Comparison of the two methods ordinary least squares (OLS) and M robust (Tukey's biweight) of estimating the parameters of the multiple linear regression model with practical application

مقارنة بين طريقتي المربعات الصغرى الاعتيادية (OLS) وطريقة M الحصينة (Tukey's biweight) في تقدير معالم أنموذج الانحدار الخطي المتعدد مع تطبيق عملي

أ.م.د شروق عبد الرضا سعيدالسباح سيف حسين علي
جامعة كربلاء /قسم الإحصاء / كلية الإدارة والاقتصاد
بحث مستل من رسالة ماجستير في الاحصاء

المستخلص

تم في هذا البحث دراسة طريقتين لتقدير معالم أنموذج الانحدار المتعدد واحدة من الطرائق الكلاسيكية وطريقة أخرى من الطرائق حصينة وتطرقنا إلى الطريقة الحصينة لبيان أثر وجود قيم شاذة في قيم مشاهدات اذ تم سحب عينة عشوائية بسيطة حجمها (151) شخص مصاب بالعقم متمثلا بالمتغير المعتمد (حجم البيضة) والمتغيرات المستقلة (العمر ، الوزن ، نوع العقم ، هرمون LH ، هرمون FSH ، هرمون PL) بهدف أجراء المقارنة بين تلك الطرق وبيان الطريقة الأفضل في التقدير باستعمال مقاييس تعكس جودة تلك المقدرات منها معيار متوسط مربعات الخطأ MSE ومعيار معامل التحديد R^2 لقياس كفاءة الأنموذج ، ومن خلال البرنامج الإحصائي Stata توصلنا إلى أن أفضل طريقة للتقدير كانت طريقة M الحصينة (Tukey's biweight) .

Abstract

In this thesis study two methods have been estimated the parameters of robust multiple regression model ، one method of classical methods ، and another of robust methods ، Because of the presence of effect of extreme values , It has been drawn simple random sample (151) patient infected with infertility represented by dependent variables (egg size) and independent variables (age , weight , infertility kind , L.H hormone , F.S.H. hormone , P.L. hormone) .

to aim compare between these methods in order to find the best method for estimation by using measures reflect the quality and efficiency of those estimates , such as mean square error (MSE) and determination facter (R^2) to model efficiency ، by using statistical program (stata) the results show the method of **M Tukey's biweight** is the best .

1- منهجية البحث

1-1 المقدمة

تعتمد الطرائق الإحصائية على مجموعة من الفروض المهمة للحصول على نموذج انحدار دقيق وإن التوزيع الاحتمالي للبيانات يعد أهم الفروض التي هي قيد الدراسة والتي هي غالبا تتوزع توزيعا طبيعيا .

وفي بعض الأحيان تأخذ البيانات الموزعة نمطا مختلفا وقد لا تمثل بنمط معين من التوزيعات والسبب قد يعود أحيانا إلى وجود القيم الشاذة (out liers) وهو الأمر الذي يؤدي إلى خلل في فروض المربعات الصغرى وعندها ستفقد خصائصها الجيدة وعليه نقوم بإيجاد طرائق بديلة حصينة لمعالجة هذه المشكلة وتكون غير حساسة اتجاه وجود القيم الشاذة وتعطينا مقدرات كفوءة.

2-1 مشكلة البحث

أن وجود القيم الشاذة تؤثر على توزيع بواقي النموذج وسيكون توزيعا ملتويا وعليه يكون مخالفا لشروط طريقة المربعات الصغرى الاعتيادية لذلك يتم البحث عن طرائق بديلة للتقدير ونظرا لشيوع ظاهرة العقم في المجتمع تناول البحث المتغيرات المستقلة (العمر ، الوزن ، نوع العقم ، هرمون LH ، هرمون FSH ، هرمون PL) وأثرها على متغير الاستجابة (حجم البيضة) .

3-1 هدف البحث

مقارنة طريقة الانحدار البديلة (الحصينة) لدراسة تأثير القيم المتطرفة في نموذج الانحدار الخطي المتعدد وهي (M) الحصينة مع طريقة المربعات الصغرى التقليدية (OLS) .

4-1 فروض البحث

تفرض الرسالة وجود فروق معنوية بين المتغير المعتمد (حجم البيضة) والمتغيرات المستقلة (العمر ، الوزن ، نوع العقم ، هرمون LH ، هرمون FSH ، هرمون PL) في حالة وجود القيم الشاذة واستعمال الطرائق الحصينة في التقدير .

5-1 منهجية البحث

أعد البحث ضمن المنهج الاستقرائي حيث تم سحب العينة ودراستها وتحليل النتائج وتعميمها على مجتمع الدراسة .

الجانب النظري

1-2 تمهيد

سيتم في هذا الجانب التطرق إلى طريقة التقدير الكلاسيكية وهي (طريقة المربعات الصغرى الاعتيادية OLS) ونقارنها مع طريقة التقدير الحصينة وهي (M) لإيجاد أفضل تقدير معالم أنموذج الانحدار المتعدد .

