

الآية

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

{ فَبَدَأَ بِأَوْعِيَّتِهِمْ قَبْلَ وِعَاءِ أَخِيهِ ثُمَّ اسْتَخْرَجَهَا مِنْ وِعَاءِ أَخِيهِ كَذَلِكَ كِدْنَا لِيُوسُفَ مَا
كَانَ لِيَأْخُذَ أَخَاهُ فِي دِينِ الْمَلِكِ إِلَّا أَنْ يَشَاءَ اللَّهُ نَرْفَعُ دَرَجَاتٍ مَنْ نَشَاءُ وَفَوْقَ كُلِّ ذِي
عِلْمٍ عَلِيمٌ {
صدق الله العظيم

{ سورة يوسف : الآية 76 }

الإهداء

إلى روح والدتي طيب الله ثراها وأسكنها فسيح جناته
إلى القلب الكبير والدي أمد الله في عمره وأمتعته بالصحة والعافية
إلى من وقفت بجانبني و كانت خير معين زوجتي الغالية حباً ومودةً
إلى الأمل المشرق إبنى أحمد حفظه الله ورزقه الهداية وأنار عقله بالعلم والقرآن
إلى إخواني وأخواتي أصحاب القلوب الطيبة والنوايا الصادقة
إلى من عرفت معهم معنى الإخاء وصفاء الروح و رفقاء دربي طلاب الدفعة 24
إلى من أناروا لنا طريق العلم والمعرفة أساتذتي الأجلاء وفاءً وإكراماً
إلى كل من طلب العلم واجتهد في تحصيل المعرفة
إلى جميع الأهل والأحباب والعشيرة والأصدقاء و الزملاء الذين أحببتهم وأحبوني
أهدي إليكم جميعاً ثمرة هذا الجهد المتواضع

الباحث

الشكر و التقدير

الحمد لله رب العالمين حمد الشاكرين والصلاة والسلام علي من بلغ الرسالة وأدى الأمانة ونصح الأمة معلم البشرية الأولي سيدنا محمد وعلى آله وصحبه وسلم.

جزيل شكري أتقدم به إلى منارات العلم والمعرفة، جامعة السودان للعلوم والتكنولوجيا وجامعة الجزيرة وجامعة البطانة. وأتقدم بجزيل الشكر والتقدير للدكتور/ ياسر يوسف عبد الله ، الدكتور/ أبا سفيان محمد الحاج ، الدكتور/ حسين يوسف عبد الله الدكتور/ صديق عبد الصادق البدوي لهم جميعاً أسمى آيات الشكر والعرفان على تشجيعهم لي وإسهاماتهم الغالية التي ساعدتني كثيراً في طريقي هذا. ووافر الشكر والتقدير والاحترام أتقدم به إلى الدكتور/ عادل موسى يونس الذي تكرم بالإشراف على هذه الدراسة والذي كان خير معين و ما بذل علينا بوقته ولا جهده ولا نصائحه وملاحظاته وتصويباته الغالية التي أسفرت عن خروج هذه الدراسة في صورتها المشرفة. والشكر موصول أيضاً إلي أسرة مكتبة كلية العلوم وأسرة مكتبة الدراسات العليا جامعة السودان للعلوم والتكنولوجيا والي أسرة مكتبة كلية العلوم وتقانة المعلومات جامعة النيلين. والشكر أيضاً إلى كل الأساتذة الأجلاء بقسم الإحصاء بكلية العلوم جامعة السودان للعلوم والتكنولوجيا. وخالص الشكر لأسرتي الكريمة على وقوفها بجانبني وعطائها اللامحدود.

والشكر كل الشكر إلى كل من مد لي يد العون حتى يكتمل هذا الجهد المتواضع .

جميعاً لكم جزيل شكري وإمتنانى.

والشكر لله من قبل ومن بعد.

الباحث

المستخلص

تناولت الدراسة دراسة تأثير القيم الشاذة علي معلمات نموذج تحليل الانحدار الخطي المتعدد. وهدفت إلي عدة أهداف أهمها: دراسة تأثير وجود ومعالجة القيم الشاذة علي نتائج نموذج تحليل الانحدار الخطي المتعدد ، المقارنة بين طرق ومعايير تشخيص القيم الشاذة. ومن الفرضيات التي استندت عليها الدراسة: القيم الشاذة لها تأثير معنوي علي نتائج نموذج تحليل الانحدار الخطي المتعدد. واعتمدت الدراسة علي المنهج الوصفي لوصف متغيرات الدراسة والمنهج التحليلي الاستدلالي بتطبيق نموذج تحليل الانحدار الخطي المتعدد وبعض الطرق والاختبارات الخاصة بتشخيص ومعالجة القيم الشاذة. حيث تم الكشف عن القيم الشاذة في المتغيرات المستقلة بطريقة الرسم الصندوقي وعناصر القطر الرئيسي للمصفوفة Hat Matrix ، وكذلك تم الكشف عن القيم الشاذة في المتغير التابع باستخدام بواقي ستيودنت المحذوفة وطريقة الرسم الصندوقي ، أما القيم الشاذة المؤثرة فتم الكشف عنها باستخدام الطرق COOK'S ,DFBETAS , DFFITS , COVRATIO ,DISTANCE كما تمت معالجة القيم الشاذة بطريقة الحذف ومتوسط البتر. ومن أهم النتائج التي توصلت إليها الدراسة: هنالك تأثيراً معنوياً من قبل القيم الشاذة علي معلمات نموذج الانحدار الخطي المتعدد حيث تعمل علي تضخم قيمة تباين الأخطاء MSE وتقلل من قيمة R^2 وقيمة F ، إن وجود القيم الشاذة في بيانات نموذج الانحدار الخطي المتعدد يؤثر معنوياً وبصور كبيرة علي عدم تحقق فرضية التوزيع الطبيعي ، احتمال ظهور الارتباط الذاتي لبواقي النموذج عند وجود القيم الشاذة في بيانات النموذج ، معالجة القيم الشاذة بطريقة متوسط البتر أفضل من معالجتها بطريقة الحذف. وأوصت الدراسة بعدة توصيات أهمها: ضرورة تشخيص وتقليل أثر القيم الشاذة عند تطبيق نموذج تحليل الانحدار المتعدد للحصول علي نموذج أفضل.

Abstract

The study dealt with the effect of outliers upon the parameters the multiple linear regression model analysis. The study aimed several objectives most important ones are: study effect of existence and treatment of outliers upon the multiple linear regression model analysis then the comparison between methods and criteria of outliers diagnosis. One of The hypothesis that the study relied on is that the outliers have a significant effect upon the results of multiple linear regression model. The study depended on the descriptive method to describe the study variables, then the analytical inference method to apply the multiple linear regression model, then some special methods and tests to diagnose and treat the outliers, detecting that the outliers in the in dependant variables by method of Box plot and the element of the main diagonal of Hat Matrix. It was , also, detected the outliers in the dependent variable by using of Studentized deleted residual and Box plot. The Influential outliers were found out by using DFFITS, DFBETAS, COVRATIO and Cook's Distance methods. Just as the outliers were treated by deleting and Trimmed mean. The results of the study showed that there is a significant effect of outliers upon the parameters of the multiple liner regression model, where it dose magnifying value of MSE and lessening value of R^2 and F. The existence of outliers in the multiple linear regression model data effects significantly and much bigger on unfulfilling the normal distribution of hypothesis, and probability of appearance of Autocorrelation of residual of the model at the existence of outliers in the model data. Treating outliers by Trimmed mean method is more better than treating it by deleting method. The study recommended several recommendations, the important one is the necessity of diagnosis and lessening the effect of outliers when applying the multiple linear regression model to attain a better model.

فهرس المحتويات

الصفحة	المحتويات	م
I	الآية	
II	الإهداء	
III	الشكر والتقدير	
IV	المستخلص	
V	Abstract	
X	فهرس الجداول	
XII	فهرس الأشكال	
الفصل الأول: خطة الدراسة		
1	تمهيد	1-1
2	مشكلة الدراسة	2-1
2	أهمية الدراسة	3-1
3	أهداف الدراسة	4-1
3	فرضيات الدراسة	5-1
4	منهجية الدراسة	6-1
4	مجتمع وعينة الدراسة	7-1
4	الدراسات السابقة	8-1
12	التعليق علي الدراسات السابقة	9-1
16	الإطار العام للدراسة	10-1
الفصل الثاني : نموذج الانحدار الخطي المتعدد		
18	تمهيد	1-2
18	Nature of Model طبيعة النموذج	2-2
19	Model's Assumption افتراضات النموذج	3-2
21	تقدير معلمات النموذج وتباين الأخطاء	4-2
25	Standardized Regression Model نموذج الانحدار المعياري	5-2
27	تقييم معنوية المقدرات و النموذج	6-2
29	اختبار جودة النموذج	7-2

31	اختبار ثبات معاملات النموذج	8-2
32	التنبؤ Forecasting	9-2
33	تقييم أداء النموذج في التنبؤ	10-2
36	Heteroscedasticity Problem مشكلة عدم تجانس التباين	11-2
37	النتائج المترتبة علي عدم تجانس التباين	1-11-2
38	Detection of Heteroscedasticity اكتشاف عدم تجانس التباين	2-11-2
43	Remedy of Heteroscedasticity معالجة عدم تجانس التباين	3-11-2
49	Multicollinearity Problem مشكلة الارتباط الخطي المتعدد	12-2
51	النتائج المترتبة علي وجود الارتباط المتعدد	1-12-2
53	Detection of Multicollinearity اكتشاف التداخل الخطي المتعدد	2-12-2
58	Remedy of Multicollinearity معالجة الارتباط الخطي المتعدد	3-12-2
62	Autocorrelation Problem مشكلة الارتباط الذاتي	13-2
64	النتائج المترتبة علي وجود الارتباط الذاتي	1-13-2
65	Detection of Autocorrelation اكتشاف الارتباط الذاتي	2-13-2
67	Remedy of Autocorrelation معالجة الارتباط الذاتي	3-13-2
70	تقدير معامل الارتباط الذاتي من الرتبة الأولى $\hat{\rho}$	4-13-2
71	Normality of Residuals اختبار التوزيع الطبيعي للبواقي	14-2
الفصل الثالث: المشاهدات الشاذة في نموذج تحليل الانحدار الخطي المتعدد		
73	تمهيد	1-3
73	Concept of Outlier مفهوم القيم الشاذة	2-3
74	أسباب ظهور القيم الشاذة	3-3
75	طرق الكشف عن المشاهدات الشاذة	4-3
75	الطرق اللامعلمية- طريقة Tukey	1-4-3
76	اكتشاف المشاهدات الشاذة في المتغيرات المستقلة	2-4-3
80	اكتشاف المشاهدات الشاذة في المتغير التابع	3-4-3
83	المشاهدات الشاذة المؤثرة وطرق الكشف عنها	5-3

83	قياس التأثير علي القيم المقدرة - مقياس DFFITS	1-5-3
85	قياس الأثر على معاملات الانحدار - مقياس DFBETAS	2-5-3
86	قياس الأثر علي كل معاملات الانحدار - مقياس مسافة كوك	3-5-3
88	قياس الأثر على الأخطاء المعيارية - مقياس COVRATIO	4-5-3
89	بعض الحلول المقترحة لمعالجة القيم الشاذة	6-3
الفصل الرابع: الجانب التطبيقي		
91	تمهيد	1-4
91	تعريف متغيرات الدراسة	2-4
92	وصف متغيرات الدراسة	3-4
93	تحليل الانحدار الخطي المتعدد قبل معالجة القيم الشاذة	4-4
93	تقدير معلمات النموذج	1-4-4
93	اختبار المعنوية الكلية للنموذج	2-4-4
94	اختبار معنوية معاملات الانحدار الجزئية $\hat{\beta}_j$	3-4-4
95	اختبار القدرة التفسيرية للنموذج	4-4-4
96	اختبار النموذج من الناحية القياسية	5-4-4
100	اختبار النموذج لعملية التنبؤ	6-4-4
103	تشخيص القيم الشاذة في المتغيرات المستقلة والمتغير التابع	7-4-4
109	قياس أثر المشاهدات الشاذة المؤثرة علي معلمات النموذج	8-4-4
115	معالجة القيم الشاذة للنموذج المقدر	5-4
115	معالجة القيم الشاذة بأسلوب الحذف	1-5-4
120	معالجة القيم الشاذة بطريقة متوسط البتر	2-5-4
123	المقارنة بين نتائج النماذج المقدره لاختيار أفضل نموذج	6-4
124	تشخيص أفضل نموذج مقدر	7-4
الفصل الخامس: النتائج والتوصيات		
136	النتائج	1-5
138	التوصيات	2-5
139	المراجع	
145	الملاحق	

فهرس الجداول

الرقم	اسم الجدول	الصفحة
1-4	المقاييس الوصفية التي تصف طبيعة بيانات الدراسة	92
2-4	معلمات النموذج و قيم الخطأ القياسي للنموذج المقدر قبل معالجة القيم الشاذة	93
3-4	تحليل التباين للنموذج المقدر قبل معالجة القيم الشاذة	94
4-4	نتائج اختبار t للنموذج المقدر قبل معالجة القيم الشاذة	95
5-4	قيم معاملات التحديد وإحصائية D-W للنموذج المقدر قبل معالجة القيم الشاذة	95
6-4	اختبارات الكشف عن الإرتباط الخطي المتعدد بين المتغيرات المستقلة للنموذج المقدر قبل معالجة القيم الشاذة	97
7-4	اختبار جاك- بيررا للتوزيع الطبيعي للنموذج المقدر قبل معالجة القيم الشاذة	99
8-4	اختبار كلومجروف- سيمرنوف & اختبار ويلك شابيرو للتوزيع الطبيعي للنموذج المقدر قبل معالجة القيم الشاذة	100
9-4	اختبار دالة الإرتباط الذاتي والجزئي للنموذج المقدر قبل معالجة القيم الشاذة	101
10-4	اختبارات تقييم القدرة التنبؤية للنموذج المقدر قبل معالجة القيم الشاذة	102
11-4	قيم البواقي، قيم الارتفاعات، قيم معايير تشخيص المشاهدات للنموذج المقدر الشاذة قبل معالجة القيم الشاذة	106
12-4	قيم البواقي ، البواقي المحذوفة ، بواقي ستودنت المحذوفة للنموذج قبل معالجة القيم الشاذة	108
13-4	قيم $DFFITs$ ، قيم $COVRATIO$ ، قيم مسافة كوك D_i للنموذج المقدر قبل معالجة القيم الشاذة	110
14-4	قيم معاملات الانحدار حسب المقياس $DFBETAS$ قبل معالجة القيم	112

	الشاذة	
113	ملخص المشاهدات الشاذة والمؤثرة علي النموذج المقدر قبل معالجة القيم الشاذة	15-4
118	المفاضلة بين النماذج المقدره بعد حذف المشاهدات الشاذة	16-4
122	المفاضلة بين النماذج المقدره بعد معالجة المشاهدات الشاذة بطريقة متوسط البتر	17-4
123	المفاضلة بين النماذج المقدره لاختيار أفضل نموذج:	18-4
124	يوضح نتائج اختبار t لأفضل نموذج مقدر	19-4
125	اختبارات الكشف عن الإرتباط الخطي المتعدد بين المتغيرات المستقلة لأفضل نموذج مقدر	20-4
127	اختبار جاك بيرا للتوزيع الطبيعي لأفضل نموذج مقدر	21-4
128	اختبار دالة الارتباط الذاتي والجزئي لأفضل نموذج مقدر	22-4
129	اختبارات تقييم القدرة التنبؤية لأفضل نموذج مقدر	23-4
131	ملخص المشاهدات الشاذة والمؤثرة لأفضل نموذج مقدر	24-4
134	القيم المقدره لقيم المتغير التابع خلال فترة الدراسة	25-4

فهرس الأشكال

الصفحة	اسم الشكل	الرقم
67	مناطق القبول والرفض لاختبار درين- واتسون	1-2
76	مخطط الرسم الصندوقي	1-3
98	مناطق القبول والرفض لاختبار درين- واتسون للنموذج المقدر قبل معالجة القيم الشاذة	1-4
103	مقارنة منحنى القيم الفعلية مع منحنى القيم المقدرة للنموذج المقدر قبل معالجة القيم الشاذة	2-4
104	الرسم الصندوقي لقيم المتغير التابع y	3-4
104	الرسم الصندوقي لقيم المتغير المستقل x_1	4-4
105	الرسم الصندوقي لقيم المتغير المستقل X_2	5-4
105	الرسم الصندوقي لقيم المتغير المستقل X_3	6-4
114	انتشار المشاهدات علي خط الانحدار المقدر قبل معالجة القيم الشاذة	7-4
126	مناطق القبول والرفض لاختبار درين- واتسون لأفضل نموذج مقدر	8-4
130	مقارنة منحنى القيم الفعلية مع منحنى القيم المقدرة لأفضل نموذج مقدر	9-4
131	انتشار المشاهدات علي خط الانحدار المقدر لأفضل نموذج	10-4