



AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University



مقارنة بعض طرق تقدير النموذج الانحدار اللوجستي شبه المعلمي الضبابي في حالة وجود مشكلة التعدد الخطي مع تطبيق عملي

م. د. اياد حبيب شمال /كلية الإدارة والاقتصاد/قسم الإحصاء

ayadstatistic@uodiyala.edu.iq

المستخلص:

تناول هذا البحث نموذج الانحدار اللوجستي شبه المعلمي الضبابي في حالة وجود مشكلة التعدد الخطي عندما يكون نموذج الانحدار الخطي غير مناسب في تحليل البيانات، في هذا النموذج متغير الاستجابة والمعلمات ضبابية أي تمثل الرقم الضبابي المثلثي، ويتم تقدير هذا النموذج على مرحلتين في المرحلة الأولى يتم تقدير الجزء المعلمي للنموذج من خلال مقدرات المربعات الصغرى الاعتيادية الضبابية ومقدرات ليو للوجستي الضبابية ، بينما في المرحلة الثانية يتم تقدير الجزء اللامعلمي من خلال مقدرات نداريه واتسون عندما تكون دوال كيرنل تمثل دالة Gaussian والدالة المقترحة.

تم استعمال بيانات حقيقية ضبابية تمثل الفشل الكلوي في تقدير النموذج، إذ كان متغير الاستجابة الضبابي يمثل نسبة الإصابة بالفشل الكلوي، والمتغيرات التوضيحية التي تعاني من مشكلة التعدد الخطي تمثل العوامل التي تؤثر على الإصابة بالفشل الكلوي، ومن خلال متوسط مربعات الخطأ كانت النتائج تدل على افضلية مقدرات ليو للوجستي الضبابية على مقدرات المربعات الصغرى الاعتيادية الضبابية في تقدير نموذج الانحدار اللوجستي شبه المعلمي الضبابي في حالة وجود مشكلة التعدد الخطي عند استعمال دالة كيرنل المقترحة.

الكلمات المفتاحية: نموذج الانحدار اللوجستي شبه المعلمي الضبابي ، مشكلة التعدد الخطي، الرقم الضبابي المثلثي، مقدرات المربعات الصغرى الاعتيادية الضبابية ، مقدرات ليو للوجستي الضبابية.

Comparison of some methods of Estimating Fuzzy Semi-Parametric Logistic Regression Model In Case of Multicollinearity Problem with Practical Application

Abstract:

This paper dealt the fuzzy semi-parametric logistic regression model in the case of a multicollinearity problem when the linear regression model is not suitable for data

analysis. In this model, the response variable and the parameters were fuzzy, i.e. they represent the fuzzy triangular number. The model is estimated in two stages. In the first stage, the parametric part of the model is estimated through fuzzy ordinary least squares estimators and fuzzy Liu's logistic estimators, while in the second stage the non-parametric part is estimated through Nadaraya Watson's estimators when the kernel functions represent the Gaussian function and the proposed function.

Fuzzy real data representing renal failure were used to estimate the model, where the fuzzy response variable represented the incidence of renal failure; the explanatory variables that suffer from the problem of multicollinearity represent the factors that affect the incidence of renal failure. Through the mean squared error MSE, the results indicated the preference of FLLE best FOLSE in estimating the fuzzy semi-parametric logistic regression model in the case of a multicollinearity problem when use the proposed kernel function.

Key words: the fuzzy semi-parametric logistic regression model, multicollinearity problem, the triangular fuzzy numbers, FOLSE, FLLE

1. المقدمة: Introduction

يعرف علم الإحصاء بأنه من أهم العلوم التطبيقية في الظواهر الاجتماعية ، والطبية، والاقتصادية كما وان أحد فروع علم الإحصاء هو تحليل الانحدار الذي يدرس تأثير المتغيرات التوضيحية على متغير الاستجابة.

ان تحليل الانحدار الخطي قد لا يكون قادراً على إعطاء نموذج مناسب لجميع البيانات و لا سيما في حالة عدم معرفة المشاكل التي تعاني منها البيانات مثل الضبابية وهي جزء من عدم التأكد، فهناك نماذج يكون فيها عدم التأكد بسبب العشوائية إذ يتم استعمال النظرية الاحتمالية في معالجة عدم التأكد الناتج من العشوائية، ونماذج يكون فيها عدم التأكد بسبب الضبابية وليس العشوائية، وهناك نماذج أخرى يكون فيها عدم التأكد بسبب العشوائية والضبابية. (P: .

(Shemil & Mohammed (2022)2

عند استعمال البيانات الضبابية وتمثيلها من خلال انموذج الانحدار في أنواعها جميعاً سواء كان انموذج الانحدار الخطي أم غير خطي او انموذج معلمي او شبه معلمي او لا معلمي يقسم إلى ثلاثة أنواع هي:

- 1- بيانات المتغيرات التوضيحية $(X_1, X_2, \dots, X_K)$ ومتغير الاستجابة Y ضبابية والمعلمات $(\beta_0, \beta_1 \dots \dots \beta_K)$ اعتيادية.
- 2- المعلمات $(\beta_0, \beta_1 \dots \dots \beta_K)$ ومتغير الاستجابة Y ضبابية والمتغيرات التوضيحية $(X_1, X_2, \dots, X_K)$ اعتيادية.
- 3- بيانات المتغيرات التوضيحية $(X_1, X_2, \dots, X_K)$ والمعلمات $(\beta_0, \beta_1 \dots \dots \beta_K)$ ومتغير الاستجابة ضبابية.

إنّ انموذج الانحدار الضبابي المعلمي تم دراسته من قبل الكثيرين من الباحثين ،أما انموذج الانحدار الضبابي اللامعلمي ويُعد الباحث (Cheng & Lee) اول من قام بدراسته عام 1999 من خلال اقتراح انموذج انحدار ضبابي لامعلمي يتكون من متغير تفسيري اعتيادي ومتغير توضيحي ضبابي وسمي بأنموذج الانحدار الضبابي اللامعلمي لعدم توفر الشروط الخاص بالانموذج المعلمي من توزيع الخطأ وغيرها من الشروط.

