

الفصل الأول خطة الدراسة

1-1 تمهيد:

2-1 مشكلة الدراسة:

3-1 أهمية الدراسة:

4-1 أهداف الدراسة:

5-1 فرضيات الدراسة:

6-1 منهجية الدراسة:

7-1 مجتمع وعينة الدراسة:

8-1 الدراسات السابقة:

9-1 التعليق على الدراسات السابقة:

10-1 الإطار العام للدراسة:

1-1 تمهيد:

إن نموذج تحليل الانحدار المتعدد من الأساليب الإحصائية واسعة الانتشار والاستخدام و يعتبر أسلوباً إحصائياً مهماً، حيث إنه يحدد بوضوح العلاقة بين المتغيرات المستقلة و المتغير التابع على هيئة معادلة يستدل من تقدير معالمها على أهمية وقوة واتجاه هذه العلاقة ، كما يبين تقدير الاستجابة والتنبؤ بها بما يفيد كثيراً في التخطيط واتخاذ القرارات الرصينة حولها ، لذا فان نجاح وفشل النموذج المقدر يعتمد على اختيار طريقة تقدير معالم النموذج ومن ثم على صفات المعالم التي يتم تقديرها. إن أوسع طريقة استخداماً لتقدير معالم نموذج تحليل الانحدار المتعدد هي طريقة المربعات الصغرى لما لها من خواص ومزايا جيدة تميزت بها مقدراتها ، إلا إن هذه الطريقة تكون غير كفوءة في حالة عدم تحقق إحدى فرضياتها الأساسية أو بسبب وجود القيم الشاذة التي تؤثر علي مقدراتها وتجعلها غير دقيقة وبعيدة عن الواقع. أن سلامة البيانات وتتقيتها من القيم الشاذة التي تظهر منفصلة بوضوح عن بقية القيم الأخرى في البيانات مهمة ضرورية لسلامة النتائج والاستنتاجات المبنية عليها ، فالبدائيات الأولى لموضوع القيم الشاذة كانت تتعامل مع مثل هذا النوع من المشاهدات بأن تحذف ويجرى التحليل على بقية المشاهدات حيث أن أي نقص في البيانات سيؤدي إلى نقص في المعلومات قيد البحث وبالتالي التأثير على نتائج التحليل المعتمد خصوصاً إذا كان عدد المشاهدات الإجمالي صغيراً أو عدد المشاهدات الشاذة كبيراً أو مقارباً إلى نصف عدد المشاهدات الكلي لهذا السبب فان من الأفضل تقدير تلك القيمة أو القيم الشاذة ، علىية فان دراسة القيم الشاذة المؤثرة في نموذج تحليل الانحدار الخطي المتعدد تحتاج إلي مزيد من البحث لتقليل تأثيرها علي نتائج النموذج المقدر للحصول علي نموذج ذو صفات جيدة.

2-1 مشكلة الدراسة:

تتمثل مشكلة الدراسة في أن نموذج تحليل الانحدار الخطي المتعدد يعتمد على مجموعة من الفرضيات الأساسية والضرورية للحصول على النموذج الدقيق و في حالة مخالفة إحدى هذه الفرضيات فان نتائج التحليل لم تكون دقيقة بشكل جيد ، إن إحدى هذه الفرضيات هي أن متجه الأخطاء العشوائية المستقلة يتوزع وفق التوزيع الطبيعي متعدد المتغيرات إلا أنه أحياناً قد تأخذ البيانات الموزعة نمطاً مختلفاً أو لا تتمثل بنمط معين من التوزيعات الاحتمالية ، أي مخالفة هذه الفرضية ولعل السبب في ذلك يعود أحياناً إلى وجود القيم الشاذة. كما أن وجود القيم الشاذة المؤثرة في بيانات نموذج الانحدار الخطي المتعدد يؤدي إلى أخطاء في عمليتي التنبؤ والتقدير حيث أنها قد تؤدي إلى تضخم تباين الأخطاء مما يؤدي إلى اتساع فترات الثقة وتقليل كفاءة التقدير وبالتالي فان تقدير معاملات النموذج بطريقة المربعات الصغرى الاعتيادية في ظل وجود القيم الشاذة المؤثرة تكون غير كفوءة.

3-1 أهمية الدراسة:

تتمثل أهمية الدراسة في دراسة تأثير القيم الشاذة علي معاملات نموذج تحليل الانحدار المتعدد بالتعرف على طبيعة النموذج وطرق تقديره ، وأيضاً للدراسة أهمية أخرى التعرف على القيم الشاذة وبعض طرق الكشف عنها وتشخيصها في بيانات نموذج تحليل الانحدار الخطي المتعدد باستخدام الأساليب العلمية ، وكذلك معرفة مدى تأثير وجود القيم الشاذة في البيانات على معاملات نموذج الانحدار الخطي المتعدد ، وأيضاً دراسة بعض طرق معالجة القيم الشاذة لهذا النموذج باعتبارها طرق بديلة لتقليل أثر القيم الشاذة علي نتائج النموذج المقدر. كما تعتبر هذه الدراسة من أوائل الدراسات التي تبحث في هذا المجال المهم.

1-4 أهداف الدراسة:

تهدف هذه الدراسة إلى الآتي:

1. التعرف علي نموذج تحليل الانحدار الخطي المتعدد وأهم مشكلاته.
2. التعرف على القيم الشاذة وأهم طرق كشفها و معالجتها في بيانات نموذج تحليل الانحدار الخطي المتعدد.
3. دراسة تأثير وجود ومعالجة القيم الشاذة في البيانات على معلمات نموذج تحليل الانحدار الخطي المتعدد.
4. تقدير نموذج انحدار خطي متعدد لبيانات الدراسة يمكن من خلاله التنبؤ بقيم المتغير التابع للحصول على أدق النتائج.

1-5 فرضيات الدراسة:

من خلال هذه الدراسة نسعى لاختبار الفرضيات التالية:

1. هنالك تأثير للقيم الشاذة علي قيم المتغير التابع المقدرة ومعاملات النموذج والأخطاء المعيارية.
2. متجه الأخطاء العشوائية المستقلة يتوزع توزيعاً طبيعياً متعدد المتغيرات في وجود المشاهدات أو القيم الشاذة (البواقي تتوزع توزيعاً طبيعياً).
3. طريقة الرسم الصندوقي تعطي نتائج أفضل للكشف عن القيم الشاذة في بيانات نموذج الانحدار الخطي المتعدد.
4. طريقة بواقي ستينودنت المحذوفة تعطي نتائج أفضل للكشف عن القيم الشاذة في المتغير التابع.
5. طريقة متوسط البتر تعطي نتائج أفضل من طريقة الحذف لمعالجة القيم الشاذة.

1-6 منهجية الدراسة:

تستند الدراسة إلى المنهج الوصفي القائم على وصف متغيرات الدراسة بطرق إحصائية مثل الجداول والأشكال البيانية وبعض المقاييس الوصفية كمقاييس النزعة المركزية ومقاييس التشتت للتعرف على طبيعة بيانات الدراسة ، بالإضافة إلى المنهج التحليلي الاستدلالي القائم على اختبار فرضيات الدراسة باستخدام بعض الأساليب الإحصائية ، كأسلوب نموذج تحليل الانحدار الخطي المتعدد ، بالإضافة إلى بعض الطرق والاختبارات الخاصة بتقدير وتشخيص ومعالجة القيم الشاذة، كما تمت معالجة بيانات الدراسة باستخدام بعض البرامج الإحصائية المناسبة بواسطة الحاسوب مثل برنامج التحليل الإحصائي SPSS14 ، Eviews7.

1-7 مجتمع وعينة الدراسة:

يتمثل مجتمع الدراسة في جميع طلاب الدورة 29 الذين تم قبولهم بكلية التربية رفاعة جامعة الجزيرة للعام 2007م حيث تم اختيار عينة الدراسة من مجتمع الدراسة بحجم عينة بلغ (30) طالب وطالبة تم سحبها عشوائيا بطريقة المعاينة العشوائية المنتظمة.

1-8 الدراسات السابقة:

هنالك العديد من الدراسات والبحوث السابقة التي تناولت بالبحث المشاهدات الشاذة وطرق الكشف عنها ومدى تأثيرها علي نتائج التحليل وكذلك طرق معالجتها باستخدام الطرق والأساليب العلمية الحديثة ، ومن الأهمية بمكان الوقوف علي الدراسات والبحوث التي أجريت في هذا المجال للاسترشاد بنتائجها ومقارنتها بنتائج الدراسة الحالية ومحاولة الإضافة. وفيما يلي بعض الدراسات والبحوث العلمية السابقة التي لها علاقة أصيلة بموضوع الدراسة الحالية و تأثير القيم الشاذة علي نتائج تحليل الانحدار الخطي:

عنوان الدراسة: طرق بديلة لكشف وتقليل أثر المشاهدات الشاذة علي نموذج الانحدار الخطي

البسيط:

هدفت هذه الدراسة إلي التوصل لطريقة أو طرق جديدة يمكن أن يعتمد عليها الباحثين في الكشف وتقليل أثر المشاهدات الشاذة التي تتضمنها بيانات نموذج الانحدار الخطي البسيط. وتوصلت الدراسة إلي اقتراح طريقة جديدة لاكتشاف وتقليل أثر المشاهدات الشاذة في بيانات نموذج الانحدار الخطي البسيط باستخدام المعادلة $Y_i - \hat{\beta}_1 X_i = Z_i$ من خلال طريقتين ، طريقة تقريبية يتم فيها إعطاء رتب Rank لقيم المتغير Z_i ويتم تحديد المشاهدات الشاذة بهذه الطريقة بتمثيل النقاط $Z_i, Rank_i$ بيانياً. وطريقة شمولية مصحوبة بمقياس بياني مقترح - ويتم تحديد المشاهدات الشاذة بهذه الطريقة بتطبيق المعادلة $Y_i - \hat{\beta}_1 X_i = Z_i$ وحذف مشاهدة شاذة في كل تكرار إلي أن يحدد المقياس البياني عدد المشاهدات الشاذة ، حيث تناظر المشاهدة المحذوفة قيمة الانحراف المطلق الأكبر لـ Z_i عن β_0 . ولتقليل أثر المشاهدات الشاذة المكتشفة بأي من الطريقتين اقترح الباحث ترجيح انحرافات Z_i عن β_0 المناظرة للمشاهدة الشاذة بالقيمة الكسرية $P/(1-P)$ حيث P هو احتمال المشاهدات الشاذة وذلك للحصول علي قيم Z_i^* جديدة تناظرها قيم (Y_i^*, X_i^*) محسوبة رياضياً تكون قيم بديلة للمشاهدات الشاذة. وتوصلت الدراسة إلي عدة نتائج أهمها: أن الطريقة التقريبية كشفت عن جزء من المشاهدات الشاذة بينما كشفت الطريقة الشمولية عن العدد الكلي للمشاهدات الشاذة، حيث كانت عملية الكشف من المشاهدات ذات الأثر الأكبر والتي تليها والتي تليها إلي المشاهدات ذات الأثر الأقل علي نموذج الانحدار. كان لعملية ترتيب المشاهدات وفق ترتيب المتغير Z وضرب d_i المناظرة للمشاهدات الشاذة بالقيمة

¹ - علي، حامد حسين، (2005)، " طرق بديلة لكشف وتقليل أثر المشاهدات الشاذة علي نموذج الانحدار الخطي البسيط "، رسالة ماجستير غير منشورة ، جامعة السودان للعلوم والتكنولوجيا ، الخروم، السودان.

$P/(1-P)$ الأثر الكبير في تقليل أثر المشاهدات الشاذة وذلك من غير تحيز لان القيمة $P/(1-P)$ تتناسب طردياً مع عدد المشاهدات الشاذة المكتشفة.

1-8-2 دراسة (مروان وفرح، 2006)²:

عنوان الدراسة: تأثير القيم الشاذة على نتائج تحليل الانحدار مع التطبيق على المواليد

الخدج:

هدفت هذه الدراسة إلى الكشف عن القيم الشاذة التي تشير إليها نتائج تحليل الانحدار و الرسم الصندوقي بالإضافة إلى تأثير وجود ومعالجة القيم الشاذة على نتائج تحليل الانحدار. وتوصلت الدراسة إلى عدد من النتائج أهمها: تؤثر القيم الشاذة على تركيبة النموذج ونتائج تحليل الانحدار من خلال إدخال متغيرات غير مؤثرة على الظاهرة المدروسة واستبعاد بعض المتغيرات المؤثرة. لا يمكن التنبؤ بتأثير القيم الشاذة على معامل التحديد أو القيم التقديرية لتباين المجتمع ، أي تكون قيمتهما غير حقيقية في حالة وجود القيم الشاذة. القيم الشاذة التي يشير إليها تحليل الانحدار تؤثر على نتائجه والتي يكشفها الرسم الصندوقي فقط لا تؤثر على نتائج التحليل. يشير تحليل الانحدار إلى نوعين من القيم الشاذة أحدهما قيم متطرفة وهي التي يجب معالجتها وهذا النوع أيضاً يشير إليه الرسم الصندوقي والنوع الآخر شذوذ البواقي المعيارية التي يمكن التغلب عليها من خلال معالجة قيم أخرى أو تغيير في تركيبة النموذج.

1-8-3 دراسة (محمد طه أحمد، 2007)³:

عنوان الدراسة: المشاهدات غير المألوفة في نموذج الانحدار:

هدفت هذه الدراسة إلى دراسة تشخيص المشاهدات غير المألوفة في نموذج الانحدار ومدى تأثيرها على معلماته. حيث تم تشخيص المشاهدات الشاذة في المتغيرات المستقلة باستخدام

² - ديبوب، مروان عبد العزيز، بونس، فرح عبد الغني، (2006)، " تأثير القيم الشاذة علي نتائج تحليل الانحدار مع التطبيق علي المواليد الخدج"، مجلة علوم الرافدين، المجلد17، العدد1، ص:ص: 62-81 .
³ - أحمد، محمد طه، (2007)، " المشاهدات غير المألوفة في نموذج الانحدار"، مجلة تكريت للعلوم الإدارية والاقتصادية.

مصنوفة الهات، ومشاهدات المتغير التابع الشاذة تم تشخيصها باستخدام بواقي t المعيارية Studentized residuals ، أما المشاهدات المؤثرة فتم الكشف عنها بتطبيق صيغة مسافة Cook's distance. وتوصلت الدراسة إلى أن معاملات الانحدار تتأثر بالمشاهدات غير المألوفة سواءً في المتغيرات المستقلة أو المتغير التابع ويختلف تأثيرها حسب موقعها من خط الانحدار بين مؤثرة وغير مؤثرة ، بالإضافة إلى أن بعض المشاهدات غير المألوفة أو الشاذة تكون غير مؤثرة سلباً على معاملات النموذج حيث يلغى تأثير بعض المشاهدات البعض الآخر.

1-8-4 دراسة (محمد نذير إسماعيل قاسم، يونس حازم إسماعيل، 2008)⁴:

عنوان الدراسة: الكشف عن القيم الشاذة بأسلوب بيز باستخدام معاينة جيس:

هدفت هذه الدراسة إلى دراسة القيم الشاذة والكشف عنها بطريقة بيز في نموذج تحليل الانحدار الخطي البسيط بأسلوب معاينة جيس والتي هي حالة خاصة لطريقة مونتني كارلو سلسلة ماركوف وتطبيق هذا الأسلوب على بيانات حقيقية تم الحصول عليها من احدي البحوث المنشورة لأجل الكشف عن القيم الشاذة. وتوصلت الدراسة إلى عدد من النتائج أهمها: إن الكشف عن القيم الشاذة بأسلوب بيز باستخدام معاينة جيس يحل الكثير من المسائل المعقدة ولاسيما ذات النماذج المعقدة والصعبة لان معاينة جيس تقسم المسائل والنماذج المعقدة إلى مجموعة مسائل بسيطة يمكن تحليلها ومعالجتها بسهولة كما أن النتائج تكون دقيقة وذلك لان هذه الطريقة تعتمد على المعلومات الأولية للمعلمات والتي غالباً ما تكون متوفرة في معظم الدراسات، كما إن هذه الطريقة تمكننا من الحصول على احتمالية كون المشاهدة أو المشاهدات قيماً شاذة وهذا ما لم يمكن الحصول عليه في استخدام الطرائق الكلاسيكية في الكشف عن القيم الشاذة. إن معاينة جيس هي طريقة ذات كفاءة ودقيقة لأنها تبدأ بقيم ابتدائية ومن ثم يجري عليها عدد كبير من

⁴ -قاسم، محمد نذير، إسماعيل، يونس حازم، (2008)، "الكشف عن القيم الشاذة بأسلوب بيز باستخدام معاينة جيس"، المجلة العراقية للعلوم الإحصائية، العدد (14) ص، ص (68-88).

التكرارات كما أن خوارزمتها فائقة السرعة في التنفيذ. تعد معاينة جيس من الطرائق الإحصائية المهمة جداً في التطبيقات العلمية والنظرية والعلمية لأنها تعطينا التوزيع الاحتمالي للمتغيرات العشوائية ومن ثم نستطيع الحصول علي المؤشرات الرياضية لكل متغير كالتوقع والتباين وغيرها من المؤشرات الرياضية الأخرى. عند تطبيق الطريقة الكلاسيكية للكشف عن القيم الشاذة علي البيانات نفسها قورنت النتائج مع طريقة معاينة جيس فكانت متقاربة ومشجعة جداً.

1-8-5 دراسة (ألاء عبد الستار، 2009)⁵:

عنوان الدراسة: تأثير القيم الشاذة ونقطة الأصل علي نتائج تحليل الانحدار:

The effective of Outlier and the intercept on regression analysis results

هدفت الدراسة إلى التعرف علي تأثير وجود بعض القيم الشاذة علي نتائج تحليل الانحدار مقارنة بمعالجتها ، حيث تم الكشف عن وجود القيم الشاذة بطريقة الرسم الصندوقي وعولجت بطريقة متوسط البتر. وتوصلت الدراسة إلى عدة نتائج أهمها إن وجود القيم الشاذة في بعض المتغيرات أثر سلباً على نتائج تحليل الانحدار حيث أنها قللت من قيمة F المحسوبة لمعادلة الانحدار التقديرية وكذلك قيمة معامل التحديد R^2 وزادت من قيمة متوسط مربعات الخطأ MSE مقارنةً مع نتائج التحليل بعد معالجة القيم الشاذة و التي أدت إلى عكس النتيجة.

1-8-6 دراسة (مزاحم محمد يحيى، 2009)⁶:

عنوان الدراسة: المشاهدات غير العادية في الانحدار الخطي المتعدد وبعض طرق تشخيصها

مع التطبيق:

هدفت هذه الدراسة إلى تناول طرائق تشخيص ثلاثة أنواع من المشاهدات غير العادية Unusual Observation في تحليل الانحدار المتعدد، وبيان مدى تأثير المشاهدات الشاذة في التوزيع

⁵ - داوود، ألاء عبد الستار، (2010)، "تأثير القيم الشاذة ونقطة الأصل علي نتائج تحليل الانحدار"، مجلة تكريت للعلوم الصرفة، المجلد 15، العدد 1.
⁶ - يحيى، مزاحم محمد، (2009)، "المشاهدات غير العادية في الانحدار الخطي المتعدد وبعض طرائق تشخيصها مع التطبيق"، المجلة العراقية للعلوم الإحصائية، العدد (15)، ص، ص: 161-192 .

الطبيعي للبواقي. حيث تم الكشف عن المشاهدات الشاذة Outlier باستخدام الرسم الصندوقي Box plot ومقياس البواقي القياسية Studentized residuals والكشف عن المشاهدات ذات قوة الرفع Leverage Observation من خلال المصفوفة Hat Matrix ، أما المشاهدات المؤثرة فتم الكشف عنها باستخدام مقياس DFBETAS. وتوصلت الدراسة إلى أن حذف المشاهدات الشاذة أدى إلى اقتراب توزيع البيانات من التماثل تماماً، وكذلك أثر حذف المشاهدات الشاذة في دقة معاملات النموذج بانخفاض متوسط مربعات الخطأ MSE وعامل تضخم التباين VIF وارتفاع قيمة F ومعامل التحديد R^2 وتغيير إشارات بعض المعلمات المقدرة والوصول إلى نموذج انحدار يمثل بيانات الدراسة خير تمثيل.

1-8-7 دراسة (نزار مصطفى جواد و غفران إسماعيل كمال، 2009):⁷

عنوان الدراسة: مقارنة طرق تقدير معالم نموذج الانحدار في حالة ظهور مشكلة التعدد الخطي والقيم الشاذة:

هدفت هذه الدراسة إلى دراسة بعض طرائق تقدير معالم الانحدار الخطي والمقارنة فيما بينها عند اختلاف توزيع الخطأ وحجم العينة وقوة العلاقة بين المتغيرات المستقلة. حيث أستخدم في هذه الدراسة أسلوب المحاكاة لاختبار قوة وحصانة المقدرات لنموذج الانحدار الخطي المتعدد عند وجود مشاكل التداخل الخطي و الأخطاء الغير طبيعية وذلك بتطبيق عدة طرق للتقدير منها الاعتيادية و الحصينة، وهي طريقة المربعات الصغرى OLS وانحدار الريدج RID وطريقة القيمة المطلقة الصغرى RLAV وطريقة مقدرات الريدج الموزون WRID وطريقة MM ومقدرات انحدار الريدج الحصينة RMM المعتمدة علي مقدر MM. وتوصلت الدراسة إلى أنه عندما تكون درجة التداخل الخطي عالية فإن طريقة RMM هي الأفضل من طرق التقدير

⁷ - جواد، نزار مصطفى، كمال، غفران إسماعيل، (2009)، "مقارنة طرق تقدير معالم نموذج الانحدار في حالة ظهور مشكلة التعدد الخطي والقيم الشاذة"، مجلة العلوم الإدارية والاقتصادية، العدد (55)، ص:ص 153-166، العراق، بغداد.

الأخرى، وعند المقارنة بين مقدرات الريدج الحصينة RMM ، RLAV ، WRID توصلت الدراسة إلى أن مقدرات RMM هي الأفضل من المقدرات RLAV ، WRID ولمختلف افتراضات توزيع الخطأ ودرجة التداخل الخطي المتعدد.

1-8-8 دراسة (صالح مؤيد شاكر، 2009)⁸:

عنوان الدراسة: تحسين أسلوب M الحصين في تقدير معاملات نموذج الانحدار الخطي المتعدد:

هدفت هذه الدراسة إلى إيجاد تركيبة حصينة موزونة لمقدرات نموذج الانحدار الخطي المتعدد بدمج أسلوب M الحصين وطريقة المربعات الصغرى الموزونة لمعالجة القيم الشاذة في بيانات نموذج تحليل الانحدار الخطي المتعدد للحصول على مقدرات جديدة تقام أثر القيم الشاذة على نتائج التحليل و مقارنة طرائق التقدير OLS, WLS, RM, RMW بالاعتماد على معيار متوسط مربعات الخطأ MSE . حيث تم الكشف عن القيم الشاذة بطريقة الرسم الصندوقي و فحص عناصر قطر المصفوفة H كما أستخدم مقياس مربع مسافة كوك لاختبار تأثير حذف المشاهدة i علي المعلمات المقدرة، ومقياس - Deferent When Excluding the ith case DF للكشف عن وجود القيم المتطرفة في متجه الأخطاء العشوائية. وتوصلت الدراسة إلى عدة نتائج أهمها: عدم وجود فروق في دقة نتائج الطرائق المدروسة عند عدم وجود القيم الشاذة في البيانات. لا يمكن الاعتماد على مقدرات أسلوب M الحصين في حالة وجود تطرف قوى في مشاهدات المتغير (المتغيرات) المستقلة أو حتى المتغير المعتمد أو في كليهما. وأيضاً أثبتت الدراسة أن أسلوب M الحصين الموزون R.MW المقترح قد أثبت كفاءته في تقدير معاملات

⁸ - شاكر، صالح مؤيد، (2009)، "تحسين أسلوب M الحصين في تقدير معاملات نموذج الانحدار الخطي المتعدد"، المجلة العراقية للعلوم الإحصائية، العدد (16)، ص:219-242.

نموذج الانحدار الخطى وبدقة عالية عند وجود القيم الشاذة في قيم المتغير (المتغيرات) المستقلة أو في المتغير المعتمد أو في كليهما.

1-8-9 دراسة (صباح حسيب حسن ، 2009)⁹:

عنوان الدراسة: استخدام المعايير الحصينة في تحديد المتغيرات المؤثرة في نموذج الانحدار الخطى لقياس السكر في الدم:

هدفت هذه الدراسة إلى استخدام المعايير الحصينة في تحديد المتغيرات المهمة في نموذج الانحدار الخطى المتعدد في حالة وجود القيم الشاذة في البيانات وذلك من خلال تحليل بيانات السكر في الدم بوصفها متغيراً معتمداً ومتغيرات أخرى مستقلة، حيث تم تقدير جميع النماذج الجزئية الممكنة بطريقة M الحصينة، وتمت المقارنة بين هذه النماذج باستخدام المعايير الحصينة (RC,RSIC,RAIC,RAPE,RF,RSC,RVC) لتحديد المتغيرات المستقلة في النموذج الأفضل. وتوصلت الدراسة إلى عدد من النتائج أهمها أن أفضل المعايير أداءً كان المعيار RAPE حيث أن اختيار هذا المعيار للنموذج الأفضل جاء متوافقاً مع رأي المختصين في مجال الطب.

1-8-10 دراسة (تارا احمد حسن و مهدي صابر رضا، 2011)¹⁰:

عنوان الدراسة: استخدام الانحدار الحصين لإيجاد أنسب نموذج لتمثيل بيانات الأنواء الجوية في مدينة أربيل خلال الفترة 1998-2010:

هدفت هذه الدراسة إلى إيجاد أنسب نموذج انحدار خطى متعدد لتمثيل بيانات الأنواء الجوية بدراسة أهم العوامل المؤثرة في الرطوبة النسبية، ولتحقيق هذا الهدف تم تطبيق بعض طرق

⁹ - حسن، صباح حسيب،(2009)، " استخدام المعايير الحصينة في تحديد المتغيرات المؤثرة في نموذج الانحدار الخطى لقياس السكر في الدم"، المجلة العراقية للعلوم الإحصائية، العدد (16)، ص:ص 148-131.

- حسن، تارا احمد، رضا، مهدي صابر، (2011)، "استخدام الانحدار الحصين لإيجاد أنسب نموذج لتمثيل بيانات الأنواء الجوية في مدينة أربيل خلال الفترة 1998-2010"، مجلة العلوم الاقتصادية والإدارية، العدد (89)، ص:ص 213-195. الجامعة المستنصرية.

الانحدار الحصين في حالة وجود المشاهدات الشاذة منها، طريقة مربعات الوسيط الصغرى LMS وطريقة المربعات المشدومة الصغرى LTS مقارنة بطريقة المربعات الصغرى الاعتيادية OLS وتمت المقارنة بين هذه الطرق بالاعتماد علي معيار متوسط مربعات الخطأ MSE و تم الكشف عن المشاهدات الشاذة باستخدام قطر عناصر المصفوفة H وطريقة بوكس- ويسكر- بلوت Box-and-Whisker plot البيانية والمشاهدات المؤثرة تم الكشف عنها باستخدام المقياس DFFITS. وتوصلت الدراسة إلي عدد من النتائج أهمها: أن طريقة مربعات الوسيط LMS هي الأفضل لتمثيل أنسب نموذج انحدار لدراسة بيانات حالة الطقس. كما أوصت الدراسة باستخدام طرق أخرى للكشف عن المشاهدات الشاذة في دراسات مستقبلية مثل طريقة تقديرات S,L,R.

9-1 التعليق علي الدراسات السابقة:

يتضح من الدراسات السابقة التي لها صلة بموضوع الدراسة الحالية أن بعضها تناول دراسة القيم الشاذة بالتطبيق على نموذج تحليل الانحدار الخطي البسيط ، منها دراسة (حامد، 2005م) حيث توصلت هذه الدراسة إلي اقتراح طريقة جديدة لاكتشاف وتقليل أثر القيم الشاذة في بيانات نموذج تحليل الانحدار الخطي البسيط من خلال طريقتين تقريبيه وشمولية ، ومن أهم نتائج هذه الدراسة إن الطريقة الشمولية كشفت عن العدد الكلي للمشاهدات الشاذة حيث كانت عملية الكشف من المشاهدات ذات الأثر الأكبر والتي تليها والتي تليها إلي المشاهدات ذات الأثر الأقل علي نموذج الانحدار. أما دراسة (ألاء ، 2009م) كشفت هذه الدراسة عن القيم الشاذة باستخدام طريقة الرسم الصندوقي وتمت معالجتها بطريقة متوسط البتر، ومن أهم نتائجها إن وجود القيم الشاذة في بعض المتغيرات أثر سلباً علي نتائج تحليل الانحدار حيث أنها قللت من قيمة F المحسوبة لمعادلة الانحدار التقديرية وكذلك قيمة معامل التحديد R^2 وزادت من قيمة

متوسط مربعات الخطأ MSE وهذا ما توصلت إليه الدراسة الحالية أيضاً بالإضافة إلى أن طريقة الرسم الصندوقي كشفت عن بعض القيم الشاذة في المتغيرات المستقلة. أما الدراسة التي قام بها (محمد ، 2008م) كشفت عن القيم الشاذة في بيانات نموذج الانحدار الخطي البسيط بأسلوب بيز باستخدام معاينة جيبس وأهم ما توصلت إليه أن الكشف عن القيم الشاذة بهذه الطريقة أدى إلى نتائج مشجعة ومقارنة مقارنة بنتائج الطرق الكلاسيكية ولم تعالج هذه الدراسة القيم الشاذة بأي طريقة لمعرفة مدي تأثيرها علي نتائج النموذج المقدر. كما أن هنالك بعض الدراسات السابقة التي تناولت دراسة موضوع القيم الشاذة بالتطبيق علي نموذج تحليل الانحدار الخطي المتعدد وهذا ما لجأت إليه الدراسة الحالية ، ومن هذه الدراسات دراسة (مروان و فرح ، 2006م) حيث كشفت هذه الدراسة عن القيم الشاذة في بيانات نموذج تحليل الانحدار الخطي المتعدد باستخدام طريقة الرسم الصندوقي ، وأهم ما توصلت إليه هذه الدراسة أن القيم الشاذة تؤثر علي تركيبة النموذج ونتائج تحليل الانحدار من خلال إدخال متغيرات غير مؤثرة علي الظاهرة المدروسة واستبعاد بعض المتغيرات المؤثرة وتشير هذه النتيجة إلى أن القيم الشاذة تؤثر علي اختبارات المعنوية للنموذج والتي من بينها اختبار F واختبار t ، القيم الشاذة تؤثر علي قيمة معامل التحديد والقيمة التقديرية لتباين المجتمع حيث تكون قيم هذه المعالم غير حقيقة في حالة وجود القيم الشاذة ، وأيضاً من نتائج هذه الدراسة القيم التي يكشفها الانحدار تؤثر علي نتائجه والتي يكشفها الرسم الصندوقي فقط لا تؤثر علي نتائج التحليل وتتفق هذه النتائج مع نتائج الدراسة الحالية. أما دراسة (محمد طه ، 2007م) كشفت عن القيم الشاذة في المتغيرات المستقلة باستخدام عناصر القطر الرئيسي لمصفوفة Hat Matrix ومشاهدات المتغير التابع الشاذة تم تشخيصها باستخدام بواقي t المعيارية والمشاهدات الشاذة المؤثرة فتم الكشف عنها بتطبيق صيغة مسافة كوك، وأهم ما توصلت إليه هذه الدراسة أن معاملات الانحدار تتأثر بالمشاهدات الشاذة

سواءً في المتغيرات المستقلة أو في المتغير التابع ويختلف تأثيرها حسب موقعها من خط الانحدار بين مؤثرة وغير مؤثرة وتتفق أيضاً هذه النتيجة مع ما توصلت إليه الدراسة من نتائج. أما دراسة (مزاحم محمد ، 2009م) كشفت هذه الدراسة عن القيم الشاذة باستخدام طريقة الرسم الصندوقي ومقياس البواقي المعيارية و مصفوفة Hat Matrix أما المشاهدات المؤثرة فتم الكشف عنها باستخدام المقياس DFBETAS وأهم ما توصلت إليه الدراسة إن حذف المشاهدات الشاذة أدي إلي اقتراب توزيع البيانات من التماثل وكذلك أثر حذف المشاهدات الشاذة إلي دقة معالم النموذج بانخفاض متوسط مربعات الخطأ MSE وعامل تضخم التباين وارتفاع قيمة F ومعامل التحديد R^2 وتغيير إشارات بعض المعالم المقدرة والوصول إلي نموذج يمثل بيانات الدراسة خير تمثيل وهذا ما توصلت إليه الدراسة الحالية مع أن معالجة القيم الشاذة بطريقة متوسط البتر أفضل من طريقة الحذف. أما دراسة (نزار وغفران، 2009) استخدمت هذه الدراسة أسلوب المحاكاة لاختبار قوة وحصانة المقدرات لنموذج الانحدار المتعدد عند وجود التداخل الخطي والأخطاء الغير طبيعية بتطبيقه عدة طرق منها الاعتيادية والحصينة ، ومن أهم نتائج هذه الدراسة أن مقدرات انحدار الريديج الموزون RMM هي أفضل المقدرات لمختلف افتراضات توزيع الخطأ ودرجة التداخل الخطي المتعدد. أما دراسة (صالح ، 2009م) لجأت هذه الدراسة أيضاً إلي تطبيق أسلوب الطرق الحصينة لمعالجة القيم الشاذة في بيانات نموذج تحليل الانحدار الخطي المتعدد للحصول علي مقدرات جديدة تقاوم أثر القيم الشاذة ، حيث تم الكشف عن القيم الشاذة بطريقة الرسم للصندوقي وعناصر قطر مصفوفة Hat Matrix ومقياس مربع مسافة كوك لاختبار تأثير حذف المشاهدة الشاذة علي المعالم المقدرة والمقياس DF للكشف عن وجود القيم المتطرفة في متجه الأخطاء العشوائية. وأهم ما توصلت إليه هذه الدراسة أن أسلوب M الحصين الموزون R.MW قد أثبت كفاءته في تقدير معالم نموذج الانحدار الخطي وبدقة عالية عند

وجود القيم الشاذة في قيم المتغيرات المستقلة أو المتغير التابع أو كليهما وهذا ما أوصت به الدراسة الحالية أيضاً من دراسات مستقبلية. أما دراسة (تارا ومهدي ، 2011م) طبقت أيضاً هذه الدراسة أسلوب الطرق الحصينة لتقدير نموذج الانحدار الخطي المتعدد في حالة وجود القيم الشاذة في البيانات ، حيث تم الكشف عن القيم الشاذة باستخدام عناصر قطر المصفوفة Hat Matrix وطريقة Box-and-Whisker plot البيانية والمشاهدات المؤثرة فتم الكشف عنها باستخدام المقياس DFFITS. وأهم ما توصلت إليه هذه الدراسة أن طريقة مربعات الوسيط LMS هي الأفضل لتمثيل أفضل نموذج انحدار خطي متعدد لبيانات الدراسة. أما دراسة (صباح ، 2009م) لجأت هذه الدراسة أيضاً لتطبيق أسلوب الطرق الحصينة ، حيث استخدمت بعض المعايير الحصينة في تحديد المتغيرات المهمة في نموذج الانحدار الخطي المتعدد في حالة وجود القيم الشاذة في البيانات. وأهم ما يميز الدراسة الحالية عن الدراسات السابقة:

1. توصلت هذه الدراسة إلى إن طريقة الرسم الصندوقي كشف عن بعض القيم الشاذة في بيانات نموذج تحليل الانحدار الخطي المتعدد.
2. استخدمت الدراسة بواقى ستيودنت المحذوفة للكشف عن القيم الشاذة في المتغير التابع.
3. أجرت الدراسة العديد من المقارنات بين معايير تشخيص القيم الشاذة في المتغيرات المستقلة والمتغير التابع.
4. قامت الدراسة بدراسة أثر القيم الشاذة علي أهم مشاكل نموذج تحليل الانحدار الخطي المتعدد وتوصلت إلى احتمال ظهور الارتباط الذاتي لبواقى النموذج عند وجود القيم الشاذة في البيانات.
5. تطرقت الدراسة إلى دراسة تأثير القيم الشاذة المؤثرة علي قيم المتغير التابع المقدرة ومعاملات النموذج والأخطاء المعيارية وقياس الأثر علي كل معاملات النموذج.

6. طبقت الدراسة طريقة الحذف وطريقة متوسط البتر لمعالجة القيم الشاذة وتوصلت إلي أن طريقة متوسط أفضل من طريقة الحذف.

10-1 الإطار العام للدراسة:

تحتوي هذه الدراسة على ستة فصول تفصيلها كالآتي:

الفصل الأول: خطة الدراسة وتشتمل على تمهيد ، مشكلة الدراسة ، أهمية الدراسة ، أهداف الدراسة ، فرضيات الدراسة ، منهجية الدراسة ، مجتمع وعينة الدراسة ، الدراسات السابقة ، الإطار العام للدراسة.

الفصل الثاني: نموذج تحليل الانحدار الخطي المتعدد ويتضمن تمهيد، طبيعة النموذج ، افتراضات النموذج ، تقدير معاملات النموذج وتباين الأخطاء، خصائص مقدرات المربعات الصغرى ، نموذج الانحدار المعياري ، تقييم معنوية المقدرات والنموذج ، اختبار جودة النموذج ، اختبار ثبات معاملات النموذج ، التنبؤ وتقييم أداء النموذج في التنبؤ ، مشكلة عدم تجانس التباين Heteroscedasticity ، مشكلة الإرتباط الخطي المتعدد Multicollinearity ، مشكلة الإرتباط الذاتي Autocorrelation واختبار التوزيع الطبيعي للبواقي.

الفصل الثالث: المشاهدات الشاذة في نموذج تحليل الانحدار الخطي المتعدد ويحتوي على تمهيد ، مفهوم القيم الشاذة وأسباب ظهورها ، طرق الكشف عن القيم الشاذة في المتغيرات المستقلة و المتغير التابع ، المشاهدات الشاذة المؤثرة و مقاييس الكشف عنها ، بعض الحلول المستخدمة لمعالجة القيم الشاذة.

الفصل الرابع: الجانب التطبيقي وفي هذا الفصل تم التعرض لوصف متغيرات الدراسة وتطبيق نموذج تحليل الانحدار الخطي المتعدد على بيانات الدراسة والكشف عن القيم الشاذة ومعالجتها ثم المقارنة بين نتائج التحليل قبل معالجة القيم الشاذة وبعد معالجتها و التنبؤ بقيم المتغير التابع.

الفصل الخامس: النتائج والتوصيات.

الفصل الثاني نموذج تحليل الانحدار الخطي المتعدد

1-2 تمهيد:

2-2 طبيعة النموذج:

3-2 افتراضات النموذج:

4-2 تقدير معاملات النموذج وتباين الأخطاء

5-2 نموذج الانحدار المعياري:

6-2 تقييم معنوية المقدرات و النموذج:

7-2 اختبار جودة النموذج:

8-2 اختبار ثبات معاملات النموذج:

9-2 التنبؤ Forecasting :

10-2 تقييم أداء النموذج في التنبؤ:

11-2 مشكلة عدم تجانس التباين:

12-2 مشكلة الارتباط الخطي المتعدد:

13-2 مشكلة الارتباط الذاتي:

14-2 اختبار التوزيع الطبيعي للبواقي:

2-1 تمهيد:

سوف نتناول في هذا الفصل بالشرح ، خطوات توفير نموذج تحليل الانحدار الخطي المتعدد ابتداءً من التعريف بطبيعة النموذج وفرضياته و تقدير معلماته و الاختبارات الإحصائية لتقييم النموذج وقياس مدى استقرار تقديرات النموذج خلال فترة الدراسة بالإضافة إلى اختبار مقدرة النموذج على التنبؤ مستقبلاً وأهم مشكلات النموذج.

2-2 طبيعة النموذج Nature of Model [11] [12]:

يعتبر نموذج تحليل الانحدار وسيلة إحصائية تستخدم لتحليل العلاقة بين متغير مستقل Independent Variable واحد أو أكثر ومتغير تابع Dependent Variable ويعد تحليل الانحدار من أكثر الطرق الإحصائية استخداماً في مختلف العلوم لأنه يصف العلاقة بين المتغيرات المستقلة والمتغير التابع على هيئة معادلة رياضية ، فالمعادلة التي تضم متغيراً مستقلاً واحداً تسمى معادلة الانحدار الخطي البسيط Linear Regression Equation في حين تسمى المعادلة التي تضم أكثر من متغير مستقل معادلة الانحدار الخطي المتعدد Multiple Linear Regression Equation ، إن العلاقة الدالية بين المتغير التابع Y والمتغيرات المستقلة X_1, X_2, \dots, X_k في تحليل الانحدار الخطي المتعدد يمكن التعبير عنها بوصفها دالة خطية كالتالي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + U_i \quad i = 1, 2, \dots, n \quad \dots\dots(1-2)$$

حيث أن:

Y_i : قيم المتغير التابع أو مقدار الاستجابة ، $\beta_0, \beta_1, \dots, \beta_k$: ثوابت أو معالم النموذج المجهولة

المراد تقديرها ، X_1, X_2, \dots, X_k : قيم ثابتة لـ K من المتغيرات المستقلة

¹¹ - الراوي، خاشع محمود، (1987)، "المدخل إلى تحليل الانحدار"، مديرية دار الكتب للطباعة والنشر، جامعة الموصل، العراق.
¹² - إبراهيم، بسام يونس وآخرون، (2002)، "الاقتصاد القياسي"، دار عزة للنشر والتوزيع، الخرطوم، السودان، ص:151-152.

N : عدد المشاهدات أو حجم المجتمع.

والمعادلة (2-1) يمكن كتابتها بأسلوب المصفوفات كما يلي:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}_{n \times (k+1)} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{(k+1) \times 1} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_k \end{bmatrix}_{n \times 1} \quad \dots\dots\dots(2-2)$$

و باختصار كما يلي:

$$Y = X\beta + U \quad \dots\dots\dots(3-2)$$

حيث أن:

X : مصفوفة البيانات من الدرجة $n \times (k+1)$ تحتوي على مشاهدات المتغيرات المستقلة

X_1, X_2, \dots, X_k حيث يحتوي العمود الأول قيم الواحد الصحيح لتمثيل المعامل الثابت β_0 .

β : متجه عمودي من الدرجة $(k+1) \times 1$ يحتوي على معالم نموذج الانحدار المجهولة

$\beta_0, \beta_1, \dots, \beta_k$ المراد تقديرها.

U : متجه عمودي من الدرجة $(n \times 1)$ يحتوي على قيم المتغير العشوائي U_i .

