Using The Ridge Regression Method In Addressing Linear Multiplicity Problem

استعمال طريقة انحدار الحرف في معالجة مشكلة التعدد الخطي

أ. م. د. شروق عبد الرضا السباح زينب كاظم مزهر
جامعة كربلاء / كلية الادارة والاقتصاد / قسم الاحصاء
بحث مستل من رسالة ماجستير

المستخلص:-
عند بناء نموذج انحدار خطي متعدد غالبا ما تواجه الباحث مشكلة التعدد الخطي Multi Linearity Co ، فتصبح تقديرات طريقة المربعات الصغرى غير كفؤة ، وذو قدرة تنبؤيه ضعيفة ومن البدائل لتخطي هذه المشكلة هي طريقة انحدار الحرف Ridge Regression
ويهدف الوصول الى افضل طريقة تمتلك قيم تقديرية دقيقة سحبنا عينة عشوائية بسيطة حجمها 100 شخص من المرضى المصابين بالجلطة الدماغية والقلبية الراقدين في مستشفى الحسين في محافظة كربلاء وباستعمال البرامج الاحصائية ((SPSS,NCSS)) ، لتحليل البيانات والكشف عن وجود مشكلة التعدد الخطي ومعالجتها بطريقة الانحدار OLS, R R،
وقد بينت النتائج ان طريقة انحدار الحرف هي افضل من طريقة المربعات الصغرى .

Abstract-:

When constructing a linear regression model, the researcher often encounters the problem of multi-linearity. Thus, the estimation of the method of the least squares becomes inefficient, with low predictive capacity .

Thus as alternatives to overcome this problem are the regression method of the Ridge Regression

The aim is to reach the best method with estimated values, from which a simple random gamble of 100 patients from patients with stroke and cardiac arrest at Al-Hussein hospital in Karbala governorate were taken we used statistical programs (SPSS, Statghraf Minitab), for analysis the data. Results of the showed that the method ridge is better than the method of least square

.

المقدمة:

تستخدم النماذج الخطية بشكل واسع في مختلف مجالات العلم وان الانحدار الخطي المتعدد هو احد النماذج الخطية التي يكثر استعماله في تحليل بيانات العديد من البحوث الطبية والصحية والاقتصادية والاجتماعية والعلوم التطبيقية الاخرى، عند تطبيق تحليل الانحدار الخطي المتعدد تواجهنا مشكلة التعدد الخطي التي تحدث عند وجود علاقة ارتباط قوية بين المتغيرات التوضيحية ويؤدي هذا الارتباط الى زيادة تباين مقدرات معلمات الانحدار الامر الذي يجعل نتائج طريقة المربعات الصغرى الاعتيادية (OLS) غير موثوق بها والتي من احد الفرضيات الاساسية لها هي وجوب استقلالية المتغيرات التوضيحية (عدم وجود ارتباط فيما بينها).

ولحل مشكلة التعدد الخطي تم اقتراح عدة طرق من قبل الكثير من الباحثين ومنها طريقة انحدار الحرف Ridge Regression الذي اقترح من قبل (Horel&Kennard,1970) ويعد احدى الطرق البديلة لطريقة المربعات الصغرى الاعتيادية.

- مشكلة البحث:

لغرض تقدير نموذج انحدار خطي لابد ان يتمتع هذا النموذج بخصائص معينة تعتمد على عدة افتراضات وفي حالة غياب احد تلك الافتراضات فأن النموذج سيعاني جملة من المشكلات والتي تجعل عملية التقدير خاطئة او في بعض الاحيان غير ممكنة ومن هذه المشكلات هي مشكلة التعدد الخطي اي عدم وجود استقلالية بين المتغيرات التوضيحية حيث تكون البيانات مترابطة فيما بينها.

- اهمية البحث:

تكون اهمية هذا البحث محاولة الوصول الى فهم أعمق لظاهرة مشكلة التعدد الخطي وتناول طرق الكشف عنها واهم المظاهر الدالة على وجودها في نموذج الانحدار الخطي، ودراسة طريقة انحدار الحرف كصيغة مقترحة لمعالجة مشكلة التعدد الخطي.

- منهجية البحث:
تم العمل على المنهج الاستقرائي وفيه يبدأ بملاحظة المشكلة ثم وضع الفروض لها وبعد ذلك اختبارها، وقد تم استعمال الاسلوب الاحصائي وفق هذا المنهج .

- منهجية البحث:
كانت الدراسة وفق المنهج الاستقرائي الذي يقوم على سحب العينة ودراستها وتصميم نتائجها على مجتمع الدراسة

- هدف البحث:
- تشخيص وجود مشكلة التعدد الخطي ومدى تأثيرها على نموذج الانحدار.
- دراسة مدى ارتباط المتغيرات فيما بينها وتعين المتغيرات التوضيحية ذات الاهمية الاحصائية وتأثيرها على المتغير المعتمد.
- التخلص من الارتباط الموجود بين المتغيرات المستقلة (التعدد الخطي) بطريقة انحدار الحرف.

الجانـــب النظـــري:

- الجلطة الدماغية او السكتة الدماغية [1] :
السكتة الدماغية أوالجلطة الدماغية وتعرف سابقا باسم حادثة وعائية دماغية (Cerebrovascular accident – CVA) تحدث عندما يتوقف تدفق الدم إلى أحد أجزاء الدماغ مما يؤدي الى حرمان أنسجة المخّ من الأوكسجين الضروري جدا ومواد التغذية الحيوية الأخرى ومن جراء ذلك تتعرض خلايا المخ للموت خلال دقائق قليلة.
السكتة الدماغية هي حالة طوارئ طبية، والعلاج الفوري لها أمر بالغ الحيوّية والأهمية إذ يمكن من خلاله تقليل الأضرار للدماغ ومنع المضاعفات المحتملة ما بعد السكتة.

- مضاعفات الجلطة الدماغية:
المضاعفات المحتملة نتيجة للسكتة الدماغية تختلف باختلاف الجزء المتضرر من الدماغ وهذه المضاعفات تشمل:
- شلل أو فقدان القدرة على تحريك العضلات.
- الاوجاع
- فقدان الذاكرة أو مشاكل في الفهم العام
- صعوبات في الكلام أو البلع.