2-2 طريقة التقدير الكلاسيكية [5] [6]

توجد عدة طرق لتقدير معالم نموذج الانحدار وان هذه الطرق تعطي نتائج جيدة عند وجود فروض الانحدار ولكن عند الإخلال بأحد هذه الفروض أو عند وجود قيم شاذة فأن هذه الطرق ستفقد مزاياها وستؤثر سلبا على نتائج التقدير ومن الطرق الكلاسيكية الجديرة بالذكر :

3-2 طريقة المربعات الصغرى الاعتيادية OLS [2]

وهي من الطرق المهمة في الانحدار، حيث مشكلة تحديد خط مستقيم لعدد من المشاهدات معتمدة على تحديد الخطأ العشوائي (Ui)

$$Yi = \beta_0 + \beta_1 Xi + Ui$$

ويدعى في بعض الاحيان بالعنصر الإضرابي (Disturbances Terms) ويدعى بهذه التسمية لأن بسببه يكون اضطراب بالعلاقة الخطية بين كلآ من المتغير المعتمد والمتغيرات المستقلة وبالتالي سنقسم العلاقة اعلاه إلى قسمين ، حيث يكون القسم الأول $(\beta_0 + \beta_1 Xi)$ وانه يدعى بالمتغير التوضيحي (Explained Variation) في حين يكون القسم الثاني هو المتغير العشوائي (Ui) ويدعى بالمتغيرات غير الموضحة (Unexplained Variation) وإنها تكون ناتجة من انحراف القيمة التقديرية عن قيمتها الحقيقة للمتغير المعتمد ، بسبب بعض العوامل هي :-

$$\therefore b_{OLS} = (\mathrm{X'X})^{-1}\mathrm{X'}\underline{\mathrm{Y}} \quad \ldots(1)$$

4-2 طرق التقدير الحصينة [8]

إن إتباع الطرق الكلاسيكية في تقدير معالم النموذج تكون غير دقيقة في تحليل البيانات عند وجود خلل في إحدى فروض الانحدار أو وجود القيم الشاذة أو توزيع الخطأ العشوائي توزيعا غير التوزيع الطبيعي الذي يناسب الطريقة المعتمدة في التقدير وإن وجود قيم شاذة واحدة أو أكثر ستؤدي إلى خلل خصائص مقدرات المربعات الصغرى وإن المقدر الحصين هو الذي يتفق بأنه يحافظ على الخصائص المرغوب بها للمقدرات عند انتهاك بعض فرضيات النموذج وسنتطرق إلى بعض طرق التقدير الحصينة الآتية :

5-2 طريقة M الحصينة[10] [13] [14] [12]

وهي إحدى طرق التقدير لانحدار الحصين وأن الحرف (M) يعني هي طريقة تقدير الإمكان الأعظم (Maximum Liklihood) وأن المقدر هنا بطريقة M هو

$$\hat{\beta} = [\beta_n(x_1, x_2, \dots, x_n)] = \beta \dots (2)$$

وبأخذ التوقع سيكون

$$E[\beta_n(x_1, x_2, \dots, x_n)] = \beta$$

ومن المعادلة رقم (2) تبين إن المقدر

$$\hat{\beta} = [\beta_n(x_1, x_2, \dots, x_n)]$$

هو تقدير غير متحيز والتباين أقل ما يمكن وكذلك مقدر M صغير جدا بالنسبة للتباين وأقل ما يمكن مقارنة بتباين بقية المقدرات

$$var\left(\hat{\hat{\beta}}\right) \geq \frac{[\hat{\beta}.']^2}{nE[(\frac{\partial}{\partial\beta} lnf(x_i;\beta)]^2}$$

إذ أن $\left(\hat{\hat{\beta}}\right)$ تقدير خطي آخر غير متحيز ل β

أن مبدأ طريقة M في التقدير هو تقليل أو تصغير دالة البواقي P :

$$\hat{\beta}_m = min_\beta P(y_i - \sum_{j=0}^{k} x_{ij}\beta_j) \dots (3)$$

علينا إن نحل

$$min_\beta \sum_{i=1}^{n} P(u_i) = min_\beta \sum_{i=1}^{n} P\left(\frac{e_i}{\sigma}\right)$$

$$= min_\beta \sum_{i=1}^{n} P\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{\sigma}\right)$$

وللحصول على معادلة رقم (3) ، وضعنا مقدر لـ [11] σ

$$\hat{\sigma} = \frac{\text{MAD}}{0.6745} = \frac{median|ei - median(ei)|}{0.6745}$$

بالنسبة للدالة ρ تستخدم Tukeys biasquare دالة الهدف

$$p(ui) = \begin{cases} \frac{u_i^2}{2} - \frac{u_i^4}{2c^4} + \frac{u_i^6}{6c^4}, |u_i| \le c \\ \frac{c^2}{6}, |u_i| > c \end{cases}$$

وعلاوة على ذلك نحن نبحث عن أول مشتق جزئي $\hat{\beta}_m$ لـ β لذلك

$$\sum_{i=1}^{n} x_{ij}\, \psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta}{\hat{\sigma}}\right) = 0 \ldots (4) \qquad , j = 0,1,2,3, \ldots, k$$

أذ أن $\psi = p'$، x_{ij} هي الملاحظة الأولى على المتغير المستقل $\sum_{i=1}^{n} x_{i0} = 1$

$$w(e_i) = \frac{\psi\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta}{\hat{\sigma}}\right)}{\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta}{\hat{\sigma}}\right)} \quad .....(5)$$

$$W_i = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, |u_i| \le c \\ 0 \qquad\qquad , |u_i| > c \end{cases}$$

نحن نأخذ c=4.685 لدالة Tukeys biasquare المرجحة لذا فان المعادلة (4) تصبح

$$\sum_{i=1}^{n} x_{ij} w_i (y_i - \sum_{j=0}^{k} x_{ij}\beta) = 0 \qquad , j=0,1, \ldots, k \qquad \ldots (6)$$

في هذه الطريقة نفترض وجود مقدر أولي β و $\hat{\sigma}_i$ هو أرقاما وأعداد المعلمات ثم :

$$\sum_{i=1}^{n} x_{ij} w^0{}_i (y_i - \sum_{j=0}^{k} x_{ij}\beta^0) = 0 \qquad , j=0,1, \ldots, k \qquad \ldots(7)$$