3-2 افتراضات النموذج Model's Assumption [14][13]:

لاستخدام طريقة المربعات الصغرى الاعتيادية في تقدير معالم نموذج تحليل الانحدار

الخطي المتعدد فانه يعتمد على عدة فرضيات أساسية وهي كالاتي:

1-3-2 الافتراضات العامة:

1. أن المتغير التابع Y يكون دالة خطية في K من المتغيرات المستقلة.

¹³ - الطيب، عز الدين مالك، (2008)، "المدخل إلى الاقتصاد القياسي"، الطبعة الأولى، مطبعة جي تاون، الخرطوم، السودان، ص:62-65.

¹⁴ - إبراهيم، بسام يونس وآخرون، (2002)، مرجع سابق، ص:152-155.

2. عدم وجود تداخل خطى متعدد بين المتغيرات المستقلة ، وهذا يعني أن تكون أعمدة المصفوفة X مستقلة خطياً.

3. أن تكون المتغيرات المستقلة X_1, X_2, \dots, X_k خالية من أخطاء التجميع.

4. أن تكون العلاقة المراد تقديرها قد تم تحديدها وتشخيصها.

5. عدم وجود أخطاء في قياس المتغيرات المستقلة X_1, X_2, \dots, X_k .

6. أن تكون المتغيرات المستقلة X_1, X_2, \dots, X_k غير عشوائية ، أي أنها تحتوى قيماً ثابتة في

المعاينات المتكررة ولكي تكون قيم المتغيرات المستقلة غير عشوائية يجب على الباحث

التحكم فيها تجريبياً.

7. أن يتضمن النموذج المتغيرات المستقلة التي تسهم في تفسير المتغير التابع.

2-3-2 الفروض المتعلقة بحد الخطأ العشوائي:

1. متوسط حد الخطأ العشوائي U_i يجب أن يساوي صفر، أي:

$$E(U_i) = 0 \quad \text{for all } i = 1, 2, 3, \dots, n \quad \dots\dots\dots(4-2)$$

2. وجود خاصية الاستقلالية بين حد الخطأ U والمتغيرات المستقلة، أي:

$$E(U_i, X_{ij}) = 0 \quad \text{for all } i \neq j \text{ \& } i, j = 1, 2, 3, \dots, n \quad \dots\dots\dots(5-2)$$

3. ثبات تباين حد الخطأ U واستقلال قيم حدود الخطأ عن بعضها البعض الآخر، أي أن:

$$\text{Var}(U_i) = E(U_i^2) = \sigma_u^2 \quad \text{for all } i = 1, 2, 3, \dots, n \quad \dots\dots\dots(6-2)$$

$$\text{Cov}(U_i, U_j) = 0$$

4. المتغير العشوائي U_i يتوزع وفق التوزيع الطبيعي متعدد المتغيرات بمتوسط صفر وتباين ثابت

أي: $\sigma_u^2 I_n$

$$U \sim N(0, \sigma_u^2 I_n) \quad \text{for all } i = 1, 2, 3, \dots, n \quad \dots\dots\dots(7-2)$$

2-3-3 الافتراضات الخاصة بتوزيع المتغير المعتمد:

1. المتغير المعتمد يتوزع التوزيع الطبيعي متعدد المتغيرات بمتوسط $X\beta$ وتباين σ_u^2 ، أي:

$$Y_i \sim N(X\beta, \sigma_u^2 I_n) \dots\dots\dots(8-2)$$

2. التباين المشترك بين Y_i, Y_j يساوى صفر، وهذا يعني أن قيم المتغير المعتمد مستقلة عن بعضها البعض الآخر، أي:

$$Cov(Y_i, Y_j) = 0 \quad \text{for all } i \neq j \text{ \& } i, j = 1, 2, 3, \dots, n \quad \dots\dots\dots(9-2)$$

3. العلاقة بين \hat{Y}, X تكون علاقة خطية تتمثل بمعادلة خط مستقيم.

2-4 تقدير معاملات النموذج وتباين الأخطاء [16][15]:

2-4-1 تقدير معاملات النموذج:

إن استخدام طريقة المربعات الصغرى في تقدير المعلمات للنموذج تتميز بأنها تختار أحسن نموذج مطابق للبيانات ، بحيث تجعل مجموع مربعات الخطأ بين القيمة الحقيقية Y والقيمة المقدرة \hat{Y} أقل ما يمكن $\sum_{i=1}^n e_i^2$ ، ويتم تقدير معالم النموذج (2-1) بحيث يكون للدالة أدناه نهاية صغرى:

$$Q = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i} - \dots - \beta_k X_{ki})^2 \quad \dots\dots\dots(10-2)$$

ومن المعادلة (10-2) يمكن الحصول علي مقدرات المربعات الصغرى لمعاملات النموذج المقدر كما يلي:

$$\hat{\beta} = (X'X)^{-1} X'Y \quad \dots\dots\dots(11-2)$$

¹⁵ - إبراهيم، بسام يونس وآخرون، (2002)، مرجع سابق، ص:156-159.
¹⁶ - الراوي، خاشع محمود، (1987)، مرجع سابق.

ومن المعادلة (11-2) أعلاه نحصل علي مقدرات معالم النموذج بشرط أن تكون رتبة المصفوفة (XX) كاملة، أي غير مفردة وذلك لإيجاد محددتها ومن ثم معكوسها $(XX)^{-1}$ والصيغة (11-2) يمكن التعبير عنها بالشكل المصفوفي كما يلي:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} & \cdots & \sum X_{ki} \\ \sum X_{1i}^2 & \sum X_{1i} & \cdots & \sum X_{1i}X_{ki} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum X_{ki}^2 & \sum X_{ki} & \cdots & \sum X_{ki}X_{ki} \end{bmatrix}^{-1} \begin{bmatrix} \sum Y_i \\ \sum X_{1i}Y_i \\ \vdots \\ \sum X_{ki}Y_i \end{bmatrix} \quad \text{.....(12-2)}$$

عليه تكون الصيغة التقديرية للنموذج الخطي العام (1-2) كما يلي:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_{ki} \quad \text{.....(13-2)}$$

وبصيغة المصفوفات:

$$\hat{Y} = X \cdot \hat{\beta} \quad \text{.....(14-2)}$$

2-4-2 تفسير المعلمات المقدرة:

تفسر $\hat{\beta}_0$ على أنها موقع تقاطع مستوى الانحدار بالمحور Y و تساوي متوسط المتغير المعتمد عندما تكون جميع قيم المتغيرات المستقلة تساوي الصفر، كما تفسر $\hat{\beta}_i$ على أنها معامل الانحدار الجزئي لـ Y على X_i عند جعل بقية المتغيرات المستقلة ثابتة وهي تمثل مقدار التغير الذي يطرأ على المتغير المعتمد نتيجة لتغير المتغير المستقل X_i وحدة واحدة بثبات باقي المتغيرات المستقلة. فمثلاً $\hat{\beta}_1$ تمثل التغير في القيمة المتوسطة للمتغير التابع الناتج عن تغير المتغير X_1 بوحدة واحدة بافتراض ثبات قيم المتغيرات الأخرى (X_2, X_3, \dots, X_k) وبمعنى آخر يقيس المعامل $\hat{\beta}_1$ الأثر المباشر أو الصافي لتغير X_1 بوحدة واحدة على القيمة المتوقعة للمتغير التابع، وهكذا تفسر بقية معاملات الانحدار الجزئية. والمعامل الموجب يشير إلي وجود علاقة طردية بين المتغير المستقل والمتغير التابع، بينما يشير المعامل السالب إلي أن العلاقة بين المتغيرين علاقة عكسية. كما يشير حجم المعامل إلي مقدار التغير الذي يحدث في المتغير

التابع الناتج عن زيادة مقدارها وحدة واحدة في هذا المتغير بافتراض ثبات قيم جميع المتغيرات المستقلة الأخرى.

2-4-3 خصائص مقدرات المربعات الصغرى¹⁷:

إذا توفرت الفروض الأساسية للنموذج (2-1) فإن التقديرات المتحصل عليها باستخدام طريقة المربعات الصغرى تتميز بأنها خطية وغير متحيزة ولها أقل تباين ، وذلك نسبة لكل المقدرات الخطية غير المتحيزة الأخرى. إن هذه الطريقة توفر أفضل المعايير للحكم علي جودة هذه المقدرات مما يجعلها قريبة من القيم الحقيقية لمعاملات النموذج، وهذه المعايير تعنى أفضل مقدرات خطية غير متحيزة Best Linear Unbiased Estimators وتفصيل هذه الخصائص فيما يلي:

2-4-3-1 الخطية Linearity:

وهي الحصول علي أفضل تشكيلة خطية من قيم مشاهدات العينة أو أن كل من المقدرات $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ على حدة هي دالة خطية في مشاهدات المتغير التابع Y_i ويمكن برهنة ذلك من خلال الصيغة العامة لمقدرات المربعات الصغرى $\hat{\beta} = (XX)^{-1}XY$ فبما أن $(XX)^{-1}X'$ مصفوفة أرقام ثابتة، فإن متجه المقدرات $\hat{\beta}$ يعتمد بصورة خطية علي متجه المتغير التابع ، ومن ثم فأن مقدرات المربعات الصغرى خطية.

2-4-3-2 عدم التحيز Unbiased ness:

يعني عدم التحيز أن القيمة المتوقعة لكل عنصر من عناصر المتجه $\hat{\beta}$ تساوي العنصر المقابل في متجه المعالم الحقيقية β ، أي أن:

$$E(\hat{\beta}) = \beta \quad \dots\dots\dots(15-2)$$

¹⁷ - إسماعيل، محمد عبد الرحمن، (2001)، " تحليل الانحدار الخطي"، معهد الإدارة العامة ، المملكة العربية السعودية، مركز البحوث ، ص:140-142.

2-3-3-3 أقل تباين - الكفاءة Best Estimators:

إن هذه الخاصية تعني أن تباين كل من مقدرات المربعات الصغرى $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ علي حدة أقل ما يمكن ، ويحدث ذلك في حالة كون تباين كل من $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ علي حدة أقل من تباين أي مقدره أخرى ، وللتحقق من ذلك نقوم بتعريف مقدرات خطية أخرى لها خاصية عدم التحيز ونقارن بين تباين المقدرين. وبافتراض أن $\hat{\beta}^*$ أي مقدر خطي آخر لـ $\hat{\beta}$ يأخذ الصيغة التالية:

$$\hat{\beta}^* = [(XX)^{-1} X' + C]Y \quad \dots\dots\dots(16-2)$$

حيث أن:

C مصفوفة من الدرجة $[(k+1) \times n]$ ويتكون عناصرها من ثوابت.

ومن المعادلة (16-2) يمكن الوصول إلي أن مصفوفة تباين- تغاير المقدر $\hat{\beta}^*$ تعطي بالصيغة التالية:

$$Ver - Cov(\hat{\beta}^*) = \sigma_u^2 (XX)^{-1} + \sigma_u^2 CC' \quad \dots\dots\dots(17-2)$$

ويتضح من المعادلة (17-2) أن مصفوفة تباين- تغاير متجه المقدر الخطي غير المتحيز $\hat{\beta}^*$ تساوي مصفوفة تباين- تغاير متجه المقدر الخطي غير المتحيز $\hat{\beta}$ والتي تساوي $\sigma_u^2 (XX)^{-1}$ زائد المقدار $\sigma_u^2 CC'$ ومن ثم فإن تباين كل عنصر من عناصر المتجه $\hat{\beta}^*$ أكبر من تباين العنصر المقابل في المتجه $\hat{\beta}$ ، أي أن $Var(\hat{\beta}^*) > Var(\hat{\beta})$ وبالتالي لا توجد مقدرات تبايناتها أقل من تباين مقدرات المربعات الصغرى ، لذلك فإن مقدرات المربعات الصغرى تتسم بالكفاءة لأنها خطية وغير متحيزة ولها أقل تباين.

2-4-4 تقدير تباين الخطأ العشوائي:

إن تقدير طريقة المربعات الصغرى لتباين الخطأ العشوائي هو $\hat{\sigma}_u^2$ ويحسب من المعادلة التالية:

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^n e_i^2}{n - k - 1} \dots\dots\dots(18-2)$$

حيث أن:

k : عدد المتغيرات المستقلة في النموذج ، n : عدد المشاهدات ، e_i : بواقي النموذج العادية.

2-5 نموذج الانحدار المعياري Standardized Regression Model [18]:

يتم تقدير معالم نموذج الانحدار المعياري عند الرغبة في تحديد الأهمية النسبية للمتغيرات المستقلة المؤثرة في نموذج الانحدار أو عندما تختلف وحدات القياس بالنسبة للمتغيرات المستقلة فيكون من الصعب جداً مقارنة قيم معاملات النموذج لمعرفة أيها أكثر تأثيراً على المتغير التابع ، فمثلاً نجد أن متغير العمر لأي شخص يقاس بالسنوات ومتغير الطول يقاس بالسنتيمترات ، وأيضاً من الممكن أن يقاس متغير العمر بالأيام أو الشهور ومتغير الطول بالبوصات أو الأمتار وبالتالي نحصل على أحجام معاملات مختلفة للنموذج إذا اختلفت وحدات القياس. وللتخلص من تأثير وحدات القياس المختلفة للمتغيرات على أحجام معاملاتنا يتم تحويل قيم المتغيرات إلى متغيرات معيارية Standardized Variable ومن ثم تقدير معالم نموذج الانحدار باستخدام طريقة المربعات الاعتيادية. ونموذج تحليل الانحدار المعياري يعرف بالصيغة الآتية:

$$Y_i^* = \beta_1^* X_{1i}^* + \beta_2^* X_{2i}^* + \dots + \beta_k^* X_{ki}^* + e_i^* \dots\dots\dots(19-2)$$

حيث أن:

$$X_{ij}^* = (X_{ij} - \bar{X}_j) / S_{xj} \quad , (j = 1, 2, 3, \dots, k)$$

¹⁸ - جون فوكس 1997 ص 106

المتغير التابع المعياري: $Y_i^* = (Y_i - \bar{Y}) / S_y$.

معامل الانحدار الجزئي للمتغير المستقل المعياري رقم z : $\beta_j^* = \beta_j \frac{S_x}{S_y}$.

الباقى المعياري وله وسط حسابي يساوى الصفر وانحراف معياري يختلف عن الواحد $e_i^* = \frac{e_i}{S_y}$.

الصحيح.

S_y : الانحراف المعياري للمتغير التابع، S_{xy} تمثل الانحراف المعياري للمتغير X_j . ومن المعادلة

(2-19) نجد أن النموذج المعياري المقدر يعرف بالصيغة الآتية:

$$\hat{Y}_i^* = \hat{\beta}_1^* X_{1i}^* + \hat{\beta}_2^* X_{2i}^* + \dots + \hat{\beta}_k^* X_{ki}^* \quad \dots\dots\dots(20-2)$$

وبما أن $\hat{\beta}^*$ يمثل متجه المعلمات المعيارية المقدرة، عليه سيكون تقدير المربعات الصغرى له بالمعادلة الآتية:

$$\hat{\beta}^* = \left(X'^* X^* \right)^{-1} X'^* Y^* \quad \dots\dots\dots(21-2)$$

و تفسر قيم معاملات الانحدار المعيارية $\hat{\beta}_j^*$ بالتغير في القيمة المتوقعة للمتغير التابع Y^* بوحدات الانحراف المعياري الناتج عن تغير X_j^* بانحراف معياري واحد بافتراض ثبات قيم المتغيرات المستقلة المعيارية الأخرى ، وهكذا يمكن تفسير بقية معاملات الانحدار المعيارية الجزئية الأخرى وبالتالي من الممكن قياس الأثر النسبي للمتغيرات المستقلة ، فيمكننا القول بأن أثر المتغير X_j أكبر أو أقل من أثر المتغير X_h إذا كانت قيمة المعامل المعياري الجزئي للمتغير X_j أكبر أو أقل من قيمة المعامل المعياري للمتغير X_h وهكذا لبقية المتغيرات.

6-2 تقييم معنوية المقدرات و النموذج [19] [20]:

1-6-2 اختبار المعنوية الكلية للنموذج:

إن اختبار معنوية النموذج بصورة كلية يعنى اختبار تأثير المتغيرات المستقلة X_1, X_2, \dots, X_k مجتمعة على المتغير التابع ، وبمعنى آخر هل تؤثر كل المتغيرات المستقلة المضمنة في نموذج الانحدار بمستوي معنوي في التنبؤ بقيم المتغير التابع، وللإجابة على هذا السؤال يمكن اختبار الفرضية الآتية:

فرض العدم: $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$

الفرض البديل: $H_1 : \text{are not equal } \beta'S \text{ At Least two}$

ويتم اختبار الفرضية أعلاه باستخدام اختبار F والذي يأخذ الصيغة التالية:

$$F = \frac{SSR / K}{SSE / n - k - 1} = \frac{R^2 / K}{(1 - R^2) / (n - k - 1)} \quad \dots\dots\dots(23-2)$$

حيث أن:

SSR : يمثل مجموع مربعات الانحدار.

SSE : يمثل مجموع مربعات الخطأ.

R^2 : معامل التحديد ، K : عدد المتغيرات المستقلة ، n : عدد المشاهدات.

وبعد حساب قيمة F تقارن مع القيمة الجدولية $F_{\alpha, k, n-k-1}$ ، فإذا كانت $F \leq F_{\alpha, k, n-k-1}$

فانه تقبل فرضية العدم H_0 وهذا يعنى أنه لا يوجد تأثير معنوي من قبل المتغيرات المستقلة

مجتمعة على المتغير المعتمد. أما إذا كانت $F > F_{\alpha, k, n-k-1}$ فانه ترفض فرضية العدم وتقبل

البديلة H_1 وهذا يعنى أن المتغيرات المستقلة مجتمعة تؤثر على النموذج.

¹⁹ - إبراهيم، بسام يونس وآخرون، (2002)، مرجع سابق، ص:187-192
²⁰ - - إسماعيل، محمد عبد الرحمن، (2001)، مرجع سابق، ص:177 - 180.

2-6-2 اختبار معاملات الانحدار الجزئية $\hat{\beta}_j$:

في حالة وجود تأثير معنوي من قبل المتغيرات المستقلة مجتمعة على المتغير المعتمد، يمكن اختبار تأثير كل متغير مستقل بوجود باقي المتغيرات المستقلة الأخرى ، وهذا يعنى بدوره أن باقي المتغيرات يفترض لها تأثير ثابت على المتغير المعتمد. ولنفترض أن المطلوب اختبار عدم وجود تأثير معنوي من قبل المتغير X_j على المتغير المعتمد بافتراض أن تأثير باقي المتغيرات المستقلة ثابت، وهذا يعنى اختبار الفرضية الآتية:

$$H_0 : \beta_j = 0 \setminus \beta_1, \beta_2, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_k \quad \text{فرض العدم:}$$

$$H_1 : \beta_j \neq 0 \setminus \beta_1, \beta_2, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_k \quad \text{الفرض البديل:}$$

ويستخدم المختبر t لاختبار هذه الفرضية والذي يعرف بالصيغة الآتية:

$$t_j = \frac{\hat{\beta}_j}{S.E(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{\sqrt{Var(\hat{\beta}_j)}} \quad , \quad j = 1, 2, \dots, k \quad \dots\dots\dots(24-2)$$

حيث أن:

$$S.E(\hat{\beta}_j) : \text{يمثل الخطأ القياسي للمقدر } \hat{\beta}_j .$$

وتقارن قيمة t المحسوبة مع الجدولية $t_{n-k-1, \frac{\alpha}{2}}$ فإذا كانت المحسوبة أكبر من الجدولية ،

فهذا يعنى رفض فرضية العدم وقبول الفرضية البديلة، أي هناك تأثيراً معنوياً من قبل المتغير المستقل X_j على المتغير المعتمد. أما إذا كانت قيمة t المحسوبة أقل من الجدولية، هذا يعنى أن المتغير المستقل X_j ليس له تأثيراً معنوياً على المتغير المعتمد.

3-6-2 اختبار F الجزئي المتعدد **Multiple-Partial F-test**:

إن اختبار t السابق تم استخدامه لاختبار مساهمة إضافة متغير واحد يسهم بمستوي

معنوي في التنبؤ بقيم المتغير التابع ويسمى أحياناً باختبار F الجزئي الأحادي، أما اختبار F

الجزئي المتعدد فيستخدم لتقييم مساهمة اثنين أو أكثر من المتغيرات المستقلة للمتغيرات المضمنة أصلاً في نموذج الانحدار، أي أننا نود اختبار ما إذا كان إضافة K متغير مستقل (Z_1, Z_2, \dots, Z_k) تسهم بمستوي معنوي في التنبؤ بقيم المتغير التابع بوجود المتغيرات المستقلة (X_1, X_2, \dots, X_p) . ولإجراء هذا الاختبار يتم بناء نموذج الانحدار الكامل التالي:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \beta_1^* Z_1 + \beta_2^* Z_2 + \dots + \beta_k^* Z_k + U \quad \dots\dots\dots(25-2)$$

وبناء نموذج الانحدار المخفض التالي:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + U \quad \dots\dots\dots(26-2)$$

والفرض المراد اختباره هو:

$$H_0 : \beta_1^* = \beta_2^* = \dots = \beta_k^* = 0 \quad \text{فرض العدم:}$$

$$H_1 : \beta_1^* \neq \beta_2^* \neq \dots \neq \beta_k^* \neq 0 \quad \text{الفرض البديل:}$$

ولاختبار هذه الفرضية تستخدم إحصائية الاختبار التالية:

$$F_0(Z_1, Z_2, \dots, Z_k / X_1, X_2, \dots, X_p) = \frac{[R^2(ful) - R^2(reduced)]/k}{1 - R^2(ful)/(n - p - k - 1)} \quad \dots\dots\dots(27-2)$$

حيث أن:

$R^2(ful)$: معامل التحديد للنموذج الكامل.

$R^2(reduced)$: معامل تحديد النموذج المخفض.

K : عدد المتغيرات المضافة المراد اختبار مساهمتها في تفسير المتغير التابع.

P : عدد المتغيرات المستقلة المضمنة أصلاً في النموذج.

ويتم رفض فرض العدم H_0 إذا كانت قيمة F_0 أكبر من الجدولية، أي: $F_0 > F_{\alpha, k, n-p-k-1}$.

7-2 اختبار جودة النموذج:

1-7-2 معامل التحديد Coefficient of Determination:

يستخدم معامل التحديد للحكم علي المقدرة التفسيرية للنموذج، ويعرف بأنه نسبة مساهمة المتغيرات المستقلة في إحداث التغيرات التي تطرأ على المتغير المعتمد، ويحسب من خلال نسبة مجموع مربعات الانحدار إلى مجموع المربعات الكلي ، أي:

$$R^2 = \frac{SSR}{SST} = \frac{\hat{\beta}_1 \sum x_{1i} y_i + \hat{\beta}_2 \sum x_{2i} y_i + \dots + \hat{\beta}_k \sum x_{ki}}{\sum y_i^2} \dots\dots\dots(28-2)$$

وحيث أن $(-1 \leq R \leq 1)$ فإن $(0 \leq R^2 \leq 1)$ ومن خلال ذلك فإن $1 - R^2$ تمثل نسبة مساهمة حد الخطأ في النموذج في إحداث التغيرات التي تطرأ على المتغير المعتمد. وعندما تكون $R^2 = 0$ هذا يعني أن $SSR = 0$ ويعنى ذلك أن $SST = SSE$ ، أي أن قيم البواقي e_1, e_2, \dots, e_k تكون كبيرة جداً ويشير ذلك إلى ابتعاد القيم الحقيقية Y_i عن القيم المقدرة \hat{Y}_i وتحدث $SSR = 0$ إما في حالة كون جميع قيم المتغيرات المستقلة متساوية أو أن هذه القيم تكون أصفاراً. وعندما تكون $R^2 = 1$ فهذا يعني أن $SST = SSR$ وبالتالي فإن $SSE = 0$ وحيث أن $SSE = \sum e_i^2$ فهذه الكمية لا يمكن أن تساوى صفراً إلا في حالة كون جميع قيم البواقي مساوية للصفر وفي هذه الحالة فإن القيم الحقيقية Y_i ستطبق على القيم المقدرة \hat{Y}_i ويكون معامل الارتباط ما بين المتغير المعتمد و المتغير المستقل تماماً أي مساوياً (± 1) ، هذا يعني أن نموذج الانحدار المتناول لا يتضمن حد الخطأ U_i أي أن المعادلة الناتجة لا تمثل نموذج انحدار لان من شروط نموذج الانحدار تضمنه لحد الخطأ وهذه الحالة تمثل نقطة الفصل بين علمي الرياضيات و الإحصاء، حيث أن علم الإحصاء قائم على وجود عنصر الخطأ في التجارب أما علم الرياضيات فنكون فيه العلاقات تامة بين المتغيرات. ومن المآخذ على معامل التحديد أنه يتأثر بعدد المتغيرات

المستقلة، حيث تلاحظ أن قيمته تزداد مع إضافة أي متغير مستقل جديد للنموذج. ولتلاشى هذا القصور تعين أن نُصحح قيمة R^2 بحيث لا تتأثر بعدد المتغيرات المستقلة وذلك عن طريق أخذ درجات الحرية في الحسبان، حيث أن درجات الحرية $n - k$ تقل مع زيادة عدد المتغيرات المستقلة وثبات حجم العينة، عليه تصبح صيغة معامل التحديد المعدل الذي رمز له بـ \bar{R}^2 بهدف إزالة التضخم في قيمة R^2 كالآتي:

$$\bar{R}^2 = 1 - (1 - R^2) \left(\frac{n-1}{n-k-1} \right) \quad \dots\dots\dots(29-2)$$

2-7-2 معامل التحديد الجزئي **Coefficients of Partial Determination**:

يقيس معامل التحديد الجزئي المساهمة الحدية لمتغير مستقل واحد X_i في تفسير التباين أو التغير في المتغير التابع ، عندما تكون المتغيرات الأخرى مضمنة في النموذج. ويمكن إيجاد معامل التحديد الجزئي بين المتغير Y والمتغير المستقل X_1 بعد استبعاد أثر المتغيرات X_2, X_3, \dots, X_p كما يلي:

$$r_{YX_1 / X_2, X_3, \dots, X_p}^2 = \frac{SSR(X_2, X_3, \dots, X_p) - SSR(X_1, X_2, \dots, X_p)}{SSR(X_1, X_2, \dots, X_p)} \quad \dots\dots\dots(30-2)$$

8-2 اختبار ثبات معاملات النموذج:

يستخدم اختبار معنوية معاملات الارتباط الذاتي لمعرفة مدى استقرارية النموذج المقدر خلال فترة الدراسة، حيث تكون معاملات الارتباط في النموذج المستقر تتبع توزيع طبيعي حيث $P_k \sim N(0, 1/n)$ بمستوى معنوية 5% فان حدود فترة الثقة هي $\pm 1.96\sqrt{1/N}$ فإذا وقع \hat{P}_k خارج هذا المجال فان النموذج غير مستقر. ولاختبار المعنوية الكلية لمعاملات الارتباط الذاتي تستخدم إحصائية Ljung-Box لاختبار الفرضية التالية:

$$H_0 : P_{k1} = P_{k2} = \dots\dots\dots P_{kn} = 0 \quad \text{فرض العدم:}$$

$$H_1 : P_{k1} \neq P_{k2} \neq \dots \neq P_{kn} \neq 0$$

الفرض البديل:

وتعرف إحصائية الاختبار بالعلاقة التالية:

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{P}_k}{n-k} \quad \dots\dots\dots(31-2)$$

حيث أن:

m : درجات الحرية ، n : حجم العينة ، \hat{P}_k : معامل الارتباط الذاتي.

فإذا كانت قيمة Q المحسوبة أقل من قيمة $\chi^2_{(m)}$ الجدولية نقبل فرض العدم ونرفض

الفرض البديل ، أي أن النموذج مستقر. أما إذا كانت Q أكبر من قيمة $\chi^2_{(m)}$ نرفض فرض

العدم ونقبل الفرض البديل، أي أن النموذج غير مستقر.

9-2 التنبؤ Forecasting²¹:

بعد تقييم معاملات النموذج الخطي المتعدد واختبار صلاحية معالمته يصبح صالح للتنبؤ

بالقيم المستقبلية لقيم المتغير التابع ، وتسمى هذه المرحلة بمرحلة التنبؤ. وليكن النموذج المقدر

خلال العينة n علي الشكل $\hat{Y} = X\hat{\beta}$ ويكون التنبؤ بالفترة m في المستقبل كما يلي:

$$\hat{Y}_{n+1} = \hat{\beta}_0 + \hat{\beta}_1 X_{1(n+1)} + \dots + \hat{\beta}_k X_{k(n+1)} \quad \text{التنبؤ لفترة واحدة في المستقبل:}$$

$$\hat{Y}_{n+2} = \hat{\beta}_0 + \hat{\beta}_1 X_{1(n+2)} + \dots + \hat{\beta}_k X_{k(n+2)} \quad \text{أما التنبؤ لفترتين في المستقبل:}$$

.... ..

.... ..

.... ..

$$\hat{Y}_{n+m} = \hat{\beta}_0 + \hat{\beta}_1 X_{1(n+m)} + \dots + \hat{\beta}_k X_{k(n+m)} \quad \text{والتنبؤ للفترة } m \text{ في المستقبل:}$$

عليه يمكن كتابة النموذج الخطي العام المتنبأ به كما يلي:

²¹ - تومي صالح، (2001)، "مدخل لنظرية القياس الإقتصادي"، الجزء الثاني، ديوان المطبوعات، الجزائر.

$$Y_n^m = X_n^m \beta + U_n^m \quad \dots\dots\dots(32-2)$$

حيث أن:

Y_n^m : متجه عمود من الدرجة $m \times 1$ ، X_n^m : مصفوفة من الدرجة $m \times k$ ، U_n^m : متجه عمود من الدرجة $m \times 1$.

كما يمكن كتابة النموذج المقدر للمعادلة (32-2) بالصيغة التالية:

$$\hat{Y}_n^m = X_n^m \hat{\beta} \quad \dots\dots\dots(33-2)$$

ويكون التنبؤ بالنموذج المقدر بالمعادلة (32-2) أفضل تنبؤ خطي غير متحيز حيث أن وسطه وتباينه هما:

$$E[\hat{Y}_n^m] = X_n^m E[\hat{\beta}] = X_n^m \beta \quad \dots\dots\dots(34-2)$$

$$Var[\hat{Y}_n^m] = E\left[(\hat{Y}_n^m - X_n^m \beta)(\hat{Y}_n^m - X_n^m \beta)'\right] \quad \dots\dots\dots(35-2)$$

ويمكن التنبؤ بمجال \hat{Y}_n^m بفترة ثقة $(1-\alpha)\%$ كما يلي:

$$X_n^m \hat{\beta} \pm t_{\alpha/2} \sigma_u \sqrt{X_n^m (XX)^{-1} X_n^{t'm}} \quad \dots\dots\dots(36-2)$$

2-10 تقييم أداء النموذج في التنبؤ^[23][²²]:

Predictive Evaluation of Econometric Forecasting Models

إذا اجتاز النموذج كل الاختبارات السابقة وظهر أنه خالٍ من المشكلات القياسية التي سوف نتناولها بالشرح في الفصل القادم ، عندها يمكن استخدامه للتنبؤ بقيم المتغير التابع لفترات مختارة داخل العينة ومقارنتها بالقيم الفعلية الموجودة بالعينة، فكلما صغرت الاختلافات بين القيم الحقيقية والقيم المتنبأ بها دل ذلك على جودة النموذج في التنبؤ ، وهناك عدة مقاييس تستخدم لقياس دقة التنبؤات التي يتم الحصول عليها من النموذج القياسي منها ما يلي:

²² - هتمات، سعيد (2006)، "دراسة اقتصادية وقياسية لظاهرة التضخم في الجزائر"، رسالة ماجستير.
²³ - إبراهيم، بسام يونس وآخرون، (2002)، مرجع سابق.

2-10-1 معامل التساوي لثايل Theil's Inequality Coefficient

اقترح ثايل المعيار التالي:

$$U = \sqrt{\frac{\sum_{i=1}^{n-1} (P_i - A_i)^2}{\sum_{i=1}^{n-1} A_i^2}} \quad \dots\dots\dots(37-2)$$

حيث أن: P_i : تمثل التغير في القيم المتنبأ بها نسبةً إلى القيم الفعلية ، أي أن:

$$P_i = \frac{\hat{Y}_{i+1} - Y_i}{Y_i} \quad \dots\dots\dots(38-2)$$

A_i : تمثل التغير النسبي في القيم الفعلية ، أي أن:

$$A_i = \frac{Y_{i+1} - Y_i}{Y_i} \quad \dots\dots\dots(39-2)$$

واعتماداً على المعادلة (37-2) فإن قيمة U تأخذ المدى $0 \leq U < \infty$ وكلما كانت قيمة U

قريبة من الصفر فإن ذلك يعنى اقتراب القيم التنبؤية \hat{Y}_i من القيم الفعلية Y_i وفى هذه الحالة تكون القوة التنبؤية للنموذج عالية ، أي أن تنبؤاته تكون دقيقة جداً.

أما إذا كانت قيمة U كبيرة فهذا يعنى وجود اختلاف كبير بين قيم P_i, A_i وبالتالي تباعد

القيم التنبؤية \hat{Y}_i عن القيم الفعلية Y_i وفى هذه الحالة فإن النموذج تكون تنبؤاته سيئة. وإذا كانت

$P_i = A_i$ فإن ذلك يؤدي إلى $U = 0$ وهنا تكون التنبؤات مضبوطة ولا تعاني من أية أخطاء.

أما إذا كانت $P_i = 0$ فإن ذلك يجعل $U = 1$ وفى هذه الحالة فإن تنبؤات النموذج تكون غير

مقبولة بدرجة عالية. وإذا وجدت أن $U > 1$ فتكون التنبؤات للنموذج سيئة جداً ومرفوضة. عموماً

تقبل تنبؤات النموذج إذا كانت $0 \leq U < 1$ وكلما اقتربت قيمة U من الصفر كانت التنبؤات أكثر

دقة.

2-10-2 الجذر التربيعي لمتوسط مربع الخطأ Root Mean Square Error:

يعد مقياس الجذر التربيعي لمتوسط مربع الخطأ RMSE من أهم مؤشرات جودة مطابقة النموذج للبيانات. فإذا بلغت قيمته (0.05) فأقل دل ذلك علي أن النموذج يطابق تماماً البيانات ، أما إذا كانت القيمة محصورة بين (0.05 و 0.08) دل ذلك علي أن النموذج يطابق بدرجة كبيرة بيانات العينة ، وإذا زادت قيمته عن (0.08) في هذه الحالة يتم رفض النموذج. ويحسب هذا المعيار بالصيغة الآتية:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \dots\dots\dots(40-2)$$

3-10-2 متوسط الخطأ المطلق Mean Absolute Error (MAE):

ويمثل متوسط الفرق بين قيم المشاهدات وقيم التنبؤ لنفس الفترة الزمنية ويحسب هذا المعيار من خلال الصيغة التالية:

$$MAE = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i|}{n} \quad \dots\dots\dots(41-2)$$

وكلما اقتربت قيمة MAE من الصفر يعني هذا أن كفاءة النموذج عالية في التنبؤ.

3-10-2 متوسط الخطأ النسبي المطلق Mean Absolute Percentage Error:

يمكن حساب المؤشر السابق بشكل نسبي من خلال الصيغة التالية:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100 \quad \dots\dots\dots(42-2)$$

2-11 مشكلة عدم تجانس التباين Heteroscedasticity Problem:

من الفروض الأساسية التي تقوم عليها طريقة المربعات الصغرى OLS لتقدير معالم نموذج الانحدار الخطي هو ثبات تباين حد الخطأ العشوائي عند كل مستوي من مستويات المتغير أو المتغيرات المستقلة، يعرف هذا الفرض بفرض ثبات التباين²⁴، أي أن:

$$Var(U_i) = E[U_i - E(U_i)]^2 = E(U_i)^2 = \sigma_u^2 \quad \dots\dots\dots(43-2)$$

ومعني فرض ثبات تجانس التباين للمتغير العشوائي هو أن تقلب كل من U حول متوسطها الصغرى لا يعتمد على قيم المتغير المستقل X وأن تباين كل من U يبقي ثابتاً بغض النظر عن القيم الكبيرة أو الصغيرة للمتغير المستقل حيث أن σ_u^2 ليست دالة في المتغير المستقل أي أن:

$$\sigma_u^2 \neq f(X_i) \quad .^{25}$$

وعندما لا يتحقق هذا الفرض توجد مشكلة تعرف باختلاف التباين أو عدم ثبات التباين Heteroscedasticity حيث تكون تباينات حدود الخطأ مختلفة، أي أن²⁶:

$$Var(U_i) = E(U_i)^2 = \sigma_{u_i}^2 \quad \dots\dots\dots(44-2)$$

حيث يشير الحرف i إلي أن لكل قيمة من قيم الخطأ العشوائي U_i لها تباين مختلف عن تباينات باقي القيم. وعندما يكون تباين الخطأ σ_u^2 غير ثابت فإنه يمثل داله للمتغير (المتغيرات) المستقلة أي أن: $\sigma_{u_i}^2 = f(X_i)$.

وهناك عدة أسباب تجعل تباين الخطأ العشوائي غير متجانس منها ما يلي²⁷:

²⁴ - إسماعيل، محمد عبد الرحمن، (2001)، "تحليل الانحدار الخطي"، معهد الإدارة العامة، المملكة العربية السعودية، مركز البحوث.
²⁵ - الطيب، عز الدين مالك، (2008)، "المدخل إلى الاقتصاد القياسي"، الطبعة الأولى، مطبعة حي تاون، الخرطوم، السودان.
²⁶ - إسماعيل، محمد عبد الرحمن، (2001)، "تحليل الانحدار الخطي"، معهد الإدارة العامة، المملكة العربية السعودية، مركز البحوث.
²⁷ - إبراهيم، بسام يونس وآخرون، (2002)، "الاقتصاد القياسي"، الطبعة الأولى، دار عزة للنشر والتوزيع، الخرطوم، السودان.

1. زيادة تعلم الأفراد: إذا زاد تعلم الأفراد فإن الأخطاء التي تترتب علي سلوكهم الشخصي سوف تقل عبر الزمن، ومن ثم فإن تباين حد الخطاء سوف يقل لذلك فإن $\sigma_{u_i}^2$ تتناقص بمرور الزمن.

2. إذا زادت دخول الأفراد فإن $\sigma_{u_i}^2$ سوف يزداد وذلك لأن الأفراد سيكون لهم اختيارات متعددة حول إنفاق هذه الزيادة، فتباين الإنفاق علي الغذاء بين العائلات يمكن أن يزيد بزيادة دخل العائلة ونفس الشيء بالنسبة للريح الكبير الذي يخلق أمام الشركة خيارات عديدة أكثر من الشركة ذات الريح القليل.

3. تحسين أساليب جمع البيانات: إذا تحسنت أساليب جمع البيانات والمعلومات فإن الأخطاء سوف تقل وبالتالي فإن $\sigma_{u_i}^2$ سوف تقل، فمثلاً الأخطاء التي ترد بالمستندات في المصارف التي تتوفر فيها وسائل علمية وتقنية لتحليل البيانات تكون أقل من مثيلاتها في المصارف التي لا تتوفر فيها مثل هذه الوسائل. وتظهر مشكلة عدم تجانس التباين في البيانات المقطعية Cross-sectional data أكثر من بيانات السلاسل الزمنية Time series data.

2-11-1 النتائج المترتبة علي عدم تجانس التباين²⁸:

Consequences of Heteroscedasticity

يترتب علي وجود مشكلة عدم تجانس التباين الآتي:

1. لا يمكن الحصول علي نتائج مؤكدة لتأثير المشكلة علي اختبارات الفروض.
2. الحصول علي تباينات كبيرة للتقديرات وبالتالي فلن تتمتع التقديرات المتحصل عليها بخاصية الأقل تباين أي، لن تكون هذه التقديرات ذات كفاءة سواء في حالة العينات الصغيرة أو الكبيرة، ومن ناحية أخرى فإن كبر تباينات التقديرات يؤدي إلي اتساع فترة الثقة للمعالم.

²⁸- الطيب، عز الدين مالك، (2008)، "المدخل إلي الاقتصاد القياسي"، الطبعة الأولى، مطبعة حي تاون، الخرطوم، السودان.

3. بالرغم من تخلف الفرض الخاص بثبات تباين الأخطاء فإن التقديرات المتحصل عليها تكون غير متحيزة إحصائياً إذ أن تحيز التقدير من عدمه لا يعتمد علي هذا الفرض.

4. التنبؤ بقيمة Y عند أي قيمة معطاة للمتغير المستقل X والذي يعتمد علي التقديرات من البيانات الأصلية يكون له تباين كبير ، بمعنى أن التنبؤ يكون غير كفؤ وذلك لا تباين التنبؤ يتضمن المشكلة بالداخل.

2-11-2 اكتشاف عدم تجانس التباين:

Detection of Heteroscedasticity

هنالك عدة طرق تستخدم للكشف عن مشكلة عدم تجانس التباين ، منها ما يلي:

2-11-2-1 اختبار بارك Park Test²⁹:

يقترح بارك العلاقة التالية بين تباين حد الخطأ $\sigma_{u_i}^2$ والمتغير المستقل X_i :

$$\sigma_i^2 = \sigma^2 X_i^\beta e^{V_i} \quad \dots\dots\dots (45 - 2)$$

وبأخذ لوغاريتم الطرفين للمعادلة (45 - 2) نحصل علي:

$$\ln \sigma_i^2 = \ln \sigma^2 + \beta \ln X_i + V_i \quad \dots\dots\dots (46 - 2)$$

حيث أن V_i حد خطأ عشوائي. وبما أن قيمة σ_i^2 غير معلومة يقترح بارك استخدام مربع البواقي e_i^2 كمقرب وإجراء نموذج الانحدار التالي:

$$\ln e_i^2 = \alpha + \beta \ln X_i + V_i \quad \dots\dots\dots (47 - 2)$$

فإذا كانت β معنوية إحصائياً دل ذلك على وجود مشكلة عدم ثبات التباين. أما إذا كانت β غير معنوية إحصائياً ، فإنه يمكن القول بان تباين حد الخطأ ثابت.

²⁹ -Park, R. E.,(1969), "Estimation with Heteroscedastic Error Terms. Econometric, Vol. 34. No. 4p. 888.

ويلاحظ الآتي علي اختبار بارك:

1. يتضمن اختبار بارك خطوتين عند تطبيقه هما: الخطوة الأولى يتم فيها بناء نموذج الانحدار للحصول علي البواقي ، وفي الخطوة الثانية يتم إجراء لوغاريتم مربع البواقي علي لوغاريتم المتغير المستقل حسب المعادلة (2 - 47).
2. يعاب علي اختبار بارك أن حد الخطأ العشوائي V_i في المعادلة (2 - 46) أو (2 - 47) قد لا يستوفي فرضيات المربعات الصغرى والتي من ضمنها فرضية ثبات التباين³⁰.
3. يستخدم اختبار بارك للكشف عن عدم ثبات التباين في نموذجي الانحدار البسيط والانحدار المتعدد. وفي حالة نموذج الانحدار الخطي المتعدد يتم حساب البواقي ويجري انحدار لوغاريتم مربع البواقي مع المتغير المستقل X_i مصدر اختلاف التباين والذي يتغير معه التباين.