- الجلطة القلبية:
وتُعرف أيضا باسم النوبة القلبية وهي مرض قلبي حاد مهدِّد للحياة يحدث بسبب احتباس الدم نتيجة انسداد أحد الشرايين التاجية مما يؤدي إلى ضررٍ أو موتٍ كاملٍ لجزء من عضلة القلب غالباً ما تكون النوبة طارئاً طبياً يهدد حياة المريض ويستدعي الرعاية الطبية الفورية. تُشَخَّص الحالة بواسطة تاريخ المريض الطبي ونتائج فحص تخطيط القلب وإنزيمات القلب في الدم. أكثر الإجراءات اللازم اتخاذها فوراً هي إعادة تدفق الدم إلى القلب لذا يجب الإسراع بنقل المصاب فوراً إلى مستشفى أو إلى طبيب أو إحضار الطبيب لموقع المصاب لمعالجته؛ حيث يلعب الوقت هنا دوراً هاماً جداً فتجب السرعة كي يتم إعادة تدفق الدم في الشريان التاجي بأحد أمرين أو كليهما: إذابة الخثرة ((وهي كدرة الدم المسببة لانسداد الشريان)) بمضادات التخثر ورأب الوعاء وتسمى أيضا((توسعة الشرايين))وهو إيلاج وتد على متنه بالون إلى الوعاء الدموي المنسدّ، بحيث ينتفخ البالون حين يكون بمحاذاة الخثرة فتنحسر الخثرة نحو الجوانب ويتّسع فضاء الوعاء بعدما ضاق مُتيحاً للدم الانسياب عبره وينبغي على وحدة العناية التاجية مراقبة المريض عن كثب تحسباً لمختلف التطورات.

- مضاعفـات الجلطـة القلبيـة:
تنتج مضاعفات الجلطة القلبية جراء تلف أنسجة القلب، ومن هذه المضاعفات:
- اضطراب ضربات القلب: فعند حدوث تلف في أنسجة القلب تتكون دارات كهربائية شاذة في النسيج التالف تؤثر بدورها على التوصيل الكهربائي للقلب محدثة اضطرابا في النسق الطبيعي لضربات القلب وتكون هذه الاضطرابات خطيرة قد تؤدي إلى الوفاة.
- خلل في صمامات القلب .
- تمزق أنسجة القلب: إذ تكون الأنسجة التالفة من القلب عرضة بشكل اكبر للتمزق.
- فشل القلب: يحدث فشل القلب إذا كان الجزء المتضرر من الجلطة القلبية كبيرا وغير قادر على ضخ الدم بشكل كاف ويكون هذا الفشل إما موقتا يختفي بعد الشفاء القلب أو دائما.

- التعدد الخطي [3] :
ن مصطلح التعدد الخطي او الارتباط الخطي المتعدد هو مصطلح مركب من (Multi) متعدد و(Co) مشترك او متداخل او مرتبط و(Linearty) خطي ويشير المصطلح الى وجود علاقة خطية تامة او غير تامة بين اثنين او اكثر من المتغيرات التوضيحية في نموذج الانحدار مما يؤدي الى خرق احد فروض التحليل، اي عندما تكون مصفوفة التصميم تامة الرتبة وجميع القيم الذاتية في مصفوفة التصميم تساوي واحد ويدل هذا على ان التعامد موجود . واذا كان على الاقل واحد من القيم الذاتية مساويا للصفر او يقترب من الصفر يعني ان التعامد غير موجود.

- انواع العلاقات الخطية بين المتغيرات التنبؤية [7] :

1- العلاقة الخطية التامة:
تتحقق هذه العلاقة عندما يكون هناك علا قة خطية بين قيم اثنين او اكثر من المتغيرات التوضيحية فتكون نتيجة محدد المصفوفة تساوي صفر $|X'X| = 0$ وهذا يؤدي الى انتهاك شرط الرتبة اذ تكون المصفوفة غير كاملة الرتبة اي ان Rank(X)<P الرتبة اقل من عدد المتغيرات وعليه فانه لايمكن ايجاد معكوس مصفوفة المعلومات وبالتالي لايمكن تقدير معلمات النموذج وهنا يظهر التعدد الخطي التام.

2- العلاقة الخطية الغير تامة (الجزئية) :[6,7]
تظهر هذه الحالة عندما يكون محدد مصفوفة المعلومات لا يساوي صفر $|X'X| \approx 0$ وانما قريب منه وتميل المتغيرات للتحرك سوية بالزيادة او النقصان كما في حالة استخدام المتغيرات المرتدة ففي هذه الحالة يكون تقدير معلمات النموذج غير دقيقة وغير ممثلة لواقع المشكلة المدروسة لان تباين المعلمات سيكون كبيرا نتيجة لكبر حجم الاخطاء المعيارية وهذا يؤثر على ظهور بعض او جميع قيم إحصاءه الاختبار t-test لمعلمات النموذج صغيرة نسبيا مما يؤدي الى عدم معنوية هذه المقدرات وان كانت قيمة معامل التحديد R^2 للنموذج كبيرة.

- الكشف عن وجود مشكلة التعدد الخطي :[11]

1- مقياس تضخم تباينات المعاملات Variance Inflation Factor (VIF)
يعتبر مقياس تضخم التباين من الطرق الاساسية للكشف عن وجود مشكلة التعدد الخطي وهو يقيس مدى تضخم تباينات معاملات الانحدار المقدرة عند وجود ارتباط خطي بين المتغيرات التوضيحية ،تكون قيمته دائما اكبراوتساوي واحد ويمكن ايجاده بالاعتماد على معامل التحديد كما بالصيغة التالية :

$$VIF=1/(1-R^2_j) \qquad ....(1)$$

او من خلال البرنامج الجاهز SPSS تكون قيم معامل VIF ≥ 1 فاذا كانت القيم اكبر من 5 دل على انه توجد مشكلة تعدد خطي.

2- القيم المميزة والمتجهات المميزة Eigen Values
تعد القيم المميزة لمصفوفة الارتباطات معلمات نموذج الانحدار مؤشرات جيدة لقياس التعدد الخطي ويتم حساب القيم المميزة بالخطوات التالية

ايجاد المصفوفة Z

$$Z=(X'X) \qquad ....(2)$$

ايجاد المصفوفة القطرية S

$$S=dig(X'X)^{-\frac{1}{2}} \qquad ....(3)$$

ايجاد مصفوفة SZS

$$SZS=S(X'X)S \qquad ....(4)$$

(عندما تساوي احدى القيم الذاتية $\lambda 1$، $\lambda 2$، .. ...، λp لمصفوفة (X´X) الى الصفر او احدهم قريبة من الصفر معناه يوجد تعدد خطي).

3- رقم الحالة Condition Number (C.N)
هناك عدة صيغ لحساب العدد الشرطي التي تشير إلى درجة التعدد الخطي وابرزها هي

$$C.N=\sqrt{\frac{\text{اكبر قيمة ذاتية}}{\text{اصغر اقل ذاتية}}} \qquad ....(5)$$

اذ اقترح Johnston اذا كانت قيمة C.N تتراوح بين 20 الى 30 تكون مؤشر لوجود تعدد خطي مرتفع بينما اقترح Belsley انه ذا كانت قيمة C.N تتراوح بين 30 الى 100 يكون مؤشر الى وجود تعدد خطي مرتفع جدا.