ويمكن كتابة معادلة رقم (7) كما يلي :

$$X' W_i X\beta = X' W_i Y \qquad \ldots(8)$$

إذ إن W_i هو مصفوفة من الدرجة $n*n$ والوزن لها عناصر القطر الرئيسي وان معادلة (8) تعرف بالمربعات الصغرى الموزونة *WLS* وان الحل لهذه المعادلة يعطى تقدير لـــ β ، وبكلام اخر :

$$\hat{\beta} = (X' W_i X)^{-1} (X' W_i Y)$$

دالة *Tukey* [7]

ويسمى احيانا الوزن المزدوج (Tukey Biweight)

$$\psi(u) = \begin{cases} \mu(1-u^2)^2 & ; |u| < c \\ 0 & ; |u| \geq c \end{cases} \quad \dots (9)$$

C: ثابت يأخذ القيم ***4.685 , 6.0***

الجانب التطبيقي

6-2 مرض العقم

يعد الأطفال ديمومة الحياة والركن الأساسي في تكوين الأسرة في مجتمع لهذا تأتي الرغبة القوية من الأزواج في الإنجاب ولكن بالرغم من هذه الرغبة فان هناك عوائق عديدة تمنع عملية الحمل أهمها العوامل البيولوجية عند الأزواج والتي يتم التركيز عليها في هذه الرسالة من الناحية الطبيعية لحجم البيضة من خلال إجراء التحاليل المختبرة قد تكون موجودة أو قليلة جدا وفي بعض الحالات تكون معدومة وهي احد حالات العقم وتعد طريقة التشخيص المبكر للمشاكل التي تقف وراء انخفاض أو عدم وجود البيوض ومعالجتها يمكن إن تؤدي إلى زيادة احتمال حدوث حمل وسنقوم باستخدام الطرق الحصينة كأسلوب إحصائي مناسب لدراسة البيانات حيث تم تحديد (6) متغيرات مستقلة (العمر ، الوزن ، نوع العقم ، هرمون LH ، هرمون FSH ، هرمون PL) تؤثر على متغير الاستجابة (حجم البيضة) عينة عشوائية بسيطة مكونة من (151) شخص مصاب بالعقم وبالتالي لا يمكن إهمال أو بتر هذه النسبة من البيانات كونها تمتلك كمية من المعلومات وان إهمالها يعني خسارة كبيرة وتقدير غير دقيق.

طفل الأنابيب (الحقن المعملي) وهو أكثر التقنيات المساعدة على التناسل شيوعا ويتم خلالها جمع بويضة وحيوان منوي خارج الرحم تترك الأجنة في حاضنات خاصة لمدة من ثلاثة إلى خمسة أيام ثم يحقن في رحم المرأة ويعد الحقن المعملي أكثر التقنيات نجاحا .

7-2 وصف البيانات

تمثل عينة عشوائية بسيطة حجمها (151) شخص مصاب بالعقم في محافظة ذي قار من مستشفى الحسين التعليمي – مركز العقم حيث أخذت المتغيرات من تاريخ 1\1\2014 إلى تاريخ 1\1\2015 تتمثل بما يلي :-

المتغير المعتمد (Y) يمثل حجم البيضة .
المتغير المستقل الأول (X1) يمثل العمر .
المتغير المستقل الثاني (X2) يمثل الوزن .
المتغير المستقل الثالث (X3) يمثل نوع العقم .
المتغير المستقل الرابع (X4) يمثل هرمون .LH
المتغير المستقل الخامس (X5) يمثل هرمون .FSH
المتغير المستقل السادس (X6) يمثل هرمون .PL

8-2 حجم البيضة [4]

يختلف حجم البيضة من امرأة لأخرى لكنه في الغالب يتراوح ما بين (3-5) سنتيمتر وإن حجم المبيض له تأثير كبير على إمكانيات الحمل للسيدات اللاتي لديهن مبيض صغير يحدث عندهن مشكلة العقم أو تأخر الحمل ، حيث حجم البيضة لها تأثير على الحمل سواء كانت البويضات صغيرة أو كبيرة ، وإن حجم البيضة المناسب للتلقيح تبلغ ما بين (18-25) ملليمتر، وإن حجم البيضة القابلة للحمل بتوأم تبلغ (21) ملليمتر، في كثير من الأحيان ينتج الحمل من خلال تلقيح بيضة واحدة فقط .

وإن الحمل بتوأم قد يحدث من خلال تلقيح بيضتين أو بيضة واحدة ويحدث لها تلقيح مرتين ، مما يؤدي إلى الحمل بتوأم مشابه .

وقد يكون حجم البيضة أصغر أو أكبر من الحجم المناسب للتلقيح فأن ذلك يؤدي إلى بعض المشاكل والتي تنتج عن إجهاض الجنين ، في حين قد يكون حجم البيضة أكبر من حجم المبيض مما يؤدي إلى صعوبة إخراجها من المبيض ويصعب الحمل ، في حين آخر تكون حجم البيضة أصغر من حجم المبيض ولكنها أصغر من المعدل الطبيعي للحمل فيؤدي ذلك إلى حدوث تأخر الحمل .

9-2 العمر[9]

تعد الفئة العمرية (18-28) هو السن الأكثر خصوبة لدى المرأة إذا تبدأ الخصوبة بالتراجع قليلا بعد سن (30) ويزداد التراجع والانحدار بعد سن (35) إن حوالي ثلث النساء اللواتي يؤجلن الحمل حتى منتصف الثلاثينات سوف يعانين من مشاكل عدم الإنجاب ولقد دون في العهود الحديثة اكبر عمر لامرأة حدث الحمل هو (57) سنة .