2-2-11-2 اختبار معامل ارتباط الرتب لسبيرمان³¹:

- يعتبر هذا الاختبار من أبسط أنواع اختبارات تجانس التباين ويمكن تطبيقه في حالة العينات الصغيرة والكبيرة علي حد سواء واستخدام معامل ارتباط الرتب بدلاً من معامل ارتباط بيرسون ، يرجع إلي كون معامل ارتباط بيرسون بين الخطأ العشوائي المقدر e_i والمتغير المستقل X_i صفرًا استناداً إلي أحد افتراضات نموذج الانحدار الخطي الخاص بكون المتغير العشوائي مستقل عن المتغير المستقل. ويتم تطبيق هذا الاختبار بإتباع الخطوات التالية:
1. توفيق نموذج الانحدار الخطي المتعدد:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki} \quad \dots\dots\dots (48 - 2)$$

³⁰-Gujarati, D. N. (1988), "Basic Econometrics", (2nd edition), McGraw-Hill Book Company, New York.
³¹ إبراهيم، بسام يونس وآخرون، (2002)، "الاقتصاد القياسي"، الطبعة الأولى، دار عزة للنشر والتوزيع، الخرطوم، السودان.

2. نحسب معامل ارتباط الرتب لكل متغير مستقل مع $|e_i|$ بمعنى آخر سنحسب k من

معاملات ارتباط الرتب لـ $(e_i, X_{1i}), (e_i, X_{2i}), \dots, (e_i, X_{ki})$. حيث: $e_i = Y_i - \hat{Y}_i$.

3. نختبر كل معامل اختبار r_s حسب الصيغة الآتية:

$$t = \frac{r_s \sqrt{n - k - 1}}{\sqrt{1 - r_s^2}} \quad \dots\dots\dots (49 - 2)$$

ولاختبار فرضية العدم والبدلية التالية:

فرض العدم: تباين الخطأ العشوائي متجانس: H_0

الفرض البديل: تباين الخطأ العشوائي غير متجانس: H_1

وتقارن قيمة t المحسوبة مع الجدولية $t_{n-k-1, \alpha/2}$ فإذا ما ثبت أنه علي الأقل احدي قيم t معنوية

(رفض فرضية العدم) فهذا يشير إلي تجانس التباين. أما إذا كانت جميع قيم t غير معنوية (قبول

فرضية العدم) فذلك دليل علي كون تباين الخطأ العشوائي متجانس.

2-11-3 اختبار جولدفيلد-كوانت Goldfeld-Quandt Test³²:

يعد اختبار جولدفيلد-كوانت من الاختبارات الشائعة الاستخدام في الكشف عن وجود علاقة

طردية أو عكسية بين حد الخطأ وأحد المتغيرات المستقلة.

ويتلخص اختبار جولدفيلد-كوانت في الخطوات التالية:

1. ترتيب المشاهدات حسب قيم المتغير المستقل X_i تصاعدياً ، وفي حالة نموذج الانحدار

المتعدد يتم الترتيب حسب المتغير المستقل مصدر اختلاف التباين.

³²-Goldfeld, S.M. and Quandt, R. E.,(1965), Some Test for Homoscedasticity. Journal of the American Statistical Association, 60pp.539-547.

2. يتم استبعاد عدد d مشاهدة من الوسط في حدود ربع عدد المشاهدات ($d \cong n/4$) أو 20% ومن ثم يتم تقسيم بقية المشاهدات إلى مجموعتين بحيث يكون عدد كل مجموعة يساوي $(n-d)/2$ مشاهدة.

3. يتم بناء نموذجين، أحدهما للقيم الصغيرة للمتغير X والآخر للقيم الكبيرة ويتم الحصول علي مجموع مربعات البواقي للنموذجين RSS_1 ، RSS_2 على التوالي.

4. يستخدم اختبار جولدفيلد-كواندت للكشف عن نوعين من عدم ثبات التباين هما:

النوع الأول: تباين حد الخطأ دالة تزايدية للمتغير المستقل والفرض المراد اختباره هو:

فرض العدم: تباين حد الخطأ متجانس ، ضد الفرض البديل: تباين حد الخطأ دالة تزايدية للمتغير X_i .

ولإجراء هذا الاختبار تستخدم الإحصائية التالية:

$$F_1 = \frac{RSS_2/V_2}{RSS_1/V_1} = \frac{\sum_{i=1}^{n-d} e_{2i}^2}{\sum_{i=1}^{n-d} e_{1i}^2} \sim F_{V_2, V_1} \dots\dots\dots (50 - 2)$$

حيث أن:

RSS_1 : مجموع مربعات البواقي لنموذج مشاهدات المجموعة الولي ، RSS_2 : مجموع مربعات

البواقي لنموذج مشاهدات المجموعة الثانية ، d : عدد المتغيرات المستقلة، $V_1 = V_2$: درجات

الحرية للسط والمقام وتساوى:

$$V_1 = V_2 = (n - d)/2 - p - 1 \dots\dots\dots (51 - 2)$$

وتتوزع الإحصاءة F_1 حسب توزيع F بدرجتي حرية V_1, V_2 . فإذا كانت قيمة F_1 المحسوبة

أكبر من قيمة F الجدولية فإننا نرفض فرض العدم ويتم قبول الفرض البديل القائل بأن تباين حد

الخطأ يتزايد بزيادة قيم المتغير المستقل. أما إذا كانت قيمة F_1 المحسوبة أقل من قيمة F

الجدولية فيتم قبول فرض العدم الذي ينص علي أن تباين حد الخطأ متجانس.

النوع الثاني: تباين حد الخطأ دالة تناقصية للمتغير X_i والفرض المراد اختباره هو:

فرض العدم: تباين حد الخطأ متجانس أو ثابت ، ضد الفرض البديل: تباين حد الخطأ دالة تناقصية للمتغير X_i .

ولإجراء هذا الاختبار تستخدم الإحصائية التالية:

$$F_2 = \frac{RSS_1/V_1}{RSS_2/V_2} = \frac{\sum_{i=1}^{n-d} e_{1i}^2}{\sum_{i=1}^{n-d} e_{2i}^2} \sim F_{V_1, V_2} \quad \dots\dots\dots (52 - 2)$$

وتتوزع الإحصاءة F_2 حسب توزيع F بدرجتي حرية V_2, V_1 فإذا كانت قيمة F_2 المحسوبة أكبر من قيمة F الجدولية فإننا نرفض فرض العدم ويتم قبول الفرض البديل القائل بان تباين حد الخطأ يتناقص بزيادة قيم المتغير المستقل. أما إذا كانت قيمة F_2 المحسوبة أقل من قيمة F الجدولية فيتم قبول فرض العدم الذي ينص علي أن تباين حد الخطأ متجانس.

2-11-4 اختبار جليجر Glejser Test³³:

يعد اختبار جليجر شبيه باختبار بارك فبعد الحصول علي البواقي e_i من نموذج الانحدار الأصلي ، يقترح جليجر إجراء انحدار القيم المطلقة للبواقي ($|e_i|$) علي المتغير المستقل X_i الذي يتغير معه التباين (σ_i^2) . وقد أستخدم جليجر الأشكال الدالية التالية:

$$|e_i| = \alpha + \beta X_i + V_i \quad \dots\dots\dots (53 - 2)$$

$$|e_i| = \alpha + \beta \sqrt{X_i} + V_i \quad \dots\dots\dots (54 - 2)$$

$$|e_i| = \alpha + \beta \frac{1}{X_i} + V_i \quad \dots\dots\dots (55 - 2)$$

$$|e_i| = \alpha + \beta \frac{1}{\sqrt{X_i}} + V_i \quad \dots\dots\dots (56 - 2)$$

$$|e_i| = \sqrt{\alpha + \beta X_i} + V_i \quad \dots\dots\dots (57 - 2)$$

³³ - Glejser, H.,(1979), A New Test for Heteroscedasticity. Journal of the American Statistical Association, Vol. 64,pp.316-323.

$$|e_i| = \sqrt{\alpha + \beta X_i^2} + V_i \quad \dots\dots\dots (58 - 2)$$

حيث أن V_i حد الخطأ العشوائي.

فإذا كانت قيمة β معنوية إحصائياً دل ذلك علي وجود مشكلة عدم ثبات التباين. أما إذا كانت قيمة β غير معنوية إحصائياً فإن النموذج يستوفي شرط ثبات التباين. ويعاب علي طريقة جليجر أن القيمة المتوقعة للحد العشوائي قد تختلف عن الصفر، وكذلك قد يكون تباين حد الخطأ غير ثابت فضلاً أن المعادلتين (2-56) و (2-57) غير خطية المعالم الأمر الذي يجعل طريقة المربعات الصغرى الاعتيادية لتقدير معالمها مستحيلاً. هذا وقد وجد جليجر أن المعادلات الأربعة من (2-53) إلي (2-56) تحقق نتائج جيدة يعتد بها في الكشف عن عدم ثبات التباين.

2-11-3 معالجة عدم تجانس التباين:

Remedy of Heteroscedasticity

ذكرنا سابقاً من الآثار المترتبة علي وجود عدم ثبات تباين حد الخطأ العشوائي الحصول علي مقدرات تتميز بخاصية عدم التحيز و الاتساق إلا أنها لن تتمتع بخاصية أقل تباين أي، أنها تصبح غير ذات كفاءة ، وحسب جون فوكس³⁴ أن طريقة المربعات الصغرى تقل كفاءتها إذا كانت نسبة أكبر تباين لأقل تباين لحد الخطأ تساوي (10) فأكثر، كما يترتب علي اختلال خاصية الكفاءة أن تكون قيم الأخطاء المعيارية لمعاملات الانحدار المقدره غير صحيحة. ولمعالجة مشكلة عدم تجانس تباين حد الخطأ توجد عدد من الطرق نتناول منها بالشرح ما يلي:

³⁴ -Fox, John, (1997), Applied Regression Analysis, Linear Models, and Related Methods. Sage publications.

2-11-3-1 معالجة مشكلة عدم تجانس التباين بإجراء تحويل للنموذج الأصلي³⁵:

تتم معالجة عدم تجانس التباين بهذه الطريقة من خلال إجراء تحويل للنموذج الأصلي، ويتوقف شكل التحويل للنموذج الأصلي علي نمط عدم تجانس حد الخطأ في النموذج الأصلي المقدر. وبافتراض أن النموذج الأصلي كان يأخذ الصيغة الآتية: $Y_i = \beta_0 + \beta_1 X_i + U_i$ فهناك عدة افتراضات لعدم تجانس تباين الخطأ ويختلف النموذج المحول من افتراض آخر.

الافتراض الأول:

$$\sigma_{u_i}^2 = \sigma_u^2 X_i^2 \quad \dots\dots\dots (59 - 2)$$

لمعالجة هذه الحالة يتم قسمة النموذج الأصلي علي X_i (الجذر التربيعي لمعامل σ_u^2 في المعادلة (59 - 2) وكالآتي:

$$\frac{Y_i}{X_i} = \frac{\beta_0}{X_i} + \beta_1 + \frac{U_i}{X_i} \quad \dots\dots\dots (60 - 2)$$

نلاحظ أن حد الخطأ الجديد في النموذج (60-2) هو U_i/X_i وهو متجانس التباين لان:

$$V\left(\frac{U_i}{X_i}\right) = E\left(\frac{U_i}{X_i}\right)^2 - \left[E\left(\frac{U_i}{X_i}\right)\right]^2 = \sigma_u^2 \quad \dots\dots\dots (61 - 2)$$

بعد تقدير النموذج (60 - 2) ينتج:

$$\frac{\hat{Y}_i}{X_i} = \frac{\hat{\beta}_0}{X_i} + \hat{\beta}_1 \quad \dots\dots\dots (62 - 2)$$

نلاحظ أنه في هذا النموذج أن الحد الثابت هو $\hat{\beta}_1$ بينما الميل هو $\hat{\beta}_0$ علي عكس من النموذج

الأصلي المقدر، وللرجوع إلي النموذج الأصلي المقدر يتم ضرب النموذج (62 - 2) بـ X_i .

³⁵ - إبراهيم، بسام يونس وآخرون، (2002)، "الاقتصاد القياسي"، الطبعة الأولى، دار عزة للنشر والتوزيع، الخرطوم، السودان، ص.ص:218-221.

الافتراض الثاني:

$$\sigma_{u_i}^2 = \sigma_u^2 X_i \quad \dots\dots\dots (63 - 2)$$

لمعالجة عدم تجانس التباين في هذه الحالة يتم قسمة النموذج الأصلي علي $\sqrt{X_i}$ ويمكن التحقق من أنه تم التخلص من الظاهرة بملاحظة الآتي:

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\beta_0}{\sqrt{X_i}} + \beta_1 \sqrt{X_i} + \frac{U_i}{\sqrt{X_i}} \quad \dots\dots\dots (64 - 2)$$

$$V\left(\frac{U_i}{\sqrt{X_i}}\right) = E\left(\frac{U_i}{\sqrt{X_i}}\right)^2 - \left[E\left(\frac{U_i}{\sqrt{X_i}}\right)\right]^2 = \sigma_u^2 \quad \dots\dots\dots (65 - 2)$$

بعد تقدير النموذج (64 - 2) نحصل علي:

$$\frac{\hat{Y}_i}{\sqrt{X_i}} = \frac{\hat{\beta}_0}{\sqrt{X_i}} + \hat{\beta}_1 \sqrt{X_i} \quad \dots\dots\dots (66 - 2)$$

نلاحظ أنه في هذا النموذج يمثل نموذج انحدار خطي متعدد بدون مقطع وللرجوع إلي النموذج الأصلي المقدر يتم ضرب النموذج (66 - 2) بـ $\sqrt{X_i}$.

الافتراض الثالث:

$$\sigma_{u_i}^2 = \sigma_u^2 [E(\hat{Y}_i)]^2 \quad \dots\dots\dots (67 - 2)$$

لمعالجة عدم تجانس التباين في هذه الحالة يتم قسمة النموذج الأصلي علي $E(\hat{Y}_i)$

أي $\beta_0 + \beta_1 X_i$ فنحصل علي:

$$\frac{Y_i}{\beta_0 + \beta_1 X_i} = \frac{\beta_0}{\beta_0 + \beta_1 X_i} + \frac{\beta_1 X_i}{\beta_0 + \beta_1 X_i} + \frac{U_i}{\beta_0 + \beta_1 X_i} \quad \dots\dots\dots (68 - 2)$$

وللتأكد من أنه تم التخلص من الظاهرة نحسب تباين الخطأ في النموذج (68 - 2) فنجد أن:

$$V\left(\frac{U_i}{\beta_0 + \beta_1 X_i}\right) = E\left(\frac{U_i}{\beta_0 + \beta_1 X_i}\right)^2 - \left[E\left(\frac{U_i}{\beta_0 + \beta_1 X_i}\right)\right]^2 = \sigma_u^2 \quad \dots\dots\dots (69 - 2)$$

بعد تقدير النموذج (2-68) ولغرض الرجوع إلي النموذج الأصلي المقدر يتم الضرب بـ $(\beta_0 + \beta_1 X_i)$ فنحصل علي نموذج انحدار لا يعاني من ظاهرة عدم تجانس التباين في خطأه.

الافتراض الرابع:

$$\sigma_{u_i}^2 = \sigma_u^2 |e_i| \quad \dots\dots\dots (70 - 2)$$

لمعالجة عدم تجانس التباين في هذه الحالة يتم قسمة النموذج الأصلي علي $\sqrt{|e_i|}$ ويمكن التحقق من أنه تم التخلص من الظاهرة بملاحظة الآتي:

$$\frac{Y_i}{\sqrt{|e_i|}} = \frac{\beta_0}{\sqrt{|e_i|}} + \frac{\beta_1 X_i}{\sqrt{|e_i|}} + \frac{U_i}{\sqrt{|e_i|}} \quad \dots\dots\dots (71 - 2)$$

تباين الخطأ الجديد $\frac{U_i}{\sqrt{|e_i|}}$ هو:

$$V\left(\frac{U_i}{\sqrt{|e_i|}}\right) = E\left(\frac{U_i}{\sqrt{|e_i|}}\right)^2 - \left[E\left(\frac{U_i}{\sqrt{|e_i|}}\right)\right]^2 = \sigma_u^2 \quad \dots\dots\dots (72 - 2)$$

بعد تقدير النموذج (2-71) يتم ضربه بـ $\sqrt{|e_i|}$ فنحصل علي نموذج خالي من الظاهرة.

الافتراض الخامس:

هو عبارة عن أحد التحويلات اللوغاريتمية فكما هو معلوم فإن أخذ اللوغاريتمات للقيم يؤدي إلي تقارب هذه القيم من بعضها البعض وهذا يعني بالتالي انخفاض تباين القيم. ففي هذا الافتراض

يتم اخذ اللوغاريتمات لقيم المتغيرين Y, X فنحصل علي نموذج الانحدار الآتي:

$$\text{Log} Y_i = \beta_0 + \beta_1 \text{Log} X_i + U_i \quad \dots\dots\dots (73 - 2)$$

ويتم تقدير هذا النموذج بطريقة OLS والذي يكون خالياً من الظاهرة.

الافتراض السادس:

وهو حالة عامة ، ففي جميع الحالات السابقة فإن:

$$\sigma_{u_i}^2 = \sigma_u^2 f(X_i) \quad \dots\dots\dots (74 - 2)$$

وللتخلص من عدم تجانس تباين المتغير العشوائي U يتم قسمة جميع حدود نموذج الانحدار

الأصلي على الجذر التربيعي لمعامل σ_u^2 أي على $\sqrt{f(X_i)}$.

2-3-11-2 طريقة المربعات الصغرى المرجحة³⁶:

أولاً: استخدام طريقة المربعات الصغرى عندما تكون قيم σ_i^2 معلومة:

نعلم أن صيغة نموذج الانحدار الخطي المتعدد هي:

$$Y = X\beta + U \quad \dots\dots\dots (75 - 2)$$

وأن تباين حد الخطأ في ظل عدم ثبات التباين يتخذ الصيغة التالية:

$$E(UU^T) = \sigma^2 W \quad \dots\dots\dots (76 - 2)$$

ولعلاج مشكلة عدم ثبات التباين يتم تصحيح نموذج الانحدار الخطي المتعدد وذلك بضرب قبلي

لطرفي المعادلة (75 - 2) بالمصفوفة $W^{-1/2}$ فنحصل على:

$$Y^* = X^* \beta + U^* \quad \dots\dots\dots (77 - 2)$$

حيث أن:

$$Y^* = W^{-1/2} Y \quad \dots\dots\dots (78 - 2)$$

$$X^* = W^{-1/2} X \quad \dots\dots\dots (79 - 2)$$

$$U^* = W^{-1/2} U \quad \dots\dots\dots (80 - 2)$$

$$W^{-1/2} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{\sigma_N} \end{bmatrix} \quad \dots\dots\dots (81 - 2)$$

³⁶ - إسماعيل، محمد عبد الرحمن، (2001)، "تحليل الانحدار الخطي"، معهد الإدارة العامة، المملكة العربية السعودية، مركز البحوث، ص:392-396.

ويلاحظ أن هذه الطريقة عبارة عن تحويل المتغيرات المستقلة والمتغير التابع وذلك بقسمتها علي الانحراف المعياري لحد الخطأ ، وبعد ترجيح المتغيرات تستخدم طريقة المربعات الصغرى الاعتيادية غير المرجحة علي النحو التالي:

$$\hat{\beta}_{wls} = (X^{*T} X^*)^{-1} X^{*T} Y^* \quad \dots\dots\dots (82 - 2)$$

ويتم الحصول علي تباين معاملات نموذج الانحدار المقدر كما يلي:

$$Var(\hat{\beta}_{wls}) = \sigma^2 (X^{*T} X^*)^{-1} = \sigma^2 (X^T W^{-1} X)^{-1} \quad \dots\dots\dots (83 - 2)$$

حيث أن:

$$W^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_N^2} \end{bmatrix} \quad \dots\dots\dots (84 - 2)$$

ويلاحظ أن حد الخطأ الجديد يستوفي فرضيات طريقة المربعات الصغرى العادية التي من بينها فرضية ثبات التباين.

ثانياً: استخدام طريقة المربعات الصغرى المرجحة عندما تكون قيم σ_i^2 مجهولة:

في الواقع العملي نادراً ما تكون قيم σ_i^2 معلومة ولذلك لابد من تقديرها ويمكن أن يستخدم في هذه الحالة تباين بواقى نموذج الانحدار الأصلي كأوزان لتصحيح النموذج نفسه. حيث يتم

استبدال قيم الأوزان W_i بالقيم المقدرة لها \hat{W}_i حيث:

$$\hat{W}_i = \frac{1}{\hat{\sigma}_i^2} \quad \dots\dots\dots (85 - 2)$$

ونتخذ مصفوفة التصحيح الصيغة التالية:

$$\hat{W}^{-1} = \begin{bmatrix} \frac{1}{\hat{\sigma}_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\hat{\sigma}_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{\hat{\sigma}_n} \end{bmatrix} \dots\dots\dots (86 - 2)$$

وعليه يتم الحصول علي مقدرات المربعات الصغرى المرجحة $\hat{\beta}_{wls}$ علي النحو التالي:

$$\hat{\beta}_{wls} = (X^T \hat{W}^{-1} X)^{-1} X^T \hat{W}^{-1} Y \dots\dots\dots (87 - 2)$$

وكذلك يتم الحصول علي صيغة تباين معاملات الانحدار المقدره باستخدام طريقة المربعات الصغرى المرجحة كما يلي:

$$Var(\hat{\beta}_{wls}) = \hat{\sigma}^2 (X^T W^{-1} X)^{-1} \dots\dots\dots (88 - 2)$$

2-12 مشكلة الارتباط الخطي المتعدد Multicollinearity Problem:

ذكرنا سابقاً إن إحدى الافتراضات الأساسية التي تقوم عليها طريقة المربعات الصغرى هو عدم وجود علاقة خطية تربط بين أحد المتغيرات المستقلة ومتغير مستقل آخر أو بين أحد المتغيرات المستقلة وأية تركيب خطي بين المتغيرات المستقلة الأخرى ، أي عدم وجود ارتباط خطي متعدد Multicollinearity بين المتغيرات المستقلة. والارتباط الخطي أو التداخل الخطي بمعناه الواسع هو اتجاه كثير من المتغيرات الاقتصادية للتحرك معاً في نفس الاتجاه مع مرور الزمن وبذلك يظهر ارتباطاً قوياً بين هذه المتغيرات³⁷. وهناك نوعان من الارتباط الخطي المتعدد هما³⁸:

الارتباط الخطي المتعدد التام Perfect Multicollinearity: ويقصد به أن أحد المتغيرات

المستقلة في معادلة الانحدار له ارتباط خطي تام بواحد أو أكثر من المتغيرات الأخرى.

³⁷ - الطيب، عز الدين مالك، (2008)، مرجع سابق،ص:186.

³⁸ - إسماعيل، محمد عبد الرحمن، (2001)، "تحليل الانحدار الخطي"، معهد الإدارة العامة، المملكة العربية السعودية ، مركز البحوث.

الارتباط الخطي المتعدد المرتفع **High Imperfect Multicollinearity**: ويقصد به الحالة التي يكون فيها بين اثنين أكثر من المتغيرات المستقلة ارتباطاً قوياً ولكنه ليس تاماً مما يجعل من الصعب عزل تأثيراتها الفردية علي المتغير التابع. وتظهر مشكلة الارتباط المتعدد لعدة أسباب منها ما يلي³⁹:

1. ميل بعض المتغيرات للتحرك معاً مع مرور الزمن، فعلي سبيل المثال دخل الموظف وسنوات خبرته وعمره ومرتبته الوظيفية غالباً ما تتغير سوياً ويكون بينها ارتباط خطي قوي.

2. استخدام بعض المتغيرات المستقلة بفترات تأخير ومن الطبيعي أن القيم المتعاقبة لمتغير معين يكون بينها علاقة فالدخل في الفترة الحالية يتحدد جزئياً عن طريق قيمته في الفترة السابقة وهكذا لذا فان مشكلة الارتباط الخطي غالباً ما تكون موجودة مؤكداً في نماذج فترات التأخير.

3. قلة عدد المشاهدات مقارنة بعدد المتغيرات الموجودة في النموذج وهذه الحالة تحدث في الأبحاث الطبية والإنسانية حيث عدد الأشخاص تحت الدراسة قليل والمعلومات تجمع علي عدد كبير من المتغيرات المستقلة تحت الدراسة.

2-12-1 النتائج المترتبة علي وجود الارتباط المتعدد⁴⁰:

Consequences of Multicollinearity

يمكن تلخيص النتائج المترتبة علي وجود الارتباط الخطي المتعدد أو التداخل الخطي

المتعدد في حالة التداخل الخطي التام و التداخل الخطي المرتفع كالآتي:

³⁹ - عبد المنعم، ثروت محمد، (2005)، "الانحدار"، مكتبة الأنجلو المصرية، القاهرة، مصر، ص:506.

⁴⁰ - إبراهيم، بسام يونس وآخرون،(2002)، مرجع سابق، ص:262-265 .

أولاً: حالة التداخل الخطي التام:

ليبان النتائج المترتبة علي وجود هذه الظاهرة في هذه الحالة، نفترض أنه لدينا نموذج الانحدار

المقدر الآتي:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 \quad \dots\dots\dots (89 - 2)$$

وأن شكل العلاقة بين المتغيرين X_1, X_2 هي:

$$X_1 = cX_2 \quad \dots\dots\dots (90 - 2)$$

حيث: c ثابت موجب أو سالب.

والنتائج المترتبة علي وجود هذه الحالة يمكن إجمالها بالآتي:

1- لا يمكن تقدير معاملات النموذج (2 - 89) ولغرض تقدير للنموذج (2 - 89) يمكن استخدام

أسلوب الانحرافات حيث أن

$$\hat{\beta} = (x'x)^{-1} x'y \quad \dots\dots\dots (91 - 2)$$

$$\hat{\beta}_1 = \frac{\sum x_{1i}^2 \sum x_{2i} y_i - (\sum x_{1i} x_{2i})(\sum x_{1i} y_i)}{(\sum x_{1i}^2)(\sum x_{2i}) - (\sum x_i x_{2i})} \quad \dots\dots\dots (92 - 2)$$

$$\hat{\beta}_2 = \frac{\sum x_{1i}^2 \sum x_{2i} y_i - (\sum x_{1i} x_{2i})(\sum x_{1i} y_i)}{(\sum x_{1i}^2)(\sum x_{2i}) - (\sum x_i x_{2i})} \quad (93 - 2)$$

المعادلة (2 - 90) بدلالة الانحرافات تصبح كالآتي:

$$x_1 = cx_2 \quad \dots\dots\dots (94 - 2)$$

وبتعويض (2 - 94) في كل من (2 - 92) و (2 - 93) نحصل علي:

$$\hat{\beta}_1 = \frac{c^2 \sum x_{2i}^2 \sum x_{2i} y_i - c^2 (\sum x_{2i}^2)(\sum x_{2i} y_i)}{c^2 (\sum x_{1i}^2)^2 - c^2 (\sum x_{1i}^2)^2} = \frac{0}{0} \quad \dots\dots\dots (95 - 2)$$

وكذلك بالمثل فإن:

$$\hat{\beta}_2 = \frac{0}{0} \quad \dots\dots\dots (96 - 2)$$

عليه لا يمكن تقدير معاملات نموذج الانحدار والسبب بالأساس ناتج كون $|x'x|=0$.

2- لا يمكن إيجاد تباينات المقدرات والتغايرات المشتركة فيها وكما هو معلوم:

$$V(\hat{\beta}) = \sigma_u^2 (x'x)^{-1} \quad \dots\dots\dots (97 - 2)$$

ومما سبق واعتماداً علي المعادلة (94 - 2) فإن:

$$V(\hat{\beta}) = \frac{\sigma_u^2}{\sum x_{1i}^2 \sum x_{2i}^2 - (\sum x_{1i}x_{2i})^2} \begin{bmatrix} \sum x_{2i}^2 & - \sum x_{1i}x_{2i} \\ - \sum x_{1i}x_{2i} & \sum x_{1i}^2 \end{bmatrix}$$

عليه فإن:

$$V(\hat{\beta}_1) = \sigma_u^2 \frac{\sum x_{2i}^2}{\sum x_{1i}^2 \sum x_{2i}^2 - (\sum x_{1i}x_{2i})^2} \quad \dots\dots\dots (98 - 2)$$

$$V(\hat{\beta}_2) = \sigma_u^2 \frac{\sum x_{1i}^2}{\sum x_{1i}^2 \sum x_{2i}^2 - (\sum x_{1i}x_{2i})^2} \quad \dots\dots\dots (99 - 2)$$

$$Cov(\hat{\beta}_1, \hat{\beta}_2) = \sigma_u^2 \frac{- \sum x_{1i}x_{2i}}{\sum x_{1i}^2 \sum x_{2i}^2 - (\sum x_{1i}x_{2i})^2} \quad \dots\dots\dots (100 - 2)$$

واعتماداً علي المعادلة (94 - 2) نحصل علي:

$$V(\hat{\beta}_1) = \frac{c^2 \sigma_u^2 \sum x_{1i}^2}{c^2 (\sum x_{1i}^2)^2 - c^2 (\sum x_{1i}^2)^2} = \frac{\sigma_u^2 \sum x_{1i}^2}{0} \quad \dots\dots\dots (101 - 2)$$

وبنفس الطريقة نجد أن:

$$V(\hat{\beta}_2) = \frac{\sigma_u^2 \sum x_{1i}^2}{0} \dots\dots\dots (102 - 2)$$

$$Cov(\hat{\beta}_1, \hat{\beta}_2) = \frac{-\sigma_u^2 \sum x_{1i}^2}{0} \dots\dots\dots (103 - 2)$$

عليه لا يمكن حساب تباينات المقدرات و التغايرات المشتركة لها وهذا أيضاً ناتج بسبب كون

$$|x'x| = 0$$

ثانياً: حالة التداخل الخطي من الدرجة العليا:

عندما تكون العلاقة بين المتغيرين X_1, X_2 وتقرب من ± 1 فإن $|x'x|$ سيكون صغيراً جداً ويقرب من الصفر، ويترتب علي ذلك الآتي:

1. قيم المقدرات $\hat{\beta}_1, \hat{\beta}_2$ تكون كبيرة جداً وفي هذه الحالة تكون مقدرات متحيزة.

2. تباينات هذه المقدرات و التغايرات المشتركة تكون هي الأخرى كبيرة جداً لذا فإن المقدرات لن

تتمتع بالخصائص **Best Linear Unbiased Estimators**.

2-12-2 اكتشاف التداخل الخطي المتعدد Detection of Multicollinearity:

توجد مقاييس عديدة تستخدم للكشف عن وجود الارتباط الخطي المتعدد بين المتغيرات

المستقلة و نستعرض بعض منها فيما يلي:

2-12-2-1 عامل تضخم التباين Variance Inflation Factor:

من الطرق الأساسية الواسعة الاستخدام للكشف عن وجود الارتباط الخطي المتعدد هو

عامل تضخم التباين VIF_j وتقيس عوامل تضخم التباين مدي تضخم تباينات معاملات الانحدار

المقدرة في وجود الارتباط الخطي. ويتم حساب عامل تضخم التباين حسب الصيغة التالية:

$$VIF_j = \frac{1}{1 - R_j^2} \quad j = 1, 2, 3, \dots, p \dots\dots\dots (104 - 2)$$

حيث أن:

VIF_j : عامل تضخم التباين للمتغير المستقل رقم j ، R_j^2 : معامل التحديد لنموذج انحدار المتغير

المستقل رقم j علي بقية المتغيرات المستقلة $P-1$.

فإذا كانت قيمة عامل تضخم التباين $VIF_j > 10$ فإن ذلك يدل علي وجود ارتباط خطي متعدد

مرتفع⁴¹ وبكافئ هذا الحد عندما تكون قيمة معامل التحديد لنموذج انحدار المتغير المستقل

رقم j علي بقية المتغيرات المستقلة مساوية $R_j^2 = 0.90$ ⁴².

ملاحظات علي صيغة عامل تضخم التباين:

1. يأخذ عامل تضخم التباين قيمة غير سالبة، أي أن $VIF_j \geq 0$.

2. في حالة وجود ارتباط خطي تام بين المتغير المستقل X_j وبقية المتغيرات المستقلة فإن

$(R_j^2 = 1)$ وبالتالي فان عامل تضخم التباين يتخذ قيمةً لانتهائية، أي: $VIF_j = \infty$.

3. أما في حالة عدم وجود ارتباط خطي متعدد بين المتغير المستقل X_j وبقية المتغيرات

المستقلة $(R_j^2 = 0)$ فان قيمة عامل تضخم التباين تكون مساوية للواحد الصحيح، أي:

$VIF = 1$ وتسمى هذه الحالة بحالة التعادم.

4. يقدم متوسط قيم VIF لمعاملات الانحدار معلومات عن خطورة الارتباط الخطي المتعدد

بدلالة المسافة بين معاملات الانحدار المعيارية المقدرة b'_k وبين القيم الحقيقية B'_k ويمكن

تبيان أن القيمة المتوقعة لمجموع مربعات الخطأ هذه $(b'_k - B'_k)^2$ معطاة بالعلاقة⁴³:

$$E \left[\sum_{k=1}^{p-1} (b'_k - B'_k)^2 \right] = (\sigma')^2 \sum_{k=1}^{p-1} (VIF)_k \quad \dots\dots\dots (105 - 2)$$

⁴¹ -Kleinbaum, J, et al, (1988), "Applied Regression Analysis and other Multivariable Methods", PWS-Kent Publishing Company, Boston, second edition, p:210.

⁴² - Neter, J, et al, (1990), "Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs". (3rd edition). Irwin, Homewood, IL 60430, Boston, MA 02116.

⁴³ - جون نتر وآخرون، (2000م)، "نماذج إحصائية خطية تطبيقية"، (الجزء الثاني)، ترجمة انس إسماعيل كنجو، جامعة الملك سعود الرياض، ص:531.

وعندما لا توجد صلة خطية لأي متغير X بالمتغيرات الأخرى في نموذج الانحدار $R_k^2 = 0$ فعندئذ ، $(VIF)_k = 1$ وتصبح العلاقة:

$$E \left[\sum_{k=1}^{p-1} (b'_k - B'_k)^2 \right] = (\sigma')^2 (p-1) \quad \dots\dots\dots (106 - 2)$$

وتقدم نسبة النتيجة في (2-105)، (2-106) معلومات مفيدة عن تأثير الخطية المتعددة علي مجموع مربعات الخطأ:

$$\frac{(\sigma')^2 \sum (VIF)_k}{(\sigma')^2 (P-1)} = \frac{\sum (VIF)_k}{(P-1)} \quad \dots\dots\dots (107 - 2)$$

ونلاحظ أن هذه النسبة هي متوسط قيم VIF أي:

$$\overline{VIF} = \frac{\sum_{i=1}^{p-1} (VIF)_k}{p-1} \quad \dots\dots\dots (108 - 2)$$

فإذا كانت المتغيرات المستقلة متعامدة أي لا يوجد ارتباط خطي ، فإن المتوسط يساوي الواحد الصحيح. لذلك نجد أنه كلما زادت قيمة متوسط عوامل التباين عن الواحد دل ذلك علي وجود الارتباط الخطي المتعدد بين المتغيرات المستقلة.

2-2-12-2 قيمة التسامح Tolerance:

تستخدم بعض حزم برامج الإحصاء الجاهزة معكوس عامل تضخم التباين للكشف عن وجود الارتباط الخطي المتعدد بين المتغير X_j والمتغيرات المستقلة الأخرى وتحديد دخول المتغير للنموذج أم لا ما يعرف بالتحمل Tolerance أو التسامح وتدل قيمته علي مقدار تباين المتغير المستقل المحدد الذي لا تفسره المتغيرات المستقلة الأخرى في النموذج، ويتم حساب هذا المقياس حسب الصيغة التالية:

$$Tolerance = \frac{1}{VIF_j} = 1 - R_j^2 \quad \dots\dots\dots (109 - 2)$$

فإذا كانت قيمة Tolerance لكل متغير مستقل أقل من (0.10) فهذا يشير إلي أن الارتباط المتعدد مع المتغيرات الأخرى مرتفع، مما قد يؤدي إلي احتمال تحقق المصاحبة الخطية المتعددة⁴⁴ وقيم التحمل التي تستخدم بواسطة هذه البرامج كحد أدنى لدخول أي متغير مستقل للنموذج هي: (0.01,0.001,0.0001).

3-2-3-3 تحليل القيم الكامنة Eigen values:

يمكن استخدام القيم المميزة ($\lambda_1, \lambda_2, \dots, \lambda_k$) لمصفوفة الارتباط XX كمقياس لوجود

الإرتباط الخطي المتعدد في البيانات، وتتلخص هذه الطريقة في الآتي:

1. حساب المصفوفة Z حيث:

$$Z = (X^T X) \quad \dots\dots\dots (110 - 2)$$

2. حساب المصفوفة القطرية S ، حيث:

$$S = \text{diag} (X^T X)^{-0.5} \quad \dots\dots\dots (111 - 2)$$

حيث إن قيمة أي عنصر من عناصر المصفوفة S عبارة عن معكوس الجذر التربيعي لعنصر

المصفوفة Z المقابلة لها، أي إن:

$$S = \frac{1}{\sqrt{Z_{ii}}} \quad \dots\dots\dots (112 - 2)$$

3. حساب المصفوفة SZS التي تحتوي عناصرها القطرية علي الواحد الصحيح، حيث:

$$SZS = S(X^T X)S \quad \dots\dots\dots (113 - 2)$$

4. وباستخدام طرق المصفوفات يتم إيجاد القيم الكامنة أو الجذور المميزة للمصفوفة SZS .

⁴⁴ - جولي بلانت، (2006)، "التحليل الإحصائي باستخدام برنامج SPSS"، دار الفاروق للنشر والتوزيع، القاهرة، مصر، ص:173.

فإذا وجد أن هناك واحد أو أكثر من القيم الكامنة أو الجذور المميزة قريبة من الصفر فهذا دليل على وجود مشكلة الارتباط الخطي المتعدد⁴⁵. أما إذا كانت قيمة أي من القيم الكامنة مساوية للصفر فيعني ذلك وجود ارتباط خطي تام بين المتغيرات المستقلة.

3-3-2-4 مؤشر الحالة Condition Index:

يسمى أحياناً بدليل الحالة ويحسب من خلال الصيغة التالية:

$$CI_j = \sqrt{\frac{\lambda_{\max}}{\lambda_j}} \quad j = 1, 2, \dots, p \quad \dots\dots\dots (114 - 2)$$

حيث أن: λ_{\max} : قيمة أكبر جذر مميز ، λ_j : قيمة الجذر المميز رقم j .

فإذا زادت قيمة الدليل CI عن 15 فهذا مؤشر على إمكانية وجود مشكلة التعدد الخطي ، أما إذا زادت عن 30 فهذا مؤشر على خطورة المشكلة⁴⁶.

2-12-2-5 رقم الحالة Condition Number:

يحسب هذا المؤشر كما يلي:

$$CN = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad \dots\dots\dots (115 - 2)$$

فإذا كانت قيمة CN ما بين (20 - 30) دل ذلك على وجود ارتباط خطي مرتفع⁴⁷. أما إذا كانت قيمة CN ما بين (30 - 100) كان ذلك دلالة على وجود ارتباط خطي متعدد مرتفعاً جداً⁴⁸.

⁴⁵ - عبد المنعم، ثروت محمد، مرجع سابق، ص:515.

⁴⁶ - بخيت، حسين علي، الرفاعي، غالب عوض، (2007)، "تحليل ونمذجة البيانات باستخدام الحاسوب(تطبيق شامل للحزمة (SPSS)، الطبعة الثانية، الأهلية للنشر والتوزيع، المملكة الأردنية، عمان، ص:365.

⁴⁷ - Johnston, J. (1984), "Econometric Methods", 3rd edition, McGraw-Hill Book Company, New York, P:250.

⁴⁸ - Belsley, D. et al, (1980), "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity", Wiley, New York, p:105.

2-12-3 معالجة الارتباط الخطي المتعدد:

Remedy of Multicollinearity

يتوقف أسلوب معالجة الارتباط الخطي المتعدد بين المتغيرات المستقلة علي درجة هذا الارتباط فإذا كانت درجة الارتباط منخفضة فمن الممكن قبول هذا الارتباط ، أما إذا كانت درجة الارتباط مرتفعة فلا بد من معالجة هذا الارتباط وذلك بإتباع احدي الطرق الآتية للتصحيح⁴⁹:

1. زيادة حجم العينة حيث يؤدي ذلك إلي خفض قيم الأخطاء المعيارية لمعالم النموذج.
2. استخدام معلومات قبلية Priori Information حول العلاقات بين المتغيرات المستقلة، فمثلاً نجد أن هناك علاقة بين المرتبة الوظيفية ومدة خبرة الموظف في العمل فبدلاً من إدخال متغير المرتبة X_1 والخبرة X_2 كمتغيرين مستقلين ضمن متغيرات مستقلة أخرى للأجر الذي يتقاضاه الموظف Y_i فيمكن دمج هذين المتغيرين إلي متغير واحد إذا أمكن الحصول علي معلومات تقريبية تفيد بان قيمة معلمة الخبرة مثلاً تساوي حوالي ربع معلمة المرتبة وبالحصول علي هذه المعلومة يتم بناء النموذج كالتالي:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + U_i \quad \dots\dots\dots (116 - 2)$$

وبما أن $\beta_2 = 0.25 \beta_1$ فان النموذج أعلاه يصبح:

$$Y_i = \beta_0 + \beta_1 (X_1 + 0.25 X_2) + U_i \quad \dots\dots\dots (117 - 2)$$

ومن عيوب هذا الحل صعوبة تحديد الأثر الفردي للمتغيرين علي المتغير الفردي.

3. تقليل عدد المتغيرات المستقلة ذات الارتباط المرتفع باستخدام تحليل المكونات الأساسية

Principal Component Analysis أو التحليل العاملي Factor Analysis وتهدف

هاتان الطريقتان إلي تحويل المتغيرات المرتبطة إلي عدد أقل تسمي بالعوامل Factors في

حالة التحليل العاملي وبالمكونات الأساسية Principal Components في حالة تحليل

⁴⁹ - عبد المنعم، ثروت محمد، مرجع سابق، ص:523.

المكونات الأساسية بحيث يكون لكل عامل من هذه العوامل دالة تربطه ببعض أو كل هذه المتغيرات ومن ثم استخدام المتغيرات الجديدة غير المرتبطة مع بعضها البعض كمتغيرات مستقلة جديدة للمتغير التابع.

4. حذف أو إضافة متغير ، أحياناً يلجأ الباحث المستخدم للأسلوب القياسي إلي حذف المتغير الذي يمتاز بالارتباط العالي مع المتغيرات المستقلة الأخرى أو قد يضيف متغير جديد آخر ذو أهمية بالنسبة للنموذج⁵⁰.

5. استخدام انحدار التل Ridge Regression⁵¹ حيث يتم تقدير معالم نموذج الانحدار بإجراء تعديل علي طريقة المربعات الصغرى وباستخدام طريقة انحدار التل يتم الحصول علي مقدرات (b^R) متحيزة إلا أنها أكثر دقة من المقدرات غير المتحيزة ، وإن احتمال أن يكون قيمة (b^R) قريبة من القيمة الحقيقية للمعلمة أكبر بكثير من احتمال قرب المقدر غير المتحيز (b) لقيمة المعلمة الحقيقية ، ويتم قياس الأثر المشترك للتحيز bias والتباين بحساب القيمة المتوقعة لمربع انحراف المقدر المتحيز (b^R) عن القيمة الحقيقية للمعلمة (B) ويسمي هذا المقياس بمتوسط مربعات الخطأ، و يحسب كما يلي:

$$E(b^R - \beta)^2 = \sigma^2 \{b^R\} + (E\{b^R\} - \beta)^2 \quad \dots\dots\dots (118 - 2)$$

وتوضح المعادلة (118 - 2) أن متوسط مربعات الخطأ يساوي تباين المقدر (b^R) زائداً مربع التحيز مع ملاحظة أن متوسط مربعات الخطأ يساوي تباين المقدر إذا كان غير متحيز. ولإجراء انحدار التل المعياري Standardized Ridge Regression تتبع الخطوات التالية:

يتم أولاً تحويل المتغيرات المستقلة والتابع باستخدام تحويلة الارتباط Correlation

Transformation علي النحو التالي:

⁵⁰ - إبراهيم، بسام يونس وآخرون، مرجع سابق، ص:279.

⁵¹ - إسماعيل، محمد عبد الرحمن، مرجع سابق، ص:369-372.

$$Y'_i = \frac{1}{\sqrt{n-1}} \frac{(Y_i - \bar{Y})}{S_Y} \quad \dots\dots\dots (119 - 2)$$

$$X'_{ri} = \frac{1}{\sqrt{n-1}} \frac{X_{ri} - \bar{X}_r}{S_r} \quad r = 1, 2, \dots, p \quad \dots\dots\dots (120 - 2)$$

حيث أن:

$$\bar{Y} = \frac{\sum Y_i}{n} \quad , \quad \bar{X}_r = \frac{\sum X_{ri}}{n} \quad r = 1, 2, \dots, p$$

$$S_Y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}}$$

$$S_r = \sqrt{\frac{\sum (X_{ri} - \bar{X}_r)^2}{n-1}}$$

وبما أن $(X'^T X') = R_{XX}$ ، $(X'^T Y') = R_{XY}$ فإنه يمكن الحصول علي مقدرات التل المعيارية

علي النحو التالي:

$$b^R = (R_{XX} - cI)^{-1} R_{XY} \quad \dots\dots\dots (121 - 2)$$

حيث أن:

b^R : متجه معاملات انحدار التل المعيارية ، R_{XX} : مصفوفة معاملات الإرتباط البسيط بين أزواج المتغيرات المستقلة ، R_{XY} : متجه معاملات الإرتباط البسيط بين المتغير التابع وكل من المتغيرات المستقلة ، c : ثابت التحيز وتتراوح قيمته ما بين الصفر والواحد الصحيح ، I : مصفوفة وحدة من الرتبة $P \times P$.

ولإيجاد قيم معاملات نموذج الانحدار الأصلي تستخدم العلاقة التالية:

$$b_r = \frac{S_y}{S_r} b_r^R \quad r = 1, 2, \dots, p \quad \dots\dots\dots (122 - 2)$$

$$b_0 = \bar{Y} - b_1 \bar{X}_1 - b_2 \bar{X}_2 - \dots - b_p \bar{X}_p \quad \dots\dots\dots (123 - 2)$$

وعامل تضخم التباين لمعاملات انحدار التل يحسب من المصفوفة التالية:

$$(r_{XX} + cI)^{-1} r_{XX} (r_{XX} + cI)^{-1} \dots\dots\dots (124 - 2)$$

ومجموع مربعات البواقي Residual Sum of Squares يتم حسابه حسب الصيغة التالية:

$$RSS_R = \sum (Y'_i - \hat{Y}'_i)^2 \dots\dots\dots (125 - 2)$$

حيث أن:

$$\hat{Y}'_i = b_1^R X'_{1i} + b_2^R X'_{2i} + \dots + b_p^R X'_{pi} \dots\dots\dots (126 - 2)$$

كما يتم حساب معامل التحديد لانحدار التل حسب الصيغة التالية:

$$R_R^2 = 1 - RSS_R \dots\dots\dots (127 - 2)$$

ملاحظات علي طريقة انحدار التل:

1. تعكس قيمة الثابت (c) مقدار التحيز في المقدرات ويلاحظ أنه عندما تكون قيمة الثابت مساوية للصفر نحصل علي مقدرات المربعات الصغرى الاعتيادية. وعندما تكون قيمة الثابت أكبر من الصفر نحصل علي مقدرات متحيزة إلا أنها أكثر استقراراً من مقدرات المربعات الصغرى الاعتيادية.

2. يعاب علي طريقة انحدار التل صعوبة تحديد قيمة (c) المثلي ، ولتحديد قيمة التحيز (c) التي تعطي أفضل نموذج ، يستخدم عادة الرسم البياني لقيم معاملات انحدار التل (المحور الصادي) مع قيم مختلفة لثابت التحيز ذات مسافات متساوية (المحور الأفقي) ويعرف الشكل الناتج بـ Ridge trace. وكما يؤخذ في الاعتبار قيمة عامل تضخم التباين للتأكد من حل مشكلة الارتباط الخطي. فإذا أظهر الشكل استقراراً في قيم معاملات الانحدار وانخفاض قيم عوامل تضخم التباين عند قيم محددة لثابت التحيز، يتم اختيار أحد النماذج المناظرة بصورة تحكمية.

2-13 مشكلة الارتباط الذاتي Autocorrelation Problem:

كما ذكرنا سابقاً إن احدي الفرضيات الأساسية لتطبيق طريقة المربعات الصغرى الاعتيادية، فرض استقلال الأخطاء العشوائية عن بعضها البعض، بمعنى آخر عدم وجود ارتباط تسلسلي بين حدود الخطأ، أي:

$$Cov(U_i U_j) = E(U_i U_j) = 0 \quad \forall i \neq j \quad \dots\dots\dots (2-128)$$

وبأخذ الصيغة أعلاه علي شكل فترات زمنية فإن:

$$Cov(U_t U_{t-1}) = E(U_t U_{t-1}) = 0 \quad \forall t \neq t-1 \quad \dots\dots\dots (2-129)$$

ومضمون مفهوم الارتباط الذاتي هو "أن قيم المتغير العشوائي التي تحدث خلال فترة معينة u_t ترتبط بقيم المتغير العشوائي التي تسبقها أو تليها"⁵² وهذا يعني مخالفة فرضية استقلال حدود الخطأ العشوائي عن بعضها البعض وبالتالي يؤدي إلي ظهور مشكلة بالارتباط الذاتي Autocorrelation Problem أي:

$$Cov(U_t U_{t-1}) = E(U_t U_{t-1}) \neq 0 \quad \dots\dots\dots (2-130)$$

وهناك نوعان من الارتباط الذاتي هما:

الارتباط الذاتي الموجب: في هذه الحالة تكون معظم القيم المقدره المتتالية لحد الخطأ العشوائي لها نفس الإشارة الجبرية.

الارتباط الذاتي السالب: في هذه الحالة تكون معظم القيم المقدره المتتالية لحد الخطأ العشوائي تتبادل الإشارة بين الموجب والسالب، أي أن الأخطاء العشوائية السالبة تليها أخطاء عشوائية موجبة ثم تليها أخطاء عشوائية سالبة ... وهكذا.

⁵² - السيفو، وليد إسماعيل وآخرون، (2006)، "مشاكل الاقتصاد القياسي التحليلي"، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن، ص:123 .

ويظهر الارتباط الذاتي لعدة أسباب منها⁵³:

1. عدم إدراج المتغير التابع كمتغير متباطئ Lagged variable في الحالات التي تتطلب ذلك.

2. حذف بعض المتغيرات المستقلة من النموذج الأمر الذي يؤدي إلي جعل الحد العشوائي يتضمن تلك المتغيرات المرتبطة، وبالتالي فإن الحد العشوائي لا يعكس الخطأ في النموذج فحسب وإنما يعكس المتغيرات المحذوفة أيضاً.

3. الصياغة الرياضية الخاطئة للنموذج: إذا كانت الصيغة الرياضية للنموذج مخالفة للعلاقة الطبيعية بين المتغيرات فإن الارتباط الذاتي سيكون موجوداً بين المتغيرات العشوائية.

4. عدم توصيف المتغير العشوائي توصيفاً دقيقاً: أن العوامل العشوائية كالحرب والعواصف والإضرابات تمتد آثارها إلي أكثر من فترة زمنية واحدة مما يؤثر علي كثير من المتغيرات الاقتصادية الأخرى لفترات زمنية قادمة، ومثل هذه الحالات تؤدي إلي ارتباط قيم المتغير العشوائي خلال السلسلة الزمنية.

5. تبدأ مشكلة الارتباط الذاتي في بيانات السلاسل الزمنية بطبيعة البيانات نفسها وطرق تجميعها، فقد يترتب علي وجود أخطاء القياس في تجمع هذا النوع من البيانات أخطاء تراكمية في السنوات أو النقاط الزمنية المتتالية.

2-13-1 النتائج المترتبة علي وجود الارتباط الذاتي⁵⁴:

Consequences of Autocorrelation

1. تظل مقدرات المربعات الصغرى العادية تكون غير متحيزة في ظل وجود الارتباط الذاتي،
أي أن:

⁵³ - الطيب، عز الدين مالك، مرجع سابق، ص: 160-161.
⁵⁴ - إسماعيل، محمد عبد الرحمن، (2001م)، مرجع سابق، ص: 419-422.

$$E(\hat{\beta}) = \beta \quad \dots\dots\dots (131 - 2)$$

2. مقدرات المربعات الصغرى العادية في حالة وجود الارتباط الذاتي ليست أفضل المقدرات، أي تفقد لخاصية الكفاءة ويمكن تبيان الاختلاف بين تباين مقدرات المربعات الصغرى في ظل وجود ارتباط ذاتي بين حدود الخطأ وعدمه من خلال المعادلة التالية:

$$Var(\hat{\beta})_{AR(1)} = \frac{\sigma^2}{\sum_{i=1}^n X_i^2} \left[1 + 2\rho \frac{\sum_{i=1}^{n-1} X_i X_{i+1}}{\sum_{i=1}^n X_i^2} + 2\rho^2 \frac{\sum_{i=1}^{n-2} X_i X_{i+2}}{\sum_{i=1}^n X_i^2} + \dots + 2\rho^{n-1} \frac{\sum_{i=1} X_i X_{i+(n-1)}}{\sum_{i=1}^n X_i^2} \right] \quad \dots\dots\dots (132 - 2)$$

ويلاحظ من المعادلة أعلاه أن تباين $\hat{\beta}$ في حالة وجود الارتباط الذاتي يساوي تباين $\hat{\beta}$ في حالة عدم وجود ارتباط ذاتي زائد حد يعتمد علي معامل الارتباط الذاتي ρ وعلي التغير بين قيم المتغير المستقل X فإذا كانت قيمة الارتباط الذاتي مساوية الصفر نجد أن:

$$Var(\hat{\beta})_{AR(1)} = Var(\hat{\beta}) \quad \dots\dots\dots (133 - 2)$$

أما إذا كانت قيمة الارتباط الذاتي أكبر من الصفر فإن:

$$Var(\hat{\beta})_{AR(1)} > Var(\hat{\beta}) \quad \dots\dots\dots (134 - 2)$$

ويترتب علي هذا الاختلاف أن تكون النتائج التي نحصل عليها من الاستدلال الإحصائي مثل إحصاءات F , t وفترات الثقة والتنبؤ غير صحيحة.

2-13-2 اكتشاف الارتباط الذاتي :Detection of Autocorrelation

هنالك عدة طرق تستخدم للكشف عن وجود الارتباط الذاتي بين قيم الخطأ العشوائي من

أهمها اختبار درين - واتسون Durbin-Watson و اختبار h لدرين وسنتناول الأول بالشرح

فيما يلي:

2-13-2-1 اختبار درين - واتسون Durbin-Watson:

يعتبر اختبار درين - واتسون من أهم الاختبارات التي تستخدم لاكتشاف الارتباط الذاتي

من الدرجة الأولى، ويستخدم هذا الاختبار لاختبار الفروض التالية:

$$H_0 : \rho = 0$$

فرض العدم: و ينص علي عدم وجود الارتباط ذاتي

$$H_1 : \rho \neq 0$$

الفرض البديل: و ينص علي وجود الارتباط الذاتي

وتتلخص خطوات هذا الاختبار في الآتي:

1. تقدير نموذج الانحدار وحساب القيم المقدرة \hat{r}_t و حساب قيم الأخطاء e_t .

2. حساب إحصائية الاختبار من الصيغ التالية:

$$d^* = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} = 2(1 - \hat{\rho}) \quad \dots\dots\dots (135-2)$$

واعتماداً علي المعادلة (135-2) وبما أن مدي $\hat{\rho}$ هو $-1 \leq \hat{\rho} \leq 1$ فإن ذلك يجعل

$$0 \leq d^* \leq 4$$

3. يتم استخراج القيم الجدولية لهذا الاختبار من جداول درين- واتسون حيث يتم استخراج قيمتي

الحد الأعلى لـ d^* ويرمز له بالرمز dl والحد الأدنى لـ d^* ويرمز له بالرمز du حيث

تستخرج هذه القيم اعتماداً علي k عدد المتغيرات المستقلة و n حجم العينة و α مستوي

المعنوية المحدد.

4. تقارن قيمة d^* مع dl و du وبعد ذلك يتم اتخاذ قرار بقبول أو رفض فرض

العدم H_0 كالآتي:

إذا كانت $0 < d^* < dl$: تبين هذه الحالة وجود ارتباط ذاتي موجب، أي رفض H_0

والنموذج يعاني من الارتباط الذاتي.

إذا كانت $dl < d^* < du$ فإنه من غير المؤكد وجود الارتباط الذاتي وتسمى هذه المنطقة منطقة عدم التأكد.

إذا كانت $du < d^* < 4 - du$ في هذه الحالة يتم قبول H_0 أي لا يوجد ارتباط ذاتي بين قيم حد الخطأ العشوائي.

إذا كانت $4 - du < d^* < 4 - dl$ في هذه الحالة من غير المؤكد وجود الارتباط الذاتي.

إذا كانت $4 - dl < d^* < 4$ في هذه الحالة فإنه يتم رفض فرض عدم H_0 ويقبل الفرض البديل H_1 أي يوجد ارتباط ذاتي سالب بين قيم حد الخطأ العشوائي. ويمكن توضيح جميع هذه الحالات باستخدام الشكل التالي:

الشكل (1-2): يوضح مناطق القبول والرفض لاختبار درين- واتسون

0	dL	du	2	4-du	4-dL	4
$P>0$	-	$P=0$	$P=0$	-	$P<0$	
ارتباط ذاتي موجب	منطقة غير محددة	عدم وجود ارتباط ذاتي	عدم وجود ارتباط ذاتي	منطقة غير محددة	ارتباط ذاتي سالب	

المصدر: عبد الرحمن إسماعيل، مرجع سابق.

2-13-3 معالجة الارتباط الذاتي [55] Remedy of Autocorrelation:

هنالك عدة طرق تستخدم لمعالجة ظاهرة الارتباط الذاتي بين قيم الخطأ العشوائي أهمها

طريقة التحويل وطريقة المربعات الصغرى العمومية.

2-13-3-1 طريقة التحويل Transformation Method:

يطلق عليها أيضاً طريقة كوكرين-أوركت Cochrane-Orcutt Method وتعتبر هذه

الطريقة من أسهل الطرق استخداماً. فإذا افترضنا أن نموذج الانحدار علي الشكل التالي:

$$Y_t = \beta_0 + \beta_1 X_t + U_t \quad \dots\dots\dots (136 - 2)$$

وإذا افترضنا أن U_t يعاني من الارتباط الذاتي من الرتبة الأولى وفقاً لمعادلة ماركوف أي:

$$U_t = \rho U_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad \dots\dots\dots (137 - 2)$$

فإنه يمكن كتابة المعادلة (136 - 2) كما يلي:

$$Y_t = \beta_0 + \beta_1 X_t + \rho U_{t-1} + \varepsilon_t \quad \dots\dots\dots (138 - 2)$$

عليه يتم معالجة الارتباط الذاتي بتحويل البيانات من استخدام الطريقة التالية اعتماداً علي

المعادلة (136 - 2) فإن:

$$Y_{t-1} = \beta_0 + \beta_1 X_{t-1} + U_{t-1} \quad \dots\dots\dots (139 - 2)$$

وبضرب المعادلة (139 - 2) بـ $\hat{\rho}$ نحصل علي:

$$\hat{\rho} Y_{t-1} = \hat{\rho} \beta_0 + \hat{\rho} \beta_1 X_{t-1} + \hat{\rho} U_{t-1} \quad \dots\dots\dots (140 - 2)$$

وبطرح المعادلة (140 - 2) من (136 - 2) نحصل علي:

$$Y_t - \hat{\rho} Y_{t-1} = \beta_0 (1 - \hat{\rho}) + \beta_1 (X_t - \hat{\rho} X_{t-1}) + (U_t - \hat{\rho} U_{t-1}) \quad \dots\dots\dots (141 - 2)$$

وإذا افترضنا أن:

⁵⁵ - إبراهيم، بسام يونس وآخرون، مرجع سابق ص: 250-252.

$$Y_t^* = Y_t - \hat{\rho}Y_{t-1} \quad \dots\dots\dots (142 - 2)$$

$$X_t^* = X_t - \hat{\rho}X_{t-1} \quad \dots\dots\dots (143 - 2)$$

$$V_t = U_t - \hat{\rho}U_{t-1} \quad \dots\dots\dots (144 - 2)$$

$$\alpha_0 = \beta_0(1 - \hat{\rho}) \quad \dots\dots\dots (145 - 2)$$

$$\alpha_1 = \beta_1 \quad \dots\dots\dots (146 - 2)$$

عليه فإن المعادلة (141 - 2) تصبح بالشكل التالي:

$$Y_t^* = \alpha_0 + \alpha_1 X_t^* + V_t \quad \dots\dots\dots (147 - 2)$$

وفي هذه المعادلة نلاحظ أن هنالك فقدان للقيمة الأولى لكل من X_t^* ; Y_t^* لان عدد القيم فيها

$n - 1$ ، عليه يتم حساب هاتين القيمتين من الصيغ التالية:

$$Y_1^* = Y_1 \sqrt{1 - \hat{\rho}^2} \quad \dots\dots\dots (148 - 2)$$

$$X_1^* = X_1 \sqrt{1 - \hat{\rho}^2} \quad \dots\dots\dots (149 - 2)$$

وبعد ذلك يتم تقدير النموذج (147 - 2) فنحصل على:

$$\hat{Y}_t^* = \hat{\alpha}_0 + \hat{\alpha}_1 X_t^* \quad \dots\dots\dots (150 - 2)$$

ويعاد اختبار وجود الارتباط بالنسبة للنموذج (150 - 2) فإذا وجد الارتباط الذاتي مرة ثانية عندها

نلجأ إلي عملية تحويل من الرتبة الثانية أي نأخذ Y_{t-2} في المعادلة (139 - 2) بدلاً من Y_{t-1}

وتعاد خطوات المعالجة إلي أن يتم التخلص من الظاهرة وتسمى هذه الطريقة بطريقة الإعادة أو

التكرار. ولغرض استخدام النموذج الأصلي في التنبؤ فيجب الرجوع إلي الصيغة الأصلية

وكالاتي:

من المعادلة (141 - 2) يمكن الحصول علي النموذج المقدر التالي:

$$\hat{Y}_t = \hat{\beta}_0(1 - \hat{\rho}) + \beta_1(X_t - \hat{\rho}X_{t-1}) + \hat{\rho}\hat{Y}_{t-1} \quad \dots\dots\dots (151 - 2)$$

حيث نلاحظ أن هذه المعادلة تمثل نموذج انحدار خطي متعدد به متغيرين مستقلين هما Y_{t-1} ، $X_t - \hat{\rho}X_{t-1}$ ، ويمكن الحصول علي المشاهدة المقدرة الأولى من النموذج التحويلي حيث:

$$\hat{Y}_1^* = \hat{\alpha}_0 + \hat{\alpha}_1 X_1^* \quad \dots\dots\dots (152 - 2)$$

وبتعويض (2 - 148) ، (2 - 149) في (2 - 150) فنحصل علي:

$$\hat{Y}_1 = \frac{\hat{\beta}_0(1 - \hat{\rho})}{\sqrt{1 - \hat{\rho}^2}} + \hat{\beta}_1 X_1 \quad \dots\dots\dots (153 - 2)$$

وهو نموذج انحدار خطي بسيط فيه معلمة المقطع $\frac{\hat{\beta}_0(1 - \hat{\rho})}{\sqrt{1 - \hat{\rho}^2}}$ والميل $\hat{\beta}_1$.

2-13-4 تقدير معامل الارتباط الذاتي من الرتبة الأولى $\hat{\rho}$:

نلاحظ أن طرق معالجة الارتباط الذاتي تعتمد علي معرفة قيمة معامل الإرتباط الذاتي ρ ولكن في الواقع العلمي نادراً ما تكون معلومة وبالتالي نحتاج لتقديرها وهناك عدة طرق تستخدم لتقدير قيمة معامل الارتباط الذاتي منها ما يلي:

1- باستخدام إحصائية درين - واتسون:

يستخدم تقدير معامل الارتباط الذاتي بهذه الطريقة في حالة العينات الكبيرة $n > 30$ من

العلاقة التالية:

$$\hat{\rho} = 1 - 0.5d^* \quad \dots\dots\dots (154 - 2)$$

2- باستخدام طريقة ثيل - ناجر Theil-Nagar⁵⁶:

تستخدم هذه الطريقة لتقدير قيمة معامل الارتباط الذاتي في حالة العينات الصغيرة $n < 30$

من العلاقة التالية:

$$\hat{\rho} = \frac{n^2(1 - 0.5d^*) + p}{n^2 + p^2} \quad \dots\dots\dots (155 - 2)$$

حيث أن:

n : عدد المشاهدات ، d^* : إحصائية درين-واتسون ، p : عدد المتغيرات المستقلة.

3- باستخدام طريقة كوكرين-أوركوت Cochrane- Orcutt⁵⁷:

تعتبر هذه الطريقة من أكثر الطرق استخداماً ، ويتم الحصول علي قيمة معامل الارتباط

الذاتي باستخدام هذه الطريقة من خلال إجراء انحدار e_t علي e_{t-1} حسب النموذج المقدر

التالي:

$$e_t = \hat{\rho}e_{t-1} \quad \dots\dots\dots (156 - 2)$$

وكنتيجة لتقدير قيمة $\hat{\rho}$ باستخدام طريقة المربعات الصغرى نحصل علي:

$$\hat{\rho} = \frac{\sum_{t=2}^n e_t - e_{t-1}}{\sum_{t=2}^n e_t^2} \quad \dots\dots\dots (157 - 2)$$

4- باستخدام طريقة درين ذات المرحلتين:

أقترح درين طريقة أخرى لتقدير معامل الارتباط الذاتي بين حدود الخطأ المتتالية تعرف

بطريقة درين ذات المرحلتين، حيث تمت الاستفادة في هذه الطريقة من معادلة الفروق المعممة:

$$Y_t - \rho Y_{t-1} = \beta_0(1 - \rho) + \beta_1(X_t - \rho X_{t-1}) + (U_t - \rho U_{t-1}) \quad \dots\dots\dots (158 - 2)$$

⁵⁶ - Theil, H. and Nagar, A.L. (1961), Testing the Independence of Regression Disturbances . Journal of the American Statistical Association , vol. 56 pp:793-806.

⁵⁷ - Cochrane, D. and Orcutt, G.H., (1949), Application of Least Squares Regressions to Relations Containing Autocorrelated Error Terms. Journal of the American Statistical Association, vol.44, pp:32-61.

وبإعادة تنظيم المعادلة أعلاه، وأن $\varepsilon_t = U_t - \rho U_{t-1}$ نحصل علي:

$$Y_t = \beta_0(1 - \rho) + \beta_1 X_t - \rho\beta_1 X_{t-1} + \rho Y_{t-1} + \varepsilon_t \quad \dots\dots\dots (159 - 2)$$

حيث نلاحظ أن النموذج أعلاه هو عبارة عن نموذج انحدار خطي متعدد به ثلاثة متغيرات مستقلة هي X_t ، X_{t-1} ، Y_{t-1} وبتطبيق طريقة المربعات الصغرى لتقدير المعلمات نلاحظ أنه يمكن حساب $\hat{\rho}$ حيث تمثل معامل Y_{t-1} .

14-2 اختبار التوزيع الطبيعي للبواقي :Normality of Residuals

من الفرضيات الأساسية لنماذج الانحدار المتعدد أن المتغير العشوائي يتوزع طبيعياً، حيث تؤثر فرضية التوزيع الطبيعي علي نوعية التقديرات وأنواع التقدير التي يجب أن تستخدم في الإطار العملي للانحدار، ويعتبر اختبار التوزيع الطبيعي ذا أهمية إذا كانت التنبؤات هي الهدف حيث أن معظم فترات التنبؤات تعتمد علي فرضيات التوزيع الطبيعي. وهناك عدة اختبارات تستخدم لمعرفة مدى تبعية بواقي النموذج للتوزيع الطبيعي من أهمها اختبار جاك بيرا -Jarque-Bera Test وهو اختبار يجمع بين اختبار سكيونس Skewness واختبار كيرتوزس Kurtosis للتفرطح وتحت فرضية العدم القائلة بأن حدود الخطأ موزعة طبيعياً تكون إحصائية BJ موزعة حسب قانون كأي تريبع، وتحسب علي النحو التالي:

$$BJ = n \left[\frac{S^2}{6} + \frac{(K-3)^2}{24} \right] \sim \chi^2_{1-\alpha/2} \quad \dots\dots\dots (160 - 2)$$

فإذا كانت $BJ > \chi^2_{1-\alpha/2}$ نرفض فرضية التوزيع الطبيعي للأخطاء.

الفصل الثالث

المشاهدات الشاذة في نموذج تحليل الانحدار الخطي المتعدد

1-3 تمهيد:

2-3 مفهوم القيم الشاذة:

3-3 أسباب ظهور القيم الشاذة:

4-3 طرق الكشف عن القيم الشاذة:

1-4-3 الطرق اللامعلمية:

2-4-3 اكتشاف المشاهدات الشاذة في المتغيرات المستقلة:

3-4-3 اكتشاف المشاهدات الشاذة في المتغير التابع:

5-3 المشاهدات الشاذة المؤثرة وطرق الكشف عنها:

1-5-3 قياس التأثير على القيم المقدرة - مقياس DFFITS:

2-5-3 قياس الأثر على معاملات الانحدار - مقياس DFBETAS:

3-5-3 قياس الأثر على كل معاملات الانحدار - مقياس مسافة كوك:

4-5-3 قياس الأثر على الأخطاء المعيارية - مقياس COVRATIO:

6-3 بعض الحلول المقترحة لمعالجة القيم الشاذة:

3-1 تمهيد:

من خلال هذا الفصل سوف نتناول بشي من التفصيل مفهوم القيم الشاذة وأسباب ظهورها في البيانات وطرق الكشف عنها وتقديرها في المتغيرات المستقلة والمتغير التابع و طرق تشخيص المشاهدات المؤثرة على معلمات نموذج تحليل الانحدار الخطى المتعدد والتي تؤثر على دقة النتائج ، بالإضافة إلى بعض طرق معالجة القيم الشاذة.

3-2 مفهوم القيم الشاذة :Concept of Outlier

للقيم الشاذة تعريفات عدة نذكر منها ما يلي:

عرف ⁵⁸Bross المشاهدة الشاذة بأنها تلك المشاهدة التي تظهر منحرفة بشكل كبير عن سائر مكونات العينة التي وجدت فيها تلك العينة. أما ⁵⁹Freeman فقد عرف القيمة الشاذة بأنها أية مشاهدة لم تتولد بالطريقة العامة التي ولدت الأغلبية العظمى من مشاهدات البيانات. أما ⁶⁰Barnett فقد عرفت المشاهدة الشاذة في مجموعة البيانات هي تلك المشاهدة التي تبدو غير منطقية إذا قورنت بسائر مجموعة البيانات. وأيضاً ⁶¹Al-Jobouri عرف القيمة الشاذة بأنها تلك القيمة التي تكون غير منسجمة مع بقية بيانات المجموعة لمتغير من المتغيرات لظاهرة معينة أو مجموعة من الظواهر، أو أن القيم الشاذة هي القيم التي تأتي من مجتمع يختلف عن مجتمع العينة قيد الدراسة. في حين نص تعريف إسماعيل ⁶² على المشاهدات الشاذة هي مجموعة قليلة من المشاهدات تبعد قيمها بصورة كبيرة عن بقية قيم المشاهدات في العينة. كما عرف

⁵⁸ - Bross, I.D.J. (1961), "Outliers in Patterned Experiments: strategic Re-Appraisal". Technometrics. 3,19-102.

⁵⁹ - Freeman, P.R.(1980), "On the Number of Outliers in Data from a linear Model" In Bayesian Statistics, Ed .J.M.

⁶⁰ - Barnett, V. & Lewis, T. (1978), "Outliers in Statistical Data", Jon Wiley and Sons, New York.

⁶¹ -AL-Jobouri, S. (1976), "Test of Outliers Unpublished, M.s.c Thesis University of Baghdad.

⁶² - إسماعيل، محمد عبد الرحمن، (2001)، مرجع سابق، ص: 251.

Keller⁶³ المشاهدات الشاذة بأنها المشاهدات التي تقع بعيدة عن معادلة الانحدار ويكون لها خطأ كبير مقارنة ببقية المشاهدات الطبيعية الأخرى في مجموعة البيانات و ذلك سيكون لها تأثير في النموذج الخطي و تقديراته.

3-3 أسباب ظهور القيم الشاذة Causes of Existing the Outliers:

قد تظهر القيم الشاذة في مجموعة البيانات نتيجة لأسباب عدة منها ما يلي:

1. إن البيانات تعود إلى توزيعات غير متماثلة أي يكون فيها التواء عالٍ نحو اليمين أو نحو اليسار ، ولصياغة نموذج لهذه التوزيعات قدم Green, 1976 تصنيفاً لعوائل هذه التوزيعات، توزيعات عرضة للقيم الشاذة وهي التي تكون لها نهايات تؤول إلى الصفر ببط وهذه التوزيعات تكون عرضة للقيم الشاذة بصورة مطلقة. وتوزيعات تكون مقاومة للقيم الشاذة وهي التي تؤول نهاياتها إلى الصفر بشكل أسرع من سابقتها وتكون مقاومة للقيم الشاذة بصورة مطلقة.

2. تأتي البيانات من نوعين من التوزيعات أحدهما التوزيع الأساسي Basic Distribution والذي يولد مشاهدات جيدة بينما الآخر يسمى التوزيع الملوث Contaminating Distribution والذي يولد قيماً شاذة.

3. قد تحدث القيم الشاذة نتيجة لأسباب أخرى منها أخطاء يقع فيها الباحث عند رصد القياسات أو نتيجة إلى وجود خلل في أجهزة القياسات وخاصةً في التجارب المختبرية ، أو نتيجة أخطاء في الحسابات مما يؤدي إلى ظهور القيم الشاذة.

⁶³- Keller, G. & Brian Warrack, (2000), "Statistic For Management and Economics, 5th edition, Duxbury, Thomson Learning, U.S.A.

3-4 طرق الكشف عن المشاهدات الشاذة:

في كثير من الأحيان نجد أن البيانات المشاهدة لأي ظاهرة قد تحتوى على عدد من المشاهدات الشاذة أو المتطرفة ، والمشاهدات الشاذة كما سبق تعريفها هي مجموعة قليلة من المشاهدات التي تبعد قيمها بصورة كبيرة عن بقية قيم المشاهدات الأخرى في العينة.

أن وجود قيم شاذة في بيانات نموذج الانحدار في مشاهدات المتغير التابع أو المتغيرات المستقلة يؤثر على تقديرات معالم هذا النموذج والإحصاءات المرتبطة بها ويؤثر أيضاً على اختيار المتغيرات المؤثرة في نموذج الانحدار⁶⁴.

وسوف نتناول بشي من التفصيل استخدام بعض الطرق المعلمية والتحليلية لمعالجة موضوع الحالات الشاذة من حيث تحديدها وقياس أثرها على مقدرات المربعات الصغرى في نموذج تحليل الانحدار الخطى المتعدد فيما يلي:

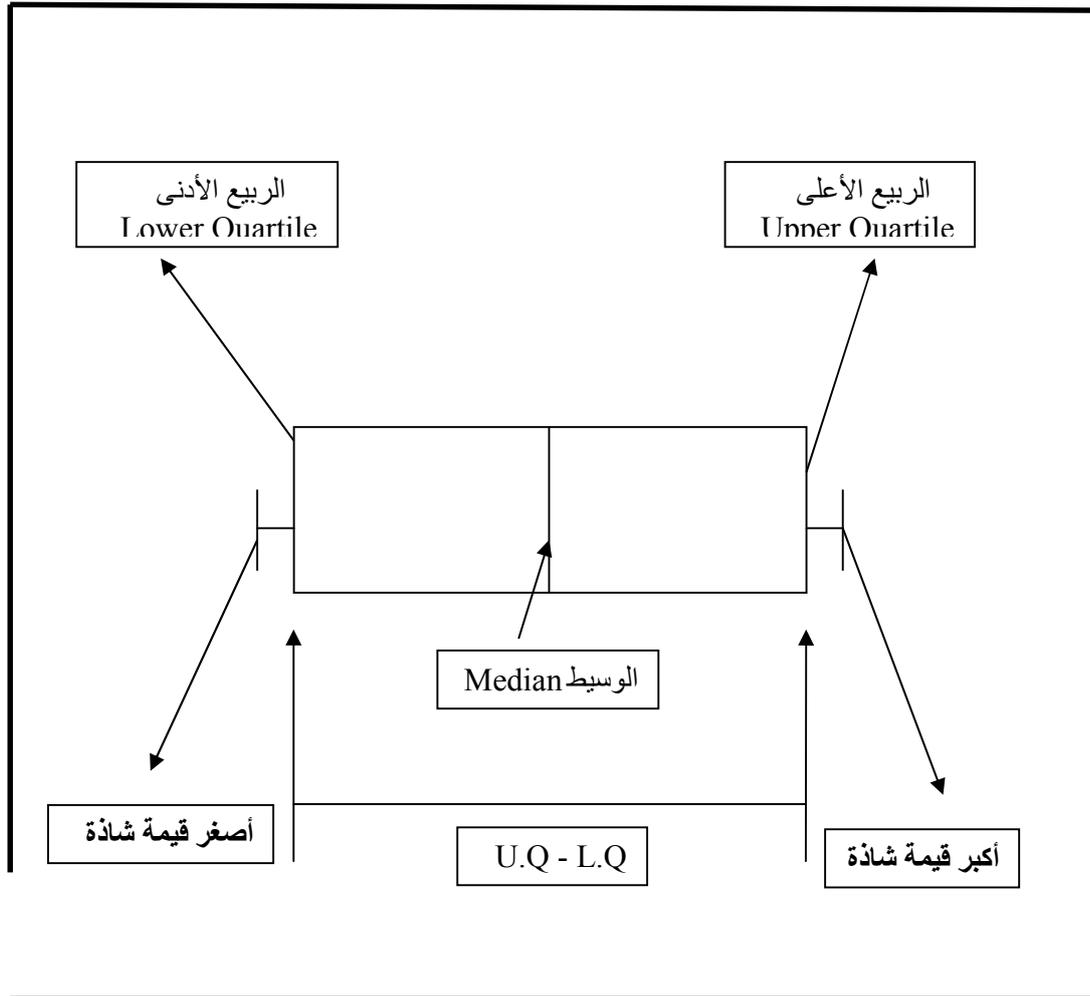
3-4-1 الطرق اللامعلمية - طريقة Tukey⁶⁵:

وهي تلك الطرائق التي لا تفترض توزيع معين علي البيانات متخطية بذلك أحد الافتراضات الأساسية التي تعتمد عليها الطرائق المعلمية ، ومن أهم هذه الطرائق طريقة الصندوق والقطع المخططة مع الملخصات الخمسة The Box and Whisker Plots with 5 Number summaries المقترحة من قبل العالم Tukey عام 1977م والتي تستخدم لفحص القيمة أو القيم الشاذة في حالة المتغير الواحد وبافتراض عدم معرفة التوزيع الاحتمالي للمتغير. ويمكن توضيح طريقة الرسم الصندوقي بالشكل (3-1) أدناه:

⁶⁴- Kleinbaum, J, et al, (1988),"Applied Regression Analysis and other Multivariable Methods", PWS-Kent Publishing Company, Boston, second edition, p:210.

⁶⁵- Tukey, J.W.(1977), "Exploratory Data Analysis", Addison-Wesley, reading, MA.

الشكل (1-3): يوضح مخطط الرسم الصندوقي:



و تعتبر المشاهدة أو المشاهدات متطرفة Extreme إذا كانت تبعد قيمها عن قمة أو قاعدة المستطيل مسافة تزيد عن ثلاثة أمثال طول الصندوق محسوبة من الربيع الأدنى أو الربيع الأعلى ، أي تكون المشاهدات أكبر من $Q_3 + 3(Q_3 - Q_1)$ أو أن تكون المشاهدات أقل من $Q_1 - 3(Q_3 - Q_1)$ وتظهر عادة في شكل نجمة (*).

أما المشاهدات الشاذة Outliers هي التي تبعد قيمها عن قمة أو قاعدة المستطيل مسافة ما بين 1.5 و 3 أمثال طول المستطيل محسوبة من الربيع الأدنى أو الربيع الأعلى وتظهر عادةً في شكل الحرف الإنجليزي (O) للإشارة على أنها شاذة.

3-4-2 اكتشاف المشاهدات الشاذة في المتغيرات المستقلة:

المشاهدات الشاذة في المتغيرات المستقلة تسمى أحياناً بنقاط الجذب Leverage points وهى تلك النقاط التي تقع بعيدة عن معظم القيم في المصفوفة X الموجودة في النموذج والتي تشمل واحده أو أكثر من المتغيرات المستقلة في تحليل الانحدار، وهذه النقاط لها تأثير كبير وقوى في مقدرات المربعات الصغرى الاعتيادية، وكذلك يقصد بنقطة الجذب هي قيمة من القيم X_i شاذة إلا إن قيم Y_i المقابلة لها تطابق النموذج بشكل جيد جداً وهذه النقطة تجذب تقديرات المربعات الصغرى نحوها.

ويتم الكشف عن المشاهدات الشاذة في المتغيرات المستقلة في نموذج الانحدار المتعدد

من خلال العناصر القطرية h_{ii} لمصفوفة القبعة Hat Matrix المعرفة بالعلاقة التالية:

$$H_{n \times n} = X(X^T X)^{-1} X^T = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix} \quad \dots\dots\dots (1-3)$$

حيث أن:

h_{ii} : هو العنصر القطري للمصفوفة H التي لها الخواص التالية:

1. أنها مصفوفة متماثلة $H = H^T$ وجامدة أيضاً $H^2 = H$.

2. العنصر القطري h_{ii} يعرف بالمعادلة التالية:

$$h_i = h_{ii} = h_i^T h_i = h_i^2 + \sum_{i \neq j} h_{ij}^2, \quad \frac{1}{n} \leq h_{ii} \leq 1$$

3. مجموع قيم العناصر القطرية يساوى عدد معالم نموذج الانحدار بما في ذلك المعامل الثابت

أي:

$$\sum_{i=1}^n h_i = p \quad \dots\dots\dots (2-3)$$

4. متوسط قيم العناصر القطرية يساوي عدد معالم نموذج الانحدار على عدد المشاهدات أي:

$$\bar{h} = \frac{p}{n} \quad \dots\dots\dots (3-3)$$

5. تسمى المصفوفة H بمصفوفة التنبؤ Predictor Matrix إذ يمكن الحصول منها على

القيم التقديرية للنموذج كما يلي:

$$\hat{Y} = H'Y = \sum_{i=1}^n h_{ij} y_j \quad \dots\dots\dots (4-3)$$

3-4-2-1 تشخيص المشاهدات الشاذة في المتغيرات المستقلة:

تعتبر العناصر القطرية h_{ii} مقياساً مفيداً في الكشف عن الحالات الشاذة في المتغيرات

المستقلة في نموذج الانحدار الخطى المتعدد ، ويسمى العنصر القطري بقيمة القبعة Hat

Value أو بالرافعة Leverage وتقيس الرافعة المسافة بين الحالة رقم i ($X_{1i}, X_{2i}, \dots, X_{pi}$)

ومتوسط قيم كل الحالات أو نقطة المنتصف $\bar{X}_{1i}, \bar{X}_{2i}, \dots, \bar{X}_{1p}$. فعندما تزداد قيم العناصر h_{ii}

تقل بالمقابل قيمة تباين البواقي e_i حتى إذا أصبحت قيمة $h_{ii} = 1$ سيكون تباين البواقي $\sigma_{ei}^2 = 0$

ونستنتج من ذلك أنه إذا كانت قيمة $h_{ii} = 1$ فإن قيمة المتغير التابع Y_i تساوى القيمة المقدرة

$$\hat{Y}_i$$

ولتشخيص الحالات الشاذة في المتغيرات المستقلة هنالك عدد من الطرق سوف نتناول

منها بالشرح ما يلي:

1. طريقة بيلسلي وآخرون (1980)⁶⁶:

تعد المشاهدات i مشاهدة شاذة أي لها نقطة انعطاف ومؤثرة في نتائج النموذج إذا كانت قوة

الرفع لها أكبر من ضعف متوسط الرافعات أي عندما تكون:

$$h_i > \frac{2(P)}{n} \quad \dots\dots\dots (5 - 3)$$

2. طريقة نيتير وآخرون (1990)⁶⁷:

اعتبر نيتير وآخرون أن الرافعة التي تزيد قيمتها عن (0.5) كبيرة و تشير إلى أن الحالة

شاذة وتستدعي دراستها، أي:

$$h_{ii} \geq 0.5 \quad \dots\dots\dots (6 - 3)$$

3. طريقة جون فوكس (1997)⁶⁸:

إن المعايير التي ذكرت من قبل لتشخيص القيم الشاذة في المتغيرات المستقلة تكون جيدة

في تحديد الحالات الشاذة في العينات الكبيرة والمتوسطة نسبةً إلى عدد المعالم ، أما في حالة

العينات الصغيرة تكون ضعيفة الأداء وغير عملية وذلك لان نسبة عدد المعالم إلى حجم العينة

$\frac{p}{n}$ سيكون كبيراً وبالتالي يتوقع ترشيح عدد كبير من الحالات الشاذة لفحصها لذلك أقترح جون

فوكس في هذا الصدد دراسة كل الحالات التي تزيد قيم رافعاتها h_i عن ثلاثة أضعاف متوسط

⁶⁶-Belsley,D. et al, (1980), "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity ", Wiley, New York, p:17.

⁶⁷ - Neter,J. et al, (1990), "Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs". (3rd edition). Irwin, Homewood, IL 60430, Boston, MA 02116, p:396.

⁶⁸- Fox, John, (1997), Applied Regression Analysis, Linear Models, and Related Methods. Sage publications, p:280.

قيم الارتفاعات بدلاً من دراسة الحالات إلى تزيد قيم ارتفاعاتها h_i عن ضعف قيم متوسط الارتفاعات كما في معيار بيلسلي وآخرون السابق ، أي تعتبر الحالة شاذة إذا كانت:

$$h_i > \frac{3(P)}{n} \quad \dots\dots\dots (7-3)$$

3-4-3 اكتشاف المشاهدات الشاذة في المتغير التابع:

للكشف عن المشاهدات الشاذة في المتغير التابع تستخدم بواقى ستيودنت المحذوفة Studentized Deleted Residual التي يتم الحصول عليها بإيجاد القيمة المعيارية للباقي المحذوف ، حيث يتم حسابه بإتباع الخطوات الآتية:

1. يتم أولاً حذف الحالة رقم i ومن ثم يتم بناء نموذج انحدار باستخدام بقية الحالات $n-1$.
2. يتم تقدير القيمة المتوقعة للمتغير التابع رقم i والتي يرمز لها بالرمز $\hat{Y}_{i(i)}$ وذلك بتعويض قيم المتغيرات المستقلة المناظرة للحالة رقم i إي $(X_{1i}, X_{2i}, \dots, X_{pi})$ في معادلة الانحدار التي تم تقديرها في الخطوة السابقة.
3. يتم حساب الباقي المحذوف بالصيغة الآتية:

$$d_i = Y_i - \hat{Y}_{i(i)} \quad \dots\dots\dots (8-3)$$

حيث أن:

d_i : الباقي المحذوف رقم i .

Y_i : القيمة الفعلية للمتغير التابع للملاحظة رقم i .

$\hat{Y}_{i(i)}$: القيمة المقدرة للمتغير التابع للملاحظة رقم i بعد حذف الحالة رقم i .

ويتضح من المعادلة (8-3) أننا نحتاج إلى بناء عدد n نموذج انحدار لحساب البواقى المحذوفة لجميع المشاهدات.

ولصعوبة هذا الإجراء تم تقدير معادلة أخرى تمكن من حساب البواقي المحذوفة دون الحاجة لبناء هذا العدد من النماذج تعطى بالصيغة الآتية⁶⁹:

$$d_i = \frac{e_i}{1 - h_{ii}} \quad \dots\dots\dots (9 - 3)$$

حيث أن:

e_i : البواقي العادية لنموذج دالة الانحدار باستخدام جميع الحالات n .

h_{ii} : العنصر القطري رقم i من المصفوفة المقدرة H باستخدام جميع المشاهدات لتقدير دالة الانحدار.

ونلاحظ من المعادلة (9-3) أعلاه أن قيمة الباقي المحذوف d_i تكون كبيرة إذا كانت قيمة h_{ii} كبيرة فهذا يجعل البواقي المحذوفة تحدد أحياناً المشاهدات الشاذة في قيم المتغير التابع التي تفشل البواقي العادية e_i في تحديدها، وفي أحيان أخرى تقود البواقي المحذوفة إلى التشخيصات نفسها التي تقود إليها البواقي العادية. لذلك تستخدم بواقي ستيودنت المحذوفة d_i^* للكشف عن المشاهدات الشاذة في المتغير التابع و التي تعطي بتطبيق المعادلة التالية:

$$d_i^* = \frac{d_i}{S.e(d_i)} \quad \dots\dots\dots (10 - 3)$$

حيث أن:

d_i^* : باقي ستيودنت المحذوف رقم i .

d_i : الباقي المحذوف رقم i .

$S.e(d_i)$: الانحراف المعياري للباقي المحذوف للمشاهدة رقم i ويساوي:

⁶⁹- Neter, J. et al, (1990), "Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs". (3rd edition). Irwin, Homewood, IL 60430, Boston, MA 02116, p:399.

$$S.e(d_i) = \sqrt{\frac{MSE_{(i)}}{1-h_{ii}}} \quad \dots\dots\dots (11-3)$$

حيث أن:

$MSE_{(i)}$: متوسط مجموع مربعات الخطأ لنموذج المقدر بعد حذف المشاهدة رقم i .

ونلاحظ من المعادلة (10-3) أننا نحتاج بناء عدد n نموذج انحدار لحساب d_i^* ولقد أثبت العلماء أن باقي ستودنت المحذوف يتبع توزيع t بدرجات حرية $n-p-2$ وتوصلوا إلى صيغة أخرى تمكن من حساب بواقي ستودنت المحذوفة دون الحاجة لبناء عدد n نموذج انحدار وذلك باستخدام المعادلة التالية:

$$d_i^* = e_i \left[\frac{n-p-1}{RSS(1-h_{ii})-e_i^2} \right]^{\frac{1}{2}} \quad \dots\dots\dots (12-3)$$

حيث أن:

e_i : الباقي العادي ، n : عدد المشاهدات ، p : عدد معالم النموذج.

RSS : مجموع مربعات البواقي لنموذج الانحدار المقدر باستخدام جميع المشاهدات n .

h_{ii} : العنصر القطري رقم i من المصفوفة المقدرة H باستخدام جميع المشاهدات لتقدير دالة الانحدار.

ولتشخيص المشاهدات الشاذة في المتغير التابع تتم مقارنة القيمة المطلقة لباقي ستودنت

المحذوف d_i^* بقيمة توزيع t عند درجات حرية $n-p-1$ ومستوى معنوية α_0 فإذا كانت:

$$|d_i^*| > t_{\alpha_0, n-p-1} \quad \dots\dots\dots (13-3)$$

تعتبر الحالة Y_i حالة شاذة لا بد من دراستها وتحديد مدى تأثيرها على مقدرات المربعات الصغرى.

3-5 المشاهدات الشاذة المؤثرة وطرق الكشف عنها:

المشاهدة المؤثرة هي التي تقع في الجهة المقابلة لوسط الخط الممثل لنموذج الانحدار إذ تكون بعيدة عن المحور y وكذلك بعيدة عن بقية المشاهدات الأخرى، تعمل هذه المشاهدة على تحريك خط الانحدار نحو الأعلى أو نحو الأدنى، بمعنى آخر باتجاه المشاهدة المؤثرة ، ويمكن أن يكون لها تأثير معنوي في نتائج نموذج الانحدار وفي القدرة على التنبؤ ، وتعتبر الحالة أو المشاهدة مؤثرة إذا كان إدراجها أو حذفها يؤدي إلى إحداث تغييرات جوهرية في قيم معاملات النموذج المقدرة والإحصاءات المرتبطة بها⁷⁰. وبعد الكشف عن الحالات الشاذة في المتغيرات المستقلة والمتغير المعتمد لابد من تحديد ما إذا كانت هذه الحالات لها تأثير على النتائج أم لا، وكما ذكرنا سابقاً تعتبر المشاهدة مؤثرة إذا كان استثنائها يسبب تغييرات جوهرية في تقدير معالم دالة الانحدار المقدرة والإحصاءات المرتبطة بها وتوجد عدة مقاييس لتحديد الحالات المؤثرة تعتمد في تحديد أثر الحالة الشاذة على قياس الفرق بين قيم مقدرات المربعات الصغرى باستخدام كل الحالات n و بإسقاط حالة واحدة أي $n-1$ ، وسوف نتناول بالشرح المفصل أكثر المقاييس المستخدمة لتحديد المشاهدات المؤثرة علي نتائج نموذج تحليل الانحدار الخطي المتعدد فيما يلي:

3-5-1 قياس التأثير علي القيم المقدرة - مقياس DFFITS⁷¹:

يستخدم هذا المقاييس لقياس أثر الحالة رقم i على القيمة المقدرة \hat{Y}_i ولقد طور بيلسلي

وآخرون المعادلة التالية لقياس أثر الحالة i علي القيمة المقدرة \hat{Y}_i :

⁷⁰-Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. and Black, W. C. (1998) "Multivariate Data Analysis", 5th edn. Prentice-Hall, New Jersey.

⁷¹- Belsley, D. et al, (1980), pp:14-15.

$$DFFITs_i = \frac{\hat{Y}_i - \hat{Y}_{i(i)}}{\sqrt{MSE_{(i)} h_{ii}}} \quad \dots\dots (14 - 3)$$

وتم تطوير الصيغة (14-3) إلى الصيغة التالية:

$$DFFITs_i = d_i^* \cdot \sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad \dots\dots (15 - 3)$$

حيث أن:

\hat{Y}_i : القيمة المقدرة للمتغير التابع للملاحظة i عند استخدام جميع المشاهدات.

$\hat{Y}_{i(i)}$: القيمة المقدرة للمتغير التابع بعد حذف الملاحظة رقم i .

d_i^* : باقي ستودنت المحذوف للملاحظة رقم i باستخدام جميع المشاهدات.

h_{ii} : قيم العنصر القطري رقم i من المصفوفة H المقدرة باستخدام جميع المشاهدات لتقدير

دالة الانحدار ، $MSE_{(i)}$: متوسط مربعات الخطأ عند حذف الملاحظة رقم i .

ونلاحظ من العلاقة (15-3) أن قيم $DFFITs$ تعتمد علي قيم h_{ii} و d_i^* ، أي أنه إذا

كانت الملاحظة i شاذة في المتغير X (قيمة h_{ii} لها كبيرة) وأيضاً إذا كانت الملاحظة i شاذة

في المتغير التابع Y (قيمة d_i^* لها كبيرة) فإن القيمة المطلقة لهذا المقياس تكون كبيرة أيضاً

وتكون القيمة الكبيرة كدليل نستدل منه فيما إذا كانت الملاحظة i مؤثرة أم لا.

ولتحديد أو تشخيص الحالات المؤثرة علي نتائج نموذج تحليل الانحدار الخطي المتعدد

هنالك عدد من الطرق الرسمية التي تستخدم لهذا الغرض نذكر منها ما يلي:

1. طريقة بيلسلي وآخرون: Belsley et al, 1980: p28:

أعتبر بيلسلي وآخرون الحالة مؤثرة على نتائج تحليل الانحدار الخطي المتعدد إذا كانت:

$$|DFFITs_i| > 2\sqrt{\frac{p}{n}} \quad \dots\dots (16 - 3)$$

2. طريقة نيتير وآخرون 1990: p401: Neter et al:

تستخدم هذه الطريقة في حالة العينات الكبيرة والصغيرة والمتوسطة، ففي حالة العينات

الصغيرة والمتوسطة تعتبر الحالة مؤثرة إذا كانت:

$$|DFFITS_i| > 1 \quad \dots\dots\dots (17 - 3)$$

أما في حالة العينات الكبيرة تعتبر الحالة مؤثرة إذا كانت:

$$|DFFITS_i| > 2\sqrt{\frac{p}{n}} \quad \dots\dots\dots (18 - 3)$$

3. طريقة شاترجي وهادي 1988 Chatterjee & Hadi⁷²:

أقترح كل من شاترجي و هادي معيار لمقارنة القيمة المطلقة لـ $DFFITS$ بقيمة أكبر قليلاً

من القيمة التي أقترحها بيلسلي وآخرون تعرف كالاتي:

$$|DFFITS_i| > 2\sqrt{\frac{p}{n-p}} \quad \dots\dots\dots (19 - 3)$$

3-5-2 قياس الأثر على معاملات الانحدار - مقياس $DFBETAS$ ⁷³:

يستخدم مقياس $DFBETAS$ الذي طوره بيلسلي وآخرون لقياس الفرق بين قيم معاملات

الانحدار b_k المقدر باستخدام جميع الحالات n وقيم معاملات الانحدار $b_{k(i)}$ المقدر بعد حذف

الحالة i أي باستخدام $n-1$ حالة، ويتم الحصول على هذا المقياس بتطبيق المعادلة التالية:

$$DFBETAS = \frac{b_k - b_{k(i)}}{S_{(i)} \cdot \sqrt{(X^T X)_{KK}^{-1}}} \quad \text{for } k = 0,1,2,\dots,p \quad \dots\dots\dots (20 - 3)$$

حيث أن:

b_k : معامل الانحدار رقم K المقدر باستخدام جميع الحالات n .

⁷²- Chatterjee, s. and Hadi, A. S.(1988), "Sensitivity Analysis in Linear Regression", New York: John Wiley.

⁷³- Belsley,D. et al,(1980), p13.

$S_{(i)}$: الخطأ المعياري لنموذج الانحدار الموفق باستخدام $n-1$ حالة، أي بعد إسقاط الحالة رقم i .

$b_{k(i)}$: معامل الانحدار رقم K المقدر باستخدام $n-1$ حالة بعد حذف الحالة رقم i .

$(X^T X)^{-1}_{KK}$: العنصر القطري رقم K من المصفوفة $(X^T X)^{-1}$.

ولتشخيص الحالات المؤثرة على قيمة معامل الانحدار رقم K فقد أقترح نيتز وآخرون⁷⁴

معياراً لتحديد الحالات المؤثرة في حالة العينات الصغيرة والمتوسطة والكبيرة، كالاتي:

في حالة العينات الصغيرة والمتوسطة تعتبر الحالة مؤثرة إذا كانت:

$$|DFBETAS_{k(i)}| > 1 \quad \dots\dots\dots (21 - 3)$$

أما في حالة العينات الكبيرة فتعتبر الحالة مؤثرة إذا كانت:

$$|DFBETAS_{k(i)}| > \frac{2}{\sqrt{n}} \quad \dots\dots\dots (22 - 3)$$

3-5-3 قياس الأثر علي كل معاملات الانحدار - مقياس مسافة كوك⁷⁵:

مقياس مسافة كوك Cook's Distance Measure هو مقياس إجمالي للتأثير المشترك

لقياس أثر حذف الحالة رقم i على جميع معاملات نموذج الانحدار المقدر، ويختلف مقياس

كوك عن مقياس $DFBETAS$ في أن الأخير يقيس أثر حذف الحالة رقم i على كل معامل من

معاملات الانحدار علي حده ويحسب مقياس كوك بتطبيق الصيغة التالية:

$$D_i = \frac{(b - b_{(i)})^T X^T X (b - b_{(i)})}{P \times MSE} \quad \dots\dots\dots (23 - 3)$$

⁷⁴ - Neter, J. et al, (1990), p403.

⁷⁵-Cook, R.D. (1977), "Detection of Influential Observation in Linear Regression. Technometrics 19: pp15-18.

حيث أن:

b : متجه معاملات نموذج الانحدار باستخدام الحالات n ، $b_{(i)}$: متجه معاملات نموذج الانحدار بعد حذف الحالة رقم i ، P : عدد معالم النموذج ، MSE : تباين نموذج الانحدار باستخدام كل الحالات n .

وعند حساب قيمة D_i حسب المعادلة (3-23) نلاحظ أنه نحتاج إلى بناء عدد n نموذج

انحدار، ولصعوبة هذا الإجراء يمكن حساب قيمة D_i بصيغة أخرى بتطبيق المعادلة التالية:

$$D_i = \frac{e_i^2}{P \times MSE} \left(\frac{h_{ii}}{(1-h_{ii})^2} \right) \quad \dots\dots\dots (24-3)$$

حيث أن:

e_i : قيمة البواقي العادية لنموذج الانحدار للملاحظة رقم i ، h_{ii} : قيم العنصر القطري رقم i من المصفوفة المقدرة H باستخدام جميع المشاهدات لتقدير دالة الانحدار.

ونلاحظ من المعادلة أعلاه (3-24) أن قيمة D_i تعتمد على كبر قيم الانعطاف h_{ii} أو

قيم البواقي e_i أو كليهما، فكلما كانت قيمة أي من h_{ii} أو e_i كبيرة تكون قيمة D_i أيضاً كبيرة.

ولتشخيص أثر الحالة رقم i على جميع معاملات نموذج الانحدار يقترح فوكس⁷⁶ أن تتم

مقارنة قيمة $(4/(n-p))$ فإذا كانت قيمة D_i أكبر من هذه القيمة تعتبر الحالة i حالة مؤثرة

على قيم معاملات نموذج الانحدار وبخلاف ذلك تكون الحالة غير مؤثرة، أي:

تعتبر الحالة مؤثرة إذا كانت:

$$D_i > \frac{4}{n-p} \quad \dots\dots\dots (25-3)$$

⁷⁶- Fox, John, (1997), P:276.

3-5-4 مقياس الأثر على الأخطاء المعيارية - مقياس COVRATIO⁷⁷:

يستخدم مقياس الأثر على الأخطاء المعيارية Influence on Standard Error الذي طوره بيلسلي وآخرون لقياس أثر أي حالة علي مصفوفة تباين وتغاير معاملات الانحدار المقدر ، وهذا المقياس عبارة عن نسبة محددة مصفوفة تباين - تغاير معاملات الانحدار المقدر بعد حذف الحالة رقم i إلى محددة مصفوفة تباين - تغاير معاملات الانحدار باستخدام جميع الحالات n ويحسب بالصيغة الآتية:

$$COVRATIO_i = \frac{\det(S_{(i)}^2 (X_{(i)}^T X_{(i)})^{-1})}{\det(S^2 (X^T X)^{-1})} \quad \dots\dots\dots (26 - 3)$$

حيث أن:

$S_{(i)}^2$: التباين المقدر بعد حذف الحالة رقم i .

$X_{(i)}$: مصفوفة البيانات بعد حذف المشاهدة رقم i من الدرجة $(p+1) \times (n-1)$.

وتم تطوير المعادلة (26-3) إلي الصيغة التالية:

$$COVRATIO_i = \frac{1}{(1-h_i) \left(\frac{n-p-1+d_i^{*2}}{n-p} \right)^p} \quad \dots\dots\dots (27 - 3)$$

حيث أن:

P : عدد معالم النموذج ، n : عدد المشاهدات ، h_{ii} : قيم العنصر القطري رقم i من المصفوفة

المقدرة H باستخدام جميع المشاهدات لتقدير دالة الانحدار ، d_i^* : باقي ستيودنت المحذوف

للمشاهدة رقم i باستخدام جميع المشاهدات.

ونلاحظ من الصيغة أعلاه أن قيمة هذا المقياس تعتمد على قيمتي الرافعة h_{ii} وباقي

ستيودنت d_i حيث تزيد قيمة المقياس بزيادة h_i وبانخفاض d_i .

⁷⁷- Belsley,D. et al,(1980), pp22-24.

ولتشخيص أثر الحالة رقم i على الأخطاء المعيارية لمعاملات الانحدار أقترح بيلسلي و آخرون مقارنة قيمة $COVRATIO$ المطلقة بالقيمة $(1+3(p)/n)$ فإذا كانت قيمة $COVRATIO$ المطلقة أكبر من هذه القيمة تعتبر الحالة i مؤثرة على قيم الأخطاء المعيارية لمعاملات الانحدار، أي تعتبر الحالة مؤثرة إذا كانت قيمة $COVRATIO$ تقع خارج المدى:

$$1 - \frac{3(p)}{n} > COVRATIO_i > 1 + \frac{3(p)}{n} \quad \dots\dots\dots (28 - 3)$$

6-3 بعض الحلول المقترحة لمعالجة القيم الشاذة:

بعد التأكد من وجود القيم الشاذة في مشاهدات المتغير التابع أو مشاهدات المتغيرات المستقلة لابد من معالجة هذه القيم لمعرفة مصادرها و الأسباب التي أدت إلى وجودها وذلك من أجل الحصول على نموذج انحدار تكون نتائجه خالية من تأثير وجود القيم الشاذة على معلمته المقدره. فإذا تبين بعد الفحص والدراسة أن وجود المشاهدات الشاذة ناتج عن خطأ في مرحلة جمع البيانات أو في مرحلة المعالجة فعلي الباحث تصحيح هذه الأخطاء وإعادة حل النموذج ، أما إذا كانت المشاهدات الشاذة بيانات حقيقية فهناك عدد من الحلول المقترحة نذكر منها ما يلي⁷⁸:

1. إعادة توصيف النموذج Model re-specification وذلك إما بإضافة أو حذف متغيرات مستقلة.
2. إجراء تحويلات إما للمتغير التابع أو المتغيرات المستقلة أو بعض منها، كتحويله اللوغاريتم والمعكوس والجذر التربيعي وغيرها.
3. حذف المشاهدات الشاذة إذا كان حجم العينة كبيراً وإعادة حل النموذج.

⁷⁸ - إسماعيل، محمد عبد الرحمن، (2001م)، مرجع سابق.

4. جمع بيانات إضافية لزيادة حجم العينة وبالتالي تقليل أثر وجود المشاهدات الشاذة إلا أن هذا الحل قد يصاحبه بعض المصاعب كارتباط الظاهرة محل الدراسة بفترة زمنية محددة مما يعنى الانتظار حني حلول الظاهرة مرة أخرى وقد تكون هذه الفترة غير معلومة كحدوث الكوارث فضلاً عن التكلفة المضافة لجمع هذه البيانات.

5. طريقة متوسط البتر Trimmed mean ونتلخص خطوات إجراء هذه الطريقة بترتيب مشاهدات العمود المشخص بأنه شاذاً تصاعدياً أو تنازلياً ، وتحذف أكبر قيمة وأصغر قيمة في بيانات العمود الشاذ ، ثم إيجاد الوسط الحسابي للقيم المتبقية أي إيجاد Trimmed mean لـ $n-2$ من القيم والذي يمثل تقديراً لهذه القيمة⁷⁹.

⁷⁹ - الجبوري، مني حسين، (1998)، "دراسة تحليلية للقيم الشاذة والقيم المفقودة لتصميم المربع اللاتيني وتصميم تام في حالة تكرار مشاهدات العينة"، رسالة ماجستير مقدمة إلى مجلس كلية الإدارة والاقتصاد، الجامعة المستنصرية، بغداد، العراق.

الفصل الرابع الجانب التطبيقي

1-4 تمهيد:

2-4 تعريف متغيرات الدراسة:

3-4 وصف متغيرات الدراسة:

4-4 تحليل الانحدار الخطي المتعدد قبل معالجة القيم الشاذة:

1-4-4 تقدير معاملات النموذج:

2-4-4 اختبار المعنوية الكلية للنموذج:

3-4-4 اختبار معنوية معاملات الانحدار الجزئية $\hat{\beta}_j$:

4-4-4 اختبار القدرة التفسيرية للنموذج:

5-4-4 اختبار النموذج من الناحية القياسية:

6-4-4 اختبار النموذج لعملية التنبؤ:

7-4-4 تشخيص القيم الشاذة في المتغيرات المستقلة والمتغير التابع:

8-4-4 قياس أثر المشاهدات الشاذة المؤثرة علي معاملات النموذج:

5-4 معالجة القيم الشاذة:

6-4 المقارنة بين نتائج النماذج المقدرة لاختيار أفضل نموذج:

7-4 تشخيص أفضل النماذج المقدرة:

1-4 تمهيد:

يتضمن هذا الفصل الجانب التطبيقي والذي من خلاله سوف نقوم بتطبيق الأسس النظرية والأدبيات المتعلقة بالدراسة التي تم التطرق إليها في الفصول السابقة وفي هذا الفصل نقوم أولاً بتقدير نموذج تحليل الانحدار الخطي المتعدد ثم تشخيصه من الناحية الإحصائية والقياسية بالإضافة إلى اختبار النموذج في عملية التنبؤ بقيم المتغير التابع وبعد ذلك نقوم بالكشف عن المشاهدات الشاذة في المتغيرات المستقلة والمتغير التابع ثم قياس أثر المشاهدات الشاذة المؤثرة علي قيم المتغير التابع المقدرة ومعاملات النموذج والأخطاء المعيارية وكذلك قياس الأثر علي كل معاملات الانحدار باستخدام بعض الطرق العلمية ثم محاولة تقليل أثر هذه المشاهدات الشاذة علي معاملات نموذج الانحدار المتعدد و أخيراً مقارنة النتائج قبل معالجة المشاهدات الشاذة وبعد المعالجة معتمدين علي بعض المعايير التي تبين كفاءة نموذج الانحدار الخطي المتعدد. وسوف نعتمد في تحليل بيانات الدراسة علي برامج التحليل الإحصائي SPSS, Eviews, Exeal.

2-4 تعريف متغيرات الدراسة:

إن البيانات التي اعتمدت عليها الدراسة في التطبيق العملي تتمثل في الآتي:

أولاً: المتغيرات المستقلة:

المعدل الفصلي الأول: يمثل المتغير المستقل (X_1) .

المعدل الفصلي الثاني: يمثل المتغير المستقل (X_2) .

المعدل الفصلي الثالث: يمثل المتغير المستقل (X_3) .

ثانياً: المتغير التابع:

المعدل التراكمي: يمثل المتغير التابع (Y_i) .

3-4 وصف متغيرات الدراسة:

الجدول رقم (4-1): يوضح بعض المقاييس الوصفية التي تصف طبيعة بيانات الدراسة:

المتغيرات Variables	الوسط الحسابي Mean	الوسيط Median	الانحراف المعياري Std. Deviation	أقل قيمة Minimum	أكبر قيمة Maximum
X ₁	2.099	2.105	0.64355	0.77	3.23
X ₂	2.090	2.135	0.61644	0.64	3.27
X ₃	1.989	2.035	0.65027	0.14	3.02
Y	2.076	2.115	0.56100	0.92	3.00

المصدر: إعداد الباحث باستخدام برنامج SPSS:

يتضح من الجدول (4-1) أعلاه الآتي:

أن الوسط الحسابي لقيم المتغير X₁ بلغ (2.099) وأن قيمة الوسيط (2.105) و أقل قيمة تساوي (0.77) وأكبر قيمة تساوي (3.23) وأن قيمة الانحراف المعياري بلغت (0.64355). كما بلغ الوسط الحسابي لقيم المتغير X₂ (2.090) و أن قيمة الوسيط (2.135) و أقل قيمة تساوي (0.64) وأكبر قيمة تساوي (3.27) وأن قيمة الانحراف المعياري بلغت (0.61644). كما بلغ الوسط الحسابي لقيم المتغير X₃ (2.090) وقيمة الوسيط (2.135) و أقل قيمة تساوي (0.64) وأكبر قيمة تساوي (3.27) وأن قيمة الانحراف المعياري بلغت (0.61644). كما بلغ الوسط الحسابي لقيم المتغير التابع Y_i (2.076) وقيمة الوسيط (2.115) و أقل قيمة تساوي (0.92) وأكبر قيمة تساوي (3.00) وأن قيمة الانحراف المعياري بلغت (0.56100).

4-4 تحليل الانحدار الخطي المتعدد قبل معالجة القيم الشاذة:

1-4-4 تقدير معاملات النموذج:

بتطبيق طريقة المربعات الصغرى الاعتيادية المعرفة بالصيغة (2-11) يمكن الحصول

علي معادلة نموذج الانحدار الخطي المتعدد المقدر من نتائج الجدول (2-4) التالي:

الجدول (2-4): معاملات النموذج و قيم الخطأ القياسي للنموذج المقدر قبل معالجة القيم الشاذة:

معلمات النموذج Parameters	قيم المعاملات Coefficient	الخطأ القياسي Std. Error
$\hat{\beta}_1$	0.400894	0.022252
$\hat{\beta}_2$	0.240036	0.027083
$\hat{\beta}_3$	0.349354	0.022205
$\hat{\beta}_0$	0.037983	0.037545

المصدر: إعداد الباحث باستخدام برنامج SPSS:

يتضح من الجدول رقم (2-4) أعلاه أن معادلة نموذج الانحدار الخطي المتعدد المقدر

تأخذ الصيغة التالية:

$$\hat{Y}_i = 0.037983 + 0.400894 X_1 + 0.240036 X_2 + 0.349354 X_3$$

$$R^2 = 0.992 \quad MSE = 0.003 \quad F = 1063.322$$

2-4-4 اختبار المعنوية الكلية للنموذج:

لاختبار معنوية النموذج المقدر بصورة كلية ، أي اختبار تأثير المتغيرات المستقلة بصورة

كلية علي المتغير التابع نستخدم اختبار F المعروف بالصيغة (2-23) والذي يمكن الحصول

عليه من جدول تحليل التباين التالي:

الجدول (3-4): يوضح تحليل التباين للنموذج المقدر قبل معالجة القيم الشاذة:

مصادر التباين Source of Variation	درجات الحرية Degrees of freedom	مجموع المربعات Sum of Squares	متوسط مجموع المربعات Mean Square	قيمة اختبار F المحسوبة	مستوي المعنوية P-value
Regression	3	9.053	3.018	1063.322	0.000
Residual	26	0.074	0.003		
Total	29	9.127			

المصدر: إعداد الباحث باستخدام برنامج SPSS:

يتبين من جدول تحليل التباين (3-4) أعلاه أن قيمة F المحسوبة تساوي 1063.322

وأن مستوي المعنوية لها يساوي 0.000 وهي أقل من مستوي المعنوية 0.05 المعتمد في هذه الدراسة ، عليه نرفض فرض العدم القائل بان نموذج الانحدار غير معنوي ، وهذا يعني أن هنالك تأثيراً معنوياً من قبل المتغيرات المستقلة علي المتغير التابع.

3-4-4 اختبار معنوية معاملات الانحدار الجزئية $\hat{\beta}_j$:

يمكن اختبار تأثير كل متغير مستقل بوجود باقي المتغيرات المستقلة الأخرى علي المتغير

التابع باستخدام اختبار t وذلك من خلال تطبيق العلاقة (2-24) ويمكن الحصول علي نتائج

هذا الاختبار من الجدول (4-4) أدناه

الجدول (4-4): يوضح نتائج اختبار t للنموذج المقدر قبل معالجة القيم الشاذة:

معلومات النموذج Parameters	قيم المعاملات Coefficient	قيم اختبار t المحسوبة t-Statistic	مستوي المعنوية P-value
$\hat{\beta}_1$	0.400894	18.0164	0.000
$\hat{\beta}_2$	0.240036	8.863025	0.000
$\hat{\beta}_3$	0.349354	15.7329	0.000
$\hat{\beta}_0$	0.037983	1.011671	0.321

المصدر: إعداد الباحث باستخدام برنامج SPSS:

يتضح من الجدول (4-4) أعلاه أن معاملات النموذج $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ معنوية وذلك بالاعتماد على مستوى المعنوية لها 0.000 وهي أقل من مستوى المعنوية 0.05 المعتمد في هذه الدراسة ، وهذا يعني أن كل متغير مستقل علي حدا يؤثر معنوياً علي المتغير التابع ، كما نلاحظ أن قيمة المعامل الثابت $\hat{\beta}_0$ غير معنوية.

4-4-4 اختبار القدرة التفسيرية للنموذج:

للحكم علي المقدرة التفسيرية للنموذج تم حساب معامل التحديد R^2 ومعامل التحديد

المعدل \bar{R}^2 بتطبيق المعادلتين (28-2) ، (29-2) والنتائج مبينة بالجدول (5-4) التالي:

الجدول (5-4): معاملات التحديد وإحصائية D-W للنموذج المقدر قبل معالجة القيم الشاذة:

معامل الارتباط المتعدد R	معامل التحديد R^2	معامل التحديد المعدل \bar{R}^2	إحصائية D – W
0.996	0.992	0.991	2.42

المصدر: إعداد الباحث باستخدام برنامج SPSS:

نلاحظ من الجدول (4-5) أن قيمة معامل التحديد ومعامل التحديد المعدل بلغت 0.99 وهي قيمة عالية جداً ، وهذا يعنى أن المتغيرات المستقلة المضمنة في النموذج تفسر نسبة 99% من التغيرات التي تحدث في المتغير التابع ، وأن 1% من التغيرات يرجع إلي عوامل أخرى منها الخطأ العشوائي.

4-4-5 اختبار النموذج من الناحية القياسية:

بعد تقدير النموذج واختباره من الناحية الإحصائية ومعرفة القدرة التفسيرية للنموذج نقوم باختباره من الناحية القياسية لمعرفة مدى تحقق الفرضيات الخاصة به وذلك فيما يلي:

أولاً: اختبار تجانس تباين حد الخطأ العشوائي:

للكشف عن تجانس تباين حد الخطأ العشوائي هنالك عدة اختبارات منها اختبار جولدفيلد-كواندت واختبار آرش ، وبتطبيق اختبار جولدفيلد-كواندت قمنا بترتيب المشاهدات تصاعدياً حسب المتغير X_3 وتم استبعاد 20% من المشاهدات في المنتصف ، والمشاهدات التي سيتم استبعادها هي 13,14,15,16,17,18 وفي هذه الحالة سيكون لدينا مجموعتين كل مجموعة تضم (12) مشاهدة ، حيث أن المجموعة الأولى تضم المشاهدات من (1-12) والمجموعة الثانية تضم المشاهدات من (19-30) وبعد ذلك تم حساب مجموع مربعات الخطأ SSE لكل مجموعة ، حيث $SSE_1=0.2418$ و $SSE_2=0.00048$ ثم حساب قيمة F المحسوبة بتطبيق المعادلة (2-50) حيث بلغت قيمتها 0.019851 ومن الجداول الخاصة بتوزيع F تم الحصول علي قيمة $F_{12,12,0.05}$ الجدولية والتي تساوي 2.96 وبما أن قيمة F المحسوبة أقل من قيمة $F_{12,12,0.05}$ الجدولية ، عليه نقبل فرض العدم H_0 القائل بان هنالك تجانس أو ثبات في تباين الأخطاء.

ثانياً: اختبار الارتباط الخطي المتعدد بين المتغيرات المستقلة:

الجدول (4-6): يوضح نتائج اختبارات الكشف عن الارتباط الخطي المتعدد بين المتغيرات

المستقلة للنموذج المقدر قبل معالجة القيم الشاذة:

المتغيرات	قيم عامل	قيم	قيم	قيم	قيمة
Variables	تضخم	التسامح	الجزور	مؤشر	رقم
	التباين	Tolerance	الكامنة	الحالة	الحالة
	VIF		Eigen value	CI	CN
X ₁	2.095	0.477	0.052	8.652	14.314
X ₂	2.848	0.351	0.036	10.399	
X ₃	2.131	0.469	0.019	14.314	
Constant	-	-	3.893	1.000	

المصدر: إعداد الباحث باستخدام برنامج SPSS:

نلاحظ من الجدول (4-6) أعلاه أن قيم VIF لجميع المتغيرات المستقلة أقل من 10

وبالتالي لا يوجد ارتباط خطي متعدد بين المتغيرات المستقلة، كما أن قيم Tolerance لجميع

المتغيرات المستقلة أكبر من 0.10 وهذا دليل أيضاً على عدم وجود الارتباط الخطي المتعدد بين

المتغيرات المستقلة. كما نلاحظ أيضاً أن جميع قيم مؤشر الحالة لجميع المتغيرات المستقلة أقل

من 15 وهذا يعني عدم وجود الارتباط الخطي المتعدد بين المتغيرات المستقلة. كما بلغت قيمة

رقم الحالة 14.314 وهي قيمة أقل من (20-30) وهذا يؤكد أيضاً على عدم وجود الارتباط

الخطي المتعدد بين المتغيرات المستقلة.

ثالثاً: اختبار الاستقلال الذاتي للبواقي:

توجد عدة اختبارات للحكم على مدى وجود استقلال ذاتي بين بواقي النموذج احدي هذه

الاختبارات اختبار Durbin-Watson واعتماداً على المعادلة (2-135) نجد أن قيمة هذا

الاختبار تساوي 2.42 وذلك اعتماداً على نتائج الجدول (4-5) ، ومن الجداول الخاصة باختبار درين- واتسون تم استخراج قيمتي الحد الأدنى $dl=1.21$ و الحد الأعلى $du=1.65$ آخذين في الاعتبار عدد المشاهدات $n=30$ وعدد المتغيرات المستقلة $k=3$ ، ونجد أن قيم كل من dl و du تحدان مساحة ما بين 0 و 4 كما هو موضح بالشكل التالي:

الشكل (4-1): مناطق القبول والرفض لاختبار درين- واتسون للنموذج المقدر قبل معالجة القيم الشاذة:

0	dl	du	2	$4-du$	DW	$4-dl$	4
$\rho > 0$		$\rho = 0$	$\rho = 0$				$\rho < 0$
ارتباط ذاتي	منطقة	عدم وجود	عدم وجود	منطقة	منطقة	ارتباط ذاتي	ارتباط ذاتي
ذاتي	غير محددة	وجود ارتباط ذاتي	وجود ارتباط ذاتي	غير محددة	غير محددة	سالب	سالب
موجب							
	1.21	1.65	2.35	2.42	2.79		

المصدر: إعداد الباحث اعتماداً على نتائج جدول درين- واتسون وإحصائية D-W :

نلاحظ من خلال الشكل (4-1) أعلاه أن قيمة D-W المحسوبة بلغت 2.42 وهي تقع ضمن منطقة قرار غير محدد بشأن الارتباط الذاتي للبواقي لذا يتعين البحث عن اختبار آخر خلاف هذا الاختبار.

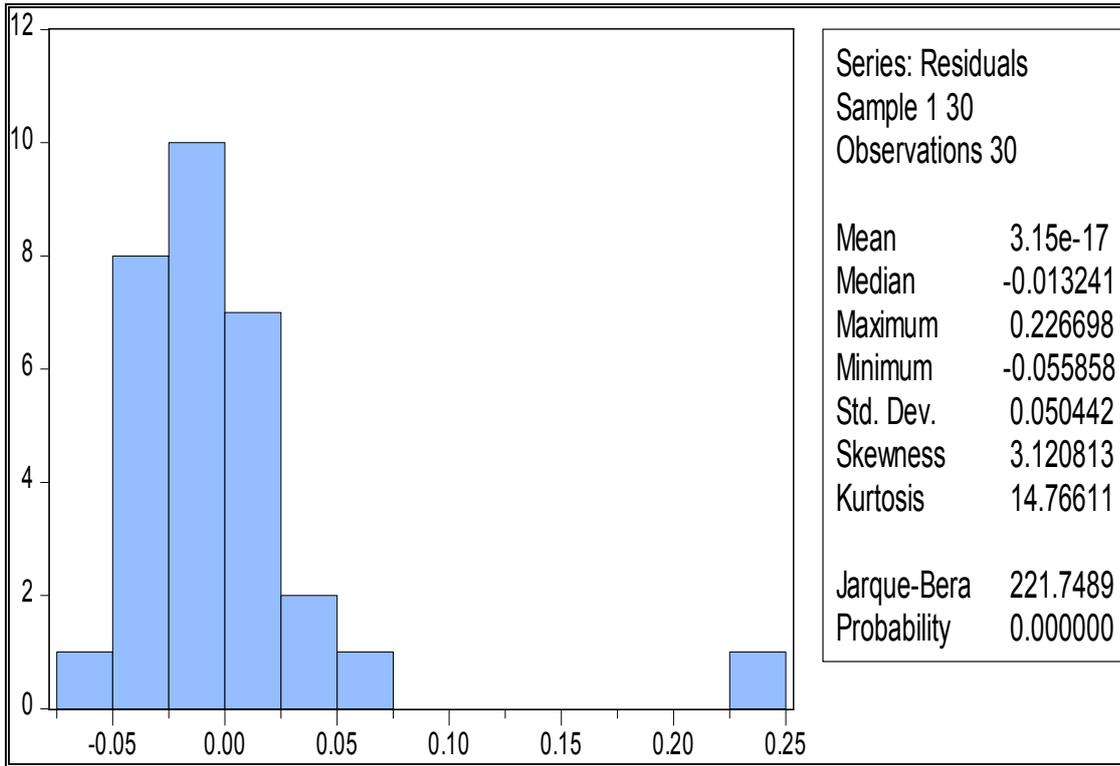
رابعاً: اختبار التوزيع الطبيعي للبقايا:

1- اختبار جاك - بيرا Jarque-Bera Test :

بتطبيق اختبار جاك بيرا Jarque-Bera Test الموضح بالمعادلة (2-160) تم

الحصول علي نتائج هذا الاختبار كما بالجدول التالي:

الجدول (4-7): اختبار جاك- بيرا للتوزيع الطبيعي للنموذج المقدر قبل معالجة القيم الشاذة:



المصدر: إعداد الباحث باستخدام برنامج Eviews :

يتضح من الجدول (4-7) أعلاه أن قيمة إحصائية اختبار Jarque-Bera بلغت

221.75 وأن مستوي المعنوية لها بلغ 0.000 وهي قيمة أقل من مستوي المعنوية 0.05، عليه

نرفض فرض عدم القائل بأن البواقي تتوزع طبيعياً. وهذا يعني مخالفة هذه الفرضية.

2- اختبار كلومجروف- سيمرنوف & اختبار ويلك شابيرو:

الجدول (4-8): يوضح اختبار كلومجروف- سيمرنوف & اختبار ويلك شابيرو للتوزيع الطبيعي

للمنموذج المقدر قبل معالجة القيم الشاذة:

الاختبار	إحصائية الاختبار	درجات الحرية	مستوي المعنوية
كلومجروف- سيمرنوف	0.224	30	0.001
ويلك شابيرو	0.68	30	0.000

المصدر: إعداد الباحث باستخدام برنامج SPSS:

يتضح من نتائج الجدول (4-8) أعلاه أن قيمة مستوى المعنوية أقل من مستوى المعنوية

0.05 في كلا الاختبارين ، عليه فأنتنا نرفض فرضية العدم ونقبل الفرضية البديلة وهذا يعني عدم

تبعية البواقي للتوزيع الطبيعي. وهذه النتيجة تدعم النتيجة التي توصلنا إليها من خلال اختبار

جاك - بيررا Jarque-Bera Test السابق.

4-4-6 اختبار النموذج لعملية التنبؤ:

إذا اجتاز النموذج المقدر كل الاختبارات السابقة وظهر أنه خال من المشكلات القياسية ،

عندها يمكن استخدامه للتنبؤ بقيم المتغير التابع لفترات مختارة داخل العينة ومقارنتها بالقيم

الفعلية الموجودة بالعينة ، وهناك عدة اختبارات تستخدم لتقييم النموذج في عملية التنبؤ منها ما

يلي.

أولاً: قياس مدى استقرار تقديرات النموذج خلال فترة الدراسة:

لاختبار استقرارية النموذج نعتمد علي دالة الارتباط الذاتي والجزئي وكما موضح بالجدول

التالي:

الجدول (4-9): اختبار دالة الارتباط الذاتي والجزئي للنموذج المقدر قبل معالجة القيم الشاذة:

Date: 03/30/15 Time: 02:23 Sample 1 30 Included observations: 30					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.220	-0.220	1.6030	0.205
		2 0.041	-0.008	1.6601	0.436
		3 -0.066	-0.062	1.8143	0.612
		4 -0.088	-0.121	2.0982	0.718
		5 -0.028	-0.078	2.276	0.831
		6 0.153	0.136	3.0646	0.801
		7 0.028	0.086	3.0963	0.876
		8 0.030	0.038	3.364	0.926
		9 -0.003	0.025	3.367	0.959
		10 -0.127	-0.093	3.9119	0.951
		11 0.079	0.060	4.2251	0.963
		12 -0.137	-0.129	5.2273	0.950
		13 0.008	-0.092	5.2312	0.970
		14 -0.060	-0.117	5.4443	0.979
		15 0.014	-0.046	5.4562	0.988
		16 0.002	0.002	5.4564	0.993

المصدر: إعداد الباحث باستخدام برنامج Eviews :

نلاحظ من الجدول (4-9) أعلاه أن رسم شكل دالة الارتباط الذاتي والجزئي لبواقي

النموذج المقدر تقع أعمدها داخل مجال الثقة ، عليه نقول أن النموذج مستقر. وللتأكد من ذلك

يمكن تطبيق اختبار Ljung-Box المعروف بالمعادلة (2-31) حيث يستخدم هذا الاختبار

لاختبار المعنوية الكلية لمعاملات دالة الارتباط الذاتي وتوافق إحصائية الاختبار المحسوبة Q

آخر قيمة في العمود Q-stat من الجدول أعلاه أي: $Q = n(n+2) \sum_{k=1}^m \frac{\hat{P}_k^2}{n-k} = 5.4564$

وبحساب قيمة $\chi^2_{(16,0.05)} = 26.296$ نجد أنها تساوي $\chi^2_{(16,0.05)}$ وبما أن القيمة $5.4564 < 26.96$ ،

عليه نقبل فرض العدم ونرفض الفرض البديل وهذا يعني أن النموذج مستقر.

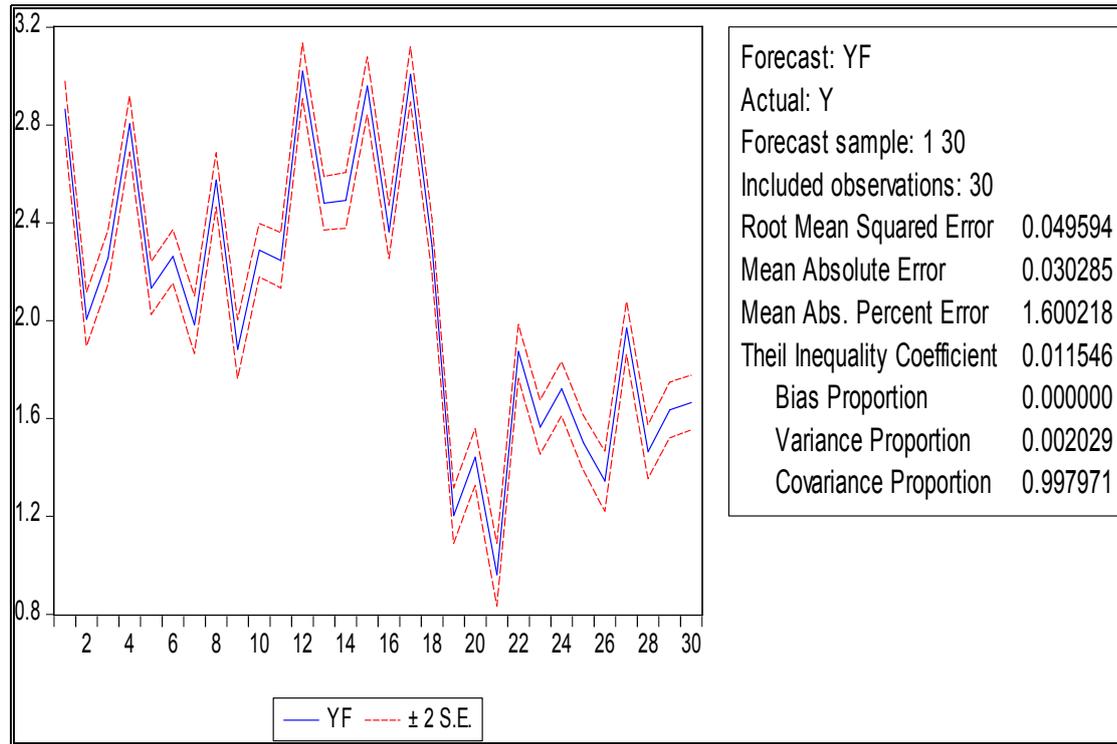
ثانياً: اختبار القدرة التنبؤية للنموذج :

هناك عدة مقاييس تستخدم لقياس دقة تنبؤات النموذج المقدر من بينها معامل التساوي

لتايل و معامل متوسط الخطأ المطلق MAE و الجذر التربيعي لمتوسط مربع الخطأ RMSE

والجدول التالي يوضح هذه الاختبارات.

الجدول (4-10): اختبارات تقييم القدرة التنبؤية للنموذج المقدر قبل معالجة القيم الشاذة:



المصدر: إعداد الباحث باستخدام برنامج Eviews :

يتضح من الجدول (4-10) أعلاه أن قيمة معامل تايل بلغت 0.011546 وهي قيمة

قريبة من الصفر، كما أن قيمة متوسط الخطأ المطلق MAE بلغت 0.030285 وهي أيضاً

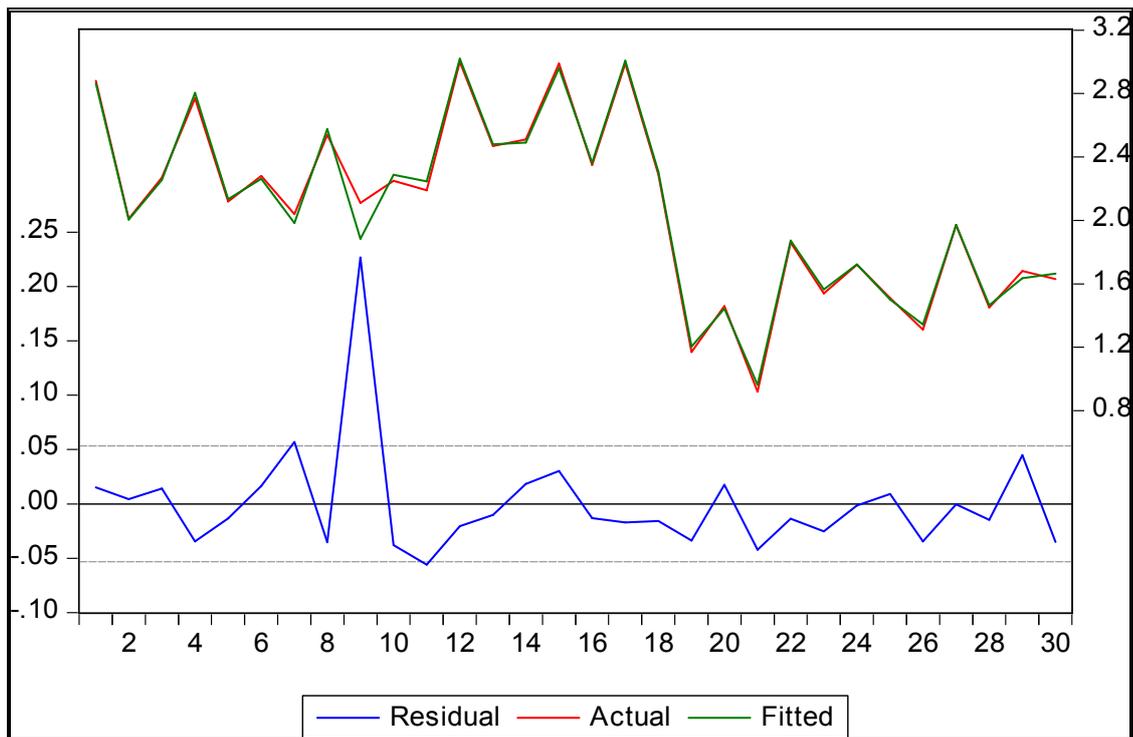
قيمة قريبة من الصفر مما يعني أن كفاءة النموذج عالية في التنبؤ بقيم المتغير التابع ، كما أن

قيمة الجذر التربيعي لمتوسط مربع الخطأ RMSE بلغت 0.049594 وكلما بلغت قيمة هذا

المؤشر 0.05 فأقل دل ذلك علي أن النموذج يطابق البيانات.

وأيضاً يمكن التحقق من مقدرة النموذج للتنبؤ من خلال مقارنة منحنى القيم الفعلية مع منحنى القيم المقدرة والذي يوضحه الشكل (2-4) التالي:

الشكل (2-4): مقارنة منحنى القيم الفعلية مع منحنى القيم المقدرة للنموذج المقدر قبل معالجة القيم الشاذة:



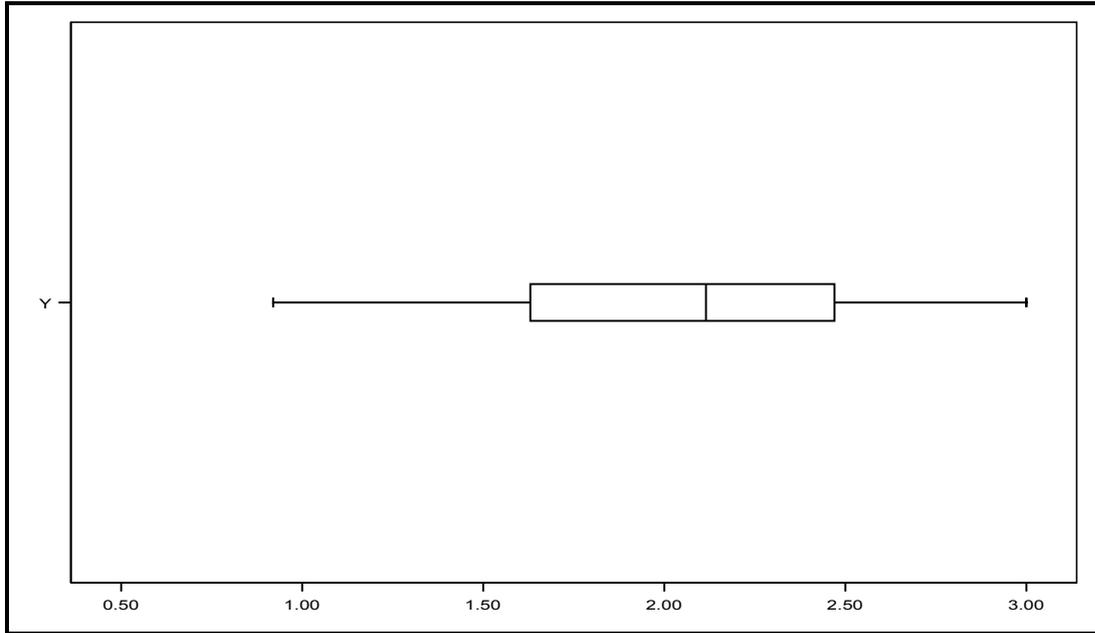
المصدر: إعداد الباحث باستخدام برنامج Eviews :

نلاحظ من الشكل (2-4) أعلاه مدى التطابق الشديد بين منحنى القيم الفعلية Actual ومنحنى القيم التقديرية Fitted في بعض القيم ، كما نلاحظ على ذات الشكل أن هنالك بعض القيم غير متطابقة وإن دل هذا إنما يدل على احتمال أن هذه القيم شاذة عند تلك النقاط وتؤثر على تنبؤات النموذج ولا بد من معالجتها لتحسين أداء النموذج في عملية التنبؤ بقيم المتغير التابع.

7-4-4-4 تشخيص القيم الشاذة في المتغيرات المستقلة والمتغير التابع:

أولاً: اكتشاف القيم الشاذة باستخدام الرسم الصندوقي:

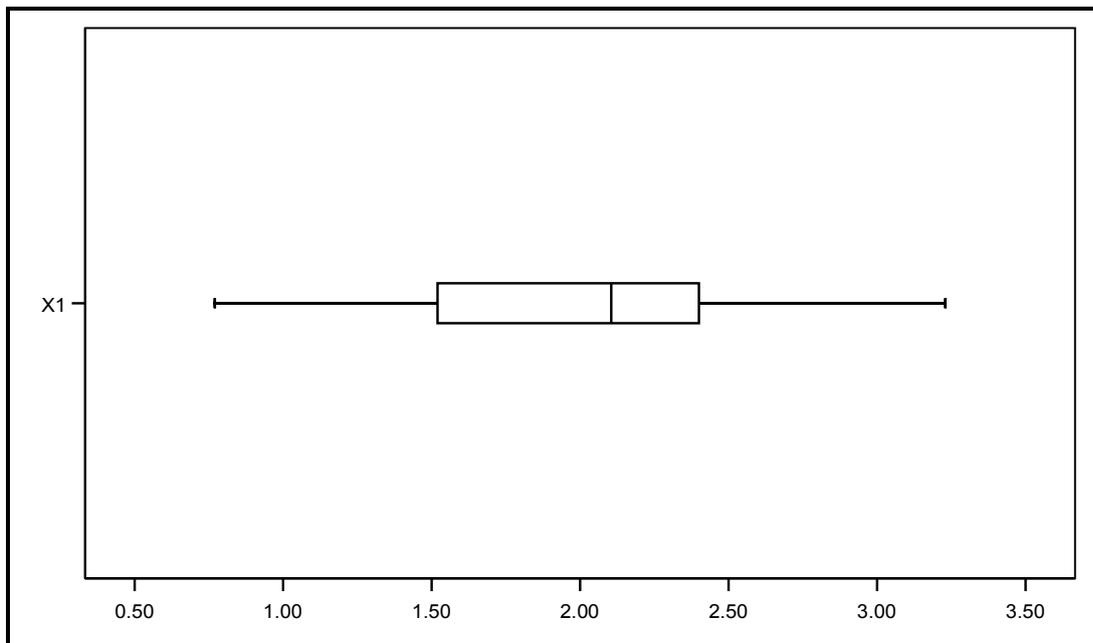
الشكل (3-4): يوضح الرسم الصندوقي لقيم المتغير التابع y:



المصدر: إعداد الباحث باستخدام البرنامج SPSS:

يتضح من الشكل (3-4) أعلاه عدم وجود قيم شاذة في بيانات المتغير التابع.

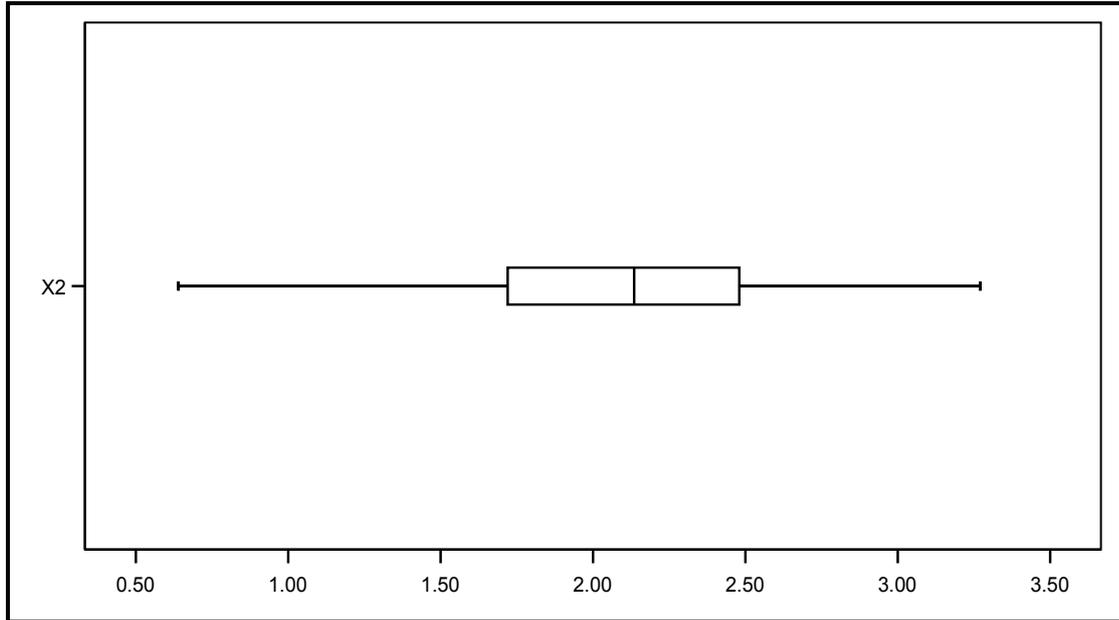
الشكل (4-4): يوضح الرسم الصندوقي لقيم المتغير المستقل X_1 :



المصدر: إعداد الباحث باستخدام البرنامج SPSS:

يتضح من الشكل (4-4) أعلاه عدم وجود قيم شاذة في بيانات المتغير المستقل X_1 .

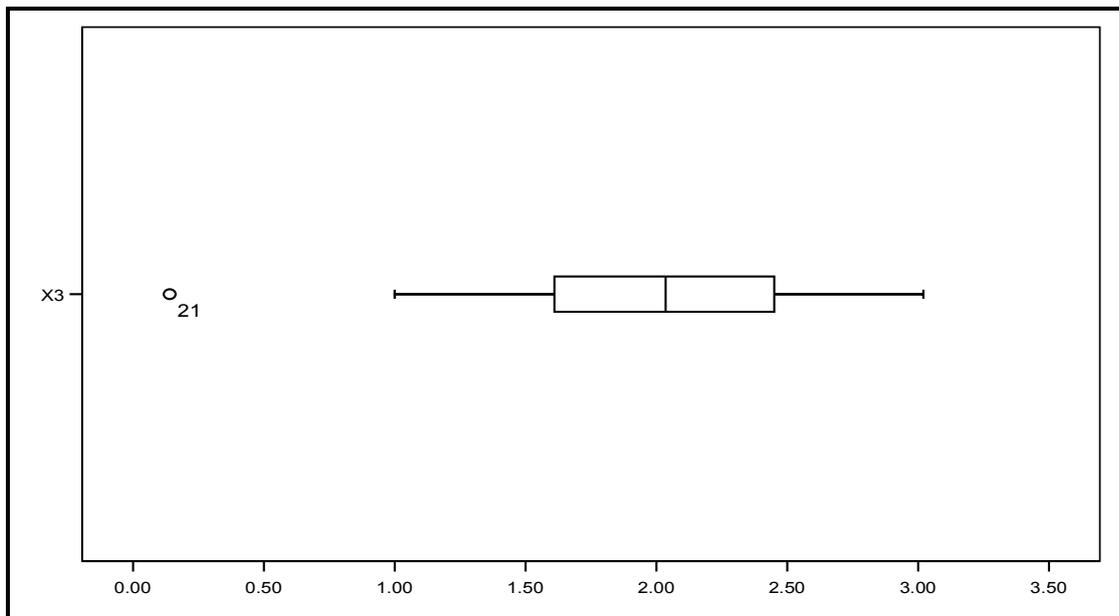
الشكل (4-5): يوضح الرسم الصندوقي لقيم المتغير المستقل X_2 :



المصدر: إعداد الباحث باستخدام البرنامج SPSS:

يتضح من الشكل (4-5) أعلاه عدم وجود قيم شاذة في بيانات المتغير المستقل X_2 .

الشكل (4-6): يوضح الرسم الصندوقي لقيم المتغير المستقل X_3 :



المصدر: إعداد الباحث باستخدام البرنامج SPSS:

يتضح من الشكل (4-6) أن المشاهدة (21) شاذة في بيانات المتغير المستقل X_3 .

ثانياً: اكتشاف القيم الشاذة في المتغيرات المستقلة باستخدام المصفوفة H:

جدول (4-11): قيم البواقي، قيم الرافعات، قيم معايير تشخيص المشاهدات الشاذة للنموذج

المقدر قبل معالجة القيم الشاذة:

Obs.	Resid. (ei)	Hat Matrix (hii)	Big(hii) hii>0.267	Big(hii) hii>0.4	Big(hii) hii>0.5
1	0.015227	0.146547			
2	0.004312	0.069849			
3	0.014166	0.076422			
4	-0.03452	0.166537			
5	-0.01328	0.037377			
6	0.016617	0.05897			
7	0.05685	0.226718			
8	-0.03532	0.086406			
9	0.226698	0.281979	0.281979		
10	-0.03785	0.056824			
11	-0.05586	0.136147			
12	-0.02051	0.138304			
13	-0.01004	0.054866			
14	0.018432	0.16425			
15	0.030268	0.216279			
16	-0.0132	0.04286			
17	-0.01704	0.13307			
18	-0.01569	0.103404			
19	-0.03374	0.121471			
20	0.017653	0.174047			
21	-0.04213	0.458989	0.458989	0.458989	
22	-0.01362	0.070181			
23	-0.02531	0.074491			
24	-0.00153	0.089176			
25	0.00915	0.114138			
26	-0.03436	0.339514	0.339514		
27	-0.0005	0.043835			
28	-0.01472	0.078807			
29	0.044906	0.137036			
30	-0.03506	0.101506			

المصدر: إعداد الباحث باستخدام برنامج Eviews:

يتم الكشف عن القيم الشاذة في المتغيرات المستقلة من خلال حساب عناصر القطر الرئيسي للمصفوفة Hat Matrix وتطبيق المعادلة (3-1) والاعتماد على برنامج Eviews تم الحصول على النتائج كما بالجدول (4-11) أعلاه. ولتشخيص المشاهدات الشاذة في المتغيرات المستقلة هنالك عدة معايير ، فحسب معيار بيلسلي وآخرون الذي يتم الحصول عليه بتطبيق المعادلة (3-5) تعتبر المشاهدة شاذة إذا تجاوزت ضعف متوسط قيم h_{ii} أي: $h_{ii} > \frac{2(p)}{n} = \frac{2 * 4}{30} = 0.267$ ، عليه نلاحظ من الجدول (4-11) أن هنالك ثلاث حالات تزيد قيم رافعاتها عن ضعف متوسط قيم الرافعات 0.267 وهي: (9) ، (21) ، (26) حيث أن قيم الرافعات المناظرة لهذه الحالات كانت 0.281979 ، 0.458989 ، 0.339514 علي التوالي.

أما إذا تم الأخذ بالمعيار الذي حدده نيتزر وآخرون $h_{ii} \geq 0.5$ فنجد أنه لا توجد حالات تزيد قيم رافعاتها عن 0.5.

أما إذا اعتمدنا علي المعيار الذي حدده جون فوكس من المعادلة (3-7) $h_{ii} > \frac{3(p)}{n} = \frac{3 * 4}{30} = 0.4$ فنجد أن هنالك حالة واحدة تزيد قيم رافعاتها عن ثلاثة أضعاف قيم الرافعات 0.4 وهي المشاهدة (21) حيث أن قيمة الرفع المناظرة لهذه الحالة بلغت 0.458989.

ثالثاً: اكتشاف القيم الشاذة في المتغير التابع:

يتم الكشف عن المشاهدات الشاذة في المتغير التابع من خلال بواقي ستيودنت المحذوفة بتطبيق المعادلة (3-12) وبالاعتماد على برنامج التحليل الإحصائي Eviews كانت النتائج بالجدول التالي:

الجدول (4-12): البواقي المحذوفة ، بواقي ستيودنت المحذوفة للنموذج المقدر قبل معالجة القيم الشاذة:

Obs.	Residual (e_i)	R Student (d_i^*)	Drop Residual (d)	$ d_i^* > t_{0.025, 24}$ $ d_i^* > 2.064$
1	0.015227	0.303956	0.017842	
2	0.004312	0.082298	0.004635	
3	0.014166	0.271717	0.015338	
4	-0.03452	-0.70289	-0.04142	
5	-0.01328	-0.24942	-0.01379	
6	0.016617	0.315936	0.017659	
7	0.05685	1.225185	0.073518	
8	-0.03532	-0.68663	-0.03866	
9	0.226698	28.43271	0.315726	28.43271
10	-0.03785	-0.72488	-0.04013	
11	-0.05586	-1.13433	-0.06466	
12	-0.02051	-0.40806	-0.0238	
13	-0.01004	-0.19019	-0.01062	
14	0.018432	0.372137	0.022054	
15	0.030268	0.634384	0.038621	
16	-0.0132	-0.24872	-0.01379	
17	-0.01704	-0.33758	-0.01965	
18	-0.01569	-0.30557	-0.0175	
19	-0.03374	-0.66856	-0.03841	
20	0.017653	0.358462	0.021373	
21	-0.04213	-1.07841	-0.07786	
22	-0.01362	-0.26035	-0.01465	
23	-0.02531	-0.48652	-0.02735	
24	-0.00153	-0.02959	-0.00168	
25	0.00915	0.179063	0.010329	
26	-0.03436	-0.78768	-0.05201	
27	-0.0005	-0.00937	-0.00052	
28	-0.01472	-0.28282	-0.01598	
29	0.044906	0.904224	0.052037	
30	-0.03506	-0.68723	-0.03902	

المصدر: إعداد الباحث باستخدام برنامج Eviews:

ولتحديد المشاهدة الشاذة في المتغير التابع يتم مقارنة قيمة باقي ستيودنت المحذوف المطلقة مع

قيمة توزيع t أي: $|d_i^*| > t_{\alpha_0, n-p-1}$ فإذا تجاوزت قيمة باقي ستيودنت المحذوف لأي مشاهدة قيمة

توزيع t تعد مشاهدة شاذة في المتغير التابع.

ويتضح من الجدول رقم (4-12) أعلاه أن بواقي ستيودنت المحذوفة التي تزيد قيمتها عن

قيمة توزيع $t_{0.025, 24} = 2.064$ هي حالة واحدة وهي الحالة رقم (9) والتي تعتبر حالة شاذة

تستدعي دراستها وقياس مدى تأثيرها علي مقدرات المربعات الصغرى. كذلك نلاحظ من نتائج

الجدولين (4-11) ، (4-12) أن الحالة رقم (9) حالة شاذة في المتغير التابع والمتغيرات

المستقلة.

8-4-4 قياس أثر المشاهدات الشاذة المؤثرة علي معاملات النموذج:

بعد تشخيص المشاهدات الشاذة في المتغيرات المستقلة و المتغير التابع سوف نقيس أثر

المشاهدات الشاذة المؤثرة علي قيم المتغير التابع المقدره ومعاملات النموذج والأخطاء المعيارية

و قياس الأثر علي كل معاملات الانحدار وبالاعتماد علي نتائج البرنامجين Eviews , SPSS

و بتطبيق المعادلات (3-15) ، (3-27) ، (3-24) تم الحصول علي النتائج كما بالجدول

أدناه:

الجدول (4-13): قيم $DFFITs$ ، قيم $COVRATIO$ ، قيم مسافة كوك D_i للنموذج المقدر قبل معالجة القيم الشاذة:

Obs	DFFITs	COVRATI	D_i	DF>0.730	DF>0.784	COV>1.4	$D_i > 0.154$
1	-0.12595	1.350658	0.00411				
2	-0.02255	1.256346	0.00013				
3	-0.07816	1.251806	0.00158				
4	0.314197	1.297929	0.02517				
5	0.049148	1.20326	0.00063				
6	-0.07909	1.223511	0.00162				
7	-0.6634	1.198122	0.10794				
8	0.211163	1.188297	0.01138				
9	-17.818	1.32E-06	2.47609	17.818	17.818	1.32E-06	2.47609
10	0.177924	1.141318	0.00806				
11	0.450322	1.107923	0.05014				
12	0.163479	1.322048	0.0069				
13	0.045824	1.230632	0.00055				
14	-0.16498	1.36918	0.00704				
15	-0.33326	1.400329	0.02842			1.400329	
16	0.052631	1.210221	0.00072				
17	0.132258	1.325099	0.00453				
18	0.103774	1.285465	0.00279				
19	0.248599	1.240489	0.01579				
20	-0.16455	1.387625	0.007				
21	0.993308	1.802765	0.24513	0.993308	0.993308	1.802765	0.24513
22	0.071526	1.244605	0.00133				
23	0.138027	1.217256	0.00491				
24	0.009258	1.284216	0.00002				
25	-0.06427	1.313834	0.00107				
26	0.564741	1.605768	0.08091			1.605768	
27	0.002006	1.223473	0				
28	0.08272	1.253816	0.00177				
29	-0.36033	1.191889	0.03269				
30	0.23099	1.20811	0.01362				

المصدر: إعداد الباحث باستخدام برنامج Eviews:

1- قياس التأثير علي القيم المقدرة - مقياس DFFITS:

لتشخيص الحالات المؤثرة علي قيم المتغير التابع المقدرة \hat{Y}_i هنالك عدة معايير ، فحسب معيار بيلسلي وآخرون 0.730 الذي تم حسابه اعتمادا علي المعادلة (3-16) يتضح من الجدول (4-13) أعلاه أن المشاهدات التي تجاوزت قيم *DFITs* المطلقة لها هذا المعيار هي المشاهدين رقم (9) ، (21) حيث بلغت قيم *DFITs* لها 17.818 ، 0.993308 علي التوالي.

أما إذا استخدمنا معيار شاترجي وهادي 0.784 والذي حُسب من المعادلة (3-19) فنجد أن المشاهدات التي تزيد قيم *DFITs* المطلقة لها عن هذه القيمة هي أيضاً المشاهدين رقم (9) ، (21) وتشير هذه النتيجة إلى نفس النتيجة التي حددها معيار بيلسلي وآخرون وهذا يعني أن المشاهدين (9,21) تؤثران معنوياً علي قيم المتغير التابع المقدرة.

2- قياس الأثر على الأخطاء المعيارية- مقياس COVRATIO:

نلاحظ من الجدول رقم (4-13) أعلاه أن هنالك أربعة حالات مؤثرة معنوياً علي قيم الأخطاء المعيارية وهي المشاهدات رقم (9) ، (15) ، (21) ، (26) حيث تزيد قيم *COVRATIO* المطلقة لها عن المعيار الذي أقترحه بيلسلي و آخرون 1.4 والذي تم حسابه من المعادلة (3-28) حيث بلغت قيم *COVRATIO* لهذه الحالات (1.32E-06)، (1.400329) ، (1.802765) ، (1.605768) علي التوالي.

3- قياس الأثر علي كل معاملات الانحدار - مقياس مسافة كوك:

يتبين من الجدول رقم (4-13) أعلاه أن هنالك حالتان مؤثرتان علي معاملات الانحدار وهي المشاهدات رقم (9) ، (21) حيث تزيد قيم مسافة كوك لها عن المعيار الذي أقترحه جون فوكس $D_i > 0.154$ والذي تم حسابه بتطبيق المعادلة (3-25).

4- قياس الأثر على معاملات الانحدار - مقياس $DFBETAS$:

الجدول (4-14): قيم معاملات الانحدار حسب المقياس $DFBETAS$ قبل معالجة القيم الشاذة:

Obs	Scaled Difference in Coefficients ($DFBETAS$)				$DFBETAS > 0.3651$			
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$
1	0.043283	0.056388	-0.0426	-0.07318				
2	0.00362	0.009185	-0.0161	0.004621				
3	-0.01277	0.055591	-0.03937	-0.00824				
4	-0.25544	0.12476	-0.03474	0.134125				
5	-0.00504	-0.01029	0.013461	-0.00607				
6	-0.03744	0.040961	0.005638	-0.00641				
7	-0.55288	0.509947	-0.07448	0.072495	0.553	0.5099		
8	-0.14115	0.075282	-0.03945	0.074529				
9	10.60349	-16.547	7.077402	5.017075	10.603	16.547	7.077	5.02
10	-0.08152	0.09175	-0.06696	0.006548				
11	-0.02281	0.291832	-0.37578	-0.02726			0.3758	
12	-0.07289	0.010542	-0.04713	0.109435				
13	-0.00245	-0.00024	-0.01793	0.015246				
14	-0.08676	0.012723	0.117407	-0.03012				
15	0.096403	0.187551	-0.1536	-0.1881				
16	-0.00724	-8.31E-05	-0.01167	0.010833				
17	-0.0355	0.000164	-0.05253	0.088559				
18	0.023859	0.034581	-0.08449	0.002313				
19	0.031619	0.032411	0.092532	-0.23078				
20	-0.0532	0.097324	-0.12635	0.08674				
21	-0.54915	0.38654	0.529031	-0.59128	0.549	0.387	0.529	0.59
22	-0.03416	0.009285	0.038049	-0.02903				
23	-0.01889	0.064607	0.01052	-0.1102				
24	0.002533	-0.00514	0.00638	-0.00437				
25	-0.04947	0.016368	0.002266	0.043192				
26	0.291739	0.232439	-0.38519	-0.30607			0.385	
27	0.000345	-0.00079	0.000852	-0.00069				
28	0.023703	-0.00801	0.032546	-0.06964				
29	-0.28284	0.224612	-0.09033	0.165525				
30	0.051239	0.128574	-0.12588	-0.14379				

المصدر: إعداد الباحث باستخدام برنامج Eviews :

وبتطبيق المعادلة (3-20) واعتماداً نتائج البرنامج Eviews تم الحصول علي قيم $DFBETAS$ كما بالجدول رقم (4-14) أعلاه ، ونلاحظ من ذات الجدول أن هنالك خمسة حالات مؤثرة علي قيم معاملات نموذج الانحدار الخطي المتعدد ، حيث أن قيم $DFBETAS$ المطلقة لها تزيد عن المعيار الذي أقترحه نيتير وآخرون $DFBETAS=0.3651$ والذي تم حسابه بتطبيق المعادلة (3-22) وهي الحالات رقم (7)، (9) ، (11) ، (21)، (26) حيث يلاحظ أن بعض هذه الحالات تؤثر معنوياً علي معامل واحد كما في الحالتين (26،11) وبعضها يؤثر علي معاملين كما في الحالة (7) والبعض الآخر يؤثر معنوياً علي جميع معاملات نموذج الانحدار كما ظهر في الحالتين (9، 21).

ونستنتج من نتائج الجدولين (4-13) و (4-14) أن الحالتان (9) ، (21) ظهرت كحالات مؤثرة معنوياً في كل المقاييس الأربعة ، وهذا يعني أن لهذه الحالات تأثير قوي جداً علي نتائج نموذج تحليل الانحدار الخطي المتعدد.

ويمكن تلخيص نتائج اكتشاف وتأثير المشاهدات الشاذة علي نموذج تحليل الانحدار

المقدر قبل معالجة المشاهدات الشاذة في الجدول التالي:

الجدول (4-15): يوضح ملخص المشاهدات الشاذة والمؤثرة علي النموذج المقدر قبل معالجة القيم الشاذة:

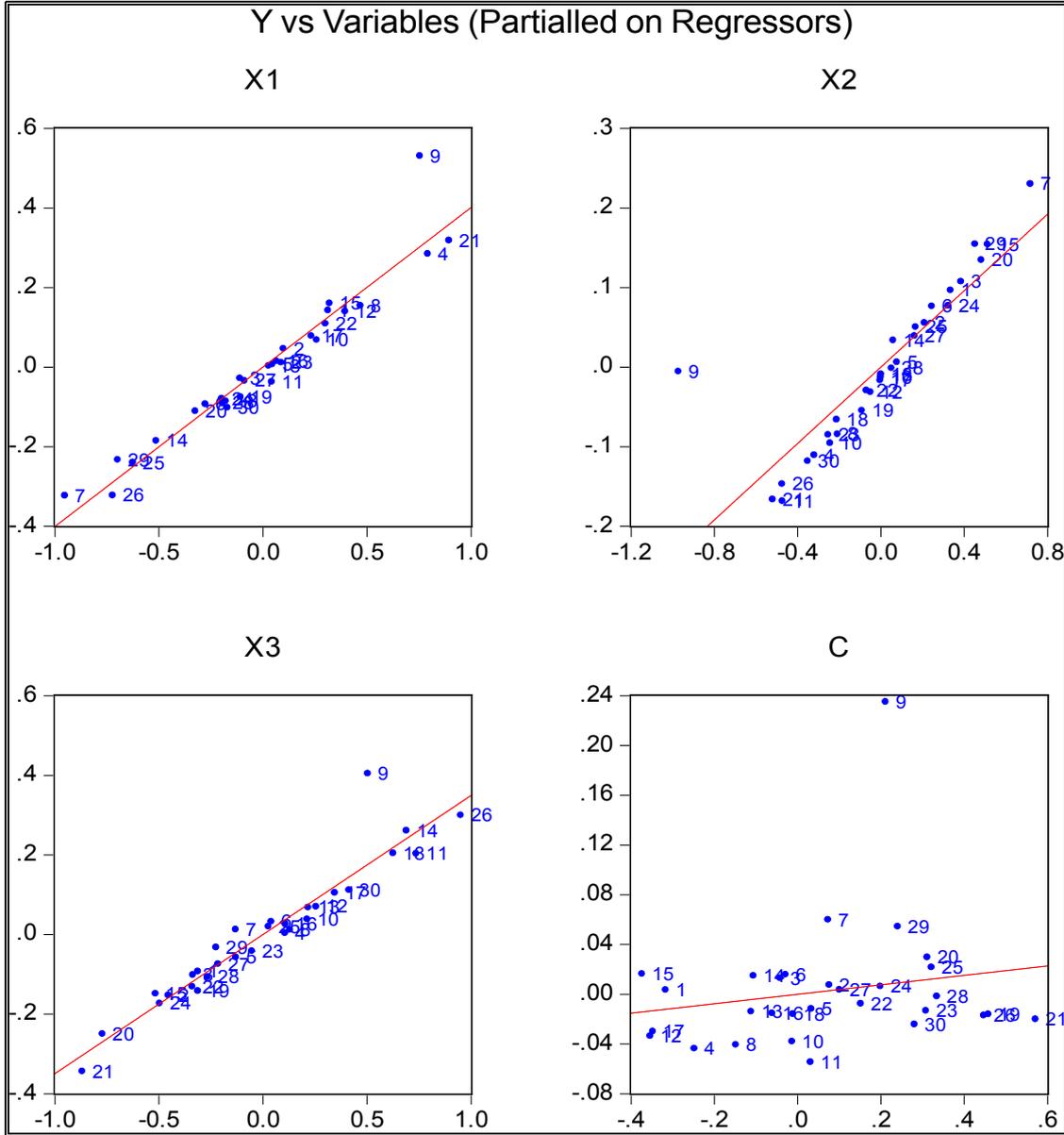
رقم	شاذة في المتغير المستقل X	شاذة في المتغير التابع Y	مؤثرة حسب المقياس $DFFITs$	مؤثرة حسب المقياس $COVRATIC$	مؤثرة حسب مقياس كوك D_i	مؤثرة حسب المقياس $DFBETAS$
7	غير شاذة	غير شاذة	غير مؤثرة	غير مؤثرة	غير مؤثرة	مؤثرة
9	شاذة	شاذة	مؤثرة	مؤثرة	مؤثرة	مؤثرة
11	غير شاذة	غير شاذة	غير مؤثرة	غير مؤثرة	غير مؤثرة	مؤثرة
15	غير شاذة	غير شاذة	غير مؤثرة	مؤثرة	غير مؤثرة	غير مؤثرة
21	شاذة	غير شاذة	مؤثرة	مؤثرة	مؤثرة	مؤثرة
26	شاذة	غير شاذة	غير مؤثرة	مؤثرة	غير مؤثرة	مؤثرة

المصدر: إعداد الباحث اعتماداً علي نتائج البرنامجين SPSS,Eviews:

كما يمكن ملاحظة شذوذ المشاهدات (7) ، (9) ، (15) ، (21) ، (26) وبعدها من خط

الانحدار المقدر من خلال الشكل (7-4) التالي:

الشكل (7-4): يوضح انتشار المشاهدات علي خط الانحدار المقدر قبل معالجة القيم الشاذة:



المصدر: إعداد الباحث اعتماداً علي نتائج البرنامج Eviews:

يتضح من الشكل (7-4) أعلاه أن المشادة رقم (9) هي أكثر المشاهدات بعداً عن خط الانحدار المقدر وهذا يعني أن هذه المشادة ذات تأثير قوي جداً علي معادلة الانحدار المقدر، كما نلاحظ من ذات الشكل أن غالبية المشاهدات تقع بصورة جيدة علي خط الانحدار المقدر.

5-4 معالجة القيم الشاذة:

بعد التأكد من وجود القيم الشاذة و المؤثرة علي معاملات النموذج المقدر لابد من تقليل أثر هذه القيم علي نتائج النموذج ، وسوف نعتمد في معالجة هذه القيم علي طريقة حذف القيم الشاذة المؤثرة بأسلوب علمي يمكن من خلاله الحصول علي نتائج أفضل مقارنةً بنتائج النموذج المقدر قبل المعالجة بالإضافة إلي تطبيق طريقة متوسط البتر وذلك فيما يلي:

1-5-4 معالجة القيم الشاذة بأسلوب الحذف:

1- نتائج النموذج المقدر بعد حذف المشاهدة رقم (7):

$$\hat{Y} = 0.035 + 0.413X_1 + 0.226X_2 + 0.351X_3$$
$$R^2 = 0.992 \quad \bar{R}^2 = 0.991 \quad MSE = 0.003 \quad F = 1084.155$$

$\lambda_0 = 3.895$	$CI_0 = 1.000$	$VIF_1 = 2.439$
$\lambda_1 = 0.054$	$CI_1 = 8.529$	$VIF_2 = 3.263$
$\lambda_2 = 0.035$	$CI_2 = 10.508$	$VIF_3 = 2.137$
$\lambda_3 = 0.017$	$CI_3 = 15.299$	

2- نتائج النموذج المقدر بعد حذف المشاهدة رقم (9):

$$\hat{Y} = 0.005 + 0.359X_1 + 0.226X_2 + 0.319X_3$$
$$R^2 = 1.000 \quad \bar{R}^2 = 1.000 \quad MSE = 8.85366039277e^{-005} \quad F = 34349.299$$

$\lambda_0 = 3.896$	$CI_0 = 1.000$	$VIF_1 = 2.382$
$\lambda_1 = 0.053$	$CI_1 = 8.604$	$VIF_2 = 3.497$
$\lambda_2 = 0.037$	$CI_2 = 10.261$	$VIF_3 = 2.261$
$\lambda_3 = 0.015$	$CI_3 = 16.223$	

3- نتائج النموذج المقدر بعد حذف المشاهدة رقم (11):

$$\hat{Y} = 0.039 + 0.401X_1 + 0.232X_2 + 0.358X_3$$

$$R^2 = 0.992 \quad \bar{R}^2 = 0.991 \quad MSE = 0.003 \quad F = 1073.880$$

$\lambda_0 = 3.892$	$CI_0 = 1.000$	$VIF_1 = 2.096$
$\lambda_1 = 0.054$	$CI_1 = 8.526$	$VIF_2 = 3.025$
$\lambda_2 = 0.036$	$CI_2 = 10.401$	$VIF_3 = 2.290$
$\lambda_3 = 0.018$	$CI_3 = 14.521$	

4- نتائج النموذج المقدر بعد حذف المشاهدة رقم (15):

$$\hat{Y} = 0.045 + 0.399X_1 + 0.235X_2 + 0.353X_3$$

$$R^2 = 0.991 \quad \bar{R}^2 = 0.990 \quad MSE = 0.003 \quad F = 939.847$$

$\lambda_0 = 3.894$	$CI_0 = 1.000$	$VIF_1 = 1.908$
$\lambda_1 = 0.052$	$CI_1 = 8.672$	$VIF_2 = 2.692$
$\lambda_2 = 0.035$	$CI_2 = 10.568$	$VIF_3 = 2.222$
$\lambda_3 = 0.020$	$CI_3 = 14.123$	

5- نتائج النموذج المقدر بعد حذف المشاهدة رقم (21):

$$\hat{Y} = 0.060 + 0.413X_1 + 0.230X_2 + 0.338X_3$$

$$R^2 = 0.991 \quad \bar{R}^2 = 0.990 \quad MSE = 0.003 \quad F = 906.988$$

$\lambda_0 = 3.911$	$CI_0 = 1.000$	$VIF_1 = 2.618$
$\lambda_1 = 0.046$	$CI_1 = 9.269$	$VIF_2 = 2.580$
$\lambda_2 = 0.027$	$CI_2 = 12.133$	$VIF_3 = 1.881$
$\lambda_3 = 0.017$	$CI_3 = 15.239$	

6- نتائج النموذج المقدر بعد حذف المشاهدة رقم (26):

$$\hat{Y} = 0.050 + 0.394 X_1 + 0.234 X_2 + 0.358 X_3$$

$$R^2 = 0.992 \quad \bar{R}^2 = 0.991 \quad MSE = 0.003 \quad F = 977.754$$

$\lambda_0 = 3.903$	$CI_0 = 1.000$	$VIF_1 = 2.020$
$\lambda_1 = 0.052$	$CI_1 = 8.636$	$VIF_2 = 2.811$
$\lambda_2 = 0.027$	$CI_2 = 12.568$	$VIF_3 = 2.637$
$\lambda_3 = 0.018$	$CI_3 = 14.709$	

7- نتائج النموذج المقدر بعد حذف المشاهدين رقم (9,21) معاً:

$$\hat{Y} = -0.008 + 0.352 X_1 + 0.327 X_2 + 0.327 X_3$$

$$R^2 = 1.000, \bar{R}^2 = 1.000, MSE = 6.0144086690 \cdot 66e^{-005}, F = 42913.747$$

$\lambda_0 = 3.914$	$CI_0 = 1.000$	$VIF_1 = 3.265$
$\lambda_1 = 0.047$	$CI_1 = 9.141$	$VIF_2 = 3.64$
$\lambda_2 = 0.027$	$CI_2 = 11.971$	$VIF_3 = 1.892$
$\lambda_3 = 0.011$	$CI_3 = 18.452$	

8- نتائج النموذج المقدر بعد حذف جميع المشاهدات الشاذة:

$$\hat{Y} = -0.004 + 0.351 X_1 + 0.323 X_2 + 0.331 X_3$$

$$R^2 = 1.000 \quad \bar{R}^2 = 1.000 \quad MSE = 6.0727924581 \cdot 2e^{-005} \quad F = 34681.018$$

$\lambda_0 = 3.929$	$CI_0 = 1.000$	$VIF_1 = 3.513$
$\lambda_1 = 0.045$	$CI_1 = 9.361$	$VIF_2 = 4.189$
$\lambda_2 = 0.018$	$CI_2 = 14.568$	$VIF_3 = 3.010$
$\lambda_3 = 0.008$	$CI_3 = 22.455$	

9- المفاضلة بين نتائج النماذج المقدره بعد حذف المشاهدات الشاذة:

الجدول (4-16): المفاضلة بين النماذج المقدره بعد حذف المشاهدات الشاذة:

Cases deleted	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	F	R^2	MSE
7	0.035	0.413	0.226	0.351	1084.155	0.992	0.003
9	0.005	0.359	0.319	0.322	34349.29	1.000	$8.85366039277e^{-005}$
11	0.039	0.401	0.232	0.358	1073.880	0.992	0.003
15	0.045	0.399	0.235	0.353	939.847	0.991	0.003
21	0.060	0.413	0.230	0.338	9006.988	0.990	0.003
26	0.050	0.394	0.234	0.358	977.754	0.992	0.003
9+21	-0.008	0.352	0.327	0.327	42913.75	1.000	$6.014408669066e^{-005}$
All Outliers	-0.004	0.351	0.323	0.331	34681.02	1.000	$6.07279245812e^{-005}$

المصدر: من إعداد الباحث اعتماداً علي نتائج البرنامج SPSS:

يتضح من الجدول (4-16) أعلاه أن هنالك تأثيراً جوهرياً من قبل القيم الشاذة علي

معلمات النموذج المقدر حيث تقل قيمة MSE عند حذف القيم الشاذة الأكثر تأثيراً وتزداد قيمة

كل من F , R^2 عند حذف هذه القيم ، وأيضاً هنالك تغيير في تركيبة النموذج.

ولاختيار النموذج الأفضل من بين النماذج المقدره بعد عملية حذف المشاهدات الشاذة نلاحظ من

الجدول (4-16) أعلاه أن حذف المشاهدين رقم (9,21) أدبي إلي تحسن كبير جداً في نتائج

النموذج وذلك بخفض قيمة MSE والذي بلغت قيمته قبل عملية الحذف $MSE=0.003$ إلى $MSE = 6.014408669066e^{-005}$ بعد عملية الحذف ، وأيضاً أدى إلى ورفع قيمة F المحسوبة من $F = 1063.322$ قبل عملية الحذف إلى $F = 42913.747$ بعد عملية الحذف وكذلك ورفع قيمة R^2 من $R^2 = 0.992$ قبل عملية الحذف إلى $R^2 = 1.000$ بعد عملية الحذف. كما أن حذف المشاهدة رقم (9) أدى أيضاً إلى تحسن كبير في نتائج النموذج وذلك بخفض قيمة MSE والذي بلغت قيمته قبل عملية الحذف $MSE = 0.003$ إلى $MSE = 8.85366039277e^{-005}$ بعد عملية الحذف ورفع قيمة F المحسوبة من $F = 1063.322$ قبل عملية الحذف إلى $F = 34349.299$ بعد عملية الحذف كما أدى أيضاً إلى رفع قيمة R^2 من $R^2 = 0.992$ قبل عملية الحذف إلى $R^2 = 1.000$ بعد عملية الحذف ، عليه لا بد من المفاضلة بين النماذج المقدره للوصول إلى قرار بشأن حذف المشاهدين (9,21) وذلك لان المشاهدة (9) كما أثبتنا سابقاً أنها شاذة في المتغير التابع والمتغيرات المستقلة وتؤثر معنوياً علي جميع معاملات النموذج وذلك حسب مقاييس قياس أثر المشاهدات الشاذة الأربعة ، كما أن المشاهدة (21) شاذة في المتغيرات المستقلة وغير شاذة في المتغير التابع وأيضاً مؤثرة علي جميع معاملات النموذج. عليه نقرر أن أفضل النماذج المقدره بعد معالجة القيم الشاذة بطريقة الحذف هو النموذج:

$$\hat{Y} = -0.008 + 0.352 X_1 + 0.327 X_2 + 0.327 X_3$$

$$R^2 = 1.000 , \bar{R}^2 = 1.000 , MSE = 6.014408669066e^{-005} , F = 42913.747$$

$$DW = 2.132$$

$\lambda_0 = 3.914$	$CI_0 = 1.000$	$VIF_1 = 3.265$
$\lambda_1 = 0.047$	$CI_1 = 9.141$	$VIF_2 = 3.64$
$\lambda_2 = 0.027$	$CI_2 = 11.971$	$VIF_3 = 1.892$
$\lambda_3 = 0.011$	$CI_3 = 18.452$	

2-5-4 معالجة القيم الشاذة بطريقة متوسط البتر:

بتطبيق هذه الطريقة سنقوم بمعالجة جميع المشاهدات الشاذة التي ظهرت في النموذج المقدر قبل معالجة القيم الشاذة ، وهي الحالات (9) ، (21) ، (26) كما تبين لنا من نتائج الجدول (5-15) وبعد معالجة هذه الحالات بطريقة متوسط البتر تم الحصول علي النماذج المقدره كما يلي:

1- نتائج النموذج المقدر بعد معالجة المشاهدات رقم (9) بطريقة متوسط البتر:

$$\hat{Y} = 0.005 + 0.359X_1 + 0.319X_2 + 0.322X_3 , \quad DW = 2.102$$
$$R^2 = 1.00 , \quad \bar{R}^2 = 1.000 , \quad MSE = 8.5302017586e^{-005} , \quad F = 35619.177$$

$\lambda_0 = 3.899$	$CI_0 = 1.000$	$VIF_1 = 2.382$
$\lambda_1 = 0.051$	$CI_1 = 8.748$	$VIF_2 = 3.495$
$\lambda_2 = 0.036$	$CI_2 = 10.429$	$VIF_3 = 2.261$
$\lambda_3 = 0.014$	$CI_3 = 16.580$	

2- نتائج النموذج المقدر بعد معالجة المشاهدات رقم (21) بطريقة متوسط البتر:

$$\hat{Y} = 0.059 + 0.413X_1 + 0.230X_2 + 0.328X_3 , \quad DW = 2.432$$
$$R^2 = 0.991 , \quad \bar{R}^2 = 0.990 , \quad MSE = 0.003 , \quad F = 940.132$$

$\lambda_0 = 3.914$	$CI_0 = 1.000$	$VIF_1 = 2.618$
$\lambda_1 = 0.044$	$CI_1 = 9.420$	$VIF_2 = 2.579$
$\lambda_2 = 0.026$	$CI_2 = 12.324$	$VIF_3 = 1.881$
$\lambda_3 = 0.016$	$CI_3 = 15.477$	

3- نتائج النموذج المقدر بعد معالجة المشاهدات رقم (26) بطريقة متوسط البتر:

$$\hat{Y} = 0.049 + 0.394X_1 + 0.234X_2 + 0.358X_3 , \quad DW = 2.394$$

$$R^2 = 0.992 , \quad \bar{R}^2 = 0.991 , \quad MSE = 0.003 , \quad F = 1013.626$$

$\lambda_0 = 3.906$	$CI_0 = 1.000$	$VIF_1 = 2.020$
$\lambda_1 = 0.051$	$CI_1 = 8.780$	$VIF_2 = 2.811$
$\lambda_2 = 0.026$	$CI_2 = 12.217$	$VIF_3 = 2.636$
$\lambda_3 = 0.017$	$CI_3 = 14.941$	

4- نتائج النموذج المقدر بعد معالجة المشاهدتين رقم (9+21) بطريقة متوسط البتر:

$$\hat{Y} = -0.008 + 0.351X_1 + 0.327X_2 + 0.327X_3 , \quad DW = 2.137$$

$$R^2 = 1.000 , \quad \bar{R}^2 = 1.000 , \quad MSE = 5.755012371742e^{-005} , \quad F = 44861.765$$

$\lambda_0 = 3.920$	$CI_0 = 1.000$	$VIF_1 = 3.259$
$\lambda_1 = 0.044$	$CI_1 = 9.446$	$VIF_2 = 3.359$
$\lambda_2 = 0.026$	$CI_2 = 12.357$	$VIF_3 = 1.892$
$\lambda_3 = 0.011$	$CI_3 = 19.018$	

5- نتائج النموذج المقدر بعد معالجة المشاهدات رقم (9+21+26) بطريقة متوسط البتر:

$$\hat{Y} = -0.008 + 0.352X_1 + 0.328X_2 + 0.326X_3 , \quad DW = 2.123$$

$$R^2 = 1.000 , \quad \bar{R}^2 = 1.000 , \quad MSE = 6.40283399471e^{-005} , \quad F = 36844.129$$

$\lambda_0 = 3.932$	$CI_0 = 1.000$	$VIF_1 = 2.966$
$\lambda_1 = 0.039$	$CI_1 = 10.092$	$VIF_2 = 3.081$
$\lambda_2 = 0.019$	$CI_2 = 14.226$	$VIF_3 = 2.365$
$\lambda_3 = 0.010$	$CI_3 = 19.536$	

6- المفاضلة بين نتائج النماذج المقدره بعد معالجة المشاهدات الشاذة بطريقة متوسط البتر :

الجدول (4-17): المفاضلة بين النماذج المقدره بعد معالجة المشاهدات الشاذة بطريقة متوسط

البتر :

المشاهدة المعالجة	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	F	R^2	MSE
9	0.005	0.359	0.319	0.322	35619.177	1.000	8.5302017586e-005
21	0.059	0.413	0.230	0.338	940.132	0.991	0.003
26	0.049	0.394	0.234	0.358	1013.626	0.992	0.003
21+9	-0.008	0.351	0.327	0.327	44861.765	1.000	5.755012371742e-005
26+21+9	-0.008	0.352	0.328	0.326	36844.129	1.000	6.402833994713e-005

المصدر: من إعداد الباحث اعتماداً علي نتائج البرنامج SPSS:

نلاحظ من نتائج الجدول (4-17) هنالك تقارب في نتائج بعض النماذج المقدره بعد معالجة القيم الشاذة بطريقة الحذف وطريقة متوسط البتر كما واضح ذلك عند معالجة المشاهدات رقم (9) والمشاهدة رقم (21) و المشاهدتين (9،21) مع تحسن في أداء النموذج المقدر عند معالجة القيم الشاذة ، ولاختيار النموذج الأفضل من بين النماذج المقدره بعد معالجة القيم الشاذة بطريقة متوسط البتر نلاحظ من ذات الجدول أن النموذج الأفضل هو النموذج المقدر بعد معالجة المشاهدتين الشاذتين (9،21) حيث بلغت قيمة F المحسوبة لهذا النموذج 44861.765 كما بلغت قيمة R^2 الواحد الصحيح وأيضاً بلغت قيمة MSE $6.402833994713e^{-005}$ و هي قيمة قليلة جداً تقترب من الصفر. ونستج من ذلك أن المشاهدتين (9،21) شاذتان وتؤثران معنوياً علي معاملات نموذج تحليل الانحدار الخطى المتعدد وهذا يؤكد لنا ما أثبتناه سابقاً.

6-4 المقارنة بين نتائج النماذج المقدرة لاختيار أفضل نموذج:

الجدول (4-18): المفاضلة بين النماذج المقدرة لاختيار أفضل نموذج:

النموذج المقدر	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	F	R^2	MSE
قبل معالجة القيم الشاذة	0.038	0.401	0.240	0.349	1063.322	0.992	0.003
بعد معالجة القيم الشاذة بطريقة الحذف	-0.008	0.352	0.327	0.327	42913.75	1.000	6.014408669066e-005
بعد معالجة القيم الشاذة بطريقة متوسط البيتر	-0.008	0.351	0.327	0.327	44861.765	1.000	5.755012371742e-005

المصدر: من إعداد الباحث اعتماداً على نتائج البرنامج SPSS

يتضح من الجدول (4-18) أعلاه أن أفضل نموذج يمكن استخدامه للتنبؤ بقيم المتغير التابع هو

النموذج الذي تم الحصول عليه بعد معالجة القيم الشاذة بطريقة متوسط البيتر وهو كالاتي:

$$\hat{Y} = -0.008 + 0.351X_1 + 0.327X_2 + 0.327X_3, \quad DW = 2.14$$

$$R^2 = 1.000, \quad \bar{R}^2 = 1.000, \quad MSE = 5.755012371742e^{-005}, \quad F = 44861.765$$

$$\lambda_0 = 3.920 \quad CI_0 = 1.000$$

$$\lambda_1 = 0.044 \quad CI_1 = 9.446 \quad VIF_1 = 3.259$$

$$\lambda_2 = 0.026 \quad CI_2 = 12.357 \quad VIF_2 = 3.359$$

$$\lambda_3 = 0.011 \quad CI_3 = 19.018 \quad VIF_3 = 1.892$$

7-4 تشخيص أفضل النماذج المقدرية:

بعد أن تم اختيار أفضل نموذج انحدار يمكن استخدامه للتنبؤ بقيم المتغير التابع لابد من القيام باختبار هذا النموذج من الناحية الإحصائية والقياسية وتشخيص المشاهدات الشاذة واختبار مقدرية هذا النموذج للتنبؤ وذلك فيما يلي:

1- اختبار معنوية النموذج من الناحية الإحصائية:

الجدول (4-19): يوضح نتائج اختبار t لأفضل نموذج مقدر:

معلمات النموذج Parameters	قيم المعاملات Coefficient	قيم اختبار t المحسوبة t-Statistic	مستوى المعنوية P-value
$\hat{\beta}_1$	0.351	88.532	0.000
$\hat{\beta}_2$	0.327	65.915	0.000
$\hat{\beta}_3$	0.327	92.371	0.000
$\hat{\beta}_0$	- 0.008	-1.273	0.214

المصدر: إعداد الباحث باستخدام برنامج SPSS:

يتضح من الجدول (5-19) أعلاه أن جميع معلمات النموذج $\beta_1, \beta_2, \beta_3$ معنوية وذلك بالاعتماد على مستوى المعنوية لها والبالغ 0.000 وهي أقل من مستوى المعنوية 0.05 المعتمد في هذه الدراسة ، كما نلاحظ عدم معنوية معلمة الحد الثابت $\hat{\beta}_0$ حيث بلغ مستوى المعنوية لها 0.214.

2- اختبار تجانس تباين حد الخطأ العشوائي:

بتطبيق خطوات اختبار جولدفيلد-كواندت تحصلنا حساب قيمة F المحسوبة والتي بلغت قيمتها 0.48 ومن الجداول الخاصة بتوزيع F تم الحصول علي قيمة $F_{12,12,0.05}$ الجدولية والتي تساوي 2.96 ، وبما أن قيمة F المحسوبة أقل من قيمة $F_{12,12,0.05}$ الجدولية ، عليه نقبل فرض العدم H_0 القائل بان هنالك تجانس أو ثبات في تباين الأخطاء.

3- اختبار الارتباط الخطي المتعدد بين المتغيرات المستقلة:

الجدول (4- 20): يوضح نتائج اختبارات الكشف عن الارتباط الخطي المتعدد بين المتغيرات المستقلة لأفضل نموذج مقدر:

المتغيرات Variables	قيم عامل تضخم التباين VIF	قيم التسامح Tolerance	قيم الجزور الكامنة Eigen value	قيم مؤشر الحالة CI	قيمة دليل الحالة CN
X ₁	3.259	0.307	3.920	9.446	18.878
X ₂	3.359	0.298	0.044	12.357	
X ₃	1.892	0.529	0.026	19.018	
Constant	-	-	0.011	1.000	

المصدر: إعداد الباحث باستخدام برنامج SPSS:

نلاحظ من الجدول (4-20) أعلاه أن قيم VIF لجميع المتغيرات المستقلة أقل من 10 وبالتالي لا يوجد ارتباط خطي متعدد بين المتغيرات المستقلة، كما أن قيم Tolerance لجميع المتغيرات المستقلة أكبر من 0.10 وهذا دليل أيضاً علي عدم وجود الارتباط الخطي المتعدد بين المتغيرات المستقلة، كما بلغت قيمة رقم الحالة CN=18.878 وهذه القيمة أقل من المعيار الذي حدده جونستون: 1984م ، عليه فان النموذج المقدر لا يعاني من مشكلة الارتباط الخطي المتعدد بين المتغيرات المستقلة.

4- اختبار الاستقلال الذاتي للبواقي:

الشكل (4-8): يبين مناطق القبول والرفض لاختبار درين- واتسون لأفضل نموذج مقدر:

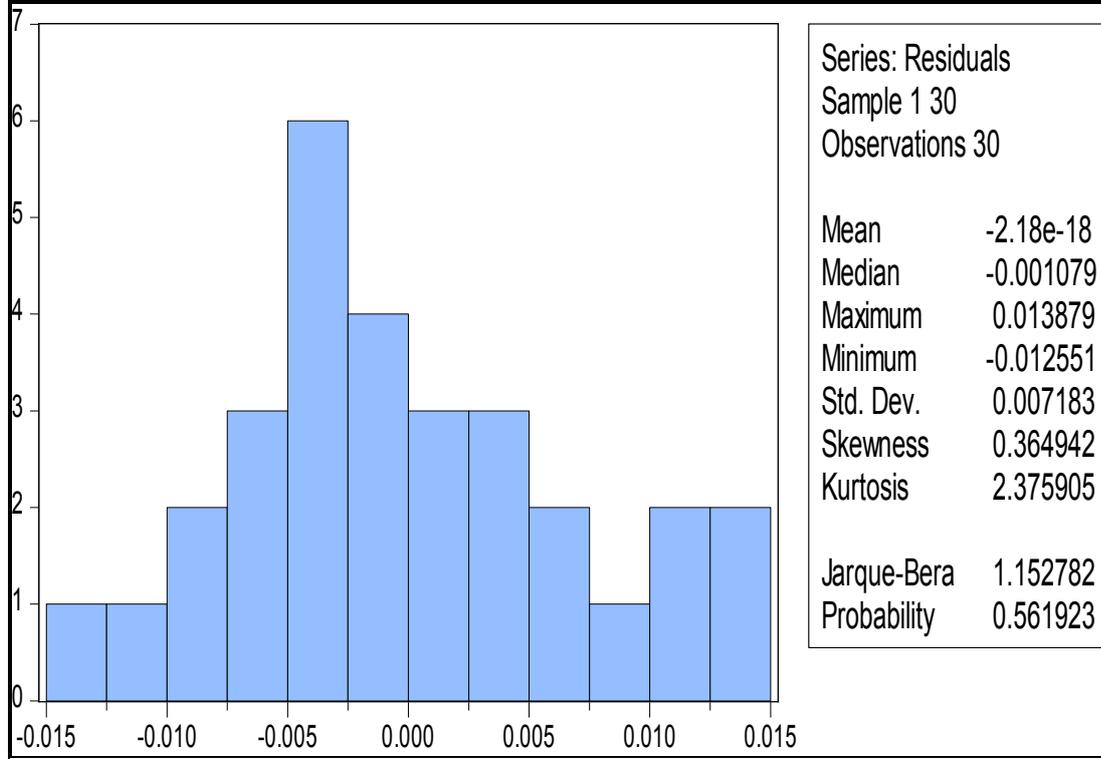
0	dl	du	2	DW	4-du	4-dl	4
$\rho > 0$		$\rho = 0$		$\rho = 0$		$\rho < 0$	
ارتباط ذاتي موجب	منطقة غير محددة	عدم وجود ارتباط ذاتي	عدم وجود ارتباط ذاتي	عدم وجود ارتباط ذاتي	منطقة غير محددة	ارتباط ذاتي سالب	
	1.21	1.65	2.14	2.35	2.79		

المصدر: إعداد الباحث:

نلاحظ من الشكل (4-8) أعلاه أن قيمة D-W المحسوبة بلغت 2.14 وهي تقع ضمن منطقة قرار عدم وجود الارتباط الذاتي للبواقي ، عليه فان النموذج المقدر يستوفي فرضية عدم وجود الارتباط الذاتي للبواقي ، وبهذه النتيجة نستنتج أن القيم الشاذة تؤثر علي فرضية الارتباط الذاتي للبواقي ويتضح ذلك من النتيجة التي تحصلنا عليها من النموذج المقدر قبل معالجة القيم الشاذة حيث أشارت نتائج اختبار D-W إلي عدم اتخاذ قرار بشأن وجود أو عدم وجود الارتباط الذاتي.

5- اختبار التوزيع الطبيعي للبواقي:

الجدول (4-21): يوضح اختبار جاك بيرا للتوزيع الطبيعي لأفضل نموذج مقدر:



المصدر: إعداد الباحث باستخدام برنامج Eviews :

يتضح من الجدول (4-21) أعلاه أن قيمة إحصائية اختبار Jarque-Bera بلغت

1.152782 وأن مستوي المعنوية لها بلغ 0.561923 وهي قيمة أكبر بكثير من مستوي المعنوية

0.05 المعتمد لهذه الدراسة ، عليه نقبل فرض العدم H_0 القائل بأن البواقي تتوزع طبيعياً ، وهي

عكس النتيجة التي حصلنا عليها قبل معالجة القيم الشاذة. وهذا يؤكد لنا أن القيم الشاذة تؤثر

معنوياً علي فرضية التوزيع الطبيعي لبواقي النموذج.

6- اختبار النموذج لعملية التنبؤ:

أولاً: قياس مدى استقرار تقديرات النموذج خلال فترة الدراسة:

الجدول (4-22): يوضح اختبار دالة الارتباط الذاتي والجزئي لأفضل نموذج مقدر:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.074	-0.074	0.1801	0.671
		2	-0.122	-0.128	0.6914	0.708
		3	-0.114	-0.137	1.1563	0.764
		4	0.070	0.033	1.3378	0.855
		5	-0.160	-0.191	2.3223	0.803
		6	-0.098	-0.142	2.7029	0.845
		7	0.146	0.095	3.5986	0.825
		8	0.184	0.140	5.0805	0.749
		9	0.015	0.071	5.0906	0.826
		10	-0.143	-0.087	6.0721	0.809
		11	-0.056	-0.082	6.2279	0.858
		12	-0.084	-0.112	6.6068	0.882
		13	-0.038	-0.035	6.6878	0.918
		14	0.116	0.129	7.5018	0.914
		15	-0.061	-0.147	7.7405	0.934
		16	0.056	-0.028	7.9575	0.950

المصدر: إعداد الباحث باستخدام برنامج Eviews :

نلاحظ من الجدول (4-22) أعلاه أن رسم شكل دالة الارتباط الذاتي والجزئي لبواقي

النموذج المقدر تقع أعمدها داخل مجال الثقة ، عليه نقول أن النموذج مستقر. وللتأكد من ذلك

يمكن تطبيق اختبار Ljung-Box المعروف بالمعادلة (2-31) حيث يستخدم هذا الاختبار

لاختبار المعنوية الكلية لمعاملات دالة الارتباط الذاتي وتوافق إحصائية الاختبار المحسوبة Q

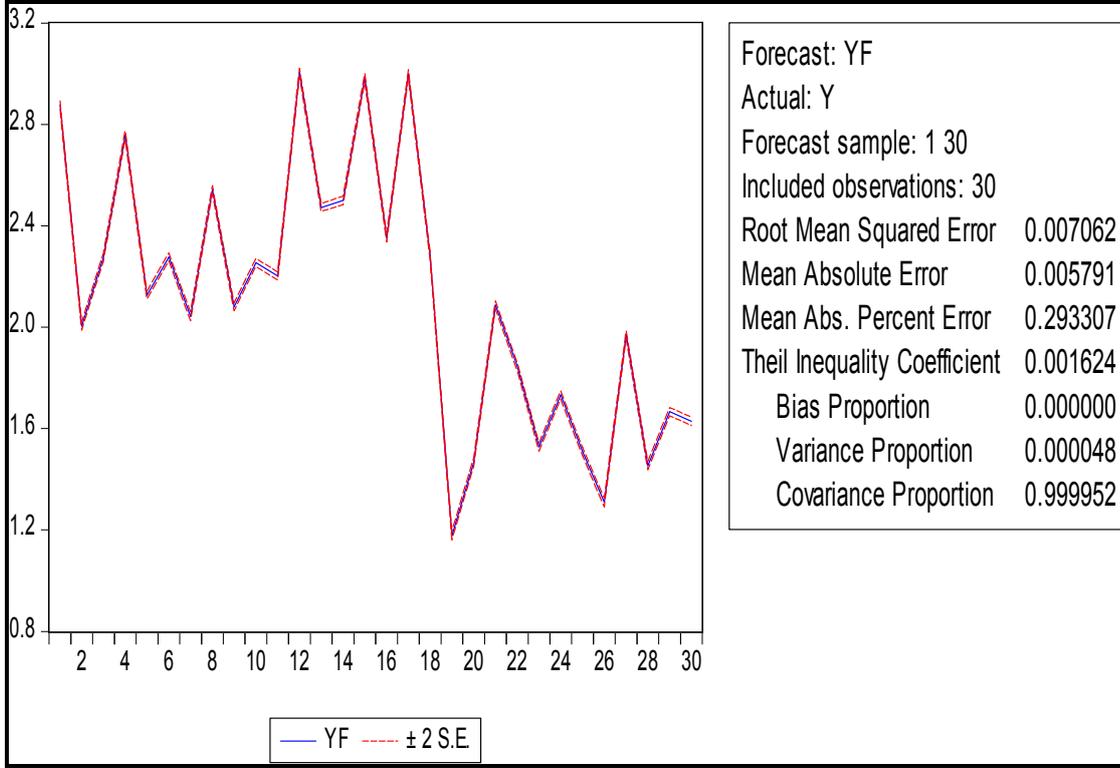
آخر قيمة في العمود Q-stat من الجدول أعلاه أي: $Q=7.9575$ وبحساب قيمة $\chi^2_{(16,0.05)}$ نجد

أنها تساوي $\chi^2_{(16,0.05)} = 26.296$ وبما أن القيمة $5.4564 < 26.96$ ، عليه نقبل فرض العدم

ونرفض الفرض البديل وهذا يعني أن النموذج مستقر.

ثانياً: اختبار القدرة التنبؤية للنموذج :

الجدول (4-23): يوضح نتائج اختبارات تقييم القدرة التنبؤية لأفضل نموذج مقدر:



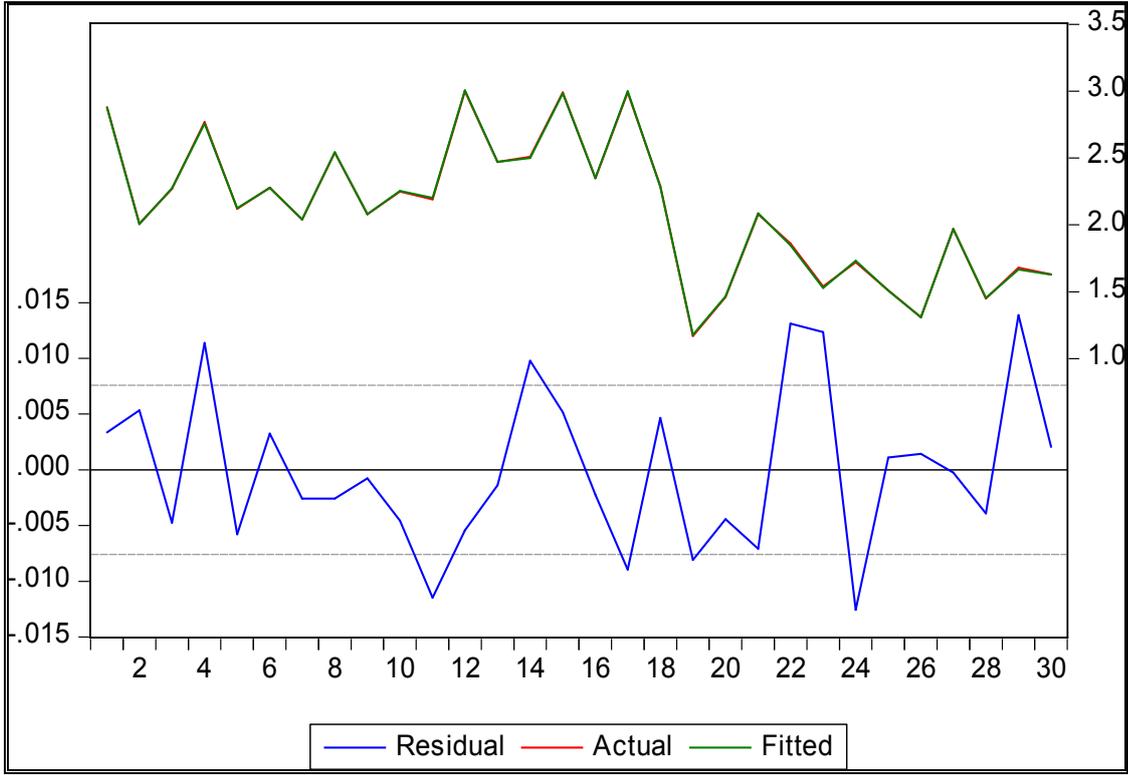
المصدر: إعداد الباحث باستخدام برنامج Eviews :

يتضح من الجدول (4-23) أعلاه أن قيمة معامل تايل بلغت 0.0021624 وهي قيمة قريبة جداً من الصفر، كما أن قيمة متوسط الخطأ المطلق MAE بلغت 0.005791 وهي أيضاً قيمة قريبة من الصفر مما يعني أن كفاءة النموذج عالية جداً في التنبؤ بقيم المتغير التابع، كما أن قيمة الجذر التربيعي لمتوسط مربع الخطأ RMSE بلغت 0.007062 وكلما بلغت قيمة هذا المؤشر 0.05 فأقل دل ذلك علي أن النموذج يطابق البيانات.

ويمكن التحقق أيضاً من مقدرة النموذج للتنبؤ من خلال مقارنة منحنى القيم الفعلية مع

منحنى القيم المقدرة والذي يوضحه الشكل (4-9) التالي:

الشكل (4-9): مقارنة منحنى القيم الفعلية مع منحنى القيم المقدرة لأفضل نموذج مقدر:



المصدر: إعداد الباحث باستخدام برنامج Eviews :

نلاحظ من الشكل (4-9) أعلاه أن هنالك تطابق تام بين منحنى القيم الفعلية ومنحنى

القيم التقديرية وإن دل هذا إنما يدل على أن معالجة القيم الشاذة بطريقة متوسط البتر أدت إلى

تحسين كبير جداً في أداء النموذج في عملية التنبؤ بقيم المتغير التابع.

7- تشخيص المشاهدات الشاذة والمؤثرة:

الجدول (4-24): يوضح ملخص المشاهدات الشاذة والمؤثرة لأفضل نموذج مقدر:

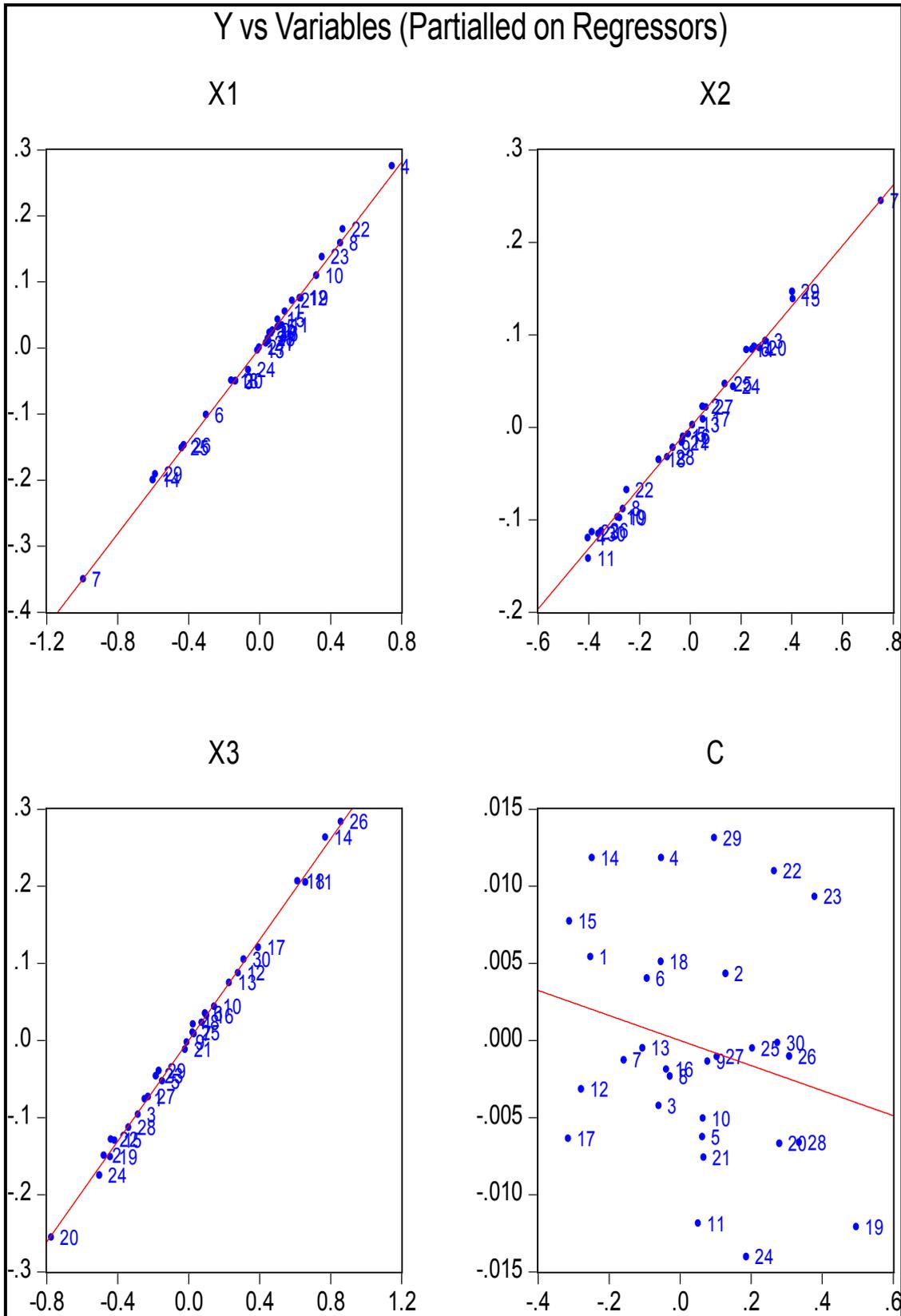
مؤثرة حسب المقياس $DFBETAS$	مؤثرة حسب مقياس كوك D_i	مؤثرة حسب المقياس $COVRATIO$	مؤثرة حسب المقياس $DFFITS$	شاذة في المتغير التابع Y	شاذة في المتغير المستقل X	رقم
مؤثرة	مؤثرة	غير مؤثرة	مؤثرة	غير شاذة	غير شاذة	4
غير مؤثرة	غير مؤثرة	مؤثرة	غير مؤثرة	غير شاذة	شاذة	7
مؤثرة	غير مؤثرة	غير مؤثرة	مؤثرة	غير شاذة	غير شاذة	11
مؤثرة	مؤثرة	غير مؤثرة	مؤثرة	غير شاذة	غير شاذة	14
مؤثرة	غير مؤثرة	غير مؤثرة	غير مؤثرة	غير شاذة	غير شاذة	17
مؤثرة	غير مؤثرة	غير مؤثرة	غير مؤثرة	غير شاذة	غير شاذة	19
مؤثرة	غير مؤثرة	غير مؤثرة	مؤثرة	غير شاذة	غير شاذة	22
مؤثرة	غير مؤثرة	غير مؤثرة	مؤثرة	غير شاذة	غير شاذة	23
مؤثرة	غير مؤثرة	غير مؤثرة	غير مؤثرة	غير شاذة	غير شاذة	24
مؤثرة	غير مؤثرة	مؤثرة	غير مؤثرة	غير شاذة	شاذة	26
مؤثرة	مؤثرة	غير مؤثرة	مؤثرة	شاذة	غير شاذة	29

المصدر: إعداد الباحث اعتماداً على نتائج البرنامجين SPSS, Eviews:

نلاحظ من الجدول (4-24) أعلاه أن معالجة القيم الشاذة أدت إلى وجود قيم شاذة أخرى وهي المشاهدتين (7،26) شاذتين في المتغيرات المستقلة وكذلك المشاهدة (29) شاذة في المتغير التابع مع وجود بعض القيم المؤثرة علي معاملات النموذج ، ولكن لا توجد قيمة واحدة شاذة في المتغيرات المستقلة والمتغير التابع معا وأيضاً تؤثر علي جميع معاملات النموذج⁸⁰ ، عليه فإن هذه القيم لا تؤثر كثيراً علي نتائج النموذج المقدر و يتضح لنا ذلك من الشكل (4-10) التالي حيث أن جميع القيم تقع علي خط الانحدار المقدر. كما نلاحظ أن المشاهدتين (7،4) تبعدان قليلاً عن بقية القيم الأخرى.

⁸⁰ - أنظر الملحق رقم (12)

الشكل (4-10): انتشار المشاهدات علي خط الانحدار المقدر لأفضل نموذج:



المصدر: إعداد الباحث باستخدام برنامج Eviews :

8- القيام بعملية التنبؤ:

الجدول (4-25): يوضح القيم المقدرة و الحقيقية لقيم المتغير التابع خلال فترة الدراسة:

Obs	القيم الفعلية Y_i	القيم المقدرة y^{\wedge}	قيم البواقي e_i
1	2.88	2.876657	0.003343
2	2.01	2.004661	0.005339
3	2.27	2.274757	-0.00476
4	2.77	2.758629	0.011371
5	2.12	2.125784	-0.00578
6	2.28	2.276758	0.003242
7	2.04	2.042595	-0.0026
8	2.54	2.542582	-0.00258
9	2.08	2.080766	-0.00077
10	2.25	2.254553	-0.00455
11	2.19	2.201481	-0.01148
12	3	3.005437	-0.00544
13	2.47	2.471392	-0.00139
14	2.51	2.500215	0.009785
15	2.99	2.984814	0.005186
16	2.35	2.352232	-0.00223
17	2.99	2.998963	-0.00896
18	2.29	2.285349	0.004651
19	1.17	1.178092	-0.00809
20	1.46	1.464435	-0.00444
21	2.08	2.087072	-0.00707
22	1.86	1.846888	0.013112
23	1.54	1.527632	0.012368
24	1.72	1.732551	-0.01255
25	1.51	1.508883	0.001117
26	1.31	1.308566	0.001434
27	1.97	1.970254	-0.00025
28	1.45	1.453924	-0.00392
29	1.68	1.666121	0.013879
30	1.63	1.627959	0.002041

المصدر: إعداد الباحث اعتماداً على نتائج البرنامج (Eviews):

نلاحظ من الجدول (4-26) أعلاه أن القيم المقدرة باستعمال النموذج الذي تم تقديره كأفضل نموذج خطي متعدد ليمثل بيانات الدراسة قريبة جداً من القيم الفعلية وهذا يتفق مع نتائج اختبارات القدرة التنبؤية التي أجريناها علي هذا النموذج.

الفصل الخامس
النتائج والتوصيات

1-5 النتائج:

2-5 التوصيات:

5-1 النتائج:

بعد معالجة بيانات الدراسة بالطرق والأساليب الإحصائية المناسبة توصلت الدراسة إلي عدة

نتائج أهمها:

1. هنالك تأثيراً معنوياً من قبل المتغيرات المستقلة مجتمعة علي المتغير التابع.
2. أن تشخيص المشاهدات الشاذة في بيانات نموذج الانحدار المتعدد باستخدام عناصر القطر الرئيسي للمصفوفة Hat Matrix وبواقي ستوبونت المحذوفة أفضل من طريقة الرسم الصندوقي التي كشفت عن بعض المشاهدات الشاذة في المتغيرات المستقلة.
3. المعيار الذي حدده جون فوكس لتحديد المشاهدات الشاذة في المتغيرات المستقلة أكثر تشدداً مقارنةً بمعيار بيلسلي وآخرون نيثير وآخرون.
4. عدم وجود اختلاف بين المعيار الذي حدده بيلسلي وآخرون و معيار شاترجي وهادي في قياس تأثير المشاهدات المؤثرة علي قيم المتغير التابع المقدر.
5. بينت مقاييس الكشف عن القيم الشاذة المؤثرة أن بعض الحالات تؤثر علي معامل واحد وبعضها يؤثر علي معاملين والبعض الآخر يؤثر علي جميع المعاملات.
6. تفشل البواقي العادية والبواقي المحذوفة في تحديد المشاهدات الشاذة.
7. أوضحت جميع مقاييس الكشف عن القيم الشاذة أن الحالتان (9) ، (21) تؤثران معنوياً علي نموذج الانحدار المقدر.
8. هنالك تأثيراً معنوياً من قبل المشاهدات الشاذة علي معاملات نموذج الانحدار الخطي المتعدد حيث تعمل علي تضخم قيمة تباين الأخطاء MSE وتقلل من قيمة R^2 وقيمة F .

9. يختلف تأثير المشاهدات الشاذة من مشاهدة إلى أخرى وذلك حسب قوة تأثير المشاهدة علي معلمات النموذج.

10. إن وجود القيم الشاذة في بيانات نموذج الانحدار الخطي المتعدد يؤثر معنوياً وبصور كبيرة علي عدم تحقق فرضية التوزيع الطبيعي.

11. احتمال ظهور الارتباط الذاتي لبواقي النموذج عند وجود القيم الشاذة في بيانات النموذج.

12. أن معالجة القيم الشاذة أدت إلي تحسن كبير جداً في أداء نموذج الانحدار الخطي المتعدد بالتنبؤ بقيم المتغير التابع.

13. معالجة القيم الشاذة بطريقة متوسط البتر أفضل من معالجتها بطريقة الحذف.

14. تؤثر القيم الشاذة علي دقة تنبؤات نموذج الانحدار المقدر.

15. النموذج المقدر له كفاءة عالية في التنبؤ بقيم المتغير التابع وذلك اعتماداً علي قيمة معامل ثايل والبالغة 0.0021624 وقيمة متوسط الخطأ المطلق $MAE=0.005791$ وهي قيم قريبة جداً من الصفر.

16. النموذج المقدر يطابق البيانات تماماً وذلك اعتماداً علي قيمة الجذر التربيعي لمتوسط مربع الخطأ حيث بلغت قيمته $RMSE=0.007062$.

5-2 التوصيات:

بناءً على النتائج التي توصلت إليها الدراسة يمكن التوصية بالآتي:

1. ضرورة تشخيص وتقليل أثر القيم الشاذة عند تطبيق نموذج تحليل الانحدار للحصول على نموذج أفضل.
2. استخدام طريقة متوسط البتر لمعالجة القيم الشاذة في بيانات نموذج الانحدار الخطي المتعدد.
3. استخدام عناصر الرئيسي للمصفوفة Hat Matrix للكشف عن القيم الشاذة في المتغيرات المستقلة وبواقي ستودنت المحذوفة للكشف عن القيم الشاذة في المتغير التابع عند تطبيق نموذج تحليل الانحدار الخطي المتعدد.
4. تقليل أثر القيم الشاذة عند مخالفة النموذج لفرضية التوزيع الطبيعي للبواقي.
5. تطبيق الطرق الحصينة في دراسات مماثلة لتقليل أثر القيم الشاذة على نتائج نموذج الانحدار الخطي المتعدد.
6. مزيد من الدراسات حول موضوع القيم الشاذة في بيانات النماذج الإحصائية الأخرى نسبةً لندرة الدراسات في هذا المجال المهم.

المراجع

أولاً: المراجع باللغة العربية:

1. إبراهيم، بسام يونس وآخرون، (2002م)، "الإقتصاد القياسي"، دار عزه للنشر والتوزيع، الخرطوم، السودان.
2. الراوي، خاشع محمود، (1987م)، "المدخل إلى تحليل الانحدار"، مديرية دار الكتب للطباعة والنشر، جامعة الموصل، العراق.
3. إسماعيل، محمد عبد الرحمن، (2001م)، "تحليل الانحدار الخطي"، مركز البحوث، الرياض.
4. الطيب، عز الدين مالك، (2008م)، "المدخل إلى الاقتصاد القياسي"، الطبعة الأولى، مطبعة جي تاون، الخرطوم، السودان.
5. السيفو، وليد إسماعيل وآخرون، (2006)، "مشاكل الاقتصاد القياسي التحليلي"، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن، ص:123.
6. أحمد، محمد طه، (2007)، "المشاهدات غير المألوفة في نموذج الانحدار"، مجلة تكريت للعلوم الإدارية والاقتصادية.
7. العباسي، عبد الحميد، (2012)، "إنحدار الريدج والانحدار البواسني"، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة.
8. الكيخيا، نجاه رشيد، (2007م)، "أساسيات الإحصائيات الإحصائي"، دار المريخ للنشر، الرياض، السعودية.
9. أمين، أسامة ربيع، (2008)، "التحليل الإحصائي للمتغيرات المتعددة (1) باستخدام برنامج SPSS"، الطبعة الأولى، كلية التجارة، جامعة المنوفية.

10. الجبوري، مني حسين، (1998م)، "دراسة تحليلية للقيم الشاذة والقيم المفقودة لتصميم المربع اللاتيني وتصميم تام في حالة تكرار مشاهدات العينة"، رسالة ماجستير مقدمة إلى مجلس كلية الإدارة والاقتصاد، الجامعة المستنصرية، بغداد، العراق.
11. بخيت، حسين علي، الرفاعي، غالب عوض، (2007)، "تحليل ونمذجة البيانات باستخدام الحاسوب (تطبيق شامل للحزمة SPSS)"، الطبعة الثانية، الأهلية للنشر والتوزيع، المملكة الأردنية، عمان، ص: 365.
12. تومي صالح، (2001)، "مدخل لنظرية القياس الإحصائي"، الجزء الثاني، ديوان المطبوعات، الجزائر، ص: 146.
13. جولي بلانت، (2006م)، "التحليل الإحصائي باستخدام برنامج SPSS"، دار الفاروق للنشر والتوزيع، مصر.
14. جون نيتز وآخرون، (2000م)، "نماذج إحصائية خطية تطبيقية"، الجزء الثاني، ترجمة انس إسماعيل كنجو، جامعة الملك سعود الرياض.
15. جواد، نزار مصطفى، كمال، غفران إسماعيل، (2009)، "مقارنة طرق تقدير معالم نموذج الانحدار في حالة ظهور مشكلة التعدد الخطي والقيم الشاذة"، مجلة العلوم الإدارية والاقتصادية، العدد (55)، ص: 153-166، العراق، بغداد.
16. حسن، تارا احمد، رضا، مهدي صابر، (2011)، "استخدام الانحدار الحصين لإيجاد أنسب نموذج لتمثيل بيانات الأنواء الجوية في مدينة أربيل خلال الفترة 1998-2010"، مجلة العلوم الاقتصادية والإدارية، العدد (89)، ص: 195-213، الجامعة المستنصرية.

17. حسن، صباح حسيب،(2009)، " استخدام المعايير الحصينة في تحديد المتغيرات المؤثرة في نموذج الانحدار الخطي لقياس السكر في الدم"، المجلة العراقية للعلوم الإحصائية، العدد (16)، ص:ص 131-148.
18. عبد المنعم، ثروت محمد، (2004م)، "مدخل حديث للإحصاء و الاحتمالات"، مكتبة العبيكان، الرياض.
19. عبد المنعم، ثروت محمد، (2005)، "الانحدار"، مكتبة الأنجلو المصرية، القاهرة، مصر، ص:506.
20. ريتشارد جونسون وآخرون، (1998م)، "التحليل الإحصائي للمتغيرات المتعددة" ، ترجمة عبد الرحمن حامد عزام، دار المريخ، السعودية.
21. دبوب، مروان عبد العزيز؛ يونس، فرح عبد الغني، (2006)، "تأثير القيم الشاذة علي نتائج تحليل الانحدار مع التطبيق علي المواليد الخدج"، مجلة علوم الرافدين، المجلد17، العدد1، ص،ص: 62-81 .
22. علي، حامد حسين، (2005)، "طرق بديلة لكشف وتقليل أثر المشاهدات الشاذة علي نموذج الانحدار الخطي البسيط"، رسالة ماجستير غير منشورة، جامعة السودان للعلوم والتكنولوجيا ، الخرطوم، السودان.
23. يحي، مزاحم محمد، (2009)، "المشاهدات غير العادية في الانحدار الخطي المتعدد وبعض طرائق تشخيصها مع التطبيق"، المجلة العراقية للعلوم الإحصائية، العدد (15)، ص، ص:161-192 .
24. داوود، ألاء عبد الستار، (2010)، "تأثير القيم الشاذة ونقطة الأصل علي نتائج تحليل الانحدار"، مجلة تكريت للعلوم الصرفة، المجلد 15، العدد1.

25. قاسم، محمد نذير، إسماعيل، يونس حازم، (2008)، "الكشف عن القيم الشاذة بأسلوب بيز

باستخدام معاينة جيس"، المجلة العراقية للعلوم الإحصائية، العدد (14) ص، ص:68-88.

26. شاكر، صالح مؤيد، (2009)، "تحسين أسلوب M الحصين في تقدير معاملات نموذج

الانحدار الخطي المتعدد"، المجلة العراقية للعلوم الإحصائية، العدد (16)، ص:ص 219-

.242

27. هتمات، سعيد (2006)، "دراسة اقتصادية وقياسية لظاهرة التضخم في الجزائر"، رسالة

ماجستير.

ثانياً: المراجع باللغة الإنجليزية:

28. Netter et al, (1996), "Applied Linear Statistical Models", WCB McGraw Hill, Fourth Edition.

29. Cook, R.D. (1977), "Detection of Influential Observation in Linear Regression. Technometrics 19:pp15-18.

30. Chatterjee.S. & H. (1986), "Influential Observation, High Leverage Points, and Outlier in Linear Regression", Statistical Science, 1(3), p379.

31. Chatterjee, s. and Hadi, A. S.(1988), "Sensitivity Analysis in Linear Regression", New York: John Wiley.

32. AL-Jobouri, S. (1976), "Test of Outliers Unpublished, M.s.c Thesis University of Baghdad.

33. Bross, I.D.J. (1961), "Outliers in Patternend Experiments: strategic Re-Appraisal". Technometrics. 3,19-102.

34. Belsley,D. et al,(1980), "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity ", Wiley, New York, p:105.

35. Barnett, V. & Lewis, T. (1978), "Outliers in Statistical Data", Jon Wiley and Sons, New York.
36. Freeman, P.R.(1980), "On the Number of Outliers in Data from a linear Model" In Bayesian Statistics, Ed .J.M.
37. Keller, G. & Brian Warrack, (2000),"Statistic For Management and Economics, 5th edition, Duxbury, Thomson Learning, U.S.A.
38. Tukey, J.W.(1977), "Exploratory Data Analysis", Addison-Wesley, reading, MA.
39. Glejser, H.,(1979), A New Test for Heteroscedasticity. Journal of the American Statistical Association, Vol. 64,pp.316-323.
40. Goldfeld, S.M. and Quandt, R. E.,(1965), Some Test for Homoscedasticity. Journal of the American Statistical Association, 60pp.539-547.
41. Gujarati, D. N. (1988), "Basic Econometrics',(2nd edition), McGraw-Hill Book Company, New York.
42. Park, R. E.,(1969), "Estimation with Heteroscedastic Error Terms. Econometrica, Vol. 34. No. 4p. 888.
43. Fox, John, (1997), Applied Regression Analysis, Linear Models, and Related Methods. Sage publications.
44. Kleinbaum, J, et al, (1988),"Applied Regression Analysis and other Multivariable Methods", *PWS-Kent* Publishing Company, Boston, second edition, p:210.
45. Neter,J. et al, (1990), "Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs". (3rd edition). Irwin, Homewood, IL 60430, Boston, MA 02116.
46. Johnston, J. (1984), "Econometric Methods", 3rd edition, McGraw-Hill Book Company, New York, P:250.

47. Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. and Black, W. C. (1998) "Multivariate Data Analysis", 5th edn. Prentice-Hall, New Jersey.
48. Theil, H. and Nagar, A.L. (1961), Testing the Independence of Regression Disturbances . Journal of the American Statistical Association , vol. 56 pp:793-806.
49. Cochrane, D. and Orcutt, G.H., (1949), Application of Least Squares Regressions to Relations Containing Autocorrelated Error Terms. Journal of the American Statistical Association, vol.44, pp:32-61.

الملاحق

ملحق رقم (1)
نتائج النموذج المقدر قبل معالجة القيم الشاذة

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
Dependent Variable: Y									
Method: Least Squares									
Date: 06/10/15 Time: 14:48									
Sample: 1 30									
Included observations: 30									
Variable	Coefficient	Std. Error	t-Statistic	Prob.					
X1	0.400894	0.022252	18.01640	0.0000					
X2	0.240036	0.027083	8.863025	0.0000					
X3	0.349354	0.022205	15.73290	0.0000					
C	0.037983	0.037545	1.011671	0.3210					
R-squared	0.991915	Mean dependent var	2.076000						
Adjusted R-squared	0.990982	S.D. dependent var	0.561000						
S.E. of regression	0.053273	Akaike info criterion	-2.903212						
Sum squared resid	0.073788	Schwarz criterion	-2.716386						
Log likelihood	47.54818	Hannan-Quinn criter.	-2.843445						
F-statistic	1063.322	Durbin-Watson stat	2.420304						
Prob(F-statistic)	0.000000								

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.038	.038		1.012	.321		
	x1	.401	.022	.460	18.016	.000	.477	2.095
	x2	.240	.027	.264	8.863	.000	.351	2.848
	x3	.349	.022	.405	15.733	.000	.469	2.131

a. Dependent Variable: y

Collinearity Diagnostics ^a							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.893	1.000	.00	.00	.00	.00
	2	.052	8.676	.94	.02	.03	.16
	3	.036	10.420	.06	.60	.01	.51
	4	.019	14.234	.00	.38	.96	.32

a. Dependent Variable: Y

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.224	30	.001	.680	30	.000

a. Lilliefors Significance Correction

المصدر: إعداد الباحث باستخدام برنامج SPSS

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats
Influence Statistics								
Date: 06/10/15 Time: 16:12								
Sample: 1 30								
Included observations: 30								
Obs.	Resid.	RStudent	Drop Resid	Hat Matrix				
1	0.015227	0.303956	0.017842	0.146547				
2	0.004312	0.082298	0.004635	0.069849				
3	0.014166	0.271717	0.015338	0.076422				
4	-0.034523	-0.702894	-0.041421	0.166537				
5	-0.013278	-0.249422	-0.013794	0.037377				
6	0.016617	0.315936	0.017659	0.058970				
7	0.056850	1.225185	0.073518	0.226718				
8	-0.035324	-0.686628	-0.038664	0.086406				
9	0.226698	28.43271	0.315726	0.281979				
10	-0.037850	-0.724881	-0.040131	0.056824				
11	-0.055858	-1.134327	-0.064661	0.136147				
12	-0.020511	-0.408057	-0.023803	0.138304				
13	-0.010038	-0.190192	-0.010621	0.054866				
14	0.018432	0.372137	0.022054	0.164250				
15	0.030268	0.634384	0.038621	0.216279				
16	-0.013203	-0.248716	-0.013794	0.042860				
17	-0.017037	-0.337580	-0.019653	0.133070				
18	-0.015690	-0.305574	-0.017500	0.103404				
19	-0.033744	-0.668560	-0.038409	0.121471				
20	0.017653	0.358462	0.021373	0.174047				
21	-0.042125	-1.078414	-0.077863	0.458989				
22	-0.013620	-0.260348	-0.014648	0.070181				
23	-0.025309	-0.486523	-0.027346	0.074491				
24	-0.001534	-0.029588	-0.001684	0.089176				
25	0.009150	0.179063	0.010329	0.114138				
26	-0.034355	-0.787684	-0.052014	0.339514				
27	-0.000498	-0.009368	-0.000520	0.043835				
28	-0.014723	-0.282817	-0.015983	0.078807				
29	0.044906	0.904224	0.052037	0.137036				
30	-0.035061	-0.687232	-0.039022	0.101506				

المصدر: إعداد الباحث باستخدام برنامج Eviews

Influence Statistics

Date: 06/10/15 Time: 16:25

Sample: 1 30

Included observations: 30

Obs.	DFFITS	COVRATIO	Di
1	-0.12595	1.350658	0.00411
2	-0.02255	1.256346	0.00013
3	-0.07816	1.251806	0.00158
4	0.314197	1.297929	0.02517
5	0.049148	1.20326	0.00063
6	-0.07909	1.223511	0.00162
7	-0.6634	1.198122	0.10794
8	0.211163	1.188297	0.01138
9	-17.818	1.32E-06	2.47609
10	0.177924	1.141318	0.00806
11	0.450322	1.107923	0.05014
12	0.163479	1.322048	0.0069
13	0.045824	1.230632	0.00055
14	-0.16498	1.36918	0.00704
15	-0.33326	1.400329	0.02842
16	0.052631	1.210221	0.00072
17	0.132258	1.325099	0.00453
18	0.103774	1.285465	0.00279
19	0.248599	1.240489	0.01579
20	-0.16455	1.387625	0.007
21	0.993308	1.802765	0.24513
22	0.071526	1.244605	0.00133
23	0.138027	1.217256	0.00491
24	0.009258	1.284216	0.00002
25	-0.06427	1.313834	0.00107
26	0.564741	1.605768	0.08091
27	0.002006	1.223473	0
28	0.08272	1.253816	0.00177
29	-0.36033	1.191889	0.03269
30	0.23099	1.20811	0.01362

المصدر: إعداد الباحث باستخدام برنامج SPSS وEviews

Scaled Difference in Coefficients (DFBETAS)

Date: 06/10/15 Time: 17:08

Sample: 1 30

Included observations: 30

Obs.	X1	X2	X3	C
1	0.043283	0.056388	-0.042604	-0.073175
2	0.003620	0.009185	-0.016097	0.004621
3	-0.012769	0.055591	-0.039369	-0.008242
4	-0.255437	0.124760	-0.034740	0.134125
5	-0.005038	-0.010293	0.013461	-0.006074
6	-0.037438	0.040961	0.005638	-0.006409
7	-0.552878	0.509947	-0.074483	0.072495
8	-0.141151	0.075282	-0.039452	0.074529
9	10.60349	-16.54697	7.077402	5.017075
10	-0.081516	0.091750	-0.066958	0.006548
11	-0.022814	0.291832	-0.375776	-0.027264
12	-0.072888	0.010542	-0.047134	0.109435
13	-0.002448	-0.000244	-0.017925	0.015246
14	-0.086755	0.012723	0.117407	-0.030118
15	0.096403	0.187551	-0.153600	-0.188098
16	-0.007241	-8.31E-05	-0.011666	0.010833
17	-0.035497	0.000164	-0.052532	0.088559
18	0.023859	0.034581	-0.084492	0.002313
19	0.031619	0.032411	0.092532	-0.230782
20	-0.053198	0.097324	-0.126351	0.086740
21	-0.549146	0.386540	0.529031	-0.591284
22	-0.034163	0.009285	0.038049	-0.029029
23	-0.018888	0.064607	0.010520	-0.110199
24	0.002533	-0.005144	0.006380	-0.004367
25	-0.049466	0.016368	0.002266	0.043192
26	0.291739	0.232439	-0.385191	-0.306069
27	0.000345	-0.000794	0.000852	-0.000687
28	0.023703	-0.008007	0.032546	-0.069639
29	-0.282842	0.224612	-0.090328	0.165525
30	0.051239	0.128574	-0.125881	-0.143790

المصدر: إعداد الباحث باستخدام برنامج Eviews

ملحق رقم (2)
نتائج النموذج المقدر بعد حذف المشاهدة رقم (7)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 09:11				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.413079	0.024180	17.08322	0.0000
X2	0.226357	0.029057	7.790226	0.0000
X3	0.350992	0.022035	15.92886	0.0000
C	0.035287	0.037254	0.947221	0.3526
R-squared	0.992372	Mean dependent var	2.077241	
Adjusted R-squared	0.991457	S.D. dependent var	0.570888	
S.E. of regression	0.052767	Akaike info criterion	-2.918425	
Sum squared resid	0.069609	Schwarz criterion	-2.729832	
Log likelihood	46.31716	Hannan-Quinn criter.	-2.859360	
F-statistic	1084.155	Durbin-Watson stat	2.483547	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistic	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.035	.037		.947	.353		
	X1	.413	.024	.466	17.083	.000	.410	2.439
	X2	.226	.029	.246	7.790	.000	.306	3.263
	X3	.351	.022	.407	15.929	.000	.468	2.137

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.895	1.000	.00	.00	.00	.00
	2	.054	8.529	.93	.01	.03	.16
	3	.035	10.508	.06	.43	.03	.61
	4	.017	15.299	.00	.55	.94	.22

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (3)

نتائج النموذج المقدر بعد حذف المشاهدة رقم (9)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 09:30				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.359220	0.004195	85.63810	0.0000
X2	0.319190	0.005535	57.67118	0.0000
X3	0.321596	0.004042	79.56917	0.0000
C	0.004713	0.006734	0.699867	0.4905
R-squared	0.999757	Mean dependent var	2.074828	
Adjusted R-squared	0.999728	S.D. dependent var	0.570893	
S.E. of regression	0.009409	Akaike info criterion	-6.366775	
Sum squared resid	0.002213	Schwarz criterion	-6.178183	
Log likelihood	96.31824	Hannan-Quinn criter.	-6.307711	
F-statistic	34349.30	Durbin-Watson stat	2.094622	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistic	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.005	.007		.700	.490		
	X1	.359	.004	.412	85.638	.000	.420	2.382
	X2	.319	.006	.336	57.671	.000	.286	3.497
	X3	.322	.004	.373	79.569	.000	.442	2.261

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.896	1.000	.00	.00	.00	.00
	2	.053	8.604	.92	.02	.02	.16
	3	.037	10.261	.05	.52	.00	.50
	4	.015	16.323	.02	.46	.98	.33

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (4)

نتائج النموذج المقدر بعد حذف المشاهدة رقم (11)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 09:41				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.401399	0.022134	18.13461	0.0000
X2	0.232176	0.027812	8.348074	0.0000
X3	0.357652	0.023264	15.37356	0.0000
C	0.039001	0.037351	1.044196	0.3064
R-squared	0.992300	Mean dependent var	2.072069	
Adjusted R-squared	0.991376	S.D. dependent var	0.570510	
S.E. of regression	0.052982	Akaike info criterion	-2.910302	
Sum squared resid	0.070176	Schwarz criterion	-2.721710	
Log likelihood	46.19938	Hannan-Quinn criter.	-2.851237	
F-statistic	1073.880	Durbin-Watson stat	2.545715	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.039	.037		1.044	.306		
	X1	.401	.022	.461	18.135	.000	.477	2.096
	X2	.232	.028	.255	8.348	.000	.331	3.025
	X3	.358	.023	.408	15.374	.000	.437	2.290

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.892	1.000	.00	.00	.00	.00
	2	.054	8.526	.94	.02	.03	.14
	3	.036	10.401	.05	.66	.00	.45
	4	.018	14.521	.00	.31	.97	.40

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (5)

نتائج النموذج المقدر بعد حذف المشاهدة رقم (15)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 10:02				
Sample: 2001 2029				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.398723	0.022770	17.51073	0.0000
X2	0.234898	0.028572	8.221270	0.0000
X3	0.352804	0.023114	15.26362	0.0000
C	0.045128	0.039619	1.139060	0.2655
R-squared	0.991211	Mean dependent var	2.044483	
Adjusted R-squared	0.990157	S.D. dependent var	0.543228	
S.E. of regression	0.053896	Akaike info criterion	-2.876085	
Sum squared resid	0.072619	Schwarz criterion	-2.687492	
Log likelihood	45.70323	Hannan-Quinn criter.	-2.817020	
F-statistic	939.8472	Durbin-Watson stat	2.421240	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.045	.040		1.139	.265		
	X1	.399	.023	.453	17.511	.000	.524	1.908
	X2	.235	.029	.253	8.221	.000	.371	2.692
	X3	.353	.023	.427	15.264	.000	.450	2.222

a. Dependent Variable: Y

Collinearity Diagnostics ^a							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.894	1.000	.00	.00	.00	.00
	2	.052	8.672	.81	.00	.02	.25
	3	.035	10.568	.16	.77	.00	.34
	4	.020	14.123	.03	.22	.97	.41

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (6)

نتائج النموذج المقدر بعد حذف المشاهدة رقم (21)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 10:21				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.413075	0.024893	16.59433	0.0000
X2	0.229600	0.028680	8.005501	0.0000
X3	0.337643	0.024656	13.69409	0.0000
C	0.060114	0.042685	1.408324	0.1714
R-squared	0.990896	Mean dependent var	2.115862	
Adjusted R-squared	0.989803	S.D. dependent var	0.525918	
S.E. of regression	0.053107	Akaike info criterion	-2.905584	
Sum squared resid	0.070508	Schwarz criterion	-2.716992	
Log likelihood	46.13097	Hannan-Quinn criter.	-2.846520	
F-statistic	906.9875	Durbin-Watson stat	2.447514	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.060	.043		1.408	.171		
	X1	.413	.025	.512	16.594	.000	.382	2.618
	X2	.230	.029	.245	8.006	.000	.388	2.580
	X3	.338	.025	.358	13.694	.000	.532	1.881

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.911	1.000	.00	.00	.00	.00
	2	.046	9.269	.78	.21	.02	.00
	3	.027	12.133	.11	.16	.07	.99
	4	.017	15.239	.11	.64	.91	.00

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (7)

نتائج النموذج المقدر بعد حذف المشاهدة رقم (26)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 10:29				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.394354	0.023904	16.49748	0.0000
X2	0.233695	0.028446	8.215426	0.0000
X3	0.357970	0.024901	14.37597	0.0000
C	0.049560	0.040577	1.221365	0.2333
R-squared	0.991549	Mean dependent var	2.102414	
Adjusted R-squared	0.990535	S.D. dependent var	0.551619	
S.E. of regression	0.053666	Akaike info criterion	-2.884630	
Sum squared resid	0.072001	Schwarz criterion	-2.696037	
Log likelihood	45.82713	Hannan-Quinn criter.	-2.825565	
F-statistic	977.7541	Durbin-Watson stat	2.394503	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficient ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.050	.041		1.221	.233		
	X1	.394	.024	.431	16.497	.000	.495	2.020
	X2	.234	.028	.253	8.215	.000	.356	2.811
	X3	.358	.025	.429	14.376	.000	.379	2.637

a. Dependent Variable: Y

Collinearity Diagnostics ^a							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.903	1.000	.00	.00	.00	.00
	2	.052	8.636	.75	.00	.02	.21
	3	.027	12.029	.18	.92	.01	.26
	4	.018	14.709	.06	.07	.96	.52

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (8)

نتائج النموذج المقدر بعد حذف المشاهدة رقم (21+9)

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 10:39				
Sample: 1 28				
Included observations: 28				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.351584	0.004063	86.53648	0.0000
X2	0.327214	0.005083	64.37153	0.0000
X3	0.326622	0.003615	90.34594	0.0000
C	-0.007711	0.006547	-1.177800	0.2504
R-squared	0.999814	Mean dependent var	2.116071	
Adjusted R-squared	0.999790	S.D. dependent var	0.535567	
S.E. of regression	0.007755	Akaike info criterion	-6.749327	
Sum squared resid	0.001443	Schwarz criterion	-6.559012	
Log likelihood	98.49057	Hannan-Quinn criter.	-6.691146	
F-statistic	42913.75	Durbin-Watson stat	2.132137	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.008	.007		-1.178	.250		
	X1	.352	.004	.436	86.536	.000	.306	3.265
	X2	.327	.005	.329	64.372	.000	.297	3.364
	X3	.327	.004	.346	90.346	.000	.529	1.892

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.914	1.000	.00	.00	.00	.00
	2	.047	9.141	.70	.18	.01	.00
	3	.027	11.971	.10	.14	.04	.98
	4	.011	18.452	.20	.67	.95	.01

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (9)

نتائج النموذج المقدر بعد حذف جميع المشاهدات الشاذة

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 10:52				
Sample: 1 24				
Included observations: 24				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.350683	0.005030	69.71976	0.0000
X2	0.322838	0.006981	46.24818	0.0000
X3	0.330710	0.004708	70.24281	0.0000
C	-0.004187	0.007626	-0.549026	0.5891
R-squared	0.999808	Mean dependent var	2.113333	
Adjusted R-squared	0.999779	S.D. dependent var	0.524178	
S.E. of regression	0.007793	Akaike info criterion	-6.720218	
Sum squared resid	0.001215	Schwarz criterion	-6.523876	
Log likelihood	84.64262	Hannan-Quinn criter.	-6.668128	
F-statistic	34681.02	Durbin-Watson stat	2.222539	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.004	.008		-.549	.589		
	X1	.351	.005	.405	69.720	.000	.285	3.513
	X2	.323	.007	.293	46.248	.000	.239	4.189
	X3	.331	.005	.378	70.243	.000	.332	3.010

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.929	1.000	.00	.00	.00	.00
	2	.045	9.361	.67	.05	.00	.11
	3	.018	14.675	.02	.53	.01	.72
	4	.008	22.455	.31	.42	.99	.17

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (10)

نتائج النموذج المقدر بعد معالجة المشاهدات رقم (9) بطريقة متوسط البتر

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 14:55				
Sample: 1 30				
Included observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.359204	0.004119	87.21151	0.0000
X2	0.319226	0.005434	58.75093	0.0000
X3	0.321575	0.003968	81.03582	0.0000
C	0.004625	0.006605	0.700264	0.4900
R-squared	0.999757	Mean dependent var	2.075000	
Adjusted R-squared	0.999729	S.D. dependent var	0.560964	
S.E. of regression	0.009240	Akaike info criterion	-6.406952	
Sum squared resid	0.002220	Schwarz criterion	-6.220126	
Log likelihood	100.1043	Hannan-Quinn criter.	-6.347185	
F-statistic	35619.18	Durbin-Watson stat	2.103075	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.005	.007		.700	.490		
	X1	.359	.004	.412	87.212	.000	.420	2.382
	X2	.319	.005	.336	58.751	.000	.286	3.495
	X3	.322	.004	.373	81.036	.000	.442	2.261

a. Dependent Variable: Y

Collinearity Diagnostics ^a							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.899	1.000	.00	.00	.00	.00
	2	.051	8.748	.92	.02	.02	.16
	3	.036	10.429	.05	.52	.00	.50
	4	.014	16.580	.02	.46	.98	.33

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (10)

نتائج النموذج المقدر بعد معالجة المشاهدة رقم 21 بطريقة متوسط البتر

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.059	.042		1.419	.168		
	X1	.413	.024	.512	16.891	.000	.382	2.618
	X2	.230	.028	.246	8.156	.000	.388	2.579
	X3	.338	.024	.358	13.944	.000	.532	1.881

a. Dependent Variable: Y

Collinearity Diagnostics ^a							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.914	1.000	.00	.00	.00	.00
	2	.044	9.420	.78	.21	.02	.00
	3	.026	12.324	.11	.16	.07	.99
	4	.016	15.477	.11	.64	.91	.00

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (11)

نتائج النموذج المقدر بعد معالجة المشاهدات رقم 26 بطريقة متوسط البتر

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.049	.040		1.230	.230		
	X1	.394	.023	.431	16.803	.000	.495	2.020
	X2	.234	.028	.253	8.370	.000	.356	2.811
	X3	.358	.024	.429	14.635	.000	.379	2.636

a. Dependent Variable: Y

Collinearity Diagnostics ^a							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.906	1.000	.00	.00	.00	.00
	2	.051	8.780	.75	.00	.02	.21
	3	.026	12.217	.18	.92	.01	.26
	4	.017	14.941	.06	.07	.96	.52

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (12)

نتائج النموذج المقدر بعد معالجة المشاهدات رقم (9+21) بطريقة متوسط البتر

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 14:48				
Sample: 1 30				
Included observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.351484	0.003970	88.53164	0.0000
X2	0.327344	0.004966	65.91460	0.0000
X3	0.326650	0.003536	92.37124	0.0000
C	-0.008121	0.006378	-1.273299	0.2142
R-squared	0.999807	Mean dependent var	2.113667	
Adjusted R-squared	0.999785	S.D. dependent var	0.516850	
S.E. of regression	0.007586	Akaike info criterion	-6.801411	
Sum squared resid	0.001496	Schwarz criterion	-6.614585	
Log likelihood	106.0212	Hannan-Quinn criter.	-6.741644	
F-statistic	44861.76	Durbin-Watson stat	2.137270	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.008	.006		-1.273	.214		
	X1	.351	.004	.436	88.532	.000	.307	3.259
	X2	.327	.005	.329	65.915	.000	.298	3.359
	X3	.327	.004	.346	92.371	.000	.529	1.892

a. Dependent Variable: Y

Collinearity Diagnostics							
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	X1	X2	X3
1	1	3.920	1.000	.00	.00	.00	.00
	2	.044	9.446	.70	.18	.01	.00
	3	.026	12.357	.10	.14	.04	.98
	4	.011	19.018	.20	.67	.95	.01

a. Dependent Variable: Y

المصدر: إعداد الباحث باستخدام برنامج SPSS

Influence Statistics

Date: 04/02/15 Time: 05:04

Sample: 1 30

Included observations: 30

Obs.	Resid.	RStudent	DFFITS	Drop Resid	COVRATIO	Hat Matrix
1	0.003343	0.470012	-0.195350	0.003920	1.324506	0.147300
2	0.005339	0.728882	-0.221286	0.005831	1.174611	0.084392
3	-0.004757	-0.646423	0.190573	-0.005171	1.189960	0.079964
4	0.011371	1.754458	-0.909498	0.014426	0.932805	0.211811
5	-0.005784	-0.771868	0.156765	-0.006022	1.108599	0.039615
6	0.003242	0.435868	-0.118206	0.003481	1.218438	0.068509
7	-0.002595	-0.413821	0.295902	-0.003922	1.720377	0.338316
8	-0.002582	-0.353644	0.121269	-0.002885	1.281583	0.105217
9	-0.000766	-0.100871	0.019366	-0.000794	1.211008	0.035550
10	-0.004553	-0.616716	0.176386	-0.004925	1.191382	0.075616
11	-0.011481	-1.711373	0.746704	-0.013667	0.894063	0.159928
12	-0.005437	-0.767926	0.313484	-0.006343	1.243285	0.142842
13	-0.001392	-0.185316	0.045173	-0.001475	1.232585	0.056087
14	0.009785	1.510124	-0.835997	0.012783	1.077914	0.234577
15	0.005186	0.770322	-0.414941	0.006690	1.374116	0.224899
16	-0.002232	-0.295221	0.061871	-0.002330	1.204358	0.042074
17	-0.008963	-1.295222	0.536212	-0.010499	1.056836	0.146313
18	0.004651	0.644690	-0.233458	0.005261	1.238809	0.115932
19	-0.008092	-1.197987	0.587776	-0.010039	1.161021	0.194019
20	-0.004435	-0.645858	0.321794	-0.005536	1.366744	0.198876
21	-0.007072	-0.946533	0.177507	-0.007320	1.051911	0.033974
22	0.013112	1.950169	-0.753074	0.015068	0.762953	0.129767
23	0.012368	1.835920	-0.739182	0.014373	0.819718	0.139492
24	-0.012551	-1.818657	0.602787	-0.013930	0.789866	0.098983
25	0.001117	0.153403	-0.054876	0.001260	1.314603	0.113448
27	-0.000254	-0.033576	0.007321	-0.000266	1.225254	0.045384
28	-0.003924	-0.538554	0.182130	-0.004373	1.244870	0.102630
29	0.013879	2.117477	-0.901720	0.016396	0.714400	0.153507
30	0.002041	0.282310	-0.106169	0.002330	1.318420	0.123907

Scaled Difference in Coefficients (DFBETAS)					
Obs.	X1	X2	X3	C	
1	0.038640	0.084770	-0.058075	-0.107574	
2	0.074081	0.025204	-0.168799	0.082910	
3	0.046842	-0.131592	0.088617	0.033418	
4	0.774541	-0.519707	0.025325	-0.086413	
5	-0.042930	0.003597	0.053788	-0.042370	
6	-0.070698	0.072599	0.020051	-0.035079	
7	0.263909	-0.251030	-0.007636	0.067254	
8	-0.089499	0.064498	-0.013370	0.008601	
9	-0.004012	0.004502	0.000340	-0.006816	
10	-0.108364	0.119570	-0.043946	-0.035563	
11	-0.124829	0.488731	-0.575021	-0.081528	
12	-0.099229	0.016397	-0.108777	0.193332	
13	0.000952	-0.001291	-0.020568	0.016813	
14	-0.541106	0.252269	0.622702	-0.358901	
15	0.048085	0.232604	-0.169414	-0.228500	
16	-0.010499	0.005231	-0.014603	0.009591	
17	-0.038972	-0.047757	-0.257928	0.370142	
18	-0.056106	-0.055212	0.196830	-0.030279	
19	-0.164012	0.241594	0.273575	-0.558291	
20	0.049691	-0.129305	0.259281	-0.170555	
21	-0.019598	0.019840	0.007571	-0.054683	
22	0.516112	-0.341716	-0.423931	0.467671	
23	0.367262	-0.499875	-0.167090	0.632942	
24	0.060988	-0.214342	0.446326	-0.302090	
25	-0.037171	0.014737	0.001962	0.027998	
26	-0.064156	-0.065937	0.115851	0.074940	
27	-1.21E-06	-0.001445	0.003605	-0.003027	
28	-0.014496	0.033069	0.088985	-0.160869	
29	-0.705285	0.607933	-0.176915	0.188881	
30	0.009158	-0.071087	0.044051	0.069951	

المصدر: إعداد الباحث باستخدام برنامج Eviews

ملحق رقم (13)

نتائج النموذج المقدر بعد معالجة المشاهدة رقم (9+21+26) بطريقة متوسط البتر

Dependent Variable: Y				
Method: Least Squares				
Date: 06/12/15 Time: 15:03				
Sample: 1 30				
Included observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.351685	0.004345	80.93321	0.0000
X2	0.327547	0.005446	60.14397	0.0000
X3	0.326370	0.004170	78.26094	0.0000
C	-0.008030	0.007062	-1.137038	0.2659
R-squared	0.999765	Mean dependent var	2.141000	
Adjusted R-squared	0.999738	S.D. dependent var	0.494064	
S.E. of regression	0.008002	Akaike info criterion	-6.694742	
Sum squared resid	0.001665	Schwarz criterion	-6.507916	
Log likelihood	104.4211	Hannan-Quinn criter.	-6.634975	
F-statistic	36844.13	Durbin-Watson stat	2.122701	
Prob(F-statistic)	0.000000			

المصدر: إعداد الباحث باستخدام برنامج Eviews

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.008	.007		-1.137	.266		
	x1	.352	.004	.419	80.933	.000	.337	2.966
	x2	.328	.005	.317	60.144	.000	.325	3.081
	x3	.326	.004	.362	78.261	.000	.423	2.365

a. Dependent Variable: y

Collinearity Diagnostics								
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	x1	x2	x3	
1	1	3.932	1.000	.00	.00	.00	.00	
	2	.039	10.092	.77	.11	.00	.08	
	3	.019	14.226	.00	.42	.03	.85	
	4	.010	19.536	.22	.47	.97	.06	

a. Dependent Variable: y

المصدر: إعداد الباحث باستخدام برنامج SPSS

ملحق رقم (14)
القيم المقدرة لأفضل نموذج لقيم المتغير التابع خلال فترة الدراسة

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
obs		Actual	Fitted	Residual	Residual Plot				
1		2.88000	2.87666	0.00334					
2		2.01000	2.00466	0.00534					
3		2.27000	2.27476	-0.00476					
4		2.77000	2.75863	0.01137					
5		2.12000	2.12578	-0.00578					
6		2.28000	2.27676	0.00324					
7		2.04000	2.04260	-0.00260					
8		2.54000	2.54258	-0.00258					
9		2.08000	2.08077	-0.00077					
10		2.25000	2.25455	-0.00455					
11		2.19000	2.20148	-0.01148					
12		3.00000	3.00544	-0.00544					
13		2.47000	2.47139	-0.00139					
14		2.51000	2.50022	0.00978					
15		2.99000	2.98481	0.00519					
16		2.35000	2.35223	-0.00223					
17		2.99000	2.99896	-0.00896					
18		2.29000	2.28535	0.00465					
19		1.17000	1.17809	-0.00809					
20		1.46000	1.46444	-0.00444					
21		2.08000	2.08707	-0.00707					
22		1.86000	1.84689	0.01311					
23		1.54000	1.52763	0.01237					
24		1.72000	1.73255	-0.01255					
25		1.51000	1.50888	0.00112					
26		1.31000	1.30857	0.00143					
27		1.97000	1.97025	-0.00025					
28		1.45000	1.45392	-0.00392					
29		1.68000	1.66612	0.01388					
30		1.63000	1.62796	0.00204					

المصدر: إعداد الباحث اعتماداً على نتائج البرنامج (Eviews):

ملحق رقم (15)

بيانات الدراسة

الترتيب	المعدل التراكمي Y	المعدل الفصلي الأول X ₁	المعدل الفصلي الثاني X ₂	المعدل الفصلي الثالث X ₃
1	2.88	3.09	3.05	2.45
2	2.01	2.20	2.18	1.61
3	2.27	2.28	2.55	1.98
4	2.77	3.22	2.48	2.52
5	2.12	2.21	2.20	1.95
6	2.28	2.10	2.46	2.27
7	2.04	1.48	2.59	2.09
8	2.54	2.81	2.37	2.41
9	2.11	2.25	1.15	1.91
10	2.25	2.40	2.09	2.25
11	2.19	2.11	1.89	2.60
12	3.00	3.15	2.90	2.93
13	2.47	2.44	2.46	2.50
14	2.51	2.02	2.54	2.96
15	2.99	3.23	3.27	2.41
16	2.35	2.38	2.35	2.31
17	2.99	3.02	2.93	3.02
18	2.29	2.06	2.15	2.65
19	1.17	1.30	1.23	1.00
20	1.46	1.50	1.89	1.00
21	0.92	1.80	0.64	0.14
22	1.86	2.17	1.84	1.50
23	1.54	1.69	1.43	1.45
24	1.72	1.77	2.02	1.40
25	1.51	1.19	1.72	1.64
26	1.31	0.77	1.10	2.10
27	1.97	2.00	2.12	1.78
28	1.45	1.48	1.58	1.30
29	1.68	1.33	2.06	1.63
30	1.63	1.52	1.46	1.91

المصدر: مكتب التسجيل و الامتحانات - كلية التربية رفاة: