وزارة التعليم العالي والبحث العلمي
جامعة القادسية
كلية الادارة والاقتصاد

اختبار حصانة معلمات الانحدار البوتسترابي الخطي المقدرة بوجود القيم الشاذة مع التطبيق

رسالة قدمها الطالب

علي كويش هارون حسن

الى

مجلس كلية الادارة والاقتصاد في جامعة القادسية وهي جزء من متطلبات نيل شهادة الماجستير في علوم الاحصاء

بإشراف

أ . م . د حسن سامي عريبي الشمري

1443هـ 2022م

إقرار الخبير اللغوي

اشهد أن اعداد الرسالة الموسومة ((اختبار حصانة معلمات الانحدار البوتسترابي الخطي المقدرة بوجود القيم الشاذة مع التطبيق)) قد جرت مراجعتها من الناحية اللغوية من قبلي وأصبحت خالية من الأخطاء اللغوية ولأجله وقعت .

التوقيع :

م.د جعفر طالب كريم

جامعة القادسية / كلية الإدارة والاقتصاد

اقرار المشرف

اشهد أن اعداد الرسالة الموسومة **(اختبار حصانة معلمات الانحدار البوتسترابي الخطي المقدرة بوجود القيم الشاذة مع التطبيق)** الذي قدمها الطالب **(علي كويش هارون حسن)** قد جرت تحت اشرافي في قسم الاحصاء كلية الادارة والاقتصاد / جامعة القادسية وهي جزء من متطلبات نيل درجة الماجستير في الاحصاء

التوقيع :

المشرف : أ . م . د حسن سامي عريبي

كلية الادارة والاقتصاد / جامعة القادسية

التاريخ / / 2022

توصية السيد رئيس قسم الاحصاء

بناًء على اقرار الاستاذ المشرف , ارشح هذه الرسالة للمناقشة

التوقيع :

رئيس قسم الاحصاء : أ. م .د مهند فائز كاظم السعدون

كلية الادارة والاقتصاد / جامعة القادسية

التاريخ / /2022

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

﴿رَبِّ اغْفِرْ لِي وَلِوَالِدَيَّ وَلِمَن دَخَلَ بَيْتِيَ مُؤْمِنًا وَلِلْمُؤْمِنِينَ وَالْمُؤْمِنَاتِ وَلا تَزِدِ الظَّالِمِينَ إِلاَّ تَبَارًا﴾

صدق الله العظيم

سورة نوح /الآية 28

الاهـــــــداء

الى ....

من طاعتي لهما وبري بهما اقر لعيني من رقدة الوسنان واثلج لصدري من شربة الضمآن ..... والديّ اطال الله عمرهما

الى ....

من كان لتشجيعهم أثر في مواصلة مسيرة حياتي العلمية اخواني واخواتي .....حباً واعتزازاً

الى ....

من عاشا معي بداية هذا الدرب بما فيه من مشاق لعلي اكون قد حققت لهما املاً طالما انتظروه ..... زوجتي واطفالي

علي

الشكـــر والتقديـــر

بسم الله الرحمن الرحيم

الحمد لله تعالى على ما أنعم والشكر له على ما الهم والثناء لله على ما قدم، والصلاة والسلام على خير خلقه محمد الامين (صلى الله عليه وآله وسلم) وعلى آله الطيبين الطاهرين ،وصحبه الغر الميامين .....وبعد.

اما وقد يسر لي الله اعداد هذه الرسالة ، ارى من واجب الاعتراف بالفضل والجميل ، ان اتقدم بالشكر والتقدير وعظيم الامتنان الى استاذي الفاضل الدكتور حسن سامي الشمري لما قدمه من جهود ومتابعة في تسهيل مهمتي ورعايته الصادقة التي ابداها لي طوال مدة اشرافه فجزاه الله عني كل خير.

واقدم شكري وامتناني لاعضاء الهيئة التدريسية في قسم الاحصاء لما ابدوه من عون ومشورة علمية لاتمام متطلبات البحث ولاسيما اساتذتي في مرحلة الكورسات في الدراسات العليا والذين كانوا ومازالوا خير وانفع الاساتذة ولم يبخلوا بعلمهم فكانوا كالشموع تنير جهدي المتواضع.

ويسعدني ان اقدم احترامي وتقديري للأساتذة رئيس لجنة المناقشة واعضائها المحترميون لما ابدوه من ملاحظات وارشادات علمية اغنت البحث.

ويقتضي واجب الاعتراف بالجميل والاعتزاز ان اتقدم بوافر الشكر والتقدير لزملائي في مرحلة الماجستير حيث كنا ومازلنا سنداً ودعماً لبعضنا فجزاهم الله خير الجزاء.

ومن الله التوفيق والسداد....

الباحث

فهـــــــــــــرس الموضـــــــــــــوع

ت	فهرس الموضـــــــــوع	رقم الصفحة
1.1	المقدمة	4-3
2.1	مشكلة الرسالة	5-4
3.1	اهداف الرسالة	5
	الفصل الثاني - المبحث الاول - عرض المفاهيم النظرية	34-6
1.1.2	تحليل الانحدار	7-6
2.1.2	مشاكل عدم تحقيق الفرضيات المتعلقة بالأخطاء العشوائية	7
3.1.2	مشكلة عدم تجانس التباين	8-7
4.1.2	اختبار فرضيات معاملات الانحدار	9-8
5.1.2	الحصانة	10
6.1.2	معايير مقدرات الحصانة	11-10
7.1.2	القيم الشاذة	11
8.1.2	أسباب ظهور القيم الشاذة	12
9.1.2	أنواع القيم الشاذة في الانحدار	14-12
10.1.2	طرق الكشف عن نقاط الجذب	15-14
11.1.2	طريقة (MCD)	17-16
12.1.2	طرق التقدير الحصينة	17
13.1.2	طريقة التقدير (M)	19-18
14.1.2	طريقة التقدير (MM)	20-19
15.1.2	طريقة (LMS)	20
16.1.2	طريقة (LTS)	21-20
17.1.2	أسلوب البوتستراب	22-21
18.1.2	تقنيات البوتستراب للانحدار الخطي	22
19.1.2	البوتستراب المعتمد على البواقي	23-22
20.1.2	البوتستراب المعتمد على المشاهدات	24

21.1.2	طرق البوتستراب الحصينة	27-24
1.2.2	الفصل الثاني - المبحث الثاني – الدراسات السابقة والطرق المقترحة	30-28
2.2.2	طريقة تقدير M المرجحه	32-30
3.2.2	طريقة البوتستراب المرجحه احتماليا	34-32
	الفصل الثالث (التطبيق العملي والمحاكاة)	63-35
1.3	بيانات القيمة السوقية للمصارف في سوق العراق للأوراق المالية	36-35
2.3	تطبيق طريقة (WM – Huber) على بيانات القيمة السوقية	38-37
3.3	تطبيق طريقة (WFBP-WM) على بيانات القيمة السوقية	40-38
1.4.3	دراسة المحاكاة لطريقة WM-Huber	51-40
2.4.3	دراسة المحاكاة لطريقة WFBP.WM	63-51
	الفصل الرابع	
1.4	الاستنتاجات	64
1.4	التوصيات والدراسات المستقبلية	64

جـــــــدول المختصـــــــرات

الاسم بالغة الانكليزية	المختصر
Bad Leverege Point	BLP
Bootstrap	Boot
Breakdown point	BP
Good Leverege Point	GLP
Hat Matrix	Hat
High Leverage Point	HLP
Least Absolute Deviation	LAD
Least median squares	LMS
Least trimmed squared	LTS
Leverege Point	LP
Mahanalobis Distance	MD
maximum likelihood	ML
Mean Square Error	MSE
Minimum covariance determinant	MCD
Ordenary Least Squares	OLS
Robust	Rob
Robust Bootstap Least trimmed squared	RBLTS
Robust Mahanalobis Distance	RMD
Weighted M-estimate	WM
Ordenary Least Squares Bootstrap	BOLS
Weighted Bootstrap probability	WBP
Iteratively reweighted least squares	IRLS
Least Absolute Value	LAV
Median absolute deviation	MAD
Re-weighted least squares	RLS

فهـــــــــــــرس الجـــــــداول

ت	فهرس الجـــــــداول	رقم الصفحة
(3-1)	تقدير طريقة M-Huber للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية	37
(3-2)	تقديرات WM-Huber للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية	38
(3-3)	تقدير طريقة WFBP.M للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية	39
(3-4)	تقدير طريقة WFBP.WM للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية	40-39
(3-5)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 45 \ , \alpha = 0.05$	41
(3-6)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 45 \ , \alpha = 0.05$	42-41
(3-7)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 65 \ , \alpha = 0.05$	42
(3-8)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 65 \ , \alpha = 0.05$	43
(3-9)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 85 \ , \alpha = 0.05$	44-43
(3-10)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 85 \ , \alpha = 0.05$	44
(3-11)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 100 \ , \alpha = 0.05$	45
(3-12)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 100 \ , \alpha = 0.05$	46-45
(3-13)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 45 \ , \alpha = 0.10$	46
(3-14)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 45 \ , \alpha = 0.10$	47
(3-15)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 65 \ , \alpha = 0.10$	48-47
(3-16)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 65 \ , \alpha = 0.10$	48
(3-17)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 85 \ , \alpha = 0.10$	49
(3-18)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 85 \ , \alpha = 0.10$	50-49
(3-19)	نتيجة المحاكاة لطريقة M-Huber حيث $n = 100 \ , \alpha = 0.10$	50
(3-20)	نتيجة المحاكاة لطريقة WM-Huber حيث $n = 100 \ , \alpha = 0.10$	51
(3-21)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 45 \ , \alpha = 0.05$	53-52
(3-22)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 45 \ , \alpha = 0.05$	53
(3-23)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 65 \ , \alpha = 0.05$	54
(3-24)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 65 \ , \alpha = 0.05$	55-54
(3-25)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 85 \ , \alpha = 0.05$	55

(3-26)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 85$, $\alpha = 0.05$	56
(3-27)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 100$, $\alpha = 0.05$	57-56
(3-28)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 100$, $\alpha = 0.05$	57
(3-29)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 45$, $\alpha = 0.10$	58
(3-30)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 45$, $\alpha = 0.10$	59-58
(3-31)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 65$, $\alpha = 0.10$	59
(3-32)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 65$, $\alpha = 0.10$	60-59
(3-33)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 85$, $\alpha = 0.10$	60
(3-34)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 85$, $\alpha = 0.10$	61
(3-35)	نتيجة المحاكاة لطريقة WFBP.M حيث $n = 100$, $\alpha = 0.10$	62-61
(3-36)	نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 100$, $\alpha = 0.10$	62

فهرست الاشكال

ت	الاسم الشكل	رقم الصفحة
1	القيمة الشاذة عموديا	13
2	نقطة قوة رافعة رديئه	13
3	نقطة قوة رافعة جيدة	14
4	العلاقة بين بواقي الانحدار والقيم المقدرة للقيمة السوقية باستخدام طريقة المربعات الصغرى.	36
5	العلاقة بين القيم المقدرة والقيمة السوقية باستخدام طريقة المربعات الصغرى	36
6	المنحني والمدرج التكراري لتوزيع بواقي الانحدار لطريقة المربعات الصغرى	36

المستخلص

عندما تحتوي بيانات الانحدار على قيمة شاذة ، فان هناك مشكلة حقيقية قد تؤدي إلى انهيار مقدرات طريقة المربعات الصغرى ، وعندها ينبغي التوصية بطرائق الانحدار الحصينة. اذا تعد طريقةM−Huber من الطرائق الحصينة التي لها شهرة واسعة اذ تملك اعلى نقطة انهيار لمقاومة القيم الشاذة تصل الى (0.50) ، ولكنها تفقد هذه المزية عند وجود نقطة رافعة واحدة (leverage point) في مصفوفة المتغيرات التوضيحيةX . بناء على ما تقدم فان هذه الرسالة تهدف أولا الى اقتراح ترجيح طريقة M−Huber عن طريق تخصيص وزن محدد لكل نقطة رافعة. لقد تم اختبار أداء الطريقة المقترحة بالمقارنة مع طريقة M−Huber الأصلية باستخدام بيانات حقيقية وعدد من دراسة المحاكاة. لقد بينت النتائج أن طريقة M−Huber المرجحة أكثر كفاءة وموثوقية من M−Huber

ان طريقة الـ Bootstrap هي إحدى طرائق سحب العينات العشوائية مع الارجاع ، والتي تم اقتراحها لمعالجة مشكلة العينات الصغيرة التي يصعب اشتقاق توزيعاتها، بحيث يكون توزيع عينات البوتستراب تجريبيًا أو حراً. قد تسبب عملية السحب العشوائي للمشاهدات مع الارجاع التي تعطي فرصا متساوية للظهر في عينة البوتستراب الى احتمال إعادة معاينة القيم الشاذة الى نسبة قد تصل او تتجاوز الـ 50% الامر الذي يصعب ايجاد طريقة حصينة لها القدرة على مقاومة مثل هكذا نسبة من للقيم الشاذة بما فيها طريقة M−Huber .