10-2 الوزن [9]

إن الوزن الزائد يؤثر على الإباضة مما يجعلها غير منتظمة وبالتالي يؤدي إلى نقص في معدل الخصوبة وذلك بسبب تأثير الوزن على مستويات الهرمونات فبالإضافة إلى البروجسترون والأستروجين وفي الغالب ينظر عليها بأنها هرمونات ذكرية تدعى الأندروجينات حيث إنها تشارك بالحيض والإباضة ، وإن السمنة الزائدة تسبب خلل في توازن هذه الهرمونات بالتالي ينتج الجسم الكثير من الاندروجين حيث لا يمكن إطلاق الهرمونات اللازمة من الغدة النخامية الهامة لإنتاج البويضات من خلال المبيضين ، إذ ينتج مرض يدعى متلازمة المبيض المتعدد الكيسات وهو أوسع انتشارا لدى النساء البدينات .

11-2 أنواع العقم

1-11-2 العقم الأولي

هو العقم الذي يصيب المرأة منذ بداية حياتها الزوجية تعود أسباب العقم الأولي عادة إلى أمراض غددية أو هرمونية أو عدم النضوج الكامل للمرأة لأسباب التكوين .

2-11-2 العقم الثانوي

هو العقم الذي يصيب المرأة بعد الإنجاب لطفل أو طفلين أو بعد إجراء عملية إجهاض لها ينجم العقم من مضاعفات الولادة أو الإجهاض وجميع الالتهابات التي تصيب الرحم وتؤدي إلى رفع نسبة العقم .

12-2 هرمون LH (Luteinizing Hormone) [3]

هو هرمون الملوتن يفرز من الغدة النخامية ومسؤول عن إنتاج الأجسام الصفراء وكذلك يتحكم بالجهاز التناسلي عند الرجال والنساء يعتبر مهما بتنظيم وظيفة الخصيتين لدى الرجال والمبيضين لدى النساء ، بالإضافة إلى دوره الفعال في عمليتي الإباضة والحيض حيث يحفز المبايض لإنتاج هرمون الإستراديول وبالتالي يزداد إنتاج هذا الهرمون في وقت قليل خلال اليوم (14) من الدورة الشهرية وعن هذه الدورة تتجدد كل (28) يوما ، وعندما يتدفق هذا الهرمون فالمبايض تنتج بويضة عند عملية الإباضة .

إن مقدار هرمون الملوتن والهرمون المنبه للجريب (FSH) ينخفضان ويرتفعان عند الدورة الشهرية في أوقات محددة منها ، وعند حدوث الإخصاب فهرمون الملوتن ينبه الجسم الأصفر ويحفزه لإنتاج البروجستيرون للحفاظ على الحمل .

يوجد هرمون LH لدى الرجال في خلايا تقع في الخصية تحفز على إنتاج (التستوستيرون) كذلك يساعد على إنتاج الحيوانات المنوية وإن (التستوستيرون) بدوره يحفز على إنتاج الحيوانات المنوية وكذلك يظهر السمات الذكرية عند الرجال مثل شدة الصوت وظهور شعر الوجه .

13-2 هرمون FSH (Follicle Stimulating Hormone) [3]

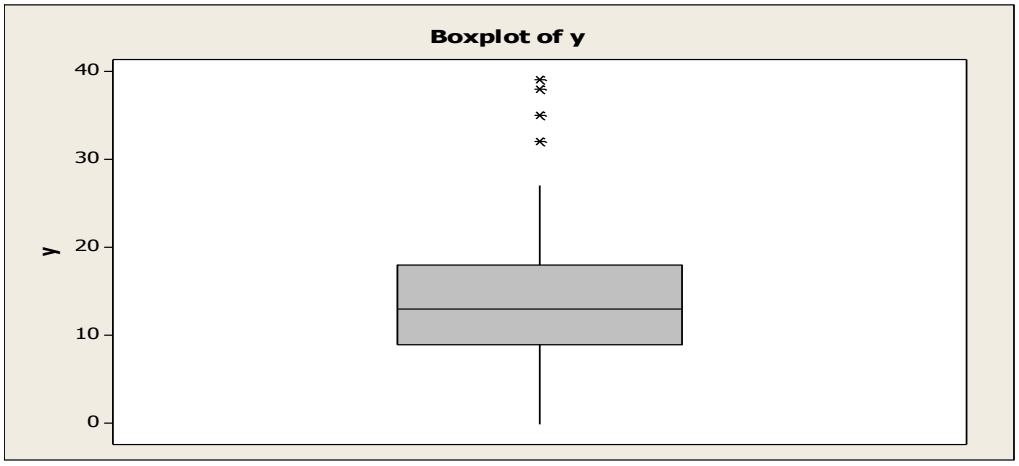
هو منشط للحوصلة حيث يعد أحد الهرمونات التناسلية السكرية ويتكون من (بيتا وألفا) وينظم عمل الغدد التناسلية عند النساء والرجال والذي يفرز من قبل الغدة النخامية الأمامية وذلك لأطلاق غدة تحت المهاد للهرمون المطلق لموجه الغدد التناسلية وإن هذا الهرمون تختلف وظيفته عند الرجال والنساء حيث يعمل عند النساء بتنظيم الدورة الشهرية وإنتاج البويضات في المبيض ، أما عند الرجال يعمل هذا الهرمون على تنظيم إنتاج الحيوانات المنوية وأن مستوياته تكون ثابتة لديهم ، بينما تختلف مستويات هرمون FSH عند النساء ويعود السبب لاختلاف المرحلة عند الدورة الشهرية حيث تصل أعلى ما يمكن خلال فترة الأباضة .

14-2 هرمون (يسمى بالبرولاكتين PL (Prolactin [1]

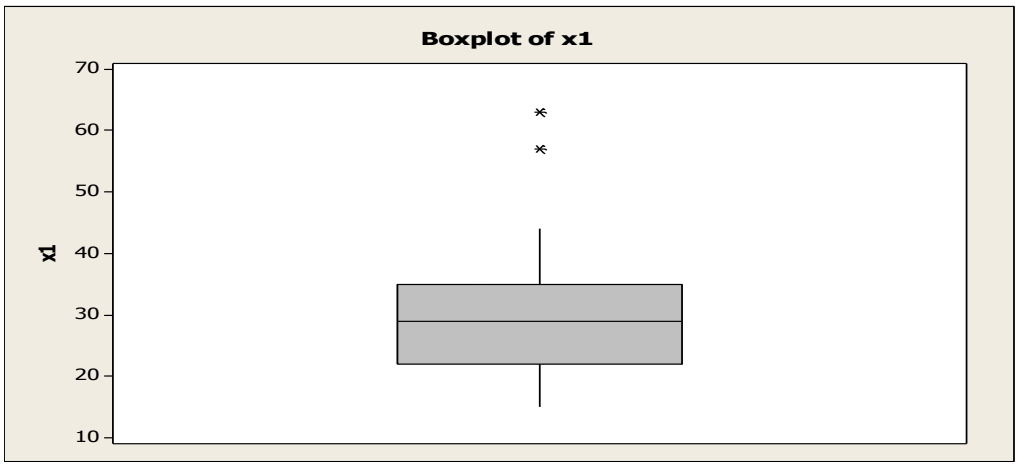
هو هرمون الحليب (البرولاكتين) وهو جزء بروتيني يفرز من قبل الغدة النخامية في الثديات وبدوره يساعد على إنتاج الحليب الذي يوجد في الغدد الثديية و كذلك يساعد على رعاية النسل ويدعى بهرمون الأمومة والأبوة الذي يصنع ويفرز من خلال غدة تدعى (غدة لاكتوتروفس) وهذه الغدة تشكل نسبة (20%) من الغدة النخامية الأمامية .

الجانب العملي :

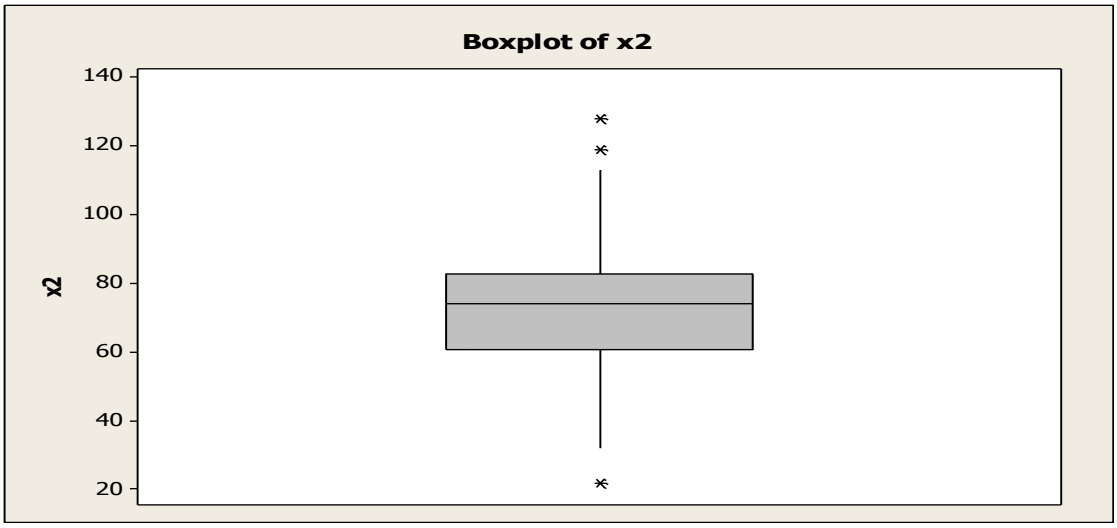
الكشف عن المشاهدة الشاذة بواسطة طريقة **Box plot**



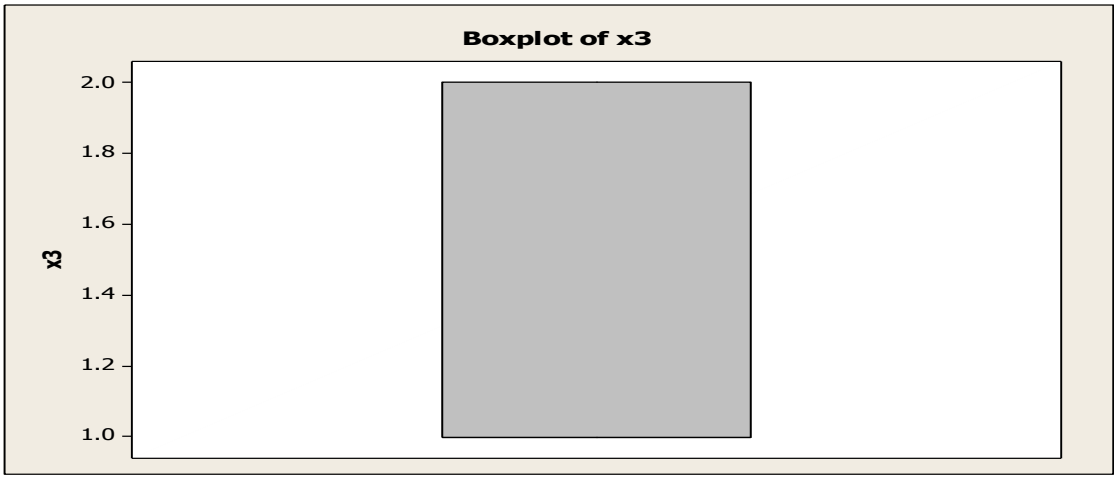
شكل رقم -1- يمثل القيم الشاذة في متغير الاستجابة Y (حجم البيضة)