في عام 2006 درس الباحثون (Donoso & et al) انموذج الانحدار الضبابي مع البرمجة التربيعية بدل من البرمجة الخطية لمعالجة مشكلة التعدد من خلال اقتراح نموذجًا يعتمد على انحدار الحرف. وفي عام 2014 درس الباحثون (Namdari & et al) المقارنة بين انموذج الانحدار اللوجستي الضبابي والانحدار اللوجستي الترتيبي من خلال التطبيق على بيانات تمثل تأثير حمض الفوليك على الشهية لدى الأطفال وكانت النتائج تدل على افضلية انموذج الانحدار اللوجستي الضبابي على انموذج الانحدار اللوجستي الترتيبي. درس الباحثون (Hesamian and et al) عام 2017 انموذج الانحدار اللوجستي بين المدخلات والمخرجات الضبابية من خلال المجموعات الضبابية الحدسية بدلاً من المجموعات الضبابية الاعتيادية واتضح أن نموذج الانحدار اللوجستي للمجموعات الحدسية المقترح يؤدي نتائج أفضل.

إنّ أحد نماذج الانحدار هو النموذج اللوجستي شبه المعلمي الذي يتكون من جزئين : أحدهما جزء معلمي والآخر جزء لا معلمي يمثل دالة التمهيد إذ أصبح النموذج شبه المعلمي أكثر تطبيقاً واهتماماً في السنوات الأخيرة على مجموعة البيانات الضبابية.

2. مشكلة البحث: Problem of Research

هناك الكثير من المشاكل التي تعاني منها نماذج الانحدار والتي تكون السبب الرئيس في عدم امكانية تقدير معلمات الانموذج أو عدم كفاءة تقدير المعلمات، من المشاكل التي يعاني منها انموذج الانحدار اللوجستي شبه المعلمي الضبابي مشكلة التعدد الخطي، فضلاً عن ذلك كما هو معروف أنّ المتغير التابع في الانموذج اللوجستي يكون ثنائي الاستجابة يأخذ القيمتين صفر و واحد (صفر يمثل الفشل) و (واحد يمثل النجاح) او متعدد الاستجابة مثل انموذج الانحدار اللوجستي متعدد الاستجابة في بعض الاحيان يكون المتغير التابع يمثل بيانات ضبابية لغوية مثل (ضعيف، متوسط، جيد، جيد جداً ، ممتاز) لذا يتم استخدام انموذج الانحدار اللوجستي شبه المعلمي الضبابي.

ان وجود العلاقة القوية او الارتباطات بين المتغيرات التفسيرية في الجزء المعلمي من أنموذج الانحدار اللوجستي المتعدد شبه المعلمي الضبابي يؤدي إلى ظهور مشكلة التعدد الخطي الذي يجعل الانموذج غير مستقر ومعلمات التقدير غير دقيقة وبالتالي لا يمكن استعمال طرائق التقدير التقليدية لتقدير الانموذج.

3. هدف البحث: Object of Research

يهدف البحث إلى الاتي:

- 1- تقدير معلمات انموذج الانحدار اللوجستي شبه المعلمي الضبابي في حالة وجود مشكلة التعدد الخطي والمقارنة بين طرائق التقدير.

2- استعمال بيانات المتغير التوضيحي الضبابي التي تمثل نسبة الفشل الكلوي في تقدير معلمات نموذج الانحدار اللوجستي المتعدد شبه المعلمي الضبابي.

4. المجموعة الضبابية ودالة الانتماء: The Fuzzy Set and Membership Function

قبل تعريف المجموعة الضبابية Fuzzy Set يمكن تعريف المجموعة الاعتيادية crisp set على انها مجموعة من المشاهدات او العناصر و كل من تلك العناصر يكون منتما ام غير منتما إلى المجموعة، في حالة العنصر ينتمي إلى المجموعة تكون درجة الانتماء له تساوي واحد وفي حالة عدم انتماء العنصر إلى المجموعة تكون درجة الانتماء له تساوي صفر ولا توجد حالة وسط بينهما. (Mohammed & Rashid (2007))

المجموعة الضبابية Fuzzy Set وهي المجموعة التي لا يمكن تعريفها بدقة متناهية، أي ان تعريفها يختلف باختلاف وجهات النظر من قبل الباحثين هي المجموعة التي تمتلك عناصرها درجة الانتماء Membership Degree او دالة انتماء. (Zadeh (1965))

فأن الاختلاف بين المجموعة الضبابية والاعتيادية يكمن في كيفية إيجاد او حساب دالة الانتماء membership function إلى المجموعة الضبابية، تم تعريف المجموعة الضبابية من قبل الباحث (Zimmerman, H, j) هي إذا كانت X مجموعة من العناصر و كل عنصر يرمز له بالرمز x فإن المجموعة الضبابية A لمجموعة العناصر X تكون بشكل ازواج مرتبة كالاتي:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x) / x \in X)\} \quad (1)$$

إذ ان $\mu_{\tilde{A}}(x)$ تمثل دالة انتماء للعنصر x في المجموعة الضبابية $\tilde{A}$.

تُعرف دالة الانتماء membership function في المجموعة الضبابية بأنها الدالة التي يُتم من خلالها حساب درجة الانتماء أي عنصر للمجموعة الضبابية أي مقدار انتماء العنصر إلى المجموعة الضبابية fuzzy set وهذه الدرجة تكون محصورة بين الصفر والوحدة [0,1].

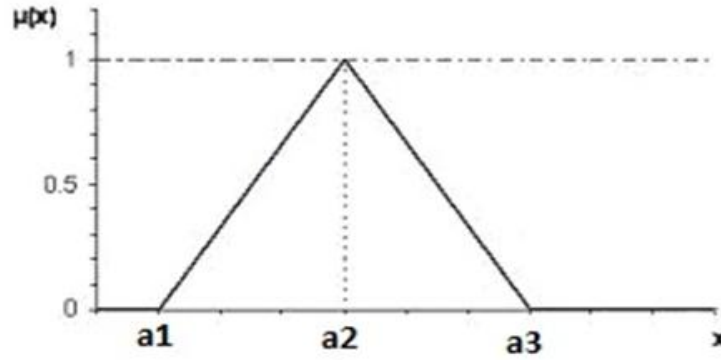
5. الرقم الضبابي المثلثي: Triangular Fuzzy Numbers

الرقم الضبابي $\bar{A}$ يدعى رقم ضبابي مثلثي في حالة وجود ارقام حقيقية $a_1, a_2, a_3 \geq 0$ ، فالرقم الضبابي المثلثي $\bar{A} = (a_1, a_2, a_3)$ يمتلك دالة انتماء تدعى دالة الانتماء المثلثية وفق المعادلة الآتية: (Abbas (2017)).

$$\mu_A(x_i) = \begin{cases} \frac{(x-a_1)}{(a_2-a_1)} & \text{if } a_1 \leq x \leq a_2 \\ \frac{(a_3-x)}{(a_3-a_2)} & \text{if } a_2 \leq x \leq a_3 \\ \text{zero} & \text{if } x < a_2, x > a_3 \end{cases} \quad (2)$$

إذ ان a_1, a_2, a_3 تنتمي إلى المجموعة الشاملة R ، ويمكن تعريف دالة الانتماء إلى المجموعة A وفق حدّ القطع α ($\alpha - cut$) إذا كانت A مجموعة ضبابية من R فإنّ ($\alpha - cut$) للمجموعة A يُرمز لها بالرمز A^α ، هي مجموعة ضبابية تمتلك كل العناصر في المجموعة الشاملة R التي درجة انتمائها أكبر أو تساوي القيمة α وتكون وفق المعدلة الآتية: (Abbas (2017)).

$$A_\alpha = \{x \in R: \mu_A(x) > \alpha\} \quad (3)$$



الشكل رقم 1: دالة انتماء الرقم الضبابي المثلثي

6. انموذج الانحدار اللوجستي شبه المعلمي الضبابي: Fuzzy Semi Parametric Logistic Regression

Model (FSPLRM)

يتكون أنموذج الانحدار اللوجستي شبه المعلمي الضبابي من تأثير متغيرات لا خطية معلمية وأخرى لا معلمية على متغير الاستجابة Y ، متغير الاستجابة Y والمعلمات $(\beta_0, \beta_1, \dots, \beta_k)$ ضبابية والمتغيرات التوضيحية $(X_1, X_2, \dots, X_K)$ ومتغير دالة التمهيد (t) اعتيادية فان دالة الاستجابة في انموذج الانحدار اللوجستي شبه المعلمي الضبابي FSPLRM تكون وفق الصيغة الآتية. (Hesamian, & Akbari. (2017))

$$\bar{p}_i = p_r(y_i = 1/X = \underline{x}, t) = \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + g(t_i)}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + g(t_i)}} \quad (4)$$

$i = 1, 2, 3, \dots, n$ تمثل عدد المشاهدات

$j = 1, 2, 3, \dots, k$ تمثل عدد المتغيرات

$\bar{p}_i$: دالة الانحدار اللوجستي الضبابية أي نها تمثل الرقم الضبابي المثلثي $(\bar{p}_i; l_{\bar{p}_i}, r_{\bar{p}_i})$

$\beta_j = (b_j; l_{b_j}, r_{b_j})$ المتجه المعلمات غير المعرفة تمثل المعلمات الضبابية المثلية $\underline{\beta} = (\beta_1, \beta_2, \dots, \beta_p)$

$g(t_i)$: تمثل دالة تمهيد اللامعلمية.

عند استعمال تحويل دالة اللوجت Logit Function للتخلص من انحناء دالة الانحدار اللوجستي شبه المعلمي الضبابي نتيجة لتأثير السلبي لهذه الانحناءات في حالة وجودها على خصائص مقدرات المعلمات . تمكن الباحث Berkson عام 1944 من إيجاد علاقة لوغاريتمية لتحويل دالة الاستجابة $\bar{p}_1$ إلى دالة خطية والتي تكون بالصيغة الآتية. (Gao & Lu (2018))

$$Z_i = \text{logit}(\bar{p}_1) = \ln \frac{\bar{p}_1}{1-\bar{p}_1} \quad (5)$$

ومن ثم يصبح انموذج الانحدار اللوجستي شبه المعلمي الضبابي FSPLRM بالصيغة الآتية:

$$Z_i = \ln \left(\frac{\bar{p}_1}{1-\bar{p}_1} \right) = \bigoplus_{j=1}^k (\underline{\beta} \otimes x_{ij}) \oplus g(t_i) \oplus \epsilon_i \quad (6)$$

ϵ_i : يمثل حد الخطأ $\otimes \oplus$ رموز الجمع الداخلي والضرب الداخلي بين المصفوفات.

عندما تمثل المخرجات الرقم الضبابي المثلي فإن انموذج الانحدار اللوجستي شبه المعلمي الضبابي بعد استخدام دالة اللوجت يكون بالصيغة الآتية: (Hesamian, & Akbari. (2017))

$$Z_i = \ln \frac{\bar{p}_1}{1-\bar{p}_1} = \left(\ln \left(\frac{\bar{p}_1}{1-\bar{p}_1} \right); \ln \frac{\bar{p}_1 + l_{\bar{p}_1}}{(1-\bar{p}_1)^2}, \ln \frac{\bar{p}_1 + r_{\bar{p}_1}}{(1-\bar{p}_1)^2} \right) \quad (7)$$

7. مشكلة التعدد الخطي: The Multicollinearity problem

مشكلة التعدد الخطي تحدث عندما تكون هناك علاقة خطية بين بعض أو كل المتغيرات التفسيرية. إن الارتباط بين هذه المتغيرات يُعرف بـمشكلة التعدد الخطي، أي تظهر مشكلة التعدد الخطي عند ما تكون قيمة أحد المتغيرات التفسيرية متساوية لجميع المشاهدات، أو عند اعتماد قيمة أحد المتغيرات التفسيرية على قيمة واحدة أو أكثر من المتغيرات التفسيرية في الأنموذج علماً بأن هذه المشكلة تواجه الباحث في حالة تجانس أو عدم تجانس التباين

إذا كانت هناك علاقة خطية بين اثنين أو أكثر من المتغيرات فإن ذلك سيؤدي انتهاك شرط الرتب لمصفوفة المعلومات ، لذا لا يمكن إيجاد معكوس مصفوفة المعلومات $(X'X)$ ومن ثم لا يمكن إيجاد تقدير المعلمات وتكون متحيزة بشكل كبير المعلمات.

