض الحزمة(h) لممهدات النواة (kernel) ، فأذا كانت قيمة (k) صغيرة فأنها ستقود الى تباين كبير وتحيز(Bias) صغير في التقديرات وعلى العكس من ذلك أي في حالة كون (k) كبيرة فأنها تؤدي الى تباين صغير وتحيز كبير وهذا بدوره سيجعل تأثيرات الحدود(Boundary Effects) لتكون مشكلة مؤثرة بصورة كبيرة جداً في التقديرات [15] .
وبافتراض مجموعة من المؤشرات والتي تُعرّف من خلالها سلسلة الوزن $w_i(x_0)$ عند$(x_0)$ وهي:
$\{x_0$ من المشاهدات الاقرب الى K تمثل مشاهدة من $x_i$ : $\} = \mathrm{J}(x_0)$
وعليه يمكن تعريف سلسلة الأوزان عند$(x_0)$ في طريقة تمهيد (K-NN) كما يأتي:[3]

$$w_i(x_0) = w_i(x_0;k) = \begin{cases} \dfrac{1}{k} \;, & if\; i \in J(x_0)\,,\;\; i = 1,2,\ldots,n \\ 0 \;, & otherwise \end{cases} \qquad \cdots (22)$$

سنقوم بتحديد قيمة المعلمة التمهيدية (k) بالأعتماد على المسافة الأقليدية $R_x$ .
وعليه فأن تقدير الدالة $g(x)$ عند$(x_0)$ في حالة العدد الضبابي من النوع (LR) يكون كما يأتي [15] :

$$\hat{g}(x_0) = \left(\frac{1}{k}\sum_{i\in J(x_0)} l_{y_i}, \frac{1}{k}\sum_{i\in J(x_0)} m_{y_i}, \frac{1}{k}\sum_{i\in J(x_0)} u_{y_i}\right)_{LR} \qquad \cdots (23)$$

## 6.2 اختيار دالة النواة (Kernel) ومعلمة التمهيد

أن أهمية الأوزان $\mathrm{K_h}(|x_i - x_o|)$ تكمن في أنها تجعل نقاط البيانات التي تكون قريبة من النقطة الافتراضية $(x_0)$ ذات مساهمة أكبر في تقدير الدالة $\hat{g}(x_0)$ من تلك النقاط التي تكون بعيدة عنها وذلك بأعطائها وزناً أكبر[15].
وهذه الدوال تكون دوال وزن كدالة كثافة احتمالية متماثلة حول الصفر ويستخدم (Kernel) لتخصيص الأوزان للمشاهدات ومن خصائص دوال (kernel) المهمة ما يأتي[11]:

1. $\int k(u)d\,u = 1$
2. $0 \le k(u) \le \infty$
3. $\mathrm{k}(-u) = k(u)$ دالة تماثلية

وكما في حالة الانحدار اللامعلمي الاحصائي لدينا في الانحدار اللامعلمي الضبابي عدة انواع من دوال (kernel) [6] .
ولقد تم اقتراح صيغتين لدالة النواة (kernel) حيث تم التوصل اليها من صيغة (Gaussian) حيث أنها تمثل دوال كثافة احتمالية تحقق خصائص دوال النواة (kernel) أعلاه وهما كما مبينة في الجدول (1).

جدول (1)
دالتي (kernel) المقترحة في الانحدار الضبابي اللامعلمي

| Kernel Function | equation | Interval |
|---|---|---|
| Sug.1 | $\mathrm{K}(\boldsymbol{u}) = \sqrt{\frac{4}{\pi}}\exp(-4\boldsymbol{u}^2)$ | $I\{\lvert \boldsymbol{u}\rvert \le \infty\}$ |
| Sug.2 | $\mathrm{K}(\boldsymbol{u}) = \sqrt{\frac{6}{\pi}}\exp(-6\boldsymbol{u}^2)$ | $I\{\lvert \boldsymbol{u}\rvert \le \infty\}$ |

أن جميع الدوال المعتمدة والمقترحة تستند على خصائص دوال النواة (Kernel) .
وسيتم في هذا البحث الاعتماد على دالتي(Kernel)المعتمدة وهما (Quartic) و(Gaussian) فضلاً عن الدالتين المقترحتين .
أن معلـمة التمهيد في صيغة الوزن $K_{\mathrm{h}}(\cdot)$ تستخدم لتعديل درجة التمهيد لتقديـرات $\hat{l}(x), \widehat{m}(x), \hat{u}(x)$
لذلك فأن الاختيار المناسب والصحيح لقيمة معلمة التمهيد هو النقطة الرئيسة والهامة في أساليب التمهيد الموضعي، حيث نلاحظ انه إذا كانت قيمة (h) كبيرة جداً، أي من المحتمل أن تذهب إلى ما لانهاية فستكون جميع المشاهدات تمتلك نفس الوزن وهذا ما يدعى بعدم التمهيد(فقدان المطابقة)، في حين أن قيمة (h) الصغيرة تجعل تقديرات$\hat{l}(x)$،$\widehat{m}(x)$،$\hat{u}(x)$ متذبذبة جداً وتقود الى فقدان المطابقة (Over Fitting) .
وعلاوة على ذلك أذا ازدادت قيمة $(h)$ فأن التحيز(Bias) للمُقدر سيزداد وبالمثل مع عملية خفض أو تقليل قيمة$(h)$ فأن الفرصة لتقليل التباين المُقدر سيتم فقدانها[15].
وعليه فأن عملية اختيار قيمة معلمة التمهيد $(h)$ يجب أن تخضع للمفاضلة (Trade off) مابين تحيز صغير وتباين صغير وهو ما يتم تحقيقه من خلال اختيار القيمة المثلى لمعلمة التمهيد. توجد عدة أساليب لاختيار القيمة المثلى لمعلمة التمهيد ومنها طريقة(Bootstrap) وطريقة العبور الشرعي المضبب (Fuzzified Cross Validation) وطريقة العبور الشرعي العام (G.C.V) وطريقة التمهيد الأمثل الطبيعي (Normal Optimal Smoothing).
سنقوم في هذا البحث بتناول طريقة التمهيد الأمثل الطبيعي بعد تضبيبها لاختيار القيمة المثلى لمعلمة التمهيد .

### 1.6.2 طريقة التمهيد الأمثل الطبيعي[2] *Normal Optimal Smoothing*

أن الصيغة المثلى لمعلمة التمهيد (*h*) تكون بحسب الصيغة الأتية:

$$\mathbf{h} = \left(\frac{4}{3n}\right)^{1/5} \sigma \qquad \dots(24)$$

إذ أن$(\sigma)$ تمثل الانحراف المعياري للمتغير المعتمد .
وبما أنه من المحتمل أن تحتوي البيانات المشاهدة على القيم الشاذة (Outliers) أو ان التوزيعات تكون طويلة الذيل (Long tailed) لذلك فمن الأفضل اللجوء الى التقدير الحصين لـلانحراف المعياري $(\sigma)$ للعينة ، وعليه سنعتمد في هذا البحث على مُقدر وسيط الانحراف المطلق (MAD) (Median Absolute Deviation) وكما يأتي :

$$\tilde{\sigma} = median\{\lvert y_i - \tilde{\mu}\rvert\}/0.6745 \qquad \dots(25)$$

إذ أن $(\tilde{\mu})$ تمثل الوسيط للعينة .
ويتم حساب قيمة (h) بحسب الصيغة الاتية :

$$h = \frac{\mathrm{median}\{\lvert y - median(y)\rvert\}}{0.6745}\left(\frac{4}{3n}\right)^{1/5} \qquad \dots(26)$$

وبما أن المتغير المعتمد ضبابياً فقد قام الباحثان بتضبيب الصيغة (25) لتتلائم مع طبيعة البحث ، أذ تم أعادة كتابتها ولقيم (الحد الأدنى والمتوسط والحد الأعلى) وعلى التتابع وكما يأتي :

$$\tilde{\sigma}_{L_{yi}} = median\left\{\left|L_{y_i} - \tilde{\mu}_{L_{yi}}\right|\right\} / 0.6745$$
$$\tilde{\sigma}_{m_{yi}} = median\left\{\left|m_{y_i} - \tilde{\mu}_{L_{yi}}\right|\right\}/0.6745 \qquad \cdots (27)$$
$$\tilde{\sigma}_{u_{yi}} = median\left\{\left|u_{y_i} - \tilde{\mu}_{L_{yi}}\right|\right\} / 0.6745$$

ويتم حساب قيمة ($h$) المثلى(للحد الادنى والمتوسط والحد الاعلى) بحسب الصيغة (26) وعلى التتابع وكما يأتي:

$$h_l = \tilde{\sigma}_{L_{yi}}\left(\frac{4}{3n}\right)^{1/5}$$
$$h_m = \tilde{\sigma}_{m_{yi}}\left(\frac{4}{3n}\right)^{1/5} \qquad \cdots (28)$$
$$h_u = \tilde{\sigma}_{u_{yi}}\left(\frac{4}{3n}\right)^{1/5}$$

## 7.2 معيار قياس أداء طرائق أساليب التمهيد اللامعلمي

أن الممهد يعد ممهداً جيداً وكفوءاً أذا أنتج خطأ تنبؤياً صغيراً وعادة ما يُقاس ذلك عن طريق معيار مربع متوسط الخطأ (M.S.E) ، أما في حالة الانحدار اللامعلمي الضبابي فلا نستطيع استخدام معيار(M.S.E) لتقييم دقة طرائق التمهيد وذلك لكون المتغير المعتمد ضبابياً والذي يتطلب استخدام قياس المسافة ، لذلك سنقوم بتعريف كمية لقياس قيمة التحيز مابين دالة الانحدار وتقديرها وكما يأتي[5] :

$$\text{Bias}(h) = \frac{1}{n}\sum_{i=1}^{n} d^2\left(g(x), \hat{g}(x;h)\right)$$
$$\text{Bias}(h) = \frac{1}{n}\sum_{i=1}^{n}\left(\left(l(x) - \hat{l}(x;h)\right)^2 + \left(m(x) - \hat{m}(x;h)\right)^2 + \left(u(x) - \hat{u}(x;h)\right)^2\right) \qquad \cdots (29)$$

إذ أن (n) تمثل عدد المشاهدات.

ولا نستطيع حساب هذه الكمية في التطبيقات العملية كون الدالة غير معلومة ولكن يتم أستخدامها في أسلوب المحاكاة لفحص أداء طرائق التمهيد اللامعلمية المستخدمة.[15]

# 3. الجانب التجريبي

أن استخدام الأسلوب التجريبي والمعروف بأسلوب المحاكاة (Simulation) يهدف الى إمكانية المقارنة مابين طرائق التقدير اللامعلمية والمستندة على أساليب التمهيد والتي تم دراستها في الجانب النظري من البحث، حيث أن أسلوب المحاكاة هو عملية افتراض للواقع الحقيقي والذي يمكننا من استخدام نماذج عدة وحجوم عينات مختلفة وبعدد من التكرارات وذلك للحصول عل نتائج تجريبية تفيد في الوصول الى اختبار للفرضيات التي تم بنائها. من اجل الوصول الى النتائج تم كتابة برنامج باستخدام (Matlab) الاصدار(2014a) للمقارنة ما بين الطرائق بالاعتماد على معيار المقارنة والمتمثل بمعيار التحيز(Bias) إذ أن القيمة الصغيرة لهذا المعيار تُدلل على جودة التقدير ودقته بالنسبة للقيم الحقيقية عند أسلوب التقدير المستخدم .

## 1.3 وصف تجربة المحاكاة

لقد تم الاعتماد على أربعة حجوم للعينات وهي (25,50,100,150) في تنفيذ تجارب المحاكاة لغرض توليد البيانات الخاصة بالمتغيرات العشوائية المستخدمة في البحث، وتم تكرار كل تجربة (1000) مرة وذلك لضمان نتائج أكثر أتساقاً (Consistent) .

تم بناء انموذج الانحدار اللامعلمي بالأعتماد على الخوارزمية الاتية:

**i. يتم توليد المتغير $x$ بصورة عشوائية من ضمن الفترة [-.25,0.25].**

ii.تم استخدام ثلاثة دوال لتوليد المتغيرات العشوائية في أنموذج الانحدار اللامعلمي الضبابي أحادي المتغير ومن خلال الدوال الرياضية الآتية :

1. $g_1(x_i) = \left(\frac{1}{5}\right)x_i^2 + 2exp\left(\frac{x_i}{10}\right)$
2. $g_2(x_i) = sin\left(\frac{x_i}{10}\right) + \left(\frac{x_i}{50}\right)^2$

3. $g_3(x_i) = 0.4x_i + 1$

iii. تم حساب المتغير المعتمد الضبابي المثلثي $(y)$ من النوع $(LR)$كما يأتي :

$y_i = g_j(x_i) + rand[-0.5, 0.5] \quad , j = 1, 2, 3 \quad i = 1, 2, \dots, n$

حيث أن قيمة $(y_i)$ تمثل متوسط العدد الضبابي $m_{yi}$ ، أما $u_{yi}$ و $l_{yi}$ فتمثل يسار العدد الضبابي(الحد الأدنى) ويمين العدد الضبابي (الحد الأعلى) على التتابع ويتم أيجادهما بعد حساب قيمة الانحراف المعياري $(\sigma_i)$ وبحسب الصيغة التالية :

$$\sigma_i = \frac{1}{4} g_j(x_i) + rand[-0.25, 0.25]$$

وعليه سنحصل على قيمة العدد الضبابي من جهة اليسار (الحد الأدنى) كما يأتي:

$$l_{yi} = y_i - \sigma_i$$

أما العدد الضبابي من جهة اليمين (الحد الأعلى) فيكون :

$$u_{yi} = y_i + \sigma_i$$

وسيتم في هذا البحث الاعتماد على دالتي(Kernel) المعتمدة وهما (Quartic) و(Gaussian) إضافة إلى الدالتين المقترحتين .

## 2.3 تنفيذ تجارب المحاكاة وتحليل نتائجها

من أجل تنفيذ تجارب المحاكاة قيد البحث فقد تم كتابة برنامج المحاكاة بلغة (Matlab) الإصدار (R2014a)، وستتم المقارنة بين أساليب التمهيد باستخدام معيار التحيز(Bias) والذي يقيس مقدار التحيز مابين دالة الانحدار الضبابي وتقديرها .

جدول (2)

قيم معيار(Bias) لطرائق التقدير المستخدمة بالنسبة للدالة الرياضية الأولى

| Bias | | Kernel Functions | | | |
|---|---|---|---|---|---|
| n | Method | Quartic | Gaussian | sug.1 | sug.2 |
| 25 | K-N-N | 0.10193 | 0.10047 | 0.10415 | 0.10694 |
| | L-L-S | 0.52483 | 0.17932 | 0.22002 | 0.24092 |
| | K-S | 0.16964 | 0.15360 | 0.20239 | 0.21442 |
| 50 | K-N-N | 0.09684 | 0.09629 | 0.09764 | 0.09861 |
| | L-L-S | 0.33784 | 0.16442 | 0.18475 | 0.19469 |
| | K-S | 0.17263 | 0.14590 | 0.17809 | 0.18600 |
| 100 | K-N-N | 0.09536 | 0.09516 | 0.09564 | 0.09597 |
| | L-L-S | 0.25672 | 0.15654 | 0.16882 | 0.173767 |
| | K-S | 0.17004 | 0.14336 | 0.16385 | 0.168407 |
| 150 | K-N-N | 0.08974 | 0.08962 | 0.08991 | 0.090107 |
| | L-L-S | 0.21203 | 0.14819 | 0.1558 | 0.15858 |
| | K-S | 0.14795 | 0.13595 | 0.15346 | 0.15663 |

جدول (3)

قيم معيار(Bias) لطرائق التقدير المستخدمة بالنسبة للدالة الرياضية الثانية

| Bias | | Kernel Functions | | | |
|---|---|---|---|---|---|
| n | Method | Quartic | Gaussian | sug.1 | sug.2 |
| 25 | K-N-N | 0.05128 | 0.0511 | 0.05165 | 0.05219 |
| | L-L-S | 0.33307 | 0.07089 | 0.09441 | 0.1041 |
| | K-S | 0.08382 | 0.06083 | 0.08942 | 0.09817 |
| 50 | K-N-N | 0.04777 | 0.04775 | 0.04783 | 0.04791 |
| | L-L-S | 0.27059 | 0.06292 | 0.07978 | 0.0857 |
| | K-S | 0.06848 | 0.05538 | 0.07473 | 0.08126 |
| 100 | K-N-N | 0.04654 | 0.04654 | 0.04655 | 0.04656 |
| | L-L-S | 0.16023 | 0.05439 | 0.06438 | 0.06802 |
| | K-S | 0.05619 | 0.05181 | 0.06099 | 0.06415 |
| 150 | K-N-N | 0.04579 | 0.04579 | 0.0458 | 0.0458 |
| | L-L-S | 0.12889 | 0.05264 | 0.0599 | 0.0623 |
| | K-S | 0.05424 | 0.0504 | 0.05723 | 0.05954 |

جدول (4)

قيم معيار(Bias) لطرائق التقدير المستخدمة بالنسبة للدالة الرياضية الثالثة

| Bias | | Kernel Functions | | | |
|---|---|---|---|---|---|
| n | Method | Quartic | Gaussian | sug.1 | sug.2 |
| 25 | K-N-N | 0.10093 | 0.09976 | 0.10257 | 0.10464 |
| | L-L-S | 0.45117 | 0.17083 | 0.19499 | 0.20496 |
| | K-S | 0.17877 | 0.14359 | 0.18786 | 0.19821 |
| 50 | K-N-N | 0.09703 | 0.09642 | 0.09786 | 0.09881 |
| | L-L-S | 0.37249 | 0.16420 | 0.1805 | 0.18700 |
| | K-S | 0.16643 | 0.14447 | 0.17385 | 0.18066 |
| 100 | K-N-N | 0.09060 | 0.09058 | 0.09065 | 0.09074 |
| | L-L-S | 0.23670 | 0.13674 | 0.14536 | 0.14872 |
| | K-S | 0.16381 | 0.12397 | 0.14252 | 0.14610 |
| 150 | K-N-N | 0.08791 | 0.08791 | 0.08793 | 0.08797 |
| | L-L-S | 0.20133 | 0.13312 | 0.14005 | 0.14264 |
| | K-S | 0.1581 | 0.12118 | 0.13701 | 0.13991 |

## 4. الاستنتاجات

نستنتج من نتائج تجربة المحاكاة والمبينة في الجداول اعلاه ما يأتي:

1. تبين للباحثين أنه كلما زاد حجم العينة أدى ذلك الى تناقص في قيمة معيار التحيز(Bias) وعند جميع الدوال الرياضية المستخدمة في المحاكاة ولجميع دوال النواة (Kernel) وذلك دليل على تحسن في نتائج تقدير أنموذج الانحدار اللامعلمي الضبابي .
2. من خلال تنفيذ تجارب المحاكاة بينت النتائج تفوق طريقة الجوار الأقرب-K(k-Nearest Neighbor) وذلك لامتلاكها اقل قيمة (Bias) ولجميع حجوم العينات والدوال الرياضية المستخدمة في الدراسة وعند جميع دوال النواة (Kernel) المعتمدة والمقترحةوالقيمة الأقل كانت عند دالة (Gaussian) .
3. أن أكبر قيمة لـ (Bias) هي لطريقة التمهيد الخطي الموضعي (L-L-S) ولجميع حجوم العينات والدوال الرياضية وعند جميع دوال النواة والقيمة الأكبر كانت عند دالة (Quartic).
4. تساوي قيمة الـ (Bias) عند حجم العينة (100) في حالة استخدام الدالة الرياضية الثانية وعند دالتي النواة (Quartic , Gaussian) في طريقة التقدير (K-NN) حيث بلغت (0.04654) .
5. تساوي قيمة الـ (Bias) عند حجم العينة (150) في حالة استخدام الدالة الرياضية الثانية حيث بلغت(0.04579) عند دالتي النواة المعتمدة (Quartic, Gaussian) وبلغت(0.04580) عند دالتي النواة المقترحة (sug.1) و(sug.2) في طريقة التقدير (K-NN).
6. عند حجم العينة (150) وفي حالة استخدام الدالة الرياضية الثالثة تساوت قيم الـ(Bias) عند دالتي النواة المعتمدة (Quartic , Gaussian) في طريقة التقدير(K-NN) حيث بلغت (0.08791).

7. أن تأثير اختيار نوع دالة النواة (Kernel) في أساليب التمهيد المستخدمة في التقدير يكون قليلاً حيث لاحظنا بأن قيم الـ (Bias) تكون متساوية أو قريبة من بعضها وعند دوال النواة المختلفة (المعتمدة والمقترحة) .

## 5. المصادر


1. **Bora,D.J.,(2014),"*Effect of different distance measures on the performance of k-Means Algorithm : An experimental study in matlab*" International journal of computer science and information technologies ,vol.5(2) ,2501-2506.**
2. **Bowman A.W.&Azzalini A., (1997),"*Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*". Clarendon Press , Oxford.**
3. **Cheng C.-B.& Lee E.S., (1999)*," Nonparametric fuzzy regressionk-NN and kernel smoothing techniques"*Computers and Mathematics with Applications,38, 239–251.**
4. **FanJ.&GijbelsI.(1996), *"Local Polynomial Modelling and Its Applications"*Chapman & Hall, London .**
5. **Farnoosh R., Ghasemian J.& Fardo S., (2012),*"A modification on ridge estimation for fuzzy nonparametric regression"*Iraninan Journal of Fuzzy systems 9, pp75-88.**
6. **Gyorfi,L., Kohler , M. ,Krzyzak, A.,& Walk, H.,(2002),"*A distribution –Free Theory of Nonparametric Regression*" Spring – verlage ,New York ,Inc**
7. **Härdle W.(1994),*"Applied Nonparametric Regression"* Humboldt University.**
8. **Hastie T. J.& Tibshirani R.J.,(1990),*" Generalized Additive Models"* .Chapman& Hall, London .**
9. **Kim B. & Bishu R.R.,(1998),*"Evaluation of fuzzy linear regression models by comparing membership functions"* Fuzzy Sets and Systems,100(1-3), pp343–352.**
10. **MackY.p.(1981),"*Local properties of K-NN regression estimates*" society for industrial and applied mathematics, 2(3), pp311–323.**
11. **Paul H.K. &Vidakovic B.(2007),"Nonparametric statistics with applications to science and engineering ",Wiley.**
12. **Peters G .,(1994),*"Fuzzy linear regression with fuzzy intervals"* Fuzzy Sets and Systems, 63(1),pp45-55**
13. **Pourahmad, S., Ayatollahi ,S. M. T Taheri, S. M. , & Agahi. Z. H. ,(2011),*"Fuzzy logistic regression based on the least squares approach with application in clinical studies".***
14. **Ullah,A.,& Viond,H.D,(1993),"*General Nonparametric Regression Estimation and testing in Econometrics*"Handbook of statistics , Vol.11.**
15. **Wang, N., Zhang, W.X. & Mei, Ch.L.(2007),"*Fuzzy nonparametric regression based on local linear smoothing technique*" Information Sciences,177, 3882 - 3900.**


……………………………………………………………………………..
……………………………………………….
……………………….