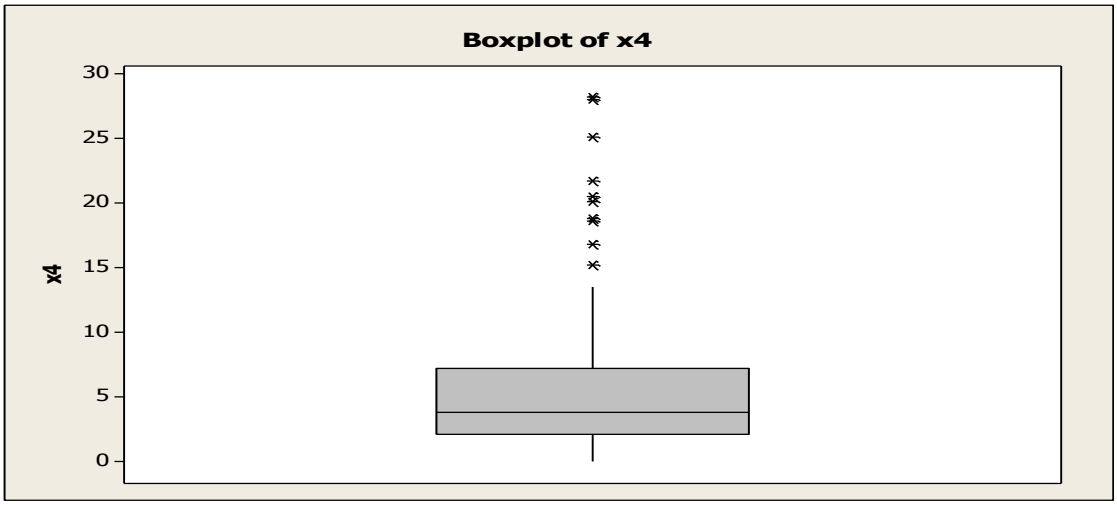
شكل رقم -2- يمثل القيم الشاذة في المتغير المستقل X1 العمر



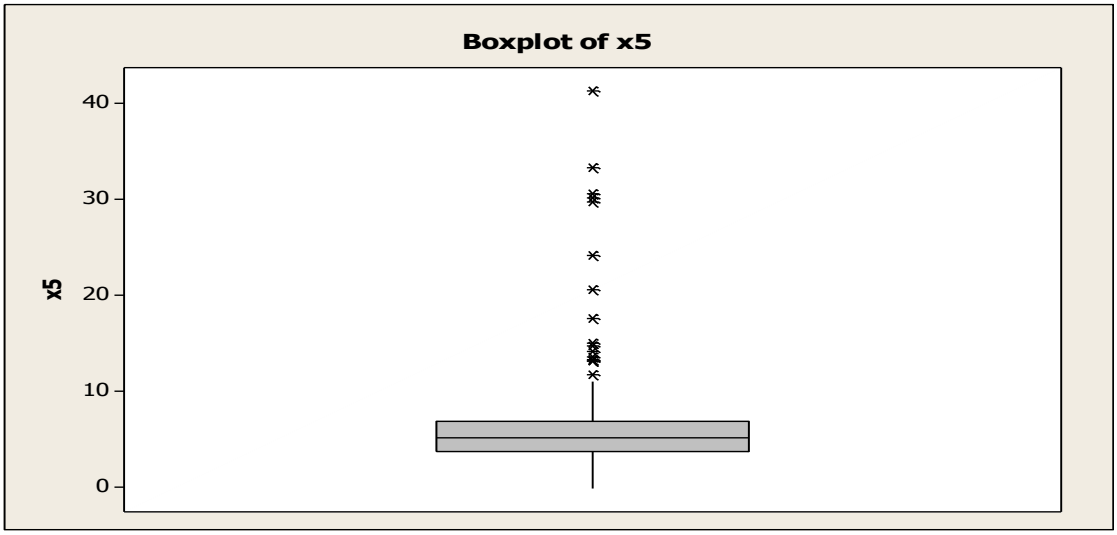
شكل رقم -3- يمثل القيم الشاذة في المتغير المستقل X2 الوزن



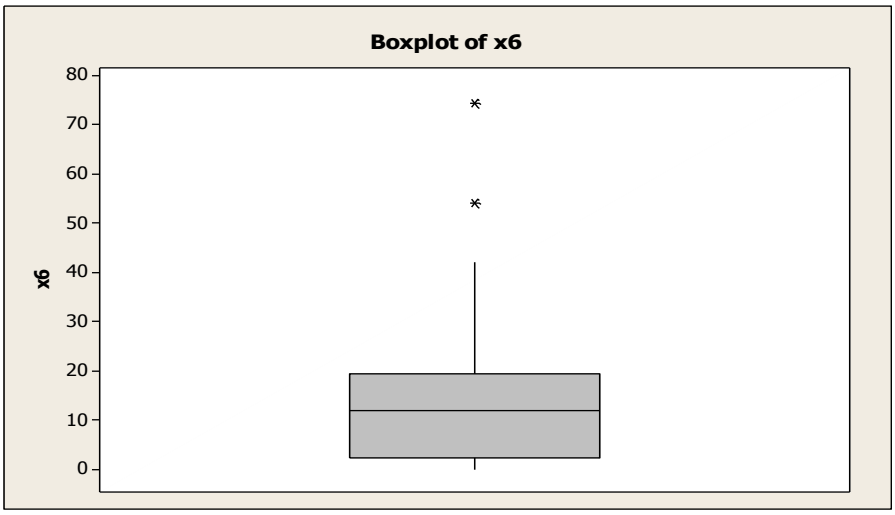
شكل رقم -4- يمثل القيم الشاذة في المتغير المستقل X3 نوع العقم