-4 مؤشر الحالة Condition Index (C.I)

يستخدم هذا المؤشر للكشف عن وجود التعدد الخطي ومن خلال القيم المميزة يتم حساب هذا المؤشر وفق الصيغة الاتية :

$$CI_j = \sqrt{\frac{\text{اكبر قيمة}}{j\ \text{القيمة}}} \qquad , \qquad j = 1,2, \ldots . p \qquad \ldots \ldots \ldots (6)$$

فاذا كانت قيمة C.I اكبر من 30 هذا دليل على وجود مشكلة تعدد خطي

-5 مصفوفة معاملات الارتباط Correlation Matrix

يعد اختبار مصفوفة معاملات ارتباط المتغيرات التوضيحية من ابسط طرق الكشف عن التعدد الخطي فاذا كان الارتباط بين المتغيرات التفسيرية $X_i\ ,\ X_j$ عالي فان القيمة المطلقة ل $|r_{12}|$ تكون قريبة من الواحد مما يعرض النموذج للتعدد الخطي ولكن ضعف الارتباط بين المتغيرات لا يعني عدم وجود تعدد خطي.

اختبار فاراروكلوبير Farrar& Glauber

وضع هذا الاختبار في عام (1967) ويستند بشكل اساسي على اختبار مربع كاي (χ^2) والصيغة الرياضية له :

$$\chi^2=-[n-1-1/6(2k+5)]\ \ln|M| \qquad ....(7)$$

اذ ان

M: تمثل مصفوفة الارتباط

K:تمثل عدد المتغيرات

n:تمثل حجم العينة

وتقارن مع القيمة الجدولية ل χ^2 ونرفض فرضية العدم اذا كانت القيمة المحسوبة هي الاكبر .

معالجة مشكلة التعدد الخطي[2]:

1. (زيادة حجم العينة) يؤدي الى تقليل الارتباط بين المتغبرات المستقلة.
2. حذف بعض المتغيرات التي يزداد ارتباطها بالمتغيرات الاخرى.
3. التحويل المعياري للمتغيرات التوضيحية.

معالجة التعدد الخطي بطريقة انحدار الحرف Ridge Regrsion :

مقدرات انحدار الحرف[4]:

تعد مقدرات انحدار الحرف احدى طرق معالجة مشكلة التعدد الخطي اي عندما يكون محدد مصفوفة المعلومات صغير جدا والمعالم المقدرة ذات تباين كبير في هذه الحالة يكون التعدد الخطي شبه تام وتكون معالجة هذه الحالة بطريقة انحدار الحرف واول من اقترح هذه الطريقة (Kennard& Hoerl عام 1970).

ويمكن تعريفه بانه اسلوب يحاول تقديم تقديرات دقيقة لمعاملات الانحدار كما يعد بانه طريقة بديلة عن المربعات الصغرى الاعتيادية حيث ان في تحليل الانحدار تكون مقدرات المربعات الصغرى غير مستقرة في حالة التعدد الخطي اي ان تأثيرعدم التعامد في المتجهات التنبؤية هو ابعاد مقدرات المربعات الصغرى الاعتيادية عن قيمتها الحقيقية التي نحاول ايجاد تقديرات قريبة منها بالإضافة الى انها قد تكون كبيرة بالقيمة المطلقة واشارتها غير منطقية وبذلك تعد غير مستقرة وقد يؤدي عدم الاستقرار من اضافة او حذف متغير معين الى اخطاء قياسية لمعلمات قد تكون مهمة ويمكن معالجة عدم الاستقرار بطريقة انحدار الحرف وذلك بإضافة كمية صغيرة موجبة تقع قيمتها بين الصفر والواحد ($0\leq k\leq 1$) الى العناصر القطرية لمصفوفة المعلومات للحصول على معلومات اكثر دقة اذ تعمل الطريقة الى فك الاعتمادية بين المتغيرات التوضيحية وكلما كان الارتباط عالي فان المعاملات تتغير بسرعة ثم تبدا بالاستقرار تدريجيا عندما تكبر قيم K الى ان تصل الى حد تستقر فيه المعلمات وكلما كان استقرار المعلمات سريعا دل ذلك ان المتغيرات التنبؤية قريبة من التعامد وبذلك تكون مهمة طريقة انحدار الحرف هي جعل المعلمات المقدرة ثابتة من حيث الاشارة والقيمة.

- التقدير بطريقة انحدار الحرف [5] :

تحدث مشكلة التعدد الخطي عندما تكون المتغيرات التنبؤية غيرمتعامدة مما يؤدي الى تضخم لتباينات المعاملات المقدرة وعدم الاستقرار فيها مما يصعب تقدير المعلمات بطريقة المربعات الصغرى الاعتيادية ذات الصيغة العامة

$$\hat{\alpha}_K = (X'X)^{-1}X'Y \qquad ....(8)$$

اذ اقترح اسلوب الحرف ليعالج هذه المشكلة في ايجاد مقدرات مستقرة ويعد هذا الاسلوب بانه تحوير لطريقة المربعات الصغرى لنحصل من خلاله على مقدرات تمتلك صفة مميزة هي صفة اقل متوسط خطأ على الرغم من انها متحيزة كذلك حدود الثقة تكون محصورة جدا وان ابرز ما يميز طريقة انحدار الحرف هي اضافة الثابت K الى عناصر المصفوفة $(X'X)$ قبل اخذ المعكوس لها ويمكن الحصول على مقدراتها باستعمال الصيغة الاتية:

$$\hat{\alpha}_k = (X'X + K)^{-1}X'Y \qquad ....(9)$$

وان قيمة K تتراوح بالمدى (0,1) ، $0 \le K \le 1$

بالتعويض عن قيمة $K=KI_p$ في المعادلة اعلاه نحصل على:

$$\hat{\alpha}_k = (X'X+KI_p)^{-1}X'Y \qquad ......(10)$$

ويمكن التعبير عن المعادلة (10) بدلالة القيم والمتجهات المميزة كما في الصيغة الاتية:

$$\hat{\alpha}_k = \sum_{j=1}^{p}(\lambda + k)^{-1} V_j V_j X'Y \qquad ......(11)$$

حيث ان $(X'X)$ هي مصفوفة قطرية والقيم المميزة للمصفوفة $(X'X)$ تمثل عناصرالقطر اي ان

$$X'X=(\text{Diag } \lambda_1\lambda_2 \ldots\ldots \lambda_p)$$

اذ ان $P = m + 1$ ، m تمثل عدد المتغيرات التوضيحية في المعادلة (10) ، تم التوصل اليها من خلال مجموعة من المعادلات الطبيعية التي تعود الى تقديرات المربعات الصغرى الاعتيادية والمتمثلة بالمعادلة (8) فعند اضافة المصفوفة القطرية kI_p ذات البعد (p×p) الى المصفوفة $(X'X)$ قبل اخذ معكوسها فانه سوف نحصل على المعادلة الاتية :

$$(X'X + KI_p)\hat{\alpha}_{(k)} = X'Y \qquad ....(12)$$

وبضرب طرفي المعادلة (9) بالمصفوفة $(X'X + kI_p)^{-1}$ نحصل على المعادلة (11) التي تمثل مقدرات انحدار الحرف . ان الفائدة الاساسية من اضافة المصفوفة (kI_p) الى المصفوفة $(X'X)$ هو التقليل من عناصر القطر الرئيسي للمصفوفة $(X'X)^{-1}$ واكبر تصغير يظهر في المتغيرات التي تعاني من تضخم في تباين معاملاتها .

طرق اختيار معلمة الحرف Methods of Choosing Ridge Parameter (K)

توجد طريقتين رئيسيتين لاختيار معلمة الحرف هما : طريقة اثر الحرف Ridge Trace وطريقة تقدير قيمة K وفيما يلي توضيح الطريقتين:

1- طريقة اثر الحرف Ridge Trace Method :

تعتمد طريقة اثر الحرف على الرسم البياني لتحديد قيمة K المثالية ويتم ذلك من خلال رسم الاحداثيين ، اذ يمثل الاحداثي الافقي قيم k المحدد بين (0,1) اما الاحداثي العمودي فيمثل المعالم المقدرة بطريقة انحدار الحرف عند كل قيمة من k ومع الزيادة التدريجية لقيمة عامل التحيز ابتداء من الصفر ضمن المدى المحدد يقوم الباحث باختيار قيمة K المناسبة لتصبح التقديرات جميعها ثابتة بزيادة k وتشهد عندها الدالة استقرارا ملحوضا.

ويمكن استخدام طريقة اثر الحرف كمعيار لاختيار المتغيرات التفسيرية وذلك بتطبيق قاعدة حذف المتغيرات من النموذج وفق الخطوات الاتية :

1. حذف المتغير التفسيري الذي تكون له معاملات غير قادرة على عكس قدرته التنبؤية وتوؤل الى الصفر.
2. حذف واحد او اكثر من المتغيرات التي لها معاملات غير مستقرة.
3. حذف المتغير وان كان مستقر اذا كان يمتلك قدرة تنبؤية صغيرة ويستدل عليه بصغر معامله المعياري.

2- طريقة تقدير Methods Of Estimate k :

تعتبر طريقة تقدير معلمة الحرف طريقة موضوعية مقابل طريقة اثر الحرف اذ ان اختيار القيمة المثلى لمعلمة الحرف يعتمد على قيم متغير الاستجابة والهدف من ايجاد قيمة Kهو خفض التباين بشكل كبير والزيادة في قيمة مربع التحيز.

انحدار الحرف واختيار المتغيرات التفسيرية [6] :

عندما تكون العلاقة بين المتغيرات متعامدة يكون اختيار افضل المتغيرات في معادلة الانحدار باحدى الطرق المعروفة وهي (اسلوب الاختيار الامامي اواسلوب الحذف الخلفي او اسلوب انحدار الخطوات المتسلسلة) ولكن عندما تكون العلاقة بين المتغيرات غير متعامدة وتوجد مشكلة التعدد الخطي بشكل كبير فان استخدام احدى الطرائق سابقة الذكر قد تكون غير مجدية ولا تعطي نتائج جيدة عند تقدير المعلمات بطريقة المربعات الصغرى الاعتيادية.

ان عملية تقريب المتغيرات التفسيرية من التعامد فيما بينها وباستخدام عامل التحيز يجعل الاختيار افضل للمتغيرات التفسيرية ذات التأثير الاقوى في متغير الاستجابة فعند الاعتماد على طريقة اثر الحرف للتغلب على تعدد العلاقة الخطية ومن خلال طريقة تعيين الثابت k يمكن الاستفادة منها في اختيار افضل المتغيرات التفسيرية ، فقد ذكر كل من Hoerl & Kennard بانه من الممكن الاعتماد على رسم الحرف بطريقة اثر الحرف كدليل لاختيار افضل المتغيرات التي تدخل في معادلة الانحدار و اقترحوا القوانين التالية :

1. يمكن التعرف على كل متغير من المتغيرات بالاعتماد على طريقة اثر الحرف (ridge trace) فالمتغيرات التوضيحية المعتمدة في النموذج تميل الى الاستقرار في الرسم عند نقطة الانقلاب.
2. عند قيمة الصفر في رسم اثر الحرف فان تقدير المعلمات لا تتقاطع مع المحور السيني.
3. قيم المعلمات المقدرة للمتغير التوضيحي الذي يعتمد في النموذج لا تكون صغيرة أي لا تقترب من الصفر.

الجانب التطبيقي:

وصف البيانات:

تم تطبيق الموضوع على بيانات واقعية ل 100 مريض من المصابين بالجلطة الدماغية والقلبية الراقدين في مستشفى الحسين في محافظة كربلاء المقدسة خلال فترة ثلاثة اشهر.

متغيرات البحث :

يمثل المتغير (اين حصول الخثرة او مكان الخثرة) هوالمتغير التابع والمتغيرات المستقلة يمكن توضيحها بالجدول الاتي رقم (1):

جدول (1) توضيح المتغيرات

الرمز	المتغيرات	الرمز	المتغيرات	الرمز	المتغيرات
C1	الجنس	C9	وقت مراجعة الطبيب	C17	المستوى التعليمي
C2	العمر	C10	حصول الخثرة	C18	عدد ساعات النوم الليل
C3	انواع علاج الضغط	C11	ساعة الاستيقاظ صباحا	C19	عدد ساعات النوم النهار
C4	وجود الضغط	C12	نوع الجلطة	C20	تناول الادوية
C5	انواع علاج السكر	C13	عامل الوراثة	C21	ممارسة النشاط الرياضة
C6	وجود السكر	C14	درجة القرابة	C22	وجود امراض اخرى
C7	تكرار حصول الجلطة	C15	نوع العمل	C23	اين وجود الجلطة
C8	وقت حصول الجلطة	C16	طبيعة مزاج المريض		

قبل اختبار البيانات يتطلب معرفة معنوية النموذج ومعامل التحديد للنموذج كما في الجدول الاتي رقم (2) الذي يبين اختبار معنوية النموذج.

جدول (2) تحليل التباين					
النموذج	مجموع مربعات	درجة الحرية	متوسط مربعات	قيمة F	القيمة الاحتمالية(sig)
الانحدار	83.792	22	3.806	83.457	.000
البواقي	3.511	77	.045		
المجموع الكلي	87.240	99			
المصدر: من اعداد الباحثة باستخدام برنامج SPSS					

نلاحظ ان قيمة ال sig هي (0.00) اقل من مستوى الدلالة الاحصائية (0.05) وهذا يعني ان النموذج معنوي وان قيمة معامل التحديدR^2 مساوية الى 0.96 ومعامل الارتباط 0.98.

الكشف عن وجود مشكلة التعدد الخطي:

تم الكشف عن وجود التعدد الخطي بين المتغيرات التوضيحية للمصابين بالجلطة القلبية والجلطة الدماغية بطرائق عدة وذلك باستخدام البرنامج الجاهز SPSS، اما الطرائق المعتمدة في الكشف عن وجود مشكلة تعدد العلاقة الخطية هي:

مقياس تضخم تباين المعاملات (VIF):

يعد هذا المقياس كما ذكر من اهم مقاييس الكشف عن وجود مشكلة التعدد الخطي اذ يمكن قياسهما بطريقتين هما:

الطريقة الاولى: بالاعتماد على قيمة معامل التحديد كما في الصيغة (1):

$$VIF=1/(1-0.96)=25$$

وبما ان قيمة VIF كبيرة تدل على وجود مشكلة التعدد الخطي.

والطريقة الثانية: من خلال قيم الVIF الجاهزة بالبرنامج اي انه نلاحظ وجود معاملات الVIF معامل تضخيم التباين اكبر من (5) فانه دليل على ان المتغير X_j يعاني من تضخم تباين معامله اي وجود مشكلة تعدد العلاقة الخطية كما موضح في الجدول (3)

الجدول (3) يمثل مقياس تضخم التباين VIF لمتغيرات المصابين بالجلطة			
المتغيرات	VIF	المتغيرات	VIF
الجنس	1.44271	نوع الجلطة	1.29584
العمر	1.16589	عامل الوراثة	1.19137
انواع علاج الضغط	1.16411	درجة القرابة	1.25721
وجود الضغط	1.10409	نوع العمل	1.13288
انواع علاج السكر	2.81879	طبيعة مزاج المريض	2.60694
وجود السكر	2.71205	المستوى التعليمي	2.52997
تكرارحصول الجلطة	1.5896	عدد ساعات النوم الليل	2.85753
وقت حصول الجلطة	1.18571	عدد ساعات النوم النهار	2.8091
وقت مراجعة الطبيب	1.60025	تناول الادوية	1.16777
حصول الخثرة	1.2747	ممارسة النشاط الرياضة	1.90493
ساعة الاستيقاظ صباحا	1.1958	وجود امراض اخرى	1.96794
* تعني ان هذه المتغيرات تعاني من تضخم تباين معاملاتها.			

وبما ان قيم تضخم تباين المعاملات VIF في الجدول (3) للمتغيرات التوضيحية لكل من وجود الضغط وانواع علاج الضغط وانواع علاج السكر و وجود السكر وحصول الخثرة وعامل الوراثة ودرجة القرابة وطبيعة مزاج المريض اكبر من (5) فان هذا يعني ان هذه المتغيرات تعاني من مشكلة تعدد العلاقة الخطية بين المتغيرات التوضيحية.

القيم المميزة او الجذور المميزة:

تعد الجذور المميزة من الاختبارات الواضحة للكشف عن وجود مشكلة تعدد خطي أي عندما يكون هناك عددا من القيم صغيرة او قريبة من الصفر هذا يعني توقع وجود مشكلة تعدد العلاقة الخطية وان مجموع هذه الجذور يكون مساوي الى عدد المتغيرات:

الجدول (4) يبين قيم الجذور المميزة

رمز الجذور المميزة	قيم الجذور المميزة	رمز الجذور المميزة	قيم الجذور المميزة
$\lambda 1$	18.522	$\lambda 12$	.072
$\lambda 2$	.942	$\lambda 13$	.059
$\lambda 3$	.807	$\lambda 14$	.050
$\lambda 4$	.497	$\lambda 15$	.040
$\lambda 5$	.435	$\lambda 16$	.037
$\lambda 6$	.348	$\lambda 17$	.034
$\lambda 7$	.334	$\lambda 18$	.027
$\lambda 8$	.243	$\lambda 19$	.022
$\lambda 9$	.228	$\lambda 20$	.011
$\lambda 10$	.147	$\lambda 21$	.006
$\lambda 11$	.132	$\lambda 22$	.005
مجموع الجذور الذاتية يساوي 23			

ومن خلال الجدول نلاحظ ان مجموع الجذور الذاتية هو 23 يكون مساوي الى عدد المتغيرات البالغ عددها23.

رقم الحالة Condition Number (C.N):

يمكن حساب العدد الشرطي او رقم الحالة بالاعتماد على القيم الذاتية في الجدول اعلاه وفق الصيغة (2) التي ذكرت سابقا.

$$CN=\sqrt{\frac{18.522}{0.002}} = 96.23$$

نلاحظ بما ان قيمة CN تقترب بكثير من 100 هذا يدل على وجود تعدد خطي كبير وخطير.

مؤشر الحالة Condition Index (C.I):

يعد مؤشر الحالة من المؤشرات التي تبرز ظاهرة مشكلة التعدد الخطي وبالاعتماد على القيم الذاتية في الجدول رقم (5) ويكون حسابه وفق الصيغة (6):

$$C.I_1 = \sqrt{\frac{18.522}{18.522}} = 1$$

$$C.I_2 = \sqrt{\frac{18.522}{.942}} = 4.433$$

و هكذا بقية القيم كما في الجدول الاتي رقم (5).

الجدول (5) يبين قيم مؤشر الحالة			
مؤشر الحالة	القيمة	مؤشر الحالة	القيمة
λ 1	1.000	λ13	17.718
λ2	4.433	λ14	19.247
λ3	4.791	λ15	21.518
λ4	6.105	λ16	22.374
λ5	6.528	λ17	23.340
λ6	7.327	λ18	26.191
λ7	7.447	λ19	29.016
λ8	8.731	λ20	41.034
λ9	9.013	λ21	55.560
λ10	11.225	λ22	60.864
λ11	11.846	λ 23	96.234
λ12	16.093		

من الجدول رقم (5) اعلاه نجد أن الجذور المميزة (من الجذر المميز رقم 1 إلى الجذر المميز رقم 19) لم تظهر قيمة لدليل الحالة المقابلة لتلك الجذور اكبر من 30 في حين أن قيم دليل الحالة من 20 الى 23 قد تجاوزت 30 مما يدل على أن البيانات للظاهرة المدروسة تعاني من مشكلة التعدد الخطي أي توجد علاقة خطية تربط بين المتغيرات التوضيحية.

اختبار فراير وكلوبير:

ان هذا الاختبار يستخدم لاختبار الفرضية الآتية :

H_0 : عدم وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية

H_1 : وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية

ومن الصيغة (7) فان قيمة مربع كاي χ^2_{cal} المحسوبة :

$$\chi^2_{cal} = -[100-1-(1/6(2(22)+5)\ln(2.81^{-05})]$$
$$= -[(99-8.1666667)(-10.4797409)]$$
$$= 951.909$$

k=23

n=100

ومن جداول توزيع مربع كاي فان القيمة الجدولية لـ مربع كاي χ^2_{tab} بدرجة حرية $k(k-1)/2$ هي (231) عند مستوى معنوية 0.01 تكون مساوية الى (183.955) عندما n=100, K=23 وبمقارنة χ^2_{cal} مع قيمة χ^2_{tab} نلاحظ ان $\chi^2_{tab} > \chi^2_{cal}$ لذا ترفض فرضية العدم وتقبل الفرضية البديلة التي تدل على وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية.

معالجة مشكلة التعدد الخطي باستعمال طريقة انحدار الحرف:

عولجت مشكلة التعدد الخطي بطريقة انحدار الحرف بالاعتماد على البرنامج الاحصائي Statghraf Minitab وذلك باستخدام اساليب مختلفة في تحديد قيمة الثابت (k) وكانت ابرز الاساليب هي (*طريقة اثر الحرف*) بهذه الطريقة تم اختيار (21) قيمة لـ k تتراوح من 0.0 الى 0.1 وقيم المعلمات التقديرية لها ب اسلوب انحدار الحرف واستخراج قيم R^2 لكل قيمة من قيم K بهدف اختيار افضل قيمة ل K كما في الجدول (6).

الجدول (6) يبين المعاملات القياسية لكل قيمة من قيم K								
المتغيرات قيم K	Col_1	Col_2	Col_3	Col_4	Col_5	Col_6	Col_7	Col_8
0.0	0.0057442	0.0040757	0.0720866	0.0524095	0.0042731	0.0365402	-0.0189769	-0.0183905
0.005	0.0061608	0.0037440	0.0687781	0.049789	0.0050441	0.036961	-0.0184476	-0.0181529
0.01	0.0065483	0.0034642	0.0659169	0.0475825	0.0055855	0.0371607	-0.0179262	-0.0180443
0.015	0.0069094	0.0032236	0.0633978	0.0456902	0.0059717	0.0372124	-0.0174117	-0.0180291
0.02	0.0072464	0.0030133	0.0611484	0.0440435	0.0062494	0.0371622	-0.016904	-0.0180828
0.025	0.0075615	0.0028269	0.0591169	0.0425935	0.0064497	0.0370405	-0.0164032	-0.0181877
0.03	0.0078568	0.0026597	0.0572647	0.0413045	0.0065935	0.0368678	-0.0159095	-0.0183312
0.035	0.0081338	0.0025082	0.0555629	0.0401493	0.0066954	0.0366584	-0.0154231	-0.0185038
0.04	0.0083942	0.0023697	0.0539888	0.039107	0.0067658	0.0364227	-0.0149442	-0.0186984
0.045	0.0086393	0.002242	0.0525247	0.038161	0.0068124	0.0361679	-0.0144731	-0.0189096
0.05	0.0088701	0.0021233	0.0511561	0.0372981	0.0068408	0.0358996	-0.0140097	-0.0191331
0.055	0.0090878	0.0020123	0.0498713	0.0365075	0.0068552	0.0356219	-0.0135542	-0.0193657
0.06	0.0092932	0.0019079	0.0486608	0.0357803	0.0068589	0.0353378	-0.0131065	-0.0196048
0.065	0.0094872	0.0018090	0.0475163	0.0351092	0.0068544	0.0350498	-0.0126668	-0.0198483
0.07	0.0096706	0.0017151	0.0464312	0.0344878	0.0068435	0.0347595	-0.012235	-0.0200946
0.075	0.0098439	0.0016253	0.0453995	0.033911	0.0068278	0.0344684	-0.011811	-0.0203424
0.08	0.0100079	0.0015392	0.0444164	0.0333742	0.0068084	0.0341776	-0.0113947	-0.0205905
0.085	0.010163	0.0014563	0.0434774	0.0328734	0.0067863	0.0338879	-0.0109861	-0.0208382
0.09	0.0103098	0.0013762	0.0425788	0.0324053	0.0067621	0.0336	-0.0105851	-0.0210848
0.095	0.0104487	0.0012985	0.0417175	0.0319669	0.0067366	0.0333143	-0.0101915	-0.0213297
0.1	0.0105802	0.0012230	0.0408903	0.0315556	0.0067101	0.0330313	-0.009805	-0.021572

تابع الى جدول (6)

المتغيرات قيمK	Col_9	Col_10	Col_11	Col_12	Col_13	Col_14	Col_15
0.0	0.0010912	-0.536552	0.0225809	0.122464	0.0112651	-0.0036646	0.0129143
0.005	0.0000732	-0.530698	0.0219135	0.122737	0.0113319	-0.0038540	0.0126634
0.01	-0.0007813	-0.525503	0.0212431	0.12294	0.0112802	-0.0041379	0.0124202
0.015	-0.0015138	-0.52083	0.0205742	0.123089	0.011154	-0.0044780	0.0121865
0.02	-0.0021528	-0.51658	0.0199102	0.123196	0.0109816	-0.0048501	0.0119631
0.025	-0.0027187	-0.512676	0.0192532	0.12327	0.0107814	-0.0052383	0.0117502
0.03	-0.0032267	-0.509061	0.0186048	0.123317	0.0105657	-0.0056325	0.0115473
0.035	-0.0036877	-0.50569	0.017966	0.123341	0.0103427	-0.0060258	0.0113541
0.04	-0.0041105	-0.502527	0.0173374	0.123347	0.0101179	-0.0064137	0.0111701
0.045	-0.0045014	-0.499542	0.0167197	0.123336	0.0098950	-0.0067934	0.0109946
0.05	-0.0048657	-0.496713	0.016113	0.123312	0.0096765	-0.0071630	0.0108273
0.055	-0.0052073	-0.49402	0.0155175	0.123275	0.0094641	-0.0075215	0.0106674
0.06	-0.0055295	-0.491447	0.0149332	0.123228	0.0092587	-0.0078682	0.0105146
0.065	-0.0058350	-0.48898	0.0143602	0.123172	0.0090610	-0.0082029	0.0103683
0.07	-0.0061258	-0.486609	0.0137984	0.123108	0.0088712	-0.0085256	0.0102281
0.075	-0.0064038	-0.484323	0.0132476	0.123036	0.0086894	-0.0088364	0.0100936
0.08	-0.00667031	-0.482114	0.0127077	0.122957	0.0085156	-0.0091356	0.0099644
0.085	-0.0069266	-0.479976	0.0121786	0.122872	0.0083497	-0.0094236	0.0098401
0.09	-0.0071736	-0.477901	0.01166	0.122782	0.0081913	-0.0097006	0.0097205
0.095	-0.0074124	-0.475885	0.0111518	0.122686	0.0080403	-0.0099672	0.0096053
0.1	-0.0076435	-0.473923	0.0106538	0.122587	0.0078963	-0.0102238	0.0094941

تابع الى الجدول (6)