يمكن تشخيص مشكلة التعدد الخطي في انموذج الانحدار اللوجستي شبه المعلمي الضبابي من خلال العدد الشرطي **condition number** ونسبة التباين **variance proportion**.

قيم العدد الشرطي تعتمد على مصفوفة الارتباطات $\emptyset$ في تشخيص مشكلة التعدد الخطي ولتي تكون وفق الصيغة الآتية (Shahmandi & et all(2013))

$$\emptyset = S'S \quad (8)$$

إذ أن S تمثل مصفوفة أبعادها (n*p) وأن كل عنصر في المصفوفة يتم إيجاده من خلال المعادلة الآتية:

$$S_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}} \quad (9)$$

لإيجاد قيم إعداد الشرطي يُتم حساب الجذور المميزة $\lambda_0, \lambda_1, \dots, \lambda_p$ إلى مصفوفة الارتباطات $\emptyset$ فإن إعداد الشرط تكون وفق الصيغة الآتية:

$$k_j = \left(\frac{\lambda_{\max}}{\lambda_j} \right)^{0.5} \quad (10)$$

إذ إن $\lambda_{\max}$ هو أكبر الجذور المميزة و λ_j الجذر المميز j لمصفوفة الارتباطات $\emptyset$. إن قيم إعداد الشرط التي تكون أكبر من 30 ($k_j > 30$) تشير إلى حالة التشخيص أي وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية

امام نسبة التباين variance proportion تستعمل للكشف عن مشكلة التعدد الخطي، لتكن M مصفوفة المتجهات المتعامدة يتم الحصول من مصفوفة الارتباطات $\emptyset$ وأن Λ مصفوفة قطرية من الجذور المميزة لمصفوفة الارتباطات $\emptyset$ التي تحقق الشرط الآتي.

$$M\emptyset M = \Lambda \quad (11)$$

نسبة التباين يمكن التعبير عنها من خلال المعادلة الآتية:

$$\omega_{uj} = \frac{m_{ju}^2 / \lambda_u}{c_{jj}} \quad (12)$$

إذ أن $j, u = 0, 1, 2, \dots, p$ المميز λ_u الجذر المميز u من مصفوفة المتجهات المميزة، m_{ju} العنصر في الصف j والعمود u من مصفوفة المتجهات المميزة، c_{jj} الجذور المميزة الصغيرة (القيم الذاتية الصغيرة نسبة إلى أكبر جذر مميز) ويحسب وفق المعادلة الآتية.

$$c_{jj} = \sum_{u=0}^p \lambda_u^{-1} m_{ju}^2 \quad (13)$$

ففي حالة وجود قيمتين من قيم نسبة التباين الموزونة كبيرتين وبالقرب من بعضهما تحدث مشكلة التعدد الخطي.

8. تقدير النموذج: Estimation Of The Model

يتم تقدير انموذج الانحدار اللوجستي شبه المعلمي الضبابي على مرحلتين المرحلة الأولى تقدير الجزء المعلمي من خلال طرائق تقدير المعلومات التي سوف يتم ذكرها في القسم رقم 9 والمرحلة الثانية تقدير الجزء اللامعلمي من خلال دالة التمهيد بالاعتماد على مقدرات نداريه واتسون وفق الصيغة الآتية: (Hesamian & Akbari. (2017))

$$g(t) = \sum_{i=1}^n w_i(t_j) \otimes \left(Z_i - \bigoplus_{j=1}^p (\underline{\beta} \otimes x_{ij}) \right) \quad (14)$$

$$g(t) = \sum_{i=1}^n w_i(t_j) \otimes Z_i - \sum_{i=1}^n w_i(t_j) \otimes (\bigoplus_{j=1}^p (\underline{\beta} \otimes x_{ij}))$$

عندما $w_i(t_j)$ الاوزان لتمهيد كيرنل يتم تقديرها وفق الصيغة الآتية: (Härdle & et al (2004))

$$w_i(t_j) = \frac{k\left(\frac{t-T_i}{h}\right)}{\sum_{i=1}^n k\left(\frac{t-T_i}{h}\right)} \quad (15)$$

عندما $k(\cdot)$ تمثل دالة كيرنل ، h تمثل معلمة تمهيد ، t يمثل المتغير المستقل اللامعلمي بالاعتماد على المعادلة رقم 6 يكون انموذج الانحدار اللوجستي الضبابي المقدر بالصيغة الآتية:

$$Z_i = \bigoplus_{j=1}^p (\underline{\beta} \otimes x_{ij}) \oplus (\sum_{i=1}^n w_i(t_j) \otimes Z_i) - \sum_{i=1}^n w_i(t_j) \otimes \left(\bigoplus_{j=1}^p (\underline{\beta} \otimes x_{ij}) \right) \quad (16)$$

$$Z_i = \bigoplus_{j=1}^p (\underline{\beta} \otimes x_{ij}) \oplus (\sum_{i=1}^n w_i(t_j) \otimes Z_i) - \left(\bigoplus_{j=1}^p (\underline{\beta} \otimes \sum_{i=1}^n w_i(t_j) x_{ij}) \right)$$

$$Z_i = (\sum_{i=1}^n w_i(t_j) \otimes Z_i) \oplus (\bigoplus_{j=1}^p \underline{\beta} \otimes (x_{ij} - \sum_{i=1}^n w_i(t_j) x_{ij}))$$

$$Z_i = (\sum_{i=1}^n w_i(t_j) \otimes Z_i) \oplus (\bigoplus_{j=1}^p \underline{\beta} \otimes x_{ij}^*)$$

$$x_{ij}^* = x_{ij} - \sum_{i=1}^n w_i(t_j) x_{ij} \quad \text{عندما}$$

يمكن إيجاد المخرجات الضبابية المقدرة في حالة متغير الاستجابة يمثل الرقم الضبابي المثلثي $\hat{Z}_i = (\hat{z}_i; l_{\hat{z}_i}, r_{\hat{z}_i})$ وفق الصيغ الآتية:

$$\begin{aligned} Z_i &= (\sum_{i=1}^n w_i(t_j) \otimes z_i) \oplus (\bigoplus_{j=1}^p b_j \otimes x_{ij}^*) \\ l_{Z_i} &= (\sum_{i=1}^n w_i(t_j) \otimes l_{z_i}) \oplus (\bigoplus_{j=1}^p l_{b_j} \otimes ss_{ji} x_{ij}^*) \ominus (\bigoplus_{j=1}^p r_{b_j} \otimes (1 - ss_{ji}) x_{ij}^*) \\ r_{Z_i} &= (\sum_{i=1}^n w_i(t_j) \otimes r_{z_i}) \oplus (\bigoplus_{j=1}^p r_{b_j} \otimes ss_{ji} x_{ij}^*) \ominus (\bigoplus_{j=1}^p l_{b_j} \otimes (1 - ss_{ji}) x_{ij}^*) \end{aligned} \quad (17)$$

$$ss_{ji} = \begin{cases} 1 & x_{ij}^* \geq 1 \\ 0 & x_{ij}^* < 0 \end{cases} \quad \text{عندما}$$

ومن ثمّ يمكن كتابة المخرجات الضبابية المقدرة التي تمثل الرقم الضبابي المثلثي $\hat{Z}_i = (\hat{z}_i; l_{\hat{z}_i}, r_{\hat{z}_i})$ بصيغة المصفوفات وفق الآتية:

$$\hat{\underline{Z}} = \left(W\underline{Z} + X^* \hat{\underline{\beta}}; W\underline{L}_Z + X_s^* \underline{L} - X_{1-s}^* \underline{R}, W\underline{R}_Z + X_s^* \underline{R} - X_{1-s}^* \underline{L}; \right) \quad (18)$$

عندما

$$W = \begin{bmatrix} w_{11} & w_{12} & . & . & . & w_{1n} \\ w_{21} & w_{22} & . & . & . & w_{2n} \\ : & : & : & : & : & : \\ . & . & . & . & . & . \\ w_{n1} & w_{n2} & . & . & . & w_{nn} \end{bmatrix} \quad X = \begin{bmatrix} x_{11} & x_{12} & . & . & . & x_{1p} \\ x_{21} & x_{22} & . & . & . & x_{2p} \\ : & : & : & : & : & : \\ . & . & . & . & . & . \\ x_{n1} & x_{n2} & . & . & . & x_{np} \end{bmatrix}$$

$$X^* = \begin{bmatrix} x_{11}^* & x_{12}^* & . & . & . & x_{1p}^* \\ x_{21}^* & x_{22}^* & . & . & . & x_{2p}^* \\ : & : & : & : & : & : \\ . & . & . & . & . & . \\ x_{n1}^* & x_{n2}^* & . & . & . & x_{np}^* \end{bmatrix} \quad X_s^* = \begin{bmatrix} s_{11}x_{11}^* & s_{12}x_{12}^* & . & . & . & s_{1p}x_{1p}^* \\ s_{21}x_{21}^* & s_{22}x_{22}^* & . & . & . & s_{2p}x_{2p}^* \\ : & : & : & : & : & : \\ . & . & . & . & . & . \\ s_{n1}x_{n1}^* & s_{n2}x_{n2}^* & . & . & . & s_{np}x_{np}^* \end{bmatrix}$$

$$X_{(1-s)}^* = \begin{bmatrix} (1-s_{11})x_{11}^* & (1-s_{12})x_{12}^* & . & . & . & (1-s_{1p})x_{1p}^* \\ (1-s_{21})x_{21}^* & s_{22}x_{22}^* & . & . & . & (1-s_{2p})x_{2p}^* \\ : & : & : & : & : & : \\ . & . & . & . & . & . \\ (1-s_{n1})x_{n1}^* & (1-s_{n2})x_{n2}^* & . & . & . & (1-s_{np})x_{np}^* \end{bmatrix}$$

$$\underline{\hat{\beta}} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_p \end{bmatrix}, \quad \underline{L} = \begin{bmatrix} l_{b_1} \\ l_{b_2} \\ . \\ . \\ l_{b_p} \end{bmatrix}, \quad \underline{R} = \begin{bmatrix} r_{b_1} \\ r_{b_2} \\ . \\ . \\ r_{b_p} \end{bmatrix}, \quad \underline{Z} = \begin{bmatrix} z_1 \\ z_2 \\ . \\ . \\ z_n \end{bmatrix}, \quad \underline{L}_Z = \begin{bmatrix} l_{z_1} \\ l_{z_2} \\ . \\ . \\ l_{z_n} \end{bmatrix}, \quad \underline{R}_Z = \begin{bmatrix} r_{z_1} \\ r_{z_2} \\ . \\ . \\ r_{z_n} \end{bmatrix}$$

يتم المقارنة بين مقدرات انموذج الانحدار اللوجستي شبة المعلمي الضبابي في حالة وجود مشكلة التعدد الخطي من خلال متوسط مربعات الخطأ MSE وفق المعادلة الآتية: (Hesamian & Akbari (2017))

$$MSE(Z) = \frac{1}{n} \sum_{i=1}^n d(\hat{Z}, Z) \quad (19)$$

عندما $d(\hat{Z}, Z)$ تمثل المسافة الاقليدية بين القيم الحقيقية Z والقيم التقديرية للأنموذج $\hat{Z}$ وتحسب وفق المعادلة الآتية:

$$d(\hat{Z}, Z) = \sqrt{\frac{1}{3} ((\hat{z}_i - z_i)^2 + (l_{\hat{z}_i} - l_{z_i})^2 + (r_{\hat{z}_i} - r_{z_i})^2)} \quad (20)$$

9. طرائق تقدير المعلمات: The Parameters estimation methods

1.9. مقدرات المربعات الصغرى الاعتيادية الضبابية Fuzzy Ordinary Least: Squares Estimators (FoLSE)

يعد الباحث spekman اول من ناقش الانحدار الجزئي شبه المعلمي عام 1988_ إذ إن الانموذج في المعادلة رقم 6 يمكن ان يكتب من خلال الصيغة الآتية: (Speckman. (1988))

$$\underline{Z} = \underline{X}\underline{\beta} + \underline{V}\underline{\gamma} + \underline{U} \quad (21)$$

إذ إن $\underline{Z}$ يمثل متجه متغير الاستجابة الضبابي و $\underline{X}\underline{\beta}$ الجزء المعلمي عندما $\underline{X}$ مصفوفة المتغيرات التوضيحية $(n \times p)$ و $\underline{\beta}$ متجه عمودي من المعلمات $(p \times 1)$ ، $\underline{V}\underline{\gamma}$ يمثل الجزء اللامعلمي إذ أن $\underline{V}$ عناصر دالة تمهيد كيرنل من الرتبة $(n \times p)$ و $\underline{\gamma}$ متجه المعلمات المضافة، $\underline{U}$ يمثل موجه حدود الخطأ، ولتقدير المعلمات وفق المربعات الصغرى كالآتي:

$$\underline{\dot{U}}\underline{U} = (\underline{Z} - \underline{X}\underline{\hat{\beta}} - \underline{V}\underline{\hat{\gamma}})' (\underline{Z} - \underline{X}\underline{\hat{\beta}} - \underline{V}\underline{\hat{\gamma}}) \quad (22)$$

$$\frac{\partial \underline{\dot{U}}\underline{U}}{\partial \underline{\beta}} = -\underline{\dot{X}} (\underline{Z} - \underline{X}\underline{\hat{\beta}} - \underline{V}\underline{\hat{\gamma}}) = 0$$

$$\frac{\partial \underline{\dot{U}}\underline{U}}{\partial \underline{\gamma}} = -\underline{\dot{V}} (\underline{Z} - \underline{X}\underline{\hat{\beta}} - \underline{V}\underline{\hat{\gamma}}) = 0$$

ومن ثم نحصل على المقدرات وفق المعادلة الآتية:

$$\underline{\hat{\beta}}_{FoLSE} = [\underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{X}]^{-1} \underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{Z} \quad (23)$$

$$\text{عندما } \underline{P}_W = \underline{V}(\underline{\dot{V}}\underline{V})^{-1} \underline{\dot{V}}$$

وبالتالي فإن مقدرات المربعات الصغرى الضبابية في حالة المخرجات تمثل الرقم الضبابي المثلثي $(z_i; l_{z_i}, r_{z_i})$ تكون وفق الصيغ الآتية:

$$\begin{aligned} \underline{\hat{\beta}}_{FoLSE} &= [\underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{X}]^{-1} \underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{Z} \\ \underline{\hat{L}}_{FoLSE} &= [\underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{X}]^{-1} \underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{L}_Z \\ \underline{\hat{R}}_{FoLSE} &= [\underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{X}]^{-1} \underline{\dot{X}}(\underline{I} - \underline{P}_W)\underline{R}_Z \end{aligned} \quad (24)$$

2.9. مقدر ليو اللوجستي الضبابية: Fuzzy Logistic Liu Estimator(FLLE)

اقترحت هذه الطريق من قبل الباحث (Liu 1993) تهدف إلى تقليل التحيز في مقدرات الإمكان الأعظم من خلال الجمع بين مقدرات اشتاين (Stien estimators) ومقدرات انحدار الحرف (ridge regression estimators) لذلك فو يمتلك خواص المقدرين. (Inan & Erdogan (2013), p:3) (Wu & Arashi (2018), p:3)

اقترح الباحث (Mansson et al 2012) مقدر ليو اللوجستي (LLE) لتصحيح مشكلة التعدد الخطي في الانحدار اللوجستي وبذلك تكون مقدرات ليو اللوجستي كالآتي: (Ebaid, & Abo-El-Hadid (2017))

$$\hat{\beta}_{FLLSE} = (\hat{X}\hat{W}X + I)^{-1}(\hat{X}\hat{W}X + \hat{D}_{Liu}I)\hat{\beta}_{FOLSE} \quad (25)$$

عندما $0 < \hat{D}_{Liu} < 1$ يمثل معامل الانكماش $\hat{W}$ و ان مصفوفة قطرية عناصرها تمثل $\hat{p}_i(1 - \hat{p}_i)$ وان I تمثل مصفوفة الوحدة هناك عدة طرائق لتقدير معامل الانكماش $\hat{D}_{Liu}$ هي:

$$\hat{D}_{Liu j} = \frac{\hat{p}_j^2 - 1}{\frac{1}{q_j} + \hat{p}_j^2} \quad (26)$$

$$\hat{p}_j = \sum_{i=0}^p \alpha_{jt} * \hat{\beta}_j = \underline{\alpha}_j * \underline{\beta} \quad (27)$$

عندما q_j الجذور المميزة، $\underline{\alpha}_j$ المتجهات المميزة (إلى مصفوفة المعلومات الموزونة القياسية $X^* \hat{W} X^*$ ، اذ يمكن حساب X^* كالآتي

$$X_p^* = \frac{(x_{ij} - \bar{x}_j)}{\{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2\}^{0.5}}$$

$$X^* = [1: X_p^*]$$

اذ ان $\underline{1}$ يمثل متجه من قيم 1، كذلك يمكن حساب معامل الانكماش من خلال المعادلة الآتية:

$$\hat{D}_{Liu j} = \max\left(0, \min\left(\frac{\hat{p}_j^2 - 1}{\frac{1}{q_j} + \hat{p}_j^2}\right)\right) \quad (28)$$

ومقدرات ليو اللوجستي الضبابية في حالة المخرجات تمثل الرقم الضبابي المثلثي $(z_i; l_{z_i}, r_{z_i})$ تكون وفق الصيغ الآتية:

$$\begin{aligned} \hat{\beta}_{Liu} &= (\hat{X}\hat{W}X + I)^{-1}(\hat{X}\hat{W}X + \hat{D}_{Liu}I)\hat{\beta}_{FOLSE} \\ \hat{\underline{L}}_{Liu} &= (\hat{X}\hat{L}\hat{W}X + I)^{-1}(\hat{X}\hat{L}\hat{W}X + \hat{\underline{L}}\hat{D}_{Liu}I)\hat{\underline{L}}_{FOLSE} \\ \hat{\underline{R}}_{Liu} &= (\hat{X}\hat{R}\hat{W}X + I)^{-1}(\hat{X}\hat{R}\hat{W}X + \hat{\underline{R}}\hat{D}_{Liu}I)\hat{\underline{R}}_{FOLSE} \end{aligned} \quad (29)$$

10. الجانب التطبيقي: The Applied Side

1.10. وصف البيانات: Description of Data

تم جمع البيانات الخاصة بهذا البحث ، والتي تمثل نسبة الإصابة بالفشل الكلوي من مستشفى الزهراء في مدينة الكوت، لعينة من الأشخاص المصابين بالفشل الكلوي قوامها 40 مريضاً، وبالاتماد على مجموعة من الأطباء إذ أكد الأطباء ان

نسبة الفشل الكلوي تكون بشكل ضبابي، وقد تم تقسيم نسبة الإصابة إلى (ضعيف، متوسط، عالية) والتي تمثل متغير الاستجابة المثلثي الضبابي.

ان من اهم العوامل التي تؤثر على الفشل الكلوي التي تمثل المتغيرات التفسيرية هي:

- 1- المتغير المستقل X_1 يمثل السكري (Diabetes) يشمل عدداً من الاضطرابات التي تتميز بوجود مشاكل في هرمون الأنسولين الذي ينتجه البنكرياس في حالة الوضع الطبيعي لمساعدة الجسم في استخدام السكر والدهون وتخزين بعضها ، وان النسبة الطبيعية للسكر في جسم الانسام من 90-120 ملليغرام/ديسيلتر.
 - 2- المتغير X_2 يمثل الهيموغلوبين الدم (Hb) نسبة الدم في الجسم يقاس بوحدة (غم/100مل من الدم) او (غم/ديسي لتر من الدم).
 - 3- المتغير X_3 يمثل عمر المصاب ويقاس بالسنوات .
 - 4- المتغير اللامعلمي $g(t)$ يمثل ضغط الدم العالي HBP للمصاب ويكون في الحالة الطبيعية 115 ميلتر زئبق.
- الجدول رقم 1 يمثل نسبة الإصابة بالفشل الكلوي والمتغيرات التفسيرية

N	نسبة الفشل الكلوي	X_1	X_2	X_3	$g(t)$
1	Low	123	9	60	130
2	High	178	7.6	67	110
3	Medium	140	10	70	130
4	High	167	9	100	110
5	High	155	6	89	125
6	Low	133	9.6	77	130
7	Medium	145	11.2	80	150
8	High	178	10.1	76	130
9	High	180	8.4	60	160
10	Medium	144	11.5	59	155
11	Medium	157	8.2	76	140
12	High	145	13	89	130

13	Medium	167	8.1	79	135
14	High	180	10	90	110
15	Medium	135	9.8	77	125
16	High	155	11.2	90	120
17	High	169	8.9	88	110
18	Medium	155	8.4	87	110
19	Low	123	12	78	135
20	Low	144	10.3	90	120
21	Medium	165	12.2	83	120
22	Low	163	9.6	79	170
23	Medium	134	8.7	92	150
24	Low	140	12	95	175
25	High	120	10.4	72	110
26	High	137	11.8	76	115
27	Low	155	9.2	89	100
28	Medium	150	11.9	95	140
29	Low	157	10.3	77	130
30	Medium	145	9.4	94	135
31	Low	122	12	81	150
32	Medium	120	8	87	140
33	Low	130	9.7	79	155
34	High	156	11	100	110

35	Low	123	7.8	77	160
36	High	175	8.5	86	180

تم تقسيم نسبة الإصابة بالفشل الكلوي إلى (ضعيف، متوسط، عالية) بالاعتماد على مجموعة من الأطباء وبالتالي فإن متغير الاستجابة الذي يمثل نسبة الإصابة بالفشل الكلوي يكون وفق الجدول الآتي:

الجدول رقم 2 يبين الرقم الضبابي المثلثي (متغير الاستجابة الضبابي)

المصطلحات الضبابية	القيم
Low	(0,0.25,0.45)
Medium	(0.35,0.55,0.75)
High	(0.65,0.85,0.95)

2.10. التحليل الإحصائي statistical analysis

يتم إجراء التحليل الإحصائي في هذا البحث بالاعتماد على البرنامج الإحصائي R وفق جداول التحليل الآتية:

الجدول رقم 3 يبين العدد الشرطي ونسبة التباين

Independent variable	Eigenvalue	Condition number	variance proportion			
X0	5.99381	1.00000	0.049	0.258	0.006	0.263
X1	0.09055	8.79410	0.029	0.531	0.073	0.214
X2	0.03431	15.1870	0.321	0.021	0.336	0.135
X3	0.00207	45.5758	0.028	0.112	0.214	0.073

من خلال الجدول رقم 3 نجد أن أعداد الشرط المبينة من المعادلة (10) تعكس نتائجها قيماً كبيره لمتغيرات الانموذج إذ كان اكبر تلك القيم الخاصة بالمتغير الأخير الذي يمثل المتغير التوضيحي الثالث إذ كانت قيمة عدد الشرط الموزون الخاص به (45.5758) وهي اكبر من القيمة 30 مما يدل على وجود مشكلة التعدد الخطي بين المتغيرات المستقلة ،

ولبيان تحديد أي المتغيرات التوضيحية يسبب مشكلة التعدد الخطي يتم من خلال حساب قيم مصفوفة نسبة التباين المبينة في المعادلة (12) اذ وجد ان قيم نسبة التباين للمتغير التوضيحي الأول والثالث كبيرتان وقريبتان من الواحد ويقعان في نفس الجذر المميز الذي يقابل اكبر عدد شرط موزون كما موضح في الجدول (3) ، ومن ذلك نستنتج بأن هناك ازدوج او وتعدد خطي بين المتغيرات المستقلة يسببه المتغير التوضيحي الأول والثالث بدرجة كبيرة جداً.

الجدول رقم 4 يبين تقدير المعلمات الضبابية من طرائق التقدير (FoLSE, FLLE)

methods	FOLSE			FLLE		
Parameter s	$\hat{L}$	$\hat{\beta}$	$\hat{R}$	$\hat{L}$	$\hat{\beta}$	$\hat{R}$
$\hat{\beta}_0$	2.73262 1	2.26607 8	2.18459 0	1.17140 4	1.13271 4	1.29096 1
$\hat{\beta}_1$	3.15511 7	3.13390 7	4.13992 8	2.03837 5	2.04102 3	2.14282 4
$\hat{\beta}_2$	0.32283 4	0.12860 7	0.32754 6	0.45317 9	0.22357 0	0.64469
$\hat{\beta}_3$	0.50839 2	0.47459 8	0.84216	0.61126 5	0.81160 1	0.98196 8

من خلال الجدول 4 الذي يمثل تقدير المعلمات الضبابية بطريقة مقدرات المربعات الصغرى الاعتيادية الضبابية ومقدرات ليو اللوجستي نجد ان المتغير التوضيحي X_1 الذي يمثل السكري له اثر كبير على الإصابة بالفشل الكلوي بعد ذلك يأتي المتغير التوضيحي X_3 الذي يمثل العمر.

الجدول رقم 5 يمثل متوسط مربعات الخطأ MSE للأنموذج المقدر

Nadaraya Watson	FOLSE	FLLE
Kernel function	MSE	MSE
Gaussian function	0.8594102	0.2178525
Suggest function	0.8632029	0.2189267

الجدول رقم 5 يمثل متوسط مربعات الخطأ لأنموذج الانحدار اللوجستي شبه المعلمي الضبابي المقدّر في حالة وجود مشكلة التعدد الخطي عندما يتم تقدير الجز المعلمي من خلال مقدرات المربعات الصغرى الاعتيادية الضبابية FOLSE ومقدرات ليو اللوجستي الضبابية FLLE التي تعالج مشكلة التعدد الخطي ، في حين الجز اللامعلمي يتم تقديره من خلال مقدرات نداريه واتسون عندما تكون دالة كيرنل هي **Gaussian function** والدالة المقترحة كالآتي:

$$k(u) = \frac{\sqrt{4}}{\sqrt{\pi}} \exp(-4u^2) \quad I(|u| < \infty) \quad (30)$$

وكانت النتائج تدل إنّ مقدرات ليو اللوجستي الضبابية أفضل من مقدرات المربعات الصغرى الاعتيادية الضبابية في تقدير الجز المعلمي بينما في تقدير الجز اللامعلمي كانت دالة كيرنل المقترحة هي الأفضل عند تقدير انموذج الانحدار اللوجستي شبه المعلمي الضبابي.

11. الاستنتاجات conclusions

- سوف يتم التعرف على اهم الاستنتاجات التي توصل اليها من خلال هذا البحث
- 1- هناك علاقة خطية تمثل مشكلة التعدد الخطي بين المتغيرات المستقلة التي لها تأثير كبير على الإصابة بالفشل الكلوي و لا سيّما بين المتغير التوضيحي X_1 الذي يمثل مرض السكري والمتغير التوضيحي X_3 الذي يمثل عمر المريض وتم تحديد هذه المشكلة من خلال العدد الشرطي الذي كانت قيمته اكبر من 30 ومصفوفة نسبة التباين.
 - 2- من خلال المعلومات الضبابية المقدرة في كلا الطريقتين (FOLSE,FLLE) نجد ان مرض السكر له دور كبير على الإصابة بالفشل الكلوي.
 - 3- من خلال متوسط مربعات الخطأ MSE الذي تم حسابه إلى انموذج الانحدار اللوجستي شبه المعلمي المقدّر نجد ان متوسط مربعات الخطأ لمقدرات ليو اللوجستي الضبابي أفضل من متوسط مربعات الخطأ لمقدرات المربعات الصغرى الاعتيادية الضبابية وذلك بسبب وجود مشكلة التعدد الخطي وعندما تكون دالة كيرنل تمثل الدالة المقترحة.

المصادر:

- 1- Abbas, Muayad S. (2017). Fuzzy nonparametric regression model estimation using some smoothing methods. Dissertation in statistics. College of Administration & Economics. University of Baghdad.
- 2- Cheng, C. B., & Lee, E. S. (1999). Nonparametric fuzzy regression k-NN and kernel smoothing techniques. Computers & Mathematics with Applications, 38(3-4), 239-251.
- 3- Donoso, S., Marín, N., & Vila, M. A. (2006, June). Quadratic programming models for fuzzy regression. In Proceedings of international conference on mathematical and statistical modeling in honor of Enrique Castillo.
- 4- Ebaid, R., Farghali, R., & Abo-El-Hadid, S. (2017). A mathematical programming approach for Liu-type estimator. Advances and Applications in Statistics, 50(4), 293.

- 5- Gao, Y., & Lu, Q. (2018). A fuzzy logistic regression model based on the least squares estimation. *Computational and Applied Mathematics*, 37(3), 3562-3579.
- 6- Härdle, W. K., Müller, M., Sperlich, S., & Werwatz, A. (2004). *Nonparametric and semiparametric models*. Springer Science & Business Media.
- 7- Hesamian, G., & Akbari, M. G. (2017). Semi-parametric logistic regression model with exact inputs and intuitionistic fuzzy outputs. *Applied Soft Computing*, 58, 517-526.
- 8- Inan, D., & Erdogan, B. E. (2013). Liu-type logistic estimator. *Communications in Statistics-Simulation and Computation*, 42(7), 1578-1586.
- 9- Mohammed, Mohammed J. & Rashid, Dhafer, H. (2007). Robust Estimations for Fuzzy Regression. The 1st Arab Statistical Conference. Amman-Jordan 12-13nov/2007.
- 10- Namdari, M., Abadi, A., Taheri, S. M., Rezaei, M., Kalantari, N., & Omidvar, N. (2014). Effect of folic acid on appetite in children: Ordinal logistic and fuzzy logistic regressions. *Nutrition*, 30(3), 274-278.
- 11- Shahmandi, M., Farmanesh, F., Gharahbeigi, M. M., & Shahmandi, L. (2013). Data Analyzing by Attention to Weighted Multicollinearity in Logistic Regression Applicable in Industrial Data. *British Journal of Applied Science & Technology*, 3(4), 748-763.
- 12- Shemail, A. H., & Mohammed, M. J. (2022). Comparison between some estimation methods for an intuitionistic fuzzy semi-parametric logistic regression model with practical application about covid-19. *International Journal of Nonlinear Analysis and Applications*, 13(1), 3723-3732.
- 13- Speckman, P. (1988). Kernel smoothing in partial linear models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 50(3), 413-436.
- 14- Wu, J., Asar, Y., & Arashi, M. (2018). On the restricted almost unbiased Liu estimator in the logistic regression model. *Communications in Statistics-Theory and Methods*, 47(18), 4389-4401.
- 15- Zadeh, L. A. (1965). Fuzzy Sets*. *Information control*, 8, 338—353.