شكل رقم -5- يمثل القيم الشاذة في المتغير المستقل X4 (هرمون LH)



شكل رقم -6- يمثل القيم الشاذة في المتغير المستقل X5 (هرمون FSH)



شكل رقم -7- يمثل القيم الشاذة في المتغير المستقل X6 (هرمون PL)

15-2 طريقة المربعات الصغرى التقليدية OLS

تعد هذه الطريقة من الطرق المهمة في تقدير معلمات أنموذج الانحدار المقدر ، ويهدف اختبار فرضية العدم :

$H_0 : B_0 = B_1 = B_2 = B_3 = B_4 = B_5 = B_6$

$H_1 : at\ least\ ar\ not\ equale$

وباستعمال البرنامج الإحصائي Stata تم تقدير أنموذج الانحدار أي تقدير واختبار معلمات الانحدار المتعدد باستعمال OLS عرضت النتائج في الجدول رقم (2-1) التالي

جدول رقم (2-1) يبين قيم معاملات الانحدار وفق طريقة OLS

المتغيرات المستقلة	معاملات الانحدار (b)	الخطأ المعياري	T	$P-value$
(العمر X_1)	.8908185	.102171	8.72	0.000
(الوزنX_2)	.1380715	.0310694	4.44	0.000
(نوع العقمX_3)	-.4444917	.442376	-1.00	0.317
(هرمون LHX_4)	.048586	.1419864	0.34	0.733
(هرمونFSHX_5)	.2792032	.0905542	3.08	0.002
(هرمون PLX_6)	-.3389045	.0705121	-4.81	0.000
الحد الثابت	-19.6469	1.438688	-13.66	0.000

يتبين من الجدول رقم (2 -3) إن قيمة معامل التحديد قد بلغت (0.9699) وهي قيمة تفسيرية عالية تعكس جودة تدقيق النموذج وتبين إن متغيرات النموذج الستة قد فسرت ما نسبته (0.9699) من إجمالي التغيرات الحاصلة بالمتغير التابع (حجم البيضة) وهي نسبة تفسيرية عالية جدا وقد بلغ متوسط مربعات الخطأ (1.288) وهي قيمة صغيرة تعكس أهمية النموذج في تفسير التغيرات الحاصلة بالمتغير التابع ولأجل بيان أي من المتغيرات كان مؤثرا بالمتغير التابع وسببا في معنوية النموذج تم إجراء اختبار t والذي يظهر بالعمود الرابع من الجدول رقم (2 -1) ومن ملاحظة قيم احصاءة اختبار t والقيم الاحتمالية المناظرة لكل احصاءة اختبار تبين لنا معنوية متغيرات العمر والوزن وهرمون FSH وهرمون PL وذلك لان قيمة (P-value) اصغر من (0.05) في حين ان نتائج الاختبار لن تظهر معنوية متغيرات نوع العقم وهرمون LH وذلك لكون ان القيم الاحتمالية المصاحبة لكل منهما قد بلغت (0.317 ، 0.733) على التوالي ولكون ان قيمة (P-value) لكل منهما اكبر من مستوى المعنوية (0.05) .

16-2 الانحدار الحصين طريقة M (Tukey'sbiweight)

وهي طريقة مهمة تستخدم لإيجاد تقدير معلمات أنموذج الانحدار المقدر ، ويهدف اختبار فرضية العدم :

$H_0 : B_0 = B_1 = B_2 = B_3 = B_4 = B_5 = B_6$

$H_1 : at\ least\ ar\ not\ equale$

وباستعمال البرنامج الاحصائي (mini tab 17) حيث كانت نتائج تقدير أنموذج الانحدار المتعدد باستعمال **M (Tukey'sbiweight)** عرضت في الجدول رقم (2-2) التالي :

جدول رقم (2-2) يبين قيم معاملات الانحدار بطريقة M (Tukey's biweight)

المتغيرات المستقلة	معاملات الانحدار (b)	الخطأ المعياري	t	$P-value$
(العمر X_1	-0.2544904	0.05126107	-4.9646	0.000003
(الوزنX_2)	0.2271079	0.01905513	11.9185	0.000000
(نوع العقمX_3)	-0.120593	0.1369224	-0.8807	0.380636
(هرمون LHX_4)	0.6092089	0.0936127	6.5078	0.000000
(هرمونFSHX_5)	-0.5324315	0.05468547	-9.7363	0.000000
(هرمون PLX_6)	0.4379432	0.02184197	20.0505	0.000000

يتبين من الجدول رقم (2-3) ان قيمة معامل التحديد بلغت (0.999653) وهي قيمة تفسيرية عالية تعكس جودة تدقيق النموذج وتبين إن متغيرات النموذج الستة قد فسرت ما نسبته (0.999653) من إجمالي التغيرات الحاصلة بالمتغير التابع (حجم البيضة) وهي نسبة تفسيرية عالية جدا وقد بلغ متوسط مربعات الخطأ (0.0551) وهي قيمة صغيرة تعكس أهمية النموذج في تفسير التغيرات الحاصلة بالمتغير التابع ولأجل بيان أي من المتغيرات كان مؤثرا بالمتغير التابع وسببا في معنوية النموذج تم إجراء اختبار t والذي يظهر بالعمود الرابع من الجدول رقم (2 -2) ومن ملاحظة قيم إحصاءه اختبار t والقيم الاحتمالية المناظرة لكل إحصاءه اختبار تبين لنا معنوية متغيرات العمر والوزن وهرمون LH وهرمون FSH وهرمون PL وذلك لان قيمة (P-value) لكل منهم اصغر من مستوى المعنوية (0.05) في حين ان نتائج الاختبار لن تظهر معنوية متغيرات نوع العقم وذلك لكون ان قيمة (P-value) المصاحبة قد بلغت (0.380636) ولكون ان قيمته اكبر من مستوى المعنوية (0.05) .