المتغيرات قيم K	Col_16	Col_17	Col_18	Col_19	Col_20	Col_21	Col_22
0.0	-0.42307	0.015712	0.013825	0.0062354	0.0194263	0.00656595	-0.0310723
0.005	-0.425907	0.015771	0.014259	0.0066351	0.0191034	0.00548875	-0.0307238
0.01	-0.428098	0.015841	0.01463	0.0070517	0.0188833	0.00447364	-0.0304044
0.015	-0.429782	0.015921	0.014950	0.0074785	0.0187419	0.00351061	-0.030111
0.02	-0.431062	0.016009	0.015230	0.0079112	0.018662	0.00259233	-0.0298406
0.025	-0.432014	0.016102	0.015476	0.0083463	0.018631	0.00171324	-0.0295904
0.03	-0.432698	0.016201	0.015694	0.0087816	0.0186395	0.00086899	-0.0293581
0.035	-0.433159	0.016303	0.015887	0.0092153	0.0186802	0.00005614	-0.0291416
0.04	-0.433434	0.016408	0.016059	0.0096460	0.0187473	-0.0007281	-0.0289389
0.045	-0.433553	0.016515	0.016214	0.0100727	0.0188364	-0.00148622	-0.0287487
0.05	-0.433538	0.016624	0.016352	0.0104946	0.0189437	-0.00222004	-0.0285696
0.055	-0.433409	0.016733	0.016476	0.0109112	0.0190662	-0.00293128	-0.0284004
0.06	-0.433183	0.016844	0.016588	0.0113219	0.0192015	-0.00362142	-0.0282401
0.065	-0.432872	0.016954	0.016688	0.0117264	0.0193476	-0.00429171	-0.0280878
0.07	-0.432488	0.017064	0.016779	0.0121246	0.0195027	-0.00494327	-0.0279429
0.075	-0.43204	0.017174	0.016860	0.0125161	0.0196653	-0.00557712	-0.0278047
0.08	-0.431537	0.017283	0.016934	0.0129009	0.0198342	-0.00619414	-0.0276725
0.085	-0.430985	0.017391	0.016999	0.0132789	0.0200083	-0.00679513	-0.0275458
0.09	-0.43039	0.017499	0.017059	0.01365	0.0201868	-0.00738084	-0.0274243
0.095	-0.429757	0.017605	0.017112	0.0140143	0.0203688	-0.00795193	-0.0273075
0.1	-0.429091	0.017710	0.017159	0.0143717	0.0205537	-0.00850902	-0.0271949

ويلاحظ ان المعلمات المقدرة بطريقة انحدار الحرف لقيم k اعلاه في الجدول رقم (6) تتغير بتغير قيم k وتكون قيمة k المثلى هي عندما تكون قيم المعاملات متقاربة وقريبة من بعضها ويلاحظ ذلك في قيمة K=0.45 مما يؤكد على ان العلاقة بين المتغيرات التوضيحية قريبة من التعامد أي اصبحت مشكلة التعدد الخطي غير كبيرة.

بعدما تم تعين معاملات المتغيرات المستقلة لكل قيمة من قيم K ومعرفة القيمة المثلى ل K يمكننا معرفة معامل التحديد لكل قيمة من قيم K وحسب الجدول الاتي رقم (7).

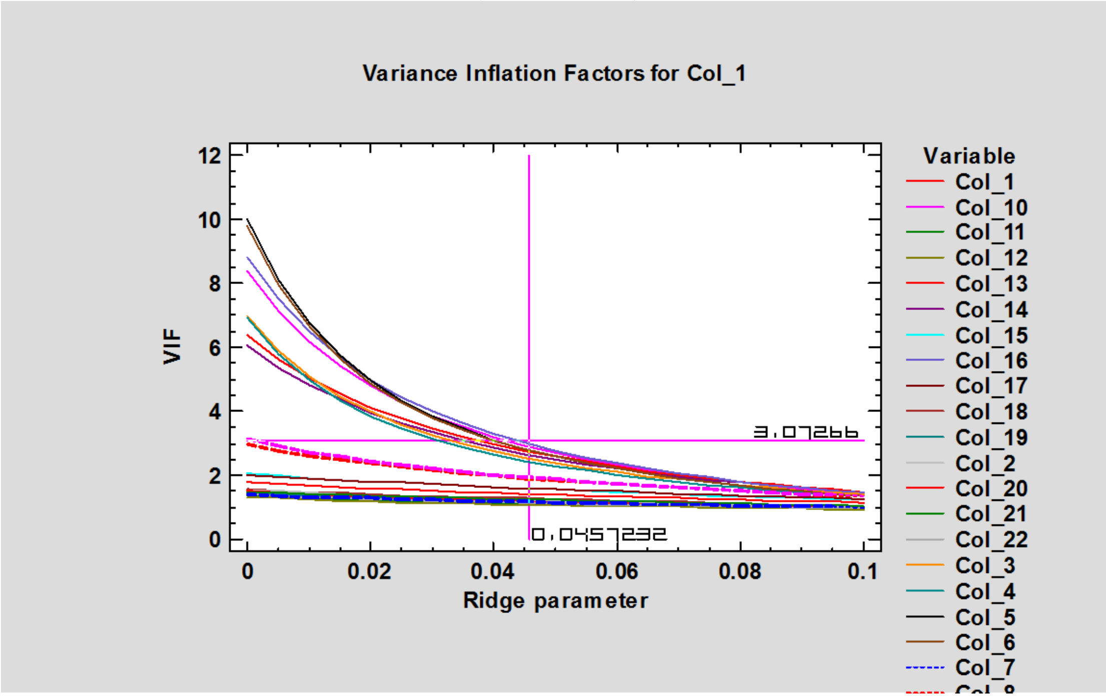
جدول رقم (7) يبين قيمة معامل التحديد لقيم K			
قيم K	معامل التحديد R^2	قيم K	معامل التحديد R^2
0.00	94.04	0.05	89.84
0.005	93.59	0.055	89.45
0.01	93.15	0.06	89.06
0.015	92.72	0.065	88.68
0.02	92.29	0.07	88.30
0.025	91.87	0.075	87.93
0.03	91.45	0.08	87.56
0.035	91.04	0.085	87.19
0.04	90.64	0.09	86.83
0.045	90.24	0.095	86.47
0.05	89.84	0.1	86.12

ان نقصان R^2 يكون مستمر بزيادة قيم k ولكن يكون التغير قليلاً عندما تكون k مناسبة في تقدير المعلمات بطريقة الحرف، ان ذلك يؤكد ان قيمة k المثلى في اثر الحرف هي K=0.045.
وبالخلاصة نستخرج قيم المعاملات القياسية لطريقة انحدار الحرف عند اختيار قيمة K مساوية الى (0.045) وحسب الجدول التالي (8)

جدول (8) يبين قيم معاملات المتغيرات المستقلة عند K=0.045			
المتغيرات	المعاملات القياسية	المتغيرات	المعاملات القياسية
C1	0.0061	C12	0.1227
C2	0.0038	C13	0.0113
C3	0.0691	C14	-0.0038
C4	0.0500	C15	0.0127
C5	0.0050	C16	-0.4257
C6	0.0369	C17	0.0158
C7	-0.0185	C18	0.0142
C8	-0.0182	C19	0.0066
C9	0.0002	C20	0.0191
C10	-0.5313	C21	0.0056
C11	0.0220	C22	-0.0308

ويمكن توضيح طريقة اثر الحرف بالرسم بالاعتماد على قيم VIF

الشكل (1) يبين قيم VIF باستخدام اسلوب اثر الحرف



نلاحظ من الشكل اعلاه ان طريقة اثر الحرف من خلال الرسم تحدد قيمة K المثلى هي 0.045 وقيمة معامل التباين VIF تساوي 3.07 وهذه القيمة اقل من (5) نستنتج ان مشكلة التعدد الخطي قد اختفت .

جدول (9) يبين قيم ال VIF لطريقة انحدار الحرف			
المتغيرات	VIF	المتغيرات	VIF
الجنس	1.770	نوع الجلطة	1.339
العمر	1.561	عامل الوراثة	*6.411
انواع علاج الضغط	*6.989	درجة القرابة	*6.201
وجود الضغط	*6.908	نوع العمل	2.046
انواع علاج السكر	*10.023	طبيعة مزاج المريض	*8.831
وجود السكر	*9.770	المستوى التعليمي	2.257
تكرارحصول الجلطة	1.486	عدد ساعات النوم الليل	1.589
وقت حصول الجلطة	2.963	عدد ساعات النوم النهار	1.456
وقت مراجعة الطبيب	3.131	تناول الادوية	1.543
حصول الخثرة	*8.409	ممارسة النشاط الرياضة	1.530
ساعة الاستيقاظ صباحا	1.347	وجود امراض اخرى	1.394

مقارنة طريقة انحدار الحرف مع طريقة المربعات الصغرى الاعتيادية بالاعتماد على قيم VIF:

نلاحظ من الجدول التالي رقم (10) ان قيم ال VIF قد انخفضت بطريقة انحدار الحرف مما كانت عليه بطريقة المربعات الصغرى الاعتيادية وهذا يدل على ان طريقة انحدار الحرف افضل من طريقة المربعات الصغرى الاعتيادية.

جدول (10) يبين الفرق بين قيم VIF لكل طريقة

المتغيرات المستقلة	OLS	RR
الجنس	1.770	1.44271
العمر	1.561	1.16589
انواع علاج الضغط	*6.989	1.16411
وجود الضغط	*6.908	1.10409
انواع علاج السكر	*10.023	2.81879
وجود السكر	*9.770	2.71205
تكرارحصول الجلطة	1.486	1.5896
وقت حصول الجلطة	2.963	1.18571
وقت مراجعة الطبيب	3.131	1.60025
حصول الخثرة	*8.409	1.2747
ساعة الاستيقاظ صباحا	1.347	1.1958
نوع الجلطة	1.339	1.29584
عامل الوراثة	*6.411	1.19137
درجة القرابة	*6.201	1.25721
نوع العمل	2.046	1.13288
طبيعة مزاج المريض	*8.831	2.60694
المستوى التعليمي	2.257	2.52997
عدد ساعات النوم الليل	1.589	2.85753
عدد ساعات النوم النهار	1.456	2.8091
تناول الادوية	1.543	1.16777
ممارسة النشاط الرياضة	1.530	1.90493
وجود امراض اخرى	1.394	1.96794

الاستنتاجات :

1. أستنتج ان مشكلة التعدد الخطي تؤثر على معنوية بعض المتغيرات المستقلة و تظهر عدم معنويتها رغم اهميتها .
2. هناك عدة اختبارات ومعايير للكشف عن وجود مشكلة التعدد الخطي بين المتغيرات التفسيرية واسلوب انحدار الحرف واحد من تلك المعايير ويعد اسلوب سهل الاستيعاب والتطبيق .
3. تبين من خلال النتائج ان طريقة انحدار الحرف افضل من طريقة المربعات الصغرى الاعتيادية في معالجة المشكلة .

التوصيات :

1. إعطاء موضوع أمراض الجلطة إهتماماً أكبر ، لكثرة انتشاره بين افراد المجتمع ، بحيث تكون هناك مستشفيات خاصة بذلك وتوفير الكوادر الطبية اللازمة .
2. يوصى بالمراجعة الدورية الى المستشفيات للفحص والتأكد من سلامة أبداننا ،وأخذ الحيطة والحذر عند ظهور أحد الاعراض التي ذكرت.
3. ضرورة استخدام الاختبارات الاحصائية للكشف عن وجود مشكلة التعدد الخطي .
4. يوصى باتباع احدى الطريقتين المقترحتين لتحيد مدى لقيمة معلمة التحيز.
5. يفضل عند استعمال طريقة انحدار الحرف الاعتماد على طرق تقدير الحرف **K (K LW, K HKB)** لانها تعطي نتائج أدق من طريقة أثر الحرف في حال زاد عدد المتغيرات التفسيرية أكثر من اربعة متغيرات.

المصادر :

1- ابو حامد ،سمير "الجلطة الدماغية " الطبعة الاولى ،دمشق خطوط النشر والتوزيع ،2009.

2- الراوي ،خاشع محمود ،"المدخل الى تحليل الانحدار " مديرية دار الكتب للطباعة والنشر ، جامعة الموصل – العراق ،1987

3- السيفو ،وليد اسماعيل ،"المدخل الى الاقتصاد القياسي "وزارة التعليم العالي والبحث العلمي ،1988**.**

4-العزاوي ،دجلة ابراهيم مهدي مع نذير عباس ابراعيم "الاقتصاد القياسي اسلوب كمي باستعمال SPSS Minitab,Eviews، الطبعة الاولى 2014.

5- النعيمي ،اسوان محمد طيب ،"اختيار المتغيرات في انحدار الحرف "رسالة ماجستير كلية الادارة والاقتصاد ،جامعة الموصل **.**2005

6- محمد ،محمد سليمان ،"التعدد الخطي اسبابه تأثيراتة والمعالجة بانحدار الحافة وانحدار المركبات الرئيسية مع التطبيق على بيانات افتراضية " اطروحة دكتوراه مقدمة الى كلية العلوم والتكنلوجيا جامعة السودان 2014.

7- Alin, A., Multicollinearity. Wiley Interdisciplinary Reviews: Computational Statistics,. **2**(3): p. 370-374(2010).

8- Hoerl, A.E. and R.W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems. Technometrics,. 12(1): p. 55-67(1970).

9- Jacquelynne R.King and Donald A.Jackson, ” variable selection in large environmental data sets using principal component analysis “,www.zoo.utoronto.ca/jackson /king%20&%20jacson.pdf, (1999).

10- Kibria, B.G ,.Performance of some new ridge regression estimators. Communications in Statistics-Simulation and Computation, 32(2): p. 419-435(2003).

11-M. El-Dereny and N. I. Rashwan, " Solving Multicollinearity Problem Using Ridge Regression Models" Vol.6,No.12.(2011).