المتعارف عليه في الادبيات الإحصائية أن طريقة البوتستراب عند ثبات المصفوفة X تتم عند إعادة أخذ عينات مع ارجاع من بواقي الانحدار التي قد تحتوي على قيم شاذة و عليه قد تحدث المشكلة المذكورة انفا. علاوة على ذلك ، فإن نقطة (نقاط) الرافعة هي قيم شاذة تظهر في بيانات مصفوفة المتغيرات التوضيحية X ، وبالتالي فإن آثارها على عينات هذا النوع من البوتستراب سيكون موجوداً أيضا لان ثبات هذه المصفوفة سيجعل من شواذها موجودة بنسبة ثابتة في كل عينات البوتستراب. إذ إن اتخاذ القرار بشأن الفرضية الصفرية لمعلمات الانحدار البوتستراپي لا يمكن الاعتماد عليه.

لذلك تذهب هذه الرسالة في هدفها الثاني الى اقتراح البوتستراب المرجح x–bootstrap احتماليا لضمان أن تكون النسبة المئوية للقيم الشاذة في عينات البوتستراب منخفضة للغاية. هذه النسب المنخفضة للقيم الشاذة مع ثبات نسبة النقاط الرافعة (High Leverge point)تتم معالجتها من خلال دمج مقترحنا الأول مع مقترح ترجيح عينات البوتستراب احتماليا و من اتخاذ قرار أكثر موثوقية بشأن اختبار فرضية معاملات الانحدار البوتسترابي الحصين. لقد تم اختبار أداء الطريقة المقترحة الثانية مع طرق أخرى نظيرة باستخدام البيانات الحقيقية والمحاكاة و قد أظهرت النتائج أن طريقتنا المقترحة أكثر كفاءة وموثوقية من الطرق الاخرى.

الــفصــل الاول

1.1.المقدمة

لقد أظهرت مدرسة الحصانة ان هناك تأثيراً واضحاً للقيم الشاذة على التقديرات الإحصائية. اذ عرفت القيمة الشاذة بأنها قيمة تتموضع بعيدا عن مركز تجمع البيانات و لا تتبع التوزيع العشوائي لمعظم البيانات (انظر Hampel,1986، Huber,198 ، Rousseuw and Zomeren,1990). هناك نوعان من القيم الشاذة يظهر في نماذج الانحدار فتلك التي تظهر في بواقي الانحدار سميت بالقيم الشاذة اما تلك التي تظهر في المتغيرات المستقلة فسميت بالـ (Leverage Points) (LP). و حيث أنموذج الانحدار يعتمد في تقدير معالمه على طريقة المربعات الصغرى التي تعد حساسة جدا لظهور مثل هكذا قيم شاذة بأنواعها فانه لا يمكن الاعتماد عليها اطلاقا ولمزيد من المعلومات انظر(Rousseeow and Leroy 1987), (Gujarati 2003).

طريقة المربعات الصغرى الاعتيادية (OLS) تعد طريقة مثالية عند تحقق فروضها كـ $\varepsilon_i \sim N(0,\sigma^2)$ لكن وجود القيم الشاذة في بواقي الانحدار سوف يخرق فرضية التوزيع الطبيعي للأخطاء العشوائية بمتوسط صفر وتباين ثابت. قد تسبب القيم الشاذة تغير قيمة المتوسط او انها قد تحدث تضخما في حجم التباين او كليهما الامر الذي يجعل نقطة انهيار هذا المقدر قريبة من الصفر. اقترح Huber and Ronchetti (1981) طريقة M−Huber لمعالجة مشكلة القيم الشاذة في بواقي الانحدار لزيادة قيمة نقطة انهيار(Breakdown point) الى ما يقارب الـ 50%. اختبار المعلمات المقدرة لـ M−Huber يتطلب اختبار كل معلمة على حدة للتاكد من تحقق الفرض العدمي لها من عدمه، بكلام اخر للتاكد من ان معلمة مجتمع للمتغير المعني مساوية بالصفر ام لا. لذلك يتطلب ذلك الحصول على قيمة t و الانحراف المعياري للمعلمة المقدرة. فاذا كانت المعلمة حصينة و الانحراف المعياري لها حصينا أيضا فان الاختبار سيقلل من حدوث مشكلة الخطأ من النوع الأول او الخطأ من النوع الثاني. اذا الفيصل هنا هو الحصانة، لكن نقطة انهيار M−Huber قد تكون قريبة من الصفر اذا ظهرت نقطة رافعة واحدة على الأقل او نقطة رافعة

عالية في مصفوفة المتغيرات المستقلة. العديد من الدراسات تناولت ترجيح مصفوفة المتغيرات المستقلة للتقليل من تأُثير النقاط الرافعة العالية و الاعتيادية، مثل

. Ellis and Morgenthaler (1992), Hubert and Rousseeuw (1997),Giloni et al. (2006a), Giloni et al. (2006b), and Arslan (2012).

هذه الدراسات اعتمدت بشكل كلي على قدرة مسافة (Mahanalobis Distances) الحصينة لتشخيص نقاط الرفع في المصفوفة (X) عند وجود نقطة رفع عالية اذ ان نقاط الرفع العالية قد تكون مصدرا لحدوث مشكلة (Masking and Swamping). اظهرت نتائج هذه الدراسات ان مسافة (Mahanalobis Distances) التقليدية تتأثر بوجود الظاهرتين آنفتي الذكر (Masking and Swamping) بحيث انها لم تتمكن من التشخيص الصحيح للعدد الكامل لنقاط الرفع تبعا لطبيعة البيانات.

في بعض الأحيان، يلجأ الباحثون الى استخدام طريقة البوتستراب للحصول على الانحراف المعياري لمعلمة الانحدار المقدرة، ذلك لان البوتستراب يعتمد على إعادة المعاينة مع الارجاع من خلال توزيع تجريبي او حر ولذلك هو لا يخضع لاي من الفرضيات الإحصائية. إذ ان هناك نوعين من طرائق البوتستراب، الأولى تعتمد على المشاهدات والأخرى تعتمد على بواقي الانحدار والأخيرة أسرع لكنهما لا تضمنان عدم تجاوز نسبة القيم الشاذة الـ 50% بالإضافة الى ان بواقي البوتستراب قد تكون غير متجانسة الى حد ما.

2.1. مشكلة الرسالة

بناء على ما تقدم يمكن تلخيص مشكلة الرسالة بالتساؤلات الاتية :

1. هل ان وجود (HLP) واحدة تتزامن مع وجود شواذ في بواقي الانحدار و (LP) في مصفوفة المتغيرات المستقلة يؤثر على مقدر M–Huber ؟
2. هل ان ترجيح مصفوفة المتغيرات المستقلة بأحدى دوال الترجيح المعروفة يقلل من تأثير كافة أنواع القيم الشاذة على مقدر M–Huber؟

3. هل ان وجود القيم الشاذة بكافة أنواعها ستؤثر في اختبار فرضيات المعلمات المقدرة للانحدار البوتسترابي المعتمد على للبواقي المعتمد تقديراتM−Huber ؟
4. هل ترجيح M−Huber سيحسن من اختبار الفرضيات حول المعلمات المقدرة للانحدار البوتسترابي المعتمد على البواقي؟

3.1.اهداف الرسالة

تهدف هذه الرسالة الى:

1. استخدام دالة الترجيح (W) المقترحة من Giloni et al. (2006a) لتقليل تأثير نقاط الرفع المكتشفة بواسطة مسافة مهالانوبيز (MD)(Mahanalobis Distances) الحصينة والمعتمدة على مصفوفة الموقع والقياس (MCD) , وذلك بهدف ترجيح (WM − Huber)

2. اقتراح طريقة حصينة الاختبار فرضيات معلمات الانحدار البوستراپي (Bootstrap) المقدرة و المعتمدة على البواقي والتي تضمن عدم وصول القيم الشاذة الى نسبة اكثر (50%).

هذه الرسالة قسمت على أربع فصول, تتضمن الفصل الاول المقدمة ومشكلة الرسالة واهداف الرسالة, اما الفصل الثاني فكان على مبحثين, الاول تضمن عرض المفاهيم النظرية بالإضافة الى طرق التقدير الحصينة لمعلمات الانحدار, اما المبحث الثاني فتناول الدراسات السابقة والطرق المقترحة (WM − Huber) , (WFBP. WM) , اما الفصل فقد الثالث تضمن دراسة المحاكاة الطرق المقترحة باستخدام بيانات القيمة السوقية لسوق العراق للأوراق المالية, واما الفصل الرابع فقد ذكر فيه الاستنتاجات والتوصيات التي توصل اليها الباحث.

الـفصـــل الثـــانـــي

المبحث الاول - المفاهيم النظرية

1.1.2.تحليل الانحدار : Regression Analysis:-

يصف تحليل الانحدار العلاقة ما بين متغير معتمد واحد و متغير واحد او عدة متغيرات مستقله بهدف التعرف على درجة العلاقة بين هذه المتغيرات ، ذلك لبناء نموذج رياضي للتنبؤ ويكتب بالصيغة الاتية (الراوي،1987) :

$$y = X\beta + \varepsilon \qquad (2.1)$$

إذ ان

y : متجه$n \times 1$ من الاستجابات المشاهدة

X : مصفوفه ذات بعد $n \times p$من القيم الثابتة للمتغيرات المستقلة

β : متجه $p \times 1$ من معلمات الانحدار المراد تقديرها غير المعلومة

ε : متجه الخطأ العشوائي يتبع التوزيع الطبيعي بمتوسط صفر وتباين ثابت

غالبا ما يمكن تقدير (β) اما بطريقة الامكان الاعظم او طريقة المربعات الصغرى. تعد طريقة المربعات الصغرى من اكثر الطرق شيوعا بسبب سهولة احتسابها. للحصول على تقديرات OLS يتم عن طريق تقليل مجموع مربعات الأخطاء الى اقل ما يمكن وفقا لعدد من الفروض التي عند تحققها تعطي الطريقة افضل التقديرات الخطية غير المتحيزة ومن بين تلك الفروض ان e هو خطأ عشوائي يتبع التوزيع الطبيعي بمتوسط حسابي $E(e) = 0$ وتباين ثابت σ^2 كذالك ان قيم الخطأ العشوائيe مستقلة عن بعضها بعض لذلك فان التباين المشترك بينهما يساوي صفراً أي :

$$Cov(e_i, e_i') = 0 \qquad (2.2)$$

اما طريقة الامكان الاعظم فهي طريقة لا يجاد تقديرات المعالم التي تجعل دالة الامكان الاعظم في نهايتها العظمى، ان كلا الطريقتين تعطي مقدرات غير متحيزة لكن الفرق هو ان تباين الامكان الاعظم يكون أكبر من تباين المربعات الصغرى. (الراوي، خاشع محمود (1987))

2 .2.1. مشاكل عدم تحقق الفرضيات المتعلقة بالأخطاء العشوائية:-

مما سبق نلاحظ أن طريقة المربعات الصغرى وطريقة الإمكان الأعظم تعتمد على فروض اساسية وإن عدم تحقق بعضها يؤدي الى صعوبة استعمال هذه الطرائق لتقدير معلمات النماذج الخطية أو الوصول الى تقديرات خاطئة وبالتالي الى استنتاجات خاطئة حول المعلمات.

3.1.2 تجانس التباين Heteroscedasticity

احدى الفرضيات الأساسية التي يعتمد عليها في تقدير معلمات الانموذج

$$var(e_i) = E\left(e_i^2\right) = \sigma_\varepsilon^2 \quad ,\forall_i, i = 1,2,\dots,n \qquad (2.3)$$

حيث ان ε تمثل الاخطاء

أي أن تباين الاخطاء العشوائية تساوي كمية ثابتة (σ^2)وتعرف هذه الفرضية بفرضية عدم تجانس تباين الخطأ . ولكن في معظم الدراسات القياسية لاسيما التي تعتمد على البيانات الاحصائية التي تأخذ شكـل البيانات المقطعية (Cross Section Data) فإن فرضية تجانس تباين الاخطاء لا تتحقق بسبب تشتت مشاهدات البيانات المقطعية الخاصة بالمتغير المعتمد لكل مستوى من مستويات المتغيرات التوضيحية. وبشكل عام تحدث مشكلة عدم تجانس التباين في حالة تقدير معالم النماذج المستندة الى بيانات مقطعية، كما هو الحال في بيانات بحوث الأسرة، التي تتضمن اسرا" متباينة بشكل كبير في مستويات دخولها. وكذلك البيانات الخاصة بالمصانع أو المؤسسات أو المدن أو المناطق المختلفة التي يوجد اختلاف كبير في قيم متغيراتها. وبذلك فإن فرضية التباين تأخذ الشكل الآتي: –

$$H_0: var(e_i) = E\left(e_i^2\right) = \sigma_i^2 \quad ,i = 1,2,\dots,n \qquad (2.4)$$

$$H_1: \sigma_1^2 \neq \sigma_2^2 \neq, \dots., \neq \sigma_n^2$$

لذا يجب التخلص من هذه المشكلة لكي نحصل على مقدرات كفوءة للمعلمات، وتوجد عدة اختبارات للكشف عن وجود مشكلة عدم تجانس التباين وتحديد المتغير التوضيحي الذي يسبب هذه المشكلة نذكر منها على سبيل المثال : اختبار ارتباط الرتب لسبيرمان

(Spearman Rank Correlation) وهذا الاختبار يعد من الاختبارات اللامعلمية ولذلك فهو لا يعتمد على شرط التوزيع الطبيعي للأخطاء العشوائية . وتوجد اختبارات اخرى تعتمد على شرط التوزيع الطبيعي للأخطاء العشوائية مثل (goldfeld and Quandt test) (Glejser Test) (Bartlett Test).(كاظم، اموري هادي 2002)

4.1.2 اختبار فرضيات معاملات الانحدار

تعد الفرضية الإحصائية من الركائز المهمة في الاختبار إذا تعرف على انها ادعاء او تصريح قد يكون صائب او خطأ حول معلمة المجتمع. ان قبولنا الفرضية الإحصائية هو ناتج عن عدم وجود ادلة كافية لرفضها من بيانات العينة ولذلك فان قبولنا لهذه الفرضية لا يعني بالضرورة كونها صحيحة اما إذا رفضنا الفرضية بناء على المعلومات الموجودة في بيانات العينة فان ذلك يعني بان الفرضية خاطئة، لذلك تجد الباحث يحاول دائما ان يضع الفرضية بشكل يأمل ان يرفضها.