جدول رقم (2 -3) يبين معامل التحديد ومتوسط مربعات الخطأ

طرق نموذج الانحدار	معيار MSE	معيار R^2
1- طريقة المربعات الصغرى التقليدية OLS	1.288	0.9699
2- الانحدار الحصين طريقة M (Tukey's biweight)	0.0551	0.9997

للمقارنة بين طرائق التقدير تم استعمال معيار متوسط مربعات الخطأ ومعيار معامل التحديد ، نتائج التقدير لكلا المعياريين لطرائق التقدير عرضت في الجدول رقم (2-3) والذي يتبين منه أفضلية طريقة الانحدار الحصين وهي طريقة M (Tukey's biweight) وذلك لكون إن قيمة معيار متوسط مربعات الخطأ قد بلغت 0.051 وهي أقل من قيمة طريقة OLS وتعززها قيمة معيار معامل التحديد 0.9997 وهي قيمة عالية جدا اكبر من قيمة طريقة OLS

16-2 الاستنتاجات والتوصيات

اولا: الاستنتاجات

بناءا على نتائج التقديرات للطرائق المستعملة في الجانب التطبيقي تم التوصل إلى بعض من الاستنتاجات وكالاتي :

1- أفضلية طريقة M (Tukey's biweight) بناءا على معيار MSE ومعيار R^2 .

2- اظهر معامل التحديد أن استعمال الطرائق الحصينة أكفاء من الطرائق الاعتيادية في تقدير معلمات أنموذج الانحدار الخطي في حالة وجود قيم شاذة وهذا ما يتطابق مع النظرية الاحصائية .

3- أظهرت طرائق التقدير معنوية تأثير العمر في العقم في حين أظهرت هذه الطرق معنوية متغير الوزن تلاه متغير هرمون PL معنوية تأثيره في متغير الاستجابة (العقم) ثم تلاه متغير هرمون FSH ويأتي من بعده متغير LH .

ثانيا: التوصيات

بناءا على الاستنتاجات اعلاه يمكن بالتوصيات الاتية :

1- استعمال طريقة M الحصينة لتوكي في حالة كون وجود قيم شاذة في مقدرات المتغير التابع وحسب نتائج التطبيق وكذلك في حال وجود قيم ذات تأثير عالي في المتغيرات التوضيحية .

2- ضرورة اعتماد توثيق الكتروني لحالات مرضى العقم في المستشفيات والعيادات وذلك من اجل توفير قاعدة بيانات لتسهيل عمل الباحثين .

3- على المؤسسات الصحية ان تأخذ بعين الاعتبار البحوث والدراسات الاحصائية المتعلقة بالأمراض والحالات الصحية لكي تستفيد من انجازات الباحثين في تطوير العمل الصحي وبيان مدى اهمية بعض المتغيرات في الحالات الصحية .

4- نوصي الجهات الطبية بالأخذ بنظر الاعتبار على اهمية تأثير متغير العمر بالدرجة الاساس ومن ثم متغير هرمون PL ومن ثم يليه متغير الوزن في تشخيص ومعالجة حالات العقم .

المصادر

أولا : المصادر العربية

1. سناء الدويكان ، "ما هو هرمون الحليب " ، مقال نشر على موقع (موضوع) على شبكة الإنترنت العالمية على الموقع الالكتروني الآتي :- https://mawdoo3.com/2018

2 . كاظم، أموري هادي ، مسلم، باسم شليبه (2002) ،" القياس الاقتصادي المتقدم (النظرية والتطبيق)"، المكتبة الوطنية ، بغداد – العراق .

3. كفاية العبادي ، " ما هو هرمون LH " ، مقال نشر على موقع (موضوع) على شبكة الإنترنت العالمية على الموقع الالكتروني الآتي :- https://mawdoo3.com/2016

4. المرسال ، " حجم البويضة الطبيعي المناسب للحمل " ، السعودية ، مقال نشر في شبكة الإنترنت العالمية على الموقع الالكتروني الآتي :- https://www.almrsal.com/2018

ثانيا : المصادر الأجنبية

5. Bickel,D.R.,(2002). . "Robust and efficient estimation of the mode of continuous data "http://www.mcg.edu/research/biostat/bickel.html.

6 . Drapper N.R. and Smith (1981) ."Applied Regression Analysis "John Wilely and Sons .

7 . HOAGLEN, D.C., MOSTELLER, F & TUKEY, J.W. (1983)
Understanding Robust and Exploratory Data Analysis. Wiley, New York.

8. Huber,P.J. (1981) ."Robust Statistics "Wiely,New York.

9. Infertility and its Treatmen , The NEW YORK task force on life and the law 1998 , NEW YORK

10. N. R. Draper and H. Smith, Applied Regression Analysis, Third Edition, Wiley Interscience Publication, United States, 1998.

11. Rousseeuw ,P.J. and Croux,C. (1993) ."Altelnatives to the Median Absolute deviation ".J.Amer.Statist.Assn.,88,1273-1283 .

12. Y. Susanti, H. Pratiwi, and T. Liana, Application of M-estimation to Predict Paddy Production in Indonesia, presented at IndoMS International Conference on Mathematics and Its Applications (IICMA), Yogyakarta, 2009.

13. Yuliana and Y. Susanti, Estimasi M dan sifat-sifatnya pada Regresi Linear Robust, Jurnal Math-Info, 1, No. 11 (2008), 8-16.

14. Y. Susanti and H. Pratiwi, Robust Regression Model for Predicting the Soybean Production in Indonesia, Canadian Journal on Scientific and Industrial Research, 2, No. 9 (2011), 318-328.