يوجد نوعان من الفرضيات , الأولى تسمى فرضية العدم التي يرمز لها بالرمز H_0 وهي تتعلق بالمعلمة موضوع الدراسة وهي تحدد القيم التي يعتقد الباحث بانها لا تعبر عن القيمة الحقيقية للمعلمة. اما الثانية فهي البديلة ويرمز لها بالرمز H_1 وهي التي تحدد قيم المعلمة التي يعتقد الباحث بانها صحيحة (اموري هادي، 2002).

تتم كتابة صيغة الفرضية كالاتي:

فرضيه العدم $H_0: \beta = 0$

فرصيه البديلة $VS. \quad H_1: \beta \neq 0$

ولأجل اختيار هذه الفرضية عند مستوى دلاله معين نرمز له بالرمز (α) فيمكن ان نتبع اسلوب اختبار t وذلك لان توزيع المعاينه الى $\frac{\hat{\beta}-\beta}{S(\hat{\beta})}$ هو توزيع t بدرجه حريه (n−2) ، وان اجراء الاختبار هذا يمكن ان يلخص بما يأتي:

نستخرج قيمه t المحسوبة ونرمز لها بالرمز t ،حيث ان قيمة t هي القيمة الخاضعة للفرضية H_0 لذلك فان قيمه t المحسوبة ,يمكن ان تكتب كما يلي :

$$t = \frac{\hat{\beta} - \beta}{S(\hat{\beta})} \qquad (2.5)$$

نستخرج قيمة t الجدوليه عند مستوى دلالة معين مقدارة (α)، وبما ان الفرضية السابقة هي فرضية بجانبين فان قيمة t الجدولية ستكون $t_{(n-2.\frac{\alpha}{2})}$

يكون قبول او رفض الفرضية H_0 تبعا الى

1. نقبل الفرضية H_0 اذا كانت قيمه t المحسوبة تقع ضمن المجال الاتي

$$-t_{\alpha/2} \leq t \leq t_{\alpha/2}$$

2. ونرفض الفرضية H_0 اذا كانت قيمه t المحسوبة تقع ضمن احد المجالين الاتي

$$t < -t_{\alpha/2} \quad \text{او} \quad t > t_{\alpha/2}$$

هذا يعني اذا كانت قيمة t المحسوبة اكبر من t الجدولية ترفض فرضية العدم , ومما تجدر الاشارة اليه هنا انه يوجد نوعان من الاخطاء في اختبار الفرضيات وهما خطأ من النوع الاول هو رفض فرضية العدم عندما تكون صحيحة تحت احتمال مقداره α ،حُيث يكون مستوى الدلالة الاحتمالية هو (0.01,0.05). وخطأ من النوع الثاني هو قبول فرضية العدم عندما تكون غير صحيحة تحت احتمال مقداره β .

5.1.2 الحصانة Robustnees

وفقا لنظرية Gauss–Marcov فأن طريقة المربعات الصغرى الاعتيادية (OLS) هي الاجراء الامثل في ظل ان الاخطاء تتوزع توزيعا طبيعيا .لا يدرك العديد من الباحثين ان اداء OLS يمكن ان يكون ضعيفا وينشأ منه توزيع ثقيل الذيل بسبب القيم الشاذة التي تنشأ من البيانات وفقا لـ (1986) Hample واخرين فأن نسبه الاعتيادية للقيم الشاذة في البيانات هي 10–1 % وهذه النسبة كبيرة جدآ للمربعات الصغرى فقد اشار Huber الى ان وجود مشاهدة واحدة شاذة قد يهدم مزايا OLS لذلك لابد من ايجاد مفهوم مناسب لحل هذه الاشكالية الا وهي الحصانة.

يعدّ مفهوم الحصانة من المفاهيم الإحصائية التي تم طرحة من قبل (Box) سنه 1953 الذي عالج مشكله القيم الشاذة وتأثيرها على البيانات إذ يقوم بإعطاء مقدر كفوء رغم وجود القيم الشاذة. تم اقتراح العديد من الأعمال حول التقدير الحصين في الأدبيات. من بينهم Edgeworth 1887 الذي اقترح مقدر القيم المطلقة الأقل (LAV) ، وكذلك Huber 1973 الذي قدم مقدرات M, وكذلك (1987) Yahai قدم مقدر MM .

6.1.2 معايير مقدرات الحصانة

لكي تعمل هذه المقدرات بصورة صحيحة هناك بعض المعايير التي تم وضعها من قبل الباحثين Huber 1972 وكذلك Hample 1974 لعمل تلك المقدرات وهي:–

أولا:– دالة التأثير Influence Function (IF)

تم اقترحها من قبل Hample لأول مرة عام (1964) فقد عمل مع مجموعة من الباحثين على تطويرها والاستدلال بها بمجال الإحصاء الحصين الى أن تم تقديمها بصيغتها الحالية. من الضروري جدا معرفة تأثير القيم الملوثة (الشاذة) على المقدر ومدى التحيز الذي تسببه تلك القيم لذلك تم وضع هذه الدالة كما في الصيغة الاتية: –

$$IF(X:T,H) = \lim_{t\to 0} \frac{T[(1-t)H + tG] - T(H)}{t} \qquad (2.6)$$

إذ أن :-

IF: دالة التأثير

T :هي المعلمة المراد تقديرها

$(1-t)H$: داله التوزيع الملوث

$tG(.)$:التوزيع النظيف

t : نسبة التلوث

ثانيا:- نقطة الانهيار Breakdown Point (BP)

تعد نقطة الانهيار من الركائز الرئيسة في الحصانة إذ تعد نقطه الفصل بين المقدر ذي الفائدة من عديم الفائدة فكلما كانت نقطه انهيار المقدر عالية كلما كان مقاوم Resistant) Estimator) للقيم الشاذة .فقد عرف Hample (1974) نقطة الانهيار بأنها اعلى نقطة مقاومة للمقدر عندما تحتوي البيانات على نسبة عالية من الشواذ.ومن الجدير بالذكر ان نقطة انهيار مقدر المربعات الصغرى هو (0) اي انها شديده التأثر بالقيم الشاذة وهنا لابد من الإشارة الى ان نقطه الانهيار لابد من ان لا تتجاوز (50%) لكي يتم التمييز بين العينة النظيفة في التوزيع وبين العينة الملوثة (محمود مهدي حسن، 2008).

7.1.2 القيم الشاذة Concept of outlier

في هذا الجزء سوف نتطرق الى بعض التعريفات المهمة عن القيم الشاذة واسبابها وانوعها في الانحدار وتأثيرها وطرق الكشف عنها . Bross سنه (1961) فقد عرفها بأنها تلك التي تكون منحرفة عن بقية المشاهدات انحرافا كبيرا . اما AL-Jobbery سنه (1978) فقد عرفها بأنها المشاهدة التي تأتي غير منسجمة من بقية المشاهدات او هي المشاهدة التي تأتي من مجتمع مختلف عن مجتمع الدراسة . Barnett سنة (1978) فقد عرفها بأنها القيمة التي لا تبدو منطقية مع بقية القيم في مجتمع الدراسة. اما Keller سنة (2000) بأنها تلك القيم التي تقع بعيدة عن القيم الطبيعية الموجودة في معادلة الانحدار والتي سوفه تؤثر على النموذج وتقديراته .

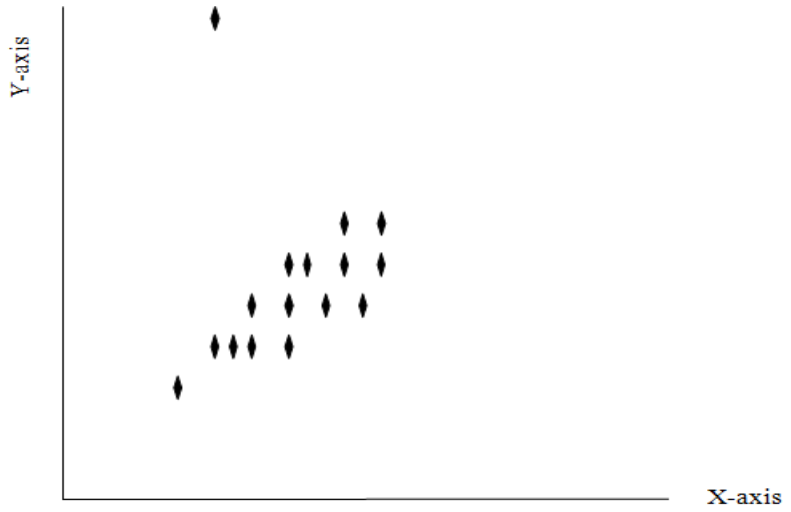
8.1.2 اسباب ظهور القيم الشاذة

ان لظهور القيم الشاذة أسباباً منها:-

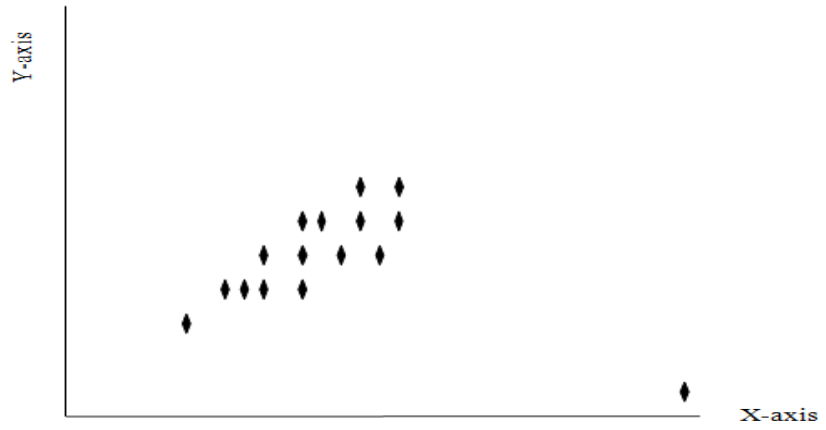
1. ان بيانات العينة تعود الى توزيعات شديده الالتواء نحو اليمين او اليسار فقد صنفت هذه التوزيعات حسب قربها من الصفر فمنها تكون شديدة التأثير بالقيم الشاذة ومنها تكون مقاومة للقيم الشاذة حسب تصنيف Greem سنه (1976)
2. تكون المشاهدات مقسمة على قسمين من التوزيعات الاول هو التوزيع الاساسي Basic Distribution الذي يعمل على تكوين مشاهدات جيدة اما التوزيع الثاني فيسمى بالتوزيع الملوث Contaminating Distribution إذ ان هذا التوزيع يكون قيما شاذة.
3. من الامور التي تسبب ظهور القيم الشاذة هي اخطاء يقع فيها الباحث الذي عنده عمل القياسات, او هناك بعض الاخطاء تحدث في الحسابات ومثل هذه الاخطاء تعمل على ظهور القيم الشاذة.

9.1.2 انواع القيم الشاذة في الانحدار

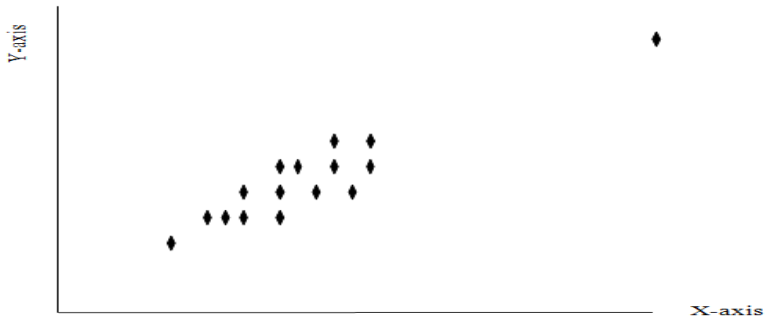
حدد Rosseeuw and Zomeren عام (1990) ان القيم الشاذة يمكن ان تحدث في ثلاثة اتجاهات مؤثرة ، قيم الشاذة عموديا (vertically outliers) ويحتوي على نقاط بيانات فيها بقايا مربعة كبيرة .القيم الشاذة افقيا (leverage points) التي توجد في المتغيرات التوضيحية X وهي نوعان نقاط جذب مؤثره (جيدة) وتسمى (Good leverage points) ونقاط جذب رديئة وتسمى (Bad leverage points) وكذلك القيم الشاذة المؤثرة في الاتجاهين X و Y وفيما يأتي رسم توضيحي لأنواع القيم الشاذة



الشكل -1- القيمة الشاذة عموديا (Y) (Uraibi, H. S.)



الشكل -2- نقطة قوة رافعة رديئة (X) (Uraibi, H. S.)



الشكل –3– نقطة قوة رافعة جيدة (X,Y) (Uraibi, H. S.)

10.1.2 طرق الكشف عن نقاط الجذب

ان لتشخيص القيم الشاذة دوراً مهماً في معرفة اين يقع تأثير تلك القيم ليتم التعامل معها ومن بين تلك الطرق هي الكشف عن نقاط الجذب (Leverage point) التي تقع في المصفوفة X ومن هذه الطرق هي :–

Hat Matrix .1

هي مصفوفة وزن تستخدم في تحديد صفوف المشاهدات التي تحتوي على القيم الشاذة ويتم الكشف عنها عن طريق العناصر القطرية h_{ii}

$$H = X(X'X)^{-1} X' \qquad (2.7)$$

عبر كل من (Iglewicz,B.y Hoaglin,D.C.1993) ان القيم التي لها نقاط جذب جيدة هي التي تكون اكبر من ($\frac{2p+2}{n}$)

إذ ان:–

p :عدد المتغيرات المستقلة

n : عدد المشاهدات الكلي

2. Mahalanobis Distance

تم اقتراحه من قبل (Mahalanobis 1936) لقياس انحرافات نقاط البيانات عن مركزها حسب الصيغة الاتية:

$$MD_i = \sqrt{(x_i - T)' C^{-1} (x_i - T)} \qquad (2.8)$$

$$i = 1,2, \dots, n$$

حيث ان C^{-1} يمثل مصفوفه التباين المشترك إذ تكمن فكرة هذا المقياس في مقارنة القيم المحسوبة لـ i مع نقاط القطع $\sqrt{\chi^2_{(0.05,p)}}$ عندما تتجاوز i قيمة القطع سوف تعتبر نقطه جذب جيدة (high leverage point) كما يمكن اعتماد قيم مصفوفة القبعة للتعبير عن MD على النحو الاتي:

$$MD_i^2 = (\mathrm{n} - 1)\left[h_{ii} - \frac{1}{n}\right] \qquad (2.9)$$

حيث تعتبر MD كدالة للعناصر القطرية $\mathrm{h_{ii}}$. تعتمد هذه الطريقة على الوسط الحسابي والتباين الذين كلاهما حساسان جداً للقيم الشاذة ولذلك لابد من تحصينها من خلال ابدال الوسط الحسابي والتباين بمعلمتي الموقع والقياس المحصنتين من أشهر البدائل المتوفرة في الادبيات الإحصائية هي مصفوفة اقل محددا للتباين المشترك (MCD)الذي يعطي مقدراً حصيناً للموقع كما سيتم شرحه لاحقا. بناء على ما تقدم يمكن اعادة كتابة مقياس مسافة (Mahalanobis) الحصينة كالاتي

$$\mathrm{RMD_i} = \sqrt{(\mathrm{x_i} - \mathrm{T_{ROb}})' \mathrm{C}^{-1}_{\mathrm{ROb}} (\mathrm{x_i} - \mathrm{T_{ROb}}),} \qquad (2.10)$$

حيث ان $\mathrm{T_{ROb}}$ *تمثل موقع حصين.* Midi el.al (2009) فضلوا مقارنة قيمة مربع المسافة MD_i^2 مع القيمة الحرجة لمربع كاي χ_p^2 لتحديد الصف الذي يحتوي على القيم الشاذة.

11.1.2 طريقة (MCD) Minimum covariance Determinan

يعد مقدر **(MCD)** الذي اقترحه (Rousseuw 1999) من المقدرات الحصينة كالموقع والقياس إذ يمتلك هذا المقدر نقطة انهيار عالية. فكرة عمل هذا المقدر هي اخذ مجموعات جزئية مثالية (h) من (n) من المشاهدات(h observation out of n) وعادة تكون (h) بين

($\frac{n}{2}+1$) و ($\frac{n+p+1}{2}$) وهي التي تعطي تباين مصفوفة العينة (Sample covariance matrix) والتي لها اصغر محدد ممكن , ويكون مقدر (MCD) للموقع والقياس هو متوسط ومصفوفة تباين تلك العينة . إذ ان ابسط طريقة حسابية للوصول الى تقريب المقدر المضبوط لـ (MCD) هو اعادة العينة بشكل عشوائي للمجموعات الفرعية من (h) بدون ارجاع لكل مجموعة العناصر , إذ يتم حساب المتوسط الحسابي ومصفوفة التباين المشتركة بالشكل الاتي :

1. يسحب مجموعة جزئية عشوائية (H_{old}) بحجم (h) من مجموعة البيانات (x) إذ ان

$$h = ([\frac{n}{2}] + 1)$$

2. حساب متجه المتوسطات $\bar{T}_{H_{OId}}$ ومصفوفة التباين والتباين المشترك $C_{H_{oId}}$ باستخدام المعادلات كما في الشكل الاتي :

$$\bar{T}_{H_{OId}} = \frac{1}{n}\sum_{i\in J} x_i$$

$$C_{H_{oId}} = \frac{1}{h-1}\sum_{i\in J}\left(x_i - \bar{T}_{H_{OId}}\right)'\left(x_i - \bar{T}_{H_{OId}}\right) \qquad (2.11)$$

إذ ان :

$C_{H_{oId}}$: مصفوفة ليست شاذة

3. اذا كان محدد المصفوفة $C_{H_{oId}}$ لا يساوي صفرا نستمر بالخوارزمية وبخلافة نتوقف .

4. واذا كان المحدد لها لا يساوي صفرا $\{\det(C_{H_{oId}}) \neq 0\}$ يتم حساب مربع مسافة مهالانوبيز (Mahanalobis Distances) على مجموعة البيانات الاصلية وباستخدام المقدرين اعلاه.

$$MD^2_{i\,oId} = \left(x_i - \bar{T}_{H_{OId}}\right)' C^{-1}_{H_{OId}}\left(x_i - \bar{T}_{H_{OId}}\right) \quad ,i = 1,2,\dots,n$$

5. ترتيب قيم مربع المسافة تصاعديا

$$MD^2_{i\,oId}\left(\pi(1)\right) \leq MD^2_{i\,oId}\left(\pi(2)\right) \leq \cdots \leq MD^2_{i\,oId}\left(\pi(n)\right), i = 1,2,\dots,n$$

إذ ان (π) تباديل

6. وهنا يتم تعريف المجموعة الجزئية جديدة H_{new} من مجموعة البيانات المشروطة وهي بالشكل الاتي :

$$H_{new} = \left\{ x_{n(1)}, x_{n(2)}, \dots, x_{n(n)} \right\}$$

وبعدها يتم حساب $\bar{T}_{H_{new}}$ و $C_{H_{new}}$ و $d^2_{new}(i)$ بنفس العمل السابق , فاذا كان $\det(C_{H_{new}}) = 0$ نعمل على اعادة الخطوات السابقة . بخلافة اذا كانت $\det(C_{H_{new}}) < \det(C_{H_{OId}})$

ونجعل $H_{OId} = H_{new}$, $\bar{T}_{H_{OId}} = \bar{T}_{H_{new}}$, $C_{H_{OId}} = C_{H_{new}}$ وبعدها نعمل على حساب $MD^2_{i\,oId}$ بنفس الاسلوب الى ان نصل الى $\det(C_{H_{new}}) = \det(C_{H_{OId}})$ وبعدها تتوقف الخوارزمية.

12.1.2 طرق التقدير الحصينة

ناقش Alma (2011) التقدير الحصين هو طريقة مهمة لتحليل البيانات الملوثة بالقيم الشاذة الهدف الرئيسي هو توفير مقدرات غير متحيزة ومرغوبة بوجود تلك القيم ، تم تصميم طرق التقدير الحصينة بحيث لا تتأثر بشكل كبير بالقيم الشاذة. سوف نتطرق لبعض طرق التقدير الاكثر شيوعا بشيء من التفصيل.

13.1.2 طريقة التقدير (M)

ناقش (2002) Fox الطريقة العامة الأكثر شيوعًا للانحدار الحصين الا وهي طريقة تقدير M–Huberالتي قدمها (1965) Huber كتعميم لتقدير دالة الإمكان الأعظم . هي تكاد تكون بنفس كفاءة OLS إذ تقوم باستبدال مجموع الأخطاء التربيعية بدالة البواقي بهدف تقليل او تصغير دالة البواقي ρ . ان دالة الهدف لـ M هي:

$$\hat{\beta}_M = min \sum_{i=1}^{n} \rho \left(\frac{e_i}{\hat{\sigma}_{MAD}} \right) \qquad 2.12$$

$where$

$$\hat{\sigma}_{MAD} = \frac{med|e_i - med(e_i)|}{0.6745}$$

حيث ان med تمثل الوسيط إذ $\hat{\sigma}_{MAD}$ وسيط الانحرافات المطلق للبواقي عن وسيطها مقسوم على الثابت 0.6745

$$w_i = \frac{\psi \left(\frac{y_i - \sum_{j=0}^{K} \acute{x}_{ij} \beta}{\hat{\sigma}_{MAD}} \right)}{\left[\frac{y_i - \sum_{j=0}^{K} \acute{x}_{ij} \beta}{\hat{\sigma}_{MAD}} \right]} \qquad (2.13)$$

ولغرض ترجيح الدالة السابقة نقوم بالإجراء الاتي إذ ان ψ هي مشتقة ρ وان $\acute{x}_{ij}$ هي المشاهدة رقم i على المتغير المستقل j .ومن ثم تطبيق طريقة المربعات الصغرى المعاد ترجيها بشكل متكرر (IRLS) . نظرًا لأن الأوزان تعتمد على المعاملات غير المعروف لـ β و σ فلا يمكننا حساب المتوسط المرجح بشكل مضبوط. لذلك يعاد ترجيح المقدرات لحين حصول التقارب. تعد طريقة الاوزان لمقدرات M هي وسيلة لبناء خوارزمية تكرارية لحساب M

1. نبدأ بالوسيط كتقدير اولي لـ β ثم نقدر $\hat{\sigma}_{MAD}$ بطريقة المربعات الصغرى
2. نحسب الاوزان لـ w_i .
3. تضرب الاوزان بالمشاهدات ويعاد حساب $\hat{\beta}$ بطريقة المربعات الصغرى
4. نحسب تقدير جديد لـ β .

5. نحسب البواقي و $\hat{\sigma}_{MAD}$ وإدخالها في معادلة (2.12) لحساب اوزان جديدة حتى تتقارب الخوارزمية. Ruckstuhl 2014

14.1.2طريقة MM

طريقة تقدير MM هي نوع خاص من طريقة تقدير M التي طورها(1987) Yohai . تهدف إلى الحصول على مقدرات ذات قيمة مقاومة عالية وأكثر كفاءة ، في ظل الاخطاء الطبيعية والتي تسمى بمقدر الانحدار MM (مقدر M المعدل)

$$\sum_{i=1}^{n} \acute{p}_1 = \left(\frac{y_i - \sum_{j=0}^{K} \acute{x}_{ij}\ \beta_j}{S_{MM}}\right) x_{ij} = 0 \qquad (2.14)$$

إذ ان S_{MM} هو الانحراف المعياري الذي تم الحصول عليه من المتبقي من طريقة التقدير M .نقطة الانهيار (BP) هي 0.5 . لايجاد التقديرات النهائية يتم حسابه كالاتي

1. يتم حساب التقديرات للمعاملات $T_\circ$ و $\theta_\circ$ التي تتمتع بنقطة انهيار تساوي 50%
2. إيجاد البواقي بالاعتماد على النتائج الخطوة الاولى وكالاتي

$$r(T_\circ, n) = y_i - T_{\circ,n}^{t} X_i \qquad 1 \le i \le n \qquad (2.15)$$

ومن ثم حساب البواقي

$$r(T_\circ, n) = S\big(r(T_\circ, n)\big)$$

باستخدام دالة ρ° التي تحقق فرضيات طريقة M−Huber بالإضافة الى استخدام الثابت b كالاتي

$$\frac{b}{a} = \frac{1}{2}$$

عندما $a = \max \rho^{\circ}(r)$

3. إيجاد مقدرات M−Huberبالاعتماد على دالة ψ

افرض ان ρ_1 دالة أخرى تحقيق فرضيات M−Huber وباستخدام الثابت b حبث ان

$$\rho_1(\mathrm{r}) \le \rho^{\circ}(r)$$

$$sup\ \rho_1(r) = sup\rho^{\circ}(r) = a$$

حيث يمكن استخدام دوال حصينة مثل Tukey للحصول على مقدرات $MM, T_{1,n}$ الذي من الممكن تعريفه اى حل ممكن للدالة الاتية :

$$\sum_{i=1}^{n} \psi_{1\left[\frac{r_i\,(t)}{sn}\right]} X_i = 0 \qquad (2.16)$$

وهذة المعادلة يجب ان تحقق

$$s(T_1, n) \leq S(T_{\circ}, n)$$

عندما

$$S(\theta) = \sum_{i=1}^{n} \rho\left(\frac{r_i(\theta)}{sn}\right) \qquad (2.17)$$

Least Median of Square Regression(LMS) 15.1.2

تعد طريقة اقل وسيط للمربعات الصغرى الذي اقترحها (Rousseuw) عام 1984 طريقة انحدار حصينة لتقدير معالم الانحدار عند وجود قيم شاذة في البيانات وذلك من خلال تخفيض وسيط مربعات البواقي بدلا من تربيعها للمجموعات الجزئية المرتبة h

$$h = \left[\frac{n}{2}\right] + [\frac{p+1}{2}]$$

حيث ان n تمثل حجم العينة وp عدد معاملات الانحدار، حيث يتم ترتيب بواقي الانحدار تصاعديا ومن ثم اختيار مجموعه من البيانات بحجم h من n من مربعات بواقي الانحدار المرتبة تصاعديا لتحقيق دالة الهدف التي ستكون معالمها حصينة

$$\beta_{min} Med(r_i^2)$$

16.1.2طريقة المربعات الصغرى المشذبة_Least Trimmed Squares

تم اقتراحها من قبل Rousseeuw عام 1987 إذ تعد من الطرق المهمة الحصينة التي تعتمد على تجاهل القيم الشاذة واعطاء نتائج مضبوطة إذ تقوم بتصغير مجموع مربعات الاخطاء بعد اتمام ترتيبها تصاعديا ويتم حسابها من خلال :

$$\min \sum_{i=1}^{h} r_i^2 \qquad (2.18)$$

حيث ان r_i^2 هي البواقي المربعة وh هي عدد المشاهدات التي يجب تحديدها من الصيغة الاتية:

$$h = [(1-\alpha) \times n] + [\alpha(p+1)] \qquad (2.19)$$

حيث ان α النسبة المئوية للقيم الشاذة او المشاهدات المشذبه, n حجم العينه ,pعدد معاملات الانحدار.ان الفكرة الاساسية من طريقة LTS هي استبعاد البواقي من المجموع وهذا ماينتج عن استبعاد القيم الشاذة للبيانات بشكل تام عن طريق قيمة h المستخرجة . حيث ذكر Rousseeuw and Leroy (2003) ان نقطة انهياره تصل الى 50% اى ان

$$BP = \frac{([(n-p)/2]+1)}{n} \qquad (2.20)$$

يكون LTS مقداراً متحيزاً وغير مرغوب فيه اذا كانت نسبه القيم الشاذة اكبر من BP .

17.1.2 أسلوب البوتستراب Bootstrap Method

تعد عملية تقدير المعلمات في نماذج الانحدار من الأسس المهمة في الاستدلال والتنبؤ بسبب عدم إمكانية معرفة القيم الحقيقية لتلك المعلمات حيث تعددت عمليات التقدير ومن اشهرها طريقة المربعات الصغرى (OLS) وطريقة الإمكان الأعظم (ML) التي تعتبر الأكثر شيوعا الا ان هذا الطرق مقيدة بافتراضات لعملها مما دفع الباحثين للبحث عن طرق اكثر مرونة وتطور.في السنوات القليلة الماضية زاد الاهتمام بدراسات اسلوب إعادة المعاينة لما لها من صفات وملاءمة النماذج الاحصائية إذ تعالج صغر حجم العينة وكذلك إيجاد اشتقاق لتوزيعاتها التي يصعب استخرجها.

قدم (Efron) عام (1979) اسلوب البوتستراب ,Bootstrap الذي يعرف بأسلوب التوزيع الحر الذي لا يخضع الى اى افتراضات حُيث يعالج صغر حجم العينة التي بصعوبة اشتقاق توزيعها ، تكمن فكرة هذا الأسلوب في انشاء عدد كبير من العينات الفرعية عن طرق سحب مع الارجاع من العينة الاصلية الصغيرة. يعد هذا الأسلوب حالة خاصه من طريقة (Monte Carlo) التي هي من طرق الاستدلال الاحصائي ، كما انها مشابه لطريقة (Jackknife) في إعادة المعاينة وتقدير تحيز الأخطاء المعيارية ومن مزايا أسلوب البوتستراب انة يستخدم في حال عدم توفر شرط الطبيعية في تقدير معاملات الانحدار كذلك يعمل هذا الأسلوب على معالجة عدم الاستقلالية في

المشاهدات التي لها دور كبير في دقة التحليل مما يعطي استنتاجات وتقديرات للخطأ في حالة البيانات غير الجيدة وبسبب عدم خضوعه للفروض لاسيما في التوزيع فان بيانات العينة المعقدة حسابيا يتم حسابها بهذا الأسلوب كالطبقية والعنقودية.

18.1.2 تقنيات البوتستراب للانحدار الخطي Bootstrap for Linear Regression

الان بعد ان تعرفنا على فكرة تقنية البوتستراب التي جاء من اجلها لابد من ربطها بموضوعنا وهو الانحدار الخطي حيث هناك طريقتان مختلفتان لإجراء البوتستراب ؛ وهي إعادة تشكيل X العشوائية وإعادة أخذ عينات X الثابتة والتي يشار إليها أيضًا باسم البوتستراب للبقايا Bootstrapping the residuals.هو الأسلوب الأكثر شيوعًا للبوتستراب في الانحدار.

19.1.2 البوتستراب المعتمد على البواقي

يتم إعطاء تقديرات OLS لـ β بواسطة $\hat{\beta} = (X'X)'X'Y$. منذ نشأته ، يعتمد bootstrap التقليدي على تقديرات OLS للحصول على بقايا البيانات الأصلية. ثم يتم الحصول على عينات bootstrap من خلال إعادة أخذ العينات المتبقية من الانحدار الأصلي. سنلخص هذه الخوارزمية على النحو الاتي

1. ملائمة OLS مع العينة الأصلية للمشاهدات للحصول عليها $\hat{\beta}_{ols}$ والقيم المناسبة $\hat{y}_i = f(X_i, \hat{\beta}_{ols})$.
2. استخرج البواقي $\varepsilon_i = y_i - \hat{y}_i$ وإعطاء احتمال $\frac{1}{n}$ لكل قيمة ε_i
3. اسحب عينة عشوائية من bootstrap مع الارجاع , أي ${\varepsilon_i}^*$ مأخود من ε_i ومرفق بها $\hat{y}_i$ للحصول على قيم البواقي ${y_i}^{*b}$ أي ان

$$ {y_i}^{*b} = f(X_i, \hat{\beta}_{ols}) + {\varepsilon_i}^{*b} \qquad (2.21) $$

4. قم بملاءمة OLS مع قيم bootstrap ${y_i}^{*b}$ على X الثابت للحصول عليها $\hat{\beta}_{ols}^{*b}$
5. كرر الخطوتين 3 و 4 للحصول على R من مرات $\hat{\beta}_{ols}^{*1}, \dots, \hat{\beta}_{ols}^{*R}$ حيث R هي مكررات bootstrap.

وفقًا Imon and Ali (2005) لا يوجد اتفاق عام بين الإحصائيين على عدد المرات R المطلوبة في البوتستراب يمكن أن يكون R صغيرًا مثل 25 لكن عند تقدير الأخطاء المعيارية يكون R عادةً في حدود 25–250. اما بالنسبة لفترات الثقة في البوتستراب ، يجب ان يكون عدد مرات R كبيرة والتي عادة ما تكون في حدود 500–10000.

أشار Kallelet (2002) إلى أنه يمكن تقدير الانحراف المعياري البوتستراب على النحو الاتي:–

$$\hat{\sigma}_{boot} = \left[\frac{1}{R-1}\sum_{r-1}^{R}(MSR^b - \mu_{boot})^2\right]^{1/2} \qquad (2.22)$$

حُيث MSR هو متوسط مربع البواقي حسب الصيغة الاتية:–

$$MSR = \sum_{i=1}^{n}(\varepsilon_i^b)^2/n \qquad (2.23)$$

كما ان البواقي يتم من خلال :–

$$\varepsilon^b = \left(y_i - \hat{y}_i^b\right) \qquad (2.24)$$

$$i = 1,2,\dots,n \qquad b = 1,2,\dots R$$

$$\mu_{boot} = \frac{1}{R}\sum_{r-1}^{R} MSR^b \qquad (2.25)$$

$$r = 1,2,\dots R$$

20.1.2 البوتستراب المعتمد على المشاهدات

ان إعادة أخذ العينات العشوائية تقدم نهجًا مختلفًا لعملية البوتستراب ، بافتراض أننا نريد أن نلائم نموذج الانحدار مع الاستجابة y والتنبؤات Xi التي تشكل عينة n من المشاهدات $\acute{z}_i = (y_i, x_i)$. فيما يأتي إجراء إعادة أخذ العينات العشوائية x

1- بيانات البوتستراب $(y_1, x_1)^*, \dots, (y_n, x_n)^*$ تؤخذ بشكل مستقل مع احتمالات متساوية $\frac{1}{n}$ من الحالات الأصلية (y_1, x_1) ,(y_n, x_n)

2- نحسب β^* لمجموعة بيانات البوتستراب $(y_1, x_1)^*, \dots, (y_n, x_n)^*$.

3- إعادة الخطوات 1,2 R من المرات للحصول على $\hat{\beta}^{*1}, \dots, \hat{\beta}^{*R}$

21.1.2 طرق البوتستراب الحصينة Robust Bootstrap technique

يمكن أن يتسبب وجود القيم الشاذة في العينة الأصلية في حدوث مشكلة خطيرة لتقنيات البوتستراب التقليدية ، لأن نسبة القيم الشاذة في عينات البوتستراب يمكن أن تكون أعلى من تلك الموجودة في مجموعة البيانات الأصلية مما يؤدي الى عدم الاستقرار العددي (Barrera and Zamar 2002).ان طبيعة عمل البوتستراب في إعادة المعاينة مع الارجاع أدت الى زيادة نقاط الانهيار (BP) لمقدر الانحدار حيث يمكن ان تزداد القيم الشاذة الى اكثر من 50% كما هو الحال بالنسبة الى المقدرات الحصينه حيث تزداد الى اكثر من 50% وعليه لايمكن تطبيق الطرق الحصينة لذلك تم في الآونة الاخيره طرح طرق من شأنها ان تعالج هذه المشكلة وهي:

أولا:- البوتستراب الديناميكي الحصين باستخدام التشذيب Dynamic Robust Bootstrap using LTS

قد تكون RBLTS بديلا جيداً لطريقة BOLS التي تعتمد على بواقي الانحدار حصرا، اذا كانت نسبة القيم الشاذة في كل عينة بوتستراب تساوي النسبة المئوية للقيم الشاذة في البيانات الاصلية لكن في الواقع العملي تكون اكبر من ذلك لذا جرت محاولات لجعل RBLTS اكثر حصانة من

خلال اجراء بعض التعديلات على عدد القيم الشاذة في البيانات الاصلية ومن ثم تحديد قيمة α التي تمثل النسبة المئوية للقيم الشاذة (BP) ، حُيث يتم حذف عينات البوتستراب ذات النسبة العالية من القيم الشاذة واستبدالها بأخرى جديدة لحين الوصول الى التكرارات المطلوبة. سميت هذه الطريقة بـ (Dynamic Robust Bootstrap using LTS) والتي تتلخص بالخطوات الأتية:-

1. يتم تحديد القيم الشاذة في البيانات بالاعتماد على طريقة (LMS) وبالتالي يتم تحديد α والتي هي نسبة القيم الشاذة في البيانات ((Rousseeuw and Leroy (2003)
2. استخدام طريقة (LTS) لتقدير معلمات النموذج $\hat{\beta}_{LTS}$ وحساب القيم $\hat{y}_i = f(x_i, \hat{\beta})$ اذ يتم تحديد القيم الشاذة كما ذكرنا في الخطوة رقم 1 .

3. الحصول على البواقي $\varepsilon_i = y_i - \hat{y}_i$ وإعطاء احتمال $\frac{1}{n}$ لكل قيمة من ε_i ونحدّد القيم المتبقية على أنها قيم شاذة إذا كانت القيمة المطلقة للبواقي أكبر من ثلاثة.
4. نسحب عينة عشوائية من n بوتستراب مع الاستبدال ، أي ε_i^* مأخوذ من ε_i ثم تضاف البواقي المسحوبة e_i^* الى الجزء الثابت من معادلة الانحدار لإيجاد قيمة

$$y_i^{*b} = f(x_i, \hat{\beta}_{LTS}) + \varepsilon_i^{*b} \qquad (2.26)$$

في هذه الخطوة، قمنا ببناء برنامج فرعي ديناميكي لاكتشاف القيم الشاذة بناءً على البواقي المعيارية. هذا البرنامج لديه القدرة على تحديد نسبة معينة من القيم الشاذة في كل عينة بوتستراب.

5. نطبق طريقة LTS مع قيم البوتستراب y_i^{*b} للحصول على $\hat{\beta}_{LTS}^{*b}$ تعتمد النسبة المئوية للقيم الشاذة التي يجب قطعها على الخطوة رقم 4.
6. كرر خطوة 3 و4 R من المرات لكي نحصل على $\hat{\beta}^{*1}, \dots, \hat{\beta}^{*R}$ حيث R عدد تكرارات البوتستراب . سيتم حذف أي عينة بوتستراب تحتوي على نسبة مئوية من القيم الشاذة أكبر من BP ولن يتم احتسابها كعينة بوتستراب وسيتم استبدالها بعينة جديدة. يتم تحديد النسبة المئوية للقيم الشاذة في كل عينة بوتستراب من الخطوة 4.

ثانيا:- البوتستراب المرجحة احتماليا (WBP)

اقترحت طريقة Bootstrap الموزونة احتماليا من قبل Ramli سنة 2008 لحل مشكلة القيم الشاذة الموجودة في عينات Bootstrap والحصول على تقديرات دقيقة دون اللجوء الى الطرق الحصينة. الفكره التي تقوم عليها هذه الطريقة هي وضع اوزان للقيم ، فالقيم الشاذة تأخذ صفر والقيم النظيفة تأخذ واحدآ. تم الاستعانة بالدراسات السابقة للوصول لنتائج هذه الدراسة فمن خلال خوارزمية (Tille,2006) Bootstrapping with Re-sampling probabilities التي تتعامل مع إعادة معاينة عينة من مجموعه عينات من خلال اعطاء احتمال (0,1) لعينات Bootstrap والخروج بعينات نظيفة

1. تشكيل مجموعة من العينات مع احتمالها $\left\{p_{(s_1)}, p_{(s_2)}, \ldots\ldots, p_{(s_j)}, \ldots\ldots, p_{(s_j)}\right\}$
2. أنشاء قيمة عشوائية مثل u تأخذ (0,1)
3. اختر عينة s_j بشرط $\sum_{i=1}^{j-1} p(s_i) \leq u \leq \sum_{i=1}^{j} p(s_i)$ لكل $1 \leq j \leq J$ واذا لم يتحقق هذا الشرط نعود الى الخطوة الثانية الى ان نحصل على مجموعة من البواقي المحذوفة. بعد ذلك نقوم بأستخراج الاحتمال للمشاهدة i^{th}

$$p_i = \frac{W_i}{\sum_{j=1}^{n} W_i} \; and \; 0 \leq p_i \leq 1 \;\; and \; i = 1,2, \ldots, n \,. \qquad (2.27)$$

ومن خلال المعادلة أعلاه يتم اختيار المشاهدة المرغوب فيها واعطاؤها أهمية أكثر من غيرها.

كما تم الاستعانة بدالة (Hampl,1971) لغرض تحديد اوزان كل مشاهدة

$$\psi_{Hampl}(\varepsilon_i) = \begin{cases} u & ,0 \leq abc(u) \leq a \\ a\, sign\,(u) & ,a \leq abc(u) \leq b \\ a\big(c - abc(u)\big)/(c-b) sign(u) & ,b \leq abc(u) \leq c \\ 0 & ,c \leq abc(u) \end{cases}$$

وبأخذ اوزان القيم الشاذة الصفر(D) نتمكن من إيجاد متجهين كالاتي

$$X = \begin{bmatrix} X_R \\ X_D \end{bmatrix}, Y = \begin{bmatrix} Y_R \\ Y_D \end{bmatrix}$$

ان خوارزمية Weighted Bootstrap with probability(WBP) تتلخص بما يأتي

1. سحب اول عينة Bootstrap عشوائيا مع الارجاع من بواقي الانحدار $\hat{\varepsilon}_i^{(-D)}$ للحصول على $\varepsilon_i^{b(-D)}$ بحجم n وبأستخدام (Tille,2006) من المعادلة اعلاه (2.27)
2. إيجاد $\hat{y}_i^{b(-D)} = f\left(x_R, \hat{\beta}^{(-D)}\right) + \hat{\varepsilon}_i^{b(-D)}$
3. حساب معاملات الانحدار للقيم للمقدرة $\hat{y}_i^{b(-D)}$ باستخدام المربعات الصغرى الاعتيادية للحصول $\hat{\beta}^{b(-D)}$
4. نكرر الخطوات (1-3) لــ R من المرات . بعد ذلك يتم استخراج معاملات الانحدار من:

$$\hat{\beta}^{(-D)}{}_{boot} = \frac{1}{R}\sum_{r=1}^{R}\hat{\beta}^{b(-D)} \qquad r = 1,2, \dots \dots R \qquad (2.28)$$

المبحث الثاني

الـــدراسات السابـقـة والطرق المقترحة

ان العديد من الباحثون اهتموا في مجال العلوم الاحصائية لمعالجة القيم الشاذة وتحصين البوتستراب (Bootstrap) إذ اقترح كل من (Barrera and Zamar 2002) تحصين معلمات الانحدار المقدرة بالبو تستراب والتي أظهرت بعض البطيء في الاحتساب وعدم الاستقرار أحيانا الا ان (Barrera 2003) اقترح خوارزمية جديدة تميزت بسرعة الاحتساب والاستقرار. اقترح ((Willems(2005) تطبيق البوتستراب مع طريقة المربعات الصغرى المشذبة (LTS) الا انه افترض عدم تجاوز نسبة القيم الشاذة في بواقي الانحدار الـ20%. (Imon and Ali 2005) قدما خوارزمية سميت بـ(Diagnostic Before Bootstrap) لحساب معالم الانحدار الخطي لمعالجة مثل هذه المشكلة، مفادها هو تشخيص القيم الشاذة وحذفها قبل اجراء اعادة المعاينة على المشاهدات المتبقية. حيث ان اجراءات تشخيص القيم الشاذة اعتمدت على طريقة (RLS) لبواقي الانحدار المقترحة من قبل (Rousseeuw and Leroy 1987) و كالاتي :

$$W_i = \begin{cases} 0, & if \quad \left|\frac{\varepsilon_i}{S}\right| > 2.5 \\ 1, & otherwise \end{cases} \qquad (2.29)$$

إذ ان :

ε_i: انها تمثل بواقي الانحدار .

S : انها تمثل تقدير قياس حصين مكافئ الى

$$S = 1.482\,(1 + 5/(n-p))\sqrt{med_i u_i^2(\hat{\theta})} \qquad (2.30)$$

إذ ان :

n : حجم العينة .

p : عدد معادلات الانحدار .

هناك مقترح اخر قدم من قبل (Ramli,2008) الا و هو تقنية (Weighted Bootstrap With probability) (WBP) المعتمدة على خوارزمية (Tille 2006) (Bootstrapping With Re-sampling probabilities). ان هذه الخوارزمية التي تقترح تخصيص احتمال لكل عينة من العينات الممكنة الظهور في عينات البوتستراب (Bootstrap) ويتم بعدها اجراء عملية المحاكاة مستندة على قيمة عشوائية تتبع توزيع منتظم. وهذه القيم العشوائية تشكل نظام تحكم في اختيار عينات أدني تحت صيغة او قاعدة محددة.

لكن الخوارزمية المقترحة عانت من بعض المشاكل بسبب اختلاف درجات الحرية وارتفاع قيمة متوسط مربعات الخطأ لها. اقترح (Uraibi 2009) تقنية اخرى تمثل خوارزمية ديناميكية حصينة حيث لها القدرة على تشخيص القيم الشاذة بداخل عينات البوتستراب (Bootstrap) من جانب وتشذيبها من جانب اخر. وان ميزه هذه التقنية رفض كل العينات المعادة التي تتضمن قيم شاذة تتجاوز الـ 50%, ولها امتياز اخر يتمثل بأنها تشذب القيم الشاذة حصرا ولا توجد هناك اية خسارة في المشاهدات الغير شاذة , وذلك من خلال استخدام تقنية المربعات الصغرى المشذبة (LST) داخل الدورات التكرارية للـ (Bootstrap) على ان نسبة التشذيب معتمدة على التشخيص الديناميكي. العيب الوحيد في هذه الخوارزمية انها تتأثر بـ (LP).

حاولت(Ramli,2008) الاستفادة أيضا من مقترح (Venables and Ripley 1999) في تشخيص القيم الشاذة من خلال البواقي بطريقة وسيط المربعات الصغرى (LMS) ومن ثم رجحت هذه البواقي بدالة (Hampel 1971) لغرض تحديد احتمال ظهور كل مشاهدة من خلال معرفة نسبة الوزن الترجيحي للمشاهدة الى مجموع كل الاوزان الترجيحية. وقد اتضح مما تقدم بانها تمكنت من تحوير الخوارزمية لـ (Tille 2006) لتخصيص احتمال مرجح للمشاهدة بدلا من العينة. وبالتأكيد ان دالة (Hampel 1971) ستخصص وزن ترجيحي مساو للصفر للقيم الشاذة وهذا يؤدي الى حرمانها من فرصة الظهور عند اعادة المعاينة مرة واحدة او اكثر. بتالي سوف

تكون عينات البوتستراب (Bootstrap) نظيفة وخالية من القيم الشاذة , ويمكنها عمل التقدير باستخدام طريقة المربعات الصغرى (OLS).

2.2.2 اقتراح طريقة تقدير M المرجحة (Weighted M−estimate)

أُقترح (Huber and Ronchetti, 1981) طريقة M−estimate وتقوم وفكرتها على اعادة ترجيح طريقة المربعات الصغرى الموزونة WLS بإحدى دوال الترجيح لغرض الحصول على معاملات انحدار حصينة .وعلى الرغم من انها طريقة حصينة في حالة وجود القيم الشاذة في متجه y او البواقي الا انها ليس لديها قدرة مقاومة نقاط الجذب (القيم الشاذة في مصفوفة x) والتي تجعل من نقطة الانهيار لمقدرها يساوي صفر (Ellis and Morgenthaler 1992) و

(Hubert and Rousseeuw 1997) و(2006a Giloni et al)و (Arslan 2012) قرروا تخفيض نقاط الجذب (Leverage point) باستخدام دوال وزن يمكن ان يحسن من نقطة الانهيار للمقدر لذلك قاموا بتحديد نقاط الجذب أولا بالاعتماد على قيم مسافة (MahalanobisDistance) الحصينة.

افترض ان مجموعة البيانات تحتوي على نقاط جذب في مصفوفة x وافترض ان توزيع ε_i ملوث بالقيم الشاذة في هذه الحالة ان طريقة المربعات الصغرى ليست هي الاختيار المثالي لتقدير معاملات الانموذج لذلك لابد من استخدام أسلوب حصين لمعالجة مثل هذه المشكلة , وهنا استخدامنا فكرة (Ellis and Morgenthaler 1992) و(Hubert and Rousseeuw 1997) و(2006a Giloni et al) و (Arslan 2012) لقليل تأثير نقاط الجذب (Leverage point) على مقدر M−Huber وحسب الخوارزمية التالية:

1. ترجيح مصفوفة التصميم X

في هذه الخطوة تم حساب MCD (Rousseeuw and Van Driessen,1999) حيث ان مقدر الموقع والقياس تم حسابهم من هذه الطريقة من اجل حساب متجه مربع مسافة (Mahalanobis Distance) RMD^2 كالاتي

$$RMD^2 = (\mathbf{x} - \hat{\mu}_{MCD})' C_{MCD}^{-1} (\mathbf{x} - \hat{\mu}_{MCD}) \qquad (2.31)$$

اذ أن i^{th} هي مشاهدة من نقاط الجذب ,عندما تكون i^{th} RMD^2 قيمة كبيرة . لذا يتم اعطاء اوزن قليلة لنقاط الجذب (LP) وعكس الشيء عندما تكون i^{th} RMD^2 قيمة صغيرة تعتبر هذه المجموعة مجموعة فرعية نظيفة (Giloni et al.,2006a)

$$\omega_i = \min\left\{1, \frac{\chi^2_{(0.05,p)}}{RMD_i^2}\right\} \qquad (2.32)$$

ويمكن كتابة المصفوفة المرجحة الجديدة كـ $x_\omega = \omega . x$ ويمكن صياغة تقديرات LS مع x_ω على النحو التالي ،

$$\hat{\beta}_\omega = (x'_\omega x_\omega)^{-1} x_\omega y \qquad (2.33)$$

$$\hat{y}_\omega = x_\omega \hat{\beta}_\omega$$

$$\hat{\varepsilon}_\omega = y - x_\omega \hat{\beta}_\omega$$

2. اعادة ترجيح المربعات الصغرى باستخدام طريقة المربعات الصغرى الموزونة تكراريا (IRLS) لمشاهدات (x_ω, y).

يتم الحصول على التقديرات $\hat{\beta}^M$ عن طريق تقليل دالة الهدف ρ والتي يمكن التعبير عنها كـ

$$\hat{\beta}^M = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{n} \rho(\hat{\varepsilon}_{\omega i}) \qquad (2.34)$$

حيث ρ هي دالة متماثلة ذات قيمة صغرى وحيدة عند الصفر. وبأخذ المشتقة الجزئية بالنسبة إلى β وجعلها تساوي صفرًا ، ينتج عن ذلك منظومة من المعادلات الطبيعية التي يمكن ان تحل مشكلة تقليل الاخطاء عن طريق مشتقة $\psi = \rho'$ سوف نحصل على

$$\sum \psi(\hat{\varepsilon}_{\omega i}) x_{\omega i} = 0 \qquad (2.35)$$

هناك عدة اختيارات لدالة ρ و ψ متاحة وسوف نختار دوال (توكي ، 1964) والتي تسمى (bisquare function)على النحو التالي ،

$$\rho(\hat{\varepsilon}_{\omega i}) = \begin{Bmatrix} \left\{1 - \left[1 - \left(\frac{\hat{\varepsilon}_{\omega i}}{k}\right)^2\right]^3\right\} & \text{if} \quad |\hat{\varepsilon}_{\omega i}| \le k \\ 1 & \text{if} \quad |\hat{\varepsilon}_{\omega i}| > k \end{Bmatrix}, \qquad k = 4.685$$

$$\psi(\hat{\varepsilon}_{\omega i}) = \hat{\varepsilon}_{\omega i}\left[1 - \left(\frac{\hat{\varepsilon}_{\omega i}}{k}\right)^2\right]^2 \qquad I(|\hat{\varepsilon}_{\omega i}| \le k),$$

$$w(\hat{\varepsilon}_{\omega i}) = \frac{\psi(\hat{\varepsilon}_{\omega i})}{\hat{\varepsilon}_{\omega i}} = \begin{Bmatrix} \left[1 - \left(\frac{\varepsilon}{k}\right)^2\right]^2 & \text{if} \quad |\hat{\varepsilon}_{\omega i}| \le k \\ 0 & \text{if} \quad |\hat{\varepsilon}_{\omega i}| > k \end{Bmatrix}$$

حيث $I(.)$ هي دالة المؤشر ، وتعرف كالتالي

$$I(\hat{\varepsilon}_{\omega i}) = \begin{cases} 1 & \text{if} \quad \hat{\varepsilon}_{\omega i} > 0 \\ 0 & \text{if} \quad \hat{\varepsilon}_{\omega i} < 0 \end{cases}$$

وبالنتيجة، يمكن كتابة معادلة التقدير كـ

$$\sum w(\hat{\varepsilon}_{\omega i}) x'_{\omega} = 0 \qquad (2.36)$$

هذه المعادلات التقديرية تقتضي تصغير $\sum w_i^2 (\hat{\varepsilon}_{\omega i})^2$ باستخدام المربعات الصغرى الموزونة تكراريا (IRLS) والتي يمكن الحصول عليها من الصيغة التالية

$$\hat{\beta}_{\omega}^{M(j)} = \left(x'_{\omega} w^{(j-1)} x_{\omega}\right)^{-1} x'_{\omega} w^{(j-1)} y \qquad (2.37)$$

في IRLS ، يتم حساب ملائمة النموذج الأولية ، ثم بعد ذلك يتم حساب المجموعة الجديدة من الأوزان بناءً على نتائج الملائمة الأولية وتعاد بشكل متكرر وتستمر حتى يحصل التقارب .

3.2.2 طريقة البوتستراب المرجحة احتمالياً Weighted Bootstrap with Probability (WFBP.WM)

تعتبر مقدرات طريقة Weighted M−estimate($\hat{\beta}_{\omega}^{M(*)}$) مقدرات حصينة ضد وجود القيم الشاذة ونقاط الجذب LP . لذاك ان بواقي طريقة WM−estimate ليست خالية من القيم الشاذة. حيث (*) تشير إلى j^{th} التكرار التي يتم فيها تحقق التقارب. وهذا بسبب ان البواقي هي الفرق بين y و $\hat{y}_{\omega}^{*} = x_{\omega}\hat{\beta}_{\omega}^{M*}$ وان الخطأ $\hat{\varepsilon}_{\omega}^{*} = y - \hat{y}_{\omega}^{*}$ ونتيجة لذلك فان أي قيمة

شاذة تظهر في البواقي$\hat{\varepsilon}_{\omega}^{*}$ يعزى سببها y تحت افتراضنا اعلاه والتي تحتوي بالأساس على قيم شاذة . ولإجراء طريقة بوتستراب خالية من القيم الشاذة، اقترحت طريقة WBP (Ramli,et.al.2009) البوتستراب الموزونة احتماليا. لذلك فإن كل i^{th} من البواقي ستحصل على احتمالية الاختيار $p_i = \frac{w_i^{(*)}}{\sum_{i=1}^{n} w_i^{(*)}}$. وفي هذه الحالة، هذه الاحتمالات ستؤثر في آلية المعاينة مع الارجاع حيث ان البواقي الشاذة ستحصل على اهمية اقل من البواقي النظيفة والغير ملوثة للبواقي (Ramli، 2009). بعبارة أخرى نحصل على $\hat{\varepsilon}_{\omega}^{(-D)}$ التي ستخضع لإجراءات البوتستراب، حيث D هو عدد البواقي التي تعطى وزن يساوي صفر وبذلك تكون احتمالية ظهورها تساوي صفر. سيسمح لنا هذا الإجراء باستخدام WLS مع عينات البوتستراب للحصول على معاملات انحدار حصين وحساب الاخطاء المعيارية. يمكن تلخيص فترات الثقة واختبار الفرضيات لخوارزمية معاملات الانحدار WBP على النحو التالي،

1. افترض ان قيم $\hat{\tau}$ المشاهد لمعاملات الانحدار الحصينة تم حسابه باستخدام مقدر MM– estimator، (Yohai,1987) كالاتي

$$\hat{\tau} = \frac{\hat{\beta}^{MM} - \beta}{SD(\hat{\beta}^{MM})} \qquad (2.38)$$

حيث β هي معلمة مجتمع الانحدار ، $\hat{\beta}^{MM}$ هي معاملات الانحدار المقدرة الحصينة , $SD(\hat{\beta}^{MM})$ الانحراف المعياري لها.

2. حساب قيم البواقي$\hat{\varepsilon}_{\omega}^{*}$ عن طريقة WM .

3– معاينة n من البواقي لعينات البوتستراب $\varepsilon_{\omega}^{(b)}$ مع احتمال من $\hat{\varepsilon}_{\omega}^{*}$ ثم قم بإرفاقه بـ $\hat{y}_{\omega}^{*}$ حيث ان $\hat{y}_{\omega}^{b} = \hat{y}_{\omega}^{*} + \varepsilon_{\omega}^{(b)}$,

4.نجري الانحدار لقيم البوتستراب لـ $\hat{y}_{\omega}^{(b)}$ على الثابت x_{ω} للحصول على $\hat{\beta}_{\omega}^{(b)} = (x_{\omega}' x_{\omega})^{-1} x_{\omega} \hat{y}_{\omega}^{(b)}$

5-يرجع السبب الى ان الاحصاءة t ، تعتبر كمية محورية تحت افتراضات التوزيع الطبيعي للبيانات. التمركز والمعيارية لصيغة اجراء البوتستراب $\hat{\beta}_{\omega}^{(b)}$ والذي سنشير لها بـ $t^{(b)}$ هو ايضا تقريبd تحت نفس الافتراضات ، حيث $t^{(b)} = \frac{\hat{\beta}_{\omega}^{(b)} - \beta}{SD(\hat{\beta}_{\omega}^{(b)})}$

6. نكرر الخطوات (3-5) R من المرات.

7. ان فترات ثقة البوتستراب المئيني يمكن استخدم في الكمية التجريبية $t^{(b)}$ لصياغة فترات الثقة لـ β حيث ان $t_{(1)}^{(b)} < t_{(2)}^{(b)} < \cdots < t_{(B)}^{(b)}$ هي ترتيب تكرارات الاحصاءه $t^{(b)}$.يمكن حساب الحد الاعلى والادنى في فترات الثقة ل β كالاتي

$$\hat{\beta}^{MM} - t_{[(B+1)\alpha/2]}^{(b)} SD(\hat{\beta}_{\omega}^{(b)}) \leq \beta \leq \hat{\beta}^{MM} + t_{[(B+1)(1-\alpha/2)]}^{(b)} SD(\hat{\beta}_{\omega}^{(b)}) \qquad (2.39)$$

حيث تشير الأقواس المربعة إلى التقريب إلى أقرب عدد صحيح.

ان اختبار $p - value$ للبوتستراب متساوي الذيل هو اختبار اخر الذي يمكن ان يؤدي الى اختبارين في الوقت نفسه ,ضد وجود القيم الموجودة في الذيل العلوي والذيل السفلي للتوزيع وكالاتي .

$$p(\hat{\tau}_j) = 2 \min\left(\frac{1}{B}\sum_{j}^{B} I\left(t_j^{(b)} \leq \hat{\tau}_j\right), \frac{1}{B}\sum_{j}^{B} I\left(t_j^{(b)} > \hat{\tau}_j\right)\right) \qquad (2.40)$$

الـفصــل الثالث

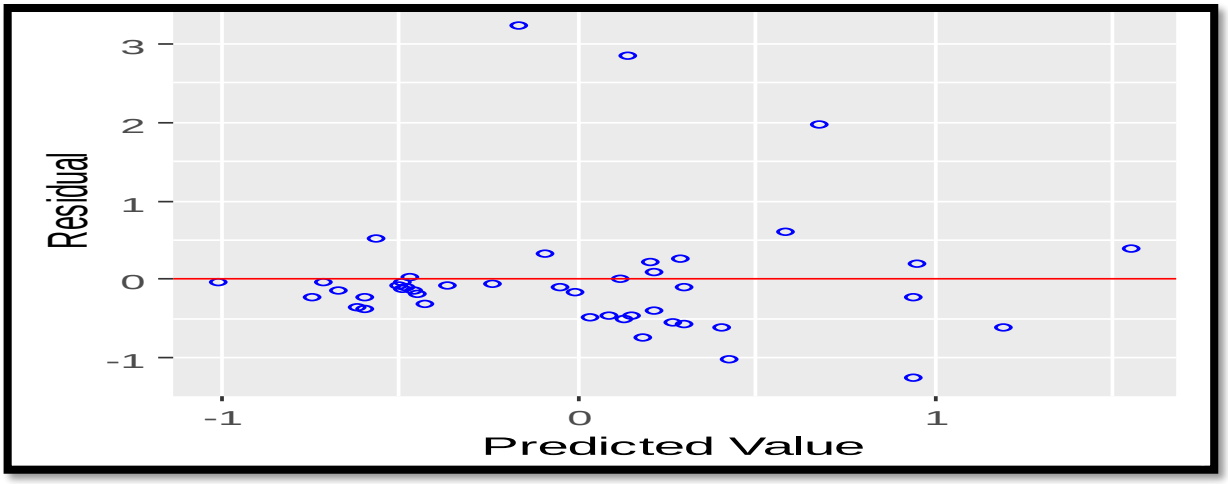
التطبيق العملي والمحاكاة

1.3. بيانات القيمة السوقية للمصارف في سوق العراق للأوراق المالية:-

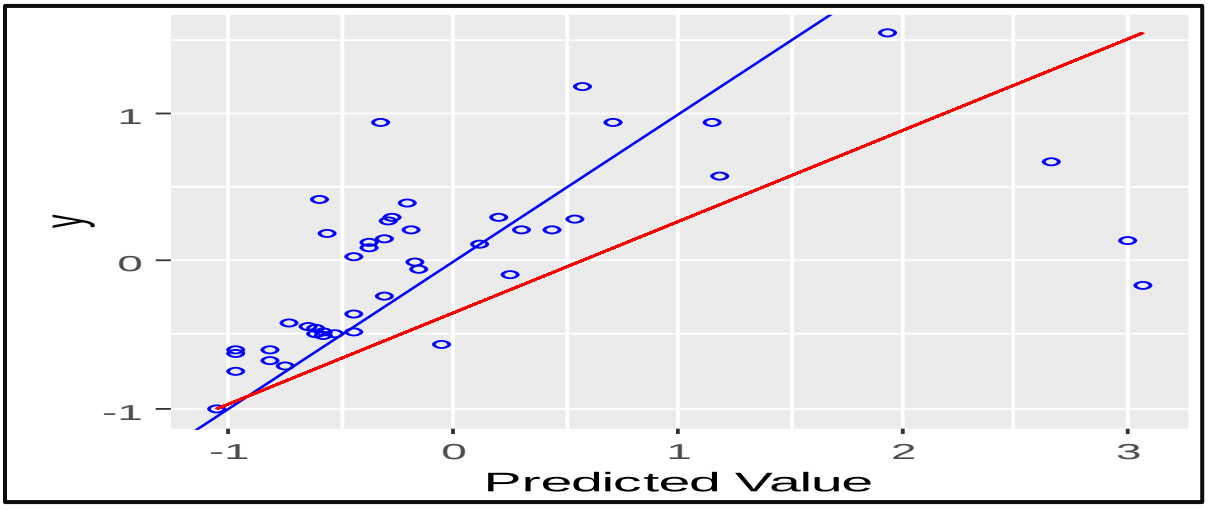
حصل الباحث على هذه البيانات من الموقع الرسمي لسوق العراق للأوراق المالية بعد تطبيق نظام (ISX60), فقد تم جمع البيانات السنوية للقيمة السوقية لتسعة من المصارف المحلية الأكثر تداولاً والمسجلة في سوق العراق للأوراق المالية للمدة من (2011 – 2015) n=45 إذ تضمنت سبعة متغيرات هي (نسبة التداول x_1, العائد على السهم x_2, نسبة دوران السهم x_3, معدل سعر السهم السنوي x_4 , الموجودات x_5, الأرباح المحتجزة x_6 , و القيمة السوقية y), إذ أعتبر الباحث ستة من هذه المتغيرات انها توضح أو تفسر حجم القيمة السوقية وفقاً لنموذج انحدار خطي متعدد يمكن وصفه بما يأتي :-

$$y = \beta_\circ + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon_i \qquad (3.1)$$

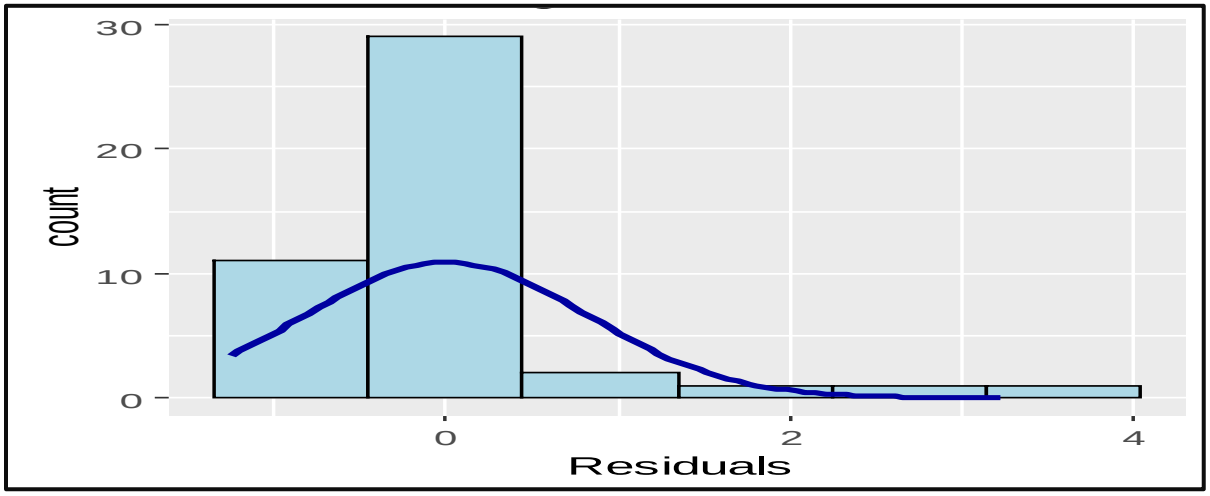
بهدف تشخيص المشاكل التي تعاني منها هذه البيانات و مدى مطابقتها مع اهداف هذه الرسالة، لقد قام الباحث بملاءمة هذه البيانات مع طريقة المربعات الصغرى اولا و من ثم الحصول على $\hat{y}$ و بواقي الانحدار. من خلال الشكل رقم (4) الذي يوضح العلاقة بين $\hat{y}$ و بواقي الانحدار ε_i يتضح جليا ان هناك مشكلة عدم تجانس بالإضافة الى وجود عدد من القيم الشاذة التي اثرت على القيم المقدرة كما نلاحظ في الشكل (5) الذي يظهر ان النموذج المقدر (الخط الاحمر) لا يلائم بيانات الدراسة المفترض (الخط الازرق) و انه متأثر بالقيم الشاذة. وجود القيم الشاذة يعني بالضرورة ان توزيع الخطأ العشوائي ليس طبيعيا و انما يتبع توزيعا ذا ذنب طويل كما نلاحظ في الشكل (6)



الشكل (4) العلاقة بين بواقي الانحدار والقيم المقدرة للقيمة السوقية باستخدام طريقة المربعات الصغرى.



الشكل (5) العلاقة بين القيم المقدرة والقيمة السوقية باستخدام طريقة المربعات الصغرى



الشكل (6) المنحني والمدرج التكراري لتوزيع بواقي الانحدار لطريقة المربعات الصغرى

2.3 **تطبيق طريقة** (WM − Huber) **على بيانات القيمة السوقية** :-

يعرض الجدول (1−3) تقديرات طريقة M−Huber للمعلمات والخطأ المعياري(SE) والانحراف المعياري للبواقي(R,S.Error) وقيمة p_value وقيمة t ومقارنتها مع تقديرات WM−Huber

الجدول(1−3) تقدير طريقة M−Huber للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية

	Value	$SE(\beta)$	t	. Value$P$	Sig	R,S.Error
Intercept	−0.247	0.040	−6.251	0.999		0.25
X1	−0.006	0.007	−0.877	0.807		
X2	−0.269	0.071	−3.807	0.999		
X3	0.003	0.040	0.080	0.468		
X4	−0.007	0.006	−1.248	0.890		
X5	0.403	0.047	8.558	1.07E−10	**	
X6	0.344	0.076	4.555	2.640E−05	**	

من الواضح أن معاملات X_5و X_6هي فقط ذات معنوية مع 0.25 للبواقي الحصينة لنموذج الانحدار نلاحظ أنه على الرغم من أن X_2 و X_3لا يمتلكان نقاط جذب وكلاهما غير معنويين ، في حين أن X_1 و X_4 اللذين لهما نقاط جذب غير معنويين أيضًا. توضح النتائج الواردة في الجدول (2−3) أن طريقة WM− Huber اعتبرت أن معاملات X_1 و X_2 معنوية لتأكيد حصانتها مقابل نقاط الجذب. من ناحية أخرى ، اتفق كل من M−Huber و WM−Huber على اعتبار المتغيرات (X_5 ، X_6) معاملات غير صفرية و (X_2 ، X_3) ليست معنوية. الأخطاء المعيارية الحصينة للبواقي في النموذج الجديد هي 0.09 وهي أقل من قيمة الأخطاء المعيارية لنموذج M−Huber.

الجدول(2−3): تقديرات WM−Huber للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية

	Value	$SE(\beta)$	t	. Value$P$	Sig	R,S.Error
Intercept	−0.138	0.021	−6.656	1.000		0.09
X1	0.088	0.023	3.742	0.0003	***	
X2	−0.143	0.072	−1.977	0.972		
X3	−0.008	0.036	0.228	0.410		
X4	0.216	0.036	6.001	2.79E−07	***	
X5	0.567	0.067	8.465	1.4E−10	***	
X6	0.181	0.052	3.506	0.0005	***	

نلاحظ في الجدول (1−3) أن ثلاثة من معاملات الانحدار سالبة، وهي الحد الثابت ، نسبة التداول ، وربح السهم ، ومتوسط السعر السنوي. يشير إلى أن هذه المتغيرات لها علاقات عكسية مع القيمة السوقية. أما باقي المتغيرات فقد حافظت على علاقة طردية مع القيمة السوقية للبنوك. يوضح الجدول (2−3) أن سعر التداول له علاقة طردية مع القيم السوقية للبنوك مثل البنوك الأخرى باستثناء ربحية السهم ومتغيرات السعر المتوسط السنوي.

3.3 تطبيق طريقة (WFBP. WM) على بيانات القيمة السوقية

يلخص الجدول (3−3) نتائج طريقة WFBP.M ، حيث نلاحظ أن فترة الثقة المئيني (AveLCI، AveUCI) و قيمة p− value وتباين المشاهدات(Ob.s) هما الأكثر معنوية للمتغيرين (الموجودات, ومتوسط السعر السنوي) من غيرهما في تحديد القيمة السوقية للبنوك في سوق العراق للأوراق المالية. في حين أن نتائج طريقة WFBP.WM الواردة في الجدول (4−3) تجد خمسة من أصل ستة متغيرات هي الأكثر معنوية في تحديد القيمة السوقية للبنوك مع استبعاد متغير سعر التداول (X1) لأنه غير معنوي.

الجدول(3−3). تقدير طريقة WFBP.M للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية

	Ob.s	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	−11.383	−0.232	−0.452	0.075		0.000	**
X1	2.786	0.033	−0.158	0.517		0.483	
X2	−5.913	0.105	−0.591	0.366		0.292	
X3	−0.623	−0.018	−0.202	0.324		0.268	
X4	10.607	0.190	0.071	0.720	**	0.090	**
X5	7.991	0.402	0.004	0.790	**	0.000	**
X6	7.430	0.061	−0.028	0.928		0.12	

من ناحية أخرى ، يوضح الجدول (4−3) هذا التشابه التام بين اختبارات C.I. والقيم p ، حيث يمكن ملاحظة حالة واحدة فقط أن القيمة الصفرية تقع خارج CI من X1 حيث تكون قيمة معامل الانحدار لها أكبر من 0.05 ،وبالتالي ، يرفض كلا الاختبارين فرضية العدم حول ال $\hat{\beta}^{MM}$.

الجدول(4−3). تقدير طريقة WFBP.WM للقيمة السوقية المعدلة لبيانات البنوك التجارية العراقية

	Ob.s	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. Value$\bar{P}$	Sig
Intercept	−11.383	−0.092	−0.374	−0.069	**	0.011	**
X1	2.786	0.024	−0.058	0.333		0.310	
X2	−5.913	−0.331	0.450	0.105	**	0.000	**
X3	−0.623	−0.262	0.125	0.181	**	0.000	**

X4	10.607	0.471	0.167	0.543	**	0.000	**
X5	7.991	0.802	0.113	0.574	**	0.000	**
X6	7.430	0.209	0.113	0.667	**	0.010	**

نلاحظ في الجدول (4-3) أن ثلاثة من معاملات الانحدار سالبة وهي التقاطع ونسبة دوران الأسهم و العائد عن الأسهم . مما يدل على ان هذه المتغيرات لها علاقة عكسية مع القيمة السوقية أما باقي المتغيرات فقد حافظت على علاقة الطردية مع القيمة السوقية للبنوك.

1.4.3 دراسة المحاكاة لطريقة WM-Huber

تم أجراء دراسات المحاكاة لمعرفة أداء طريقة WM-Huber مقارنة بطريقة.M-Huber وذلك من خلال إنشاء مصفوفة التصميم $X_{n\times 6}$ للمتغيرات الستة المستقلة التي تم توليدها بشكل عشوائي من التوزيع الطبيعي متعدد المتغيرات بمتوسط صفر وتباين $\rho^{|i-j|}$ ومصفوفة التباين والتباين المشترك $\rho = 020$. يتم بعد ذلك استبدال قيم لـ X_1 بقيم تم توليدها من توزيع $\chi^2_{(0.05,50)}$ لإنشاء نقاط جذب و $m = \alpha \times n$ من المشاهدات , كذلك ايضا $\{X_3, X_4, X_6\}$ ملوثة بنفس الطريقة السابقة لإنشاء نقاط جذب اخرى ، حيث α هي النسبة المئوية للتلوث و nحجم العينات بعد ذلك يتم تلوث متجهه الخطأ العشوائي $e_{m\times 1}$ بقيم تتبع توزيع مربع كاي بدرجة حرية 50 والجزء المتبقي من متجهه الخطأ العشوائي$e_{[(n-m)\times 1]}$ بالتوزيع الطبيعي $N(0,2)$. افترض أن معاملات الانحدار للمجتمع التي هي $\beta_{7\times 1}$,ومن ضمنها معلمة الحد الثابت والتي تأخذ القيم بالترتيب $\beta_{7\times 1} = (0.001,1,0,0,1,1,1)$ لمتغير الاستجابة $y_{n\times 1}$ على النحو الاتي :

$$y = X_{n\times 7}\,\beta_{7\times 1} + e_{n\times 1} \qquad (3.2)$$

ان سيناريو هذه المحاكاة تم اجراؤه عندما تكون $\alpha = 0.05$ لـ $n = \{45,65,85,100\}$ حيث ان n عدد العينات . تم تصميم هذه المحاكاة بحيث تحتوي على ثلاثة معاملات غير صفرية وثلاثة معاملات صفرية .ان أفضل طريقة من بين الطرق المستخدمة هي الطريقة التي تقوم بتشخيص المعاملات المعنوية وغير معنوية بصورة صحيحة وبأقل متوسط للأخطاء المعيارية الحصينة. لهذا الغرض ، تم استخدام كلتا الطريقتين للحصول على نتائج 5000 دورة للبيانات ليتم حساب متوسط المعاملات والمعايير

$\hat{\beta}, SD(\hat{\beta}), t, \overline{P}.value, and\ R, S.Error$ لمجموعات البيانات الإجمالية للطريقتين. ان اتخاذ القرار بشأن رفض أو قبول الفرضية الصفرية التي تفترض أن معاملات الانحدار تساوي الصفر وسيعتمد اتخاذ القرار على $\overline{P}.value$. حيث يجب أن ترفض الفرضية الصفرية عند $(\overline{P}.value < 0.05)$, عندما يتم قبول الفرضية البديلة لمعامل الانحدار المحدد ، فهذا يعني أن قيمة هذا المعامل اكبر من الصفر. سوف نميز المعاملات المعنوية بالرمز (***). من جهه أخرى ، أفضل نموذج هو ذلك الذي يعطي اقل خطأ معياري من بين الطرق الأخرى.

الجدول(5-3). نتيجة المحاكاة لطريقة M-Huber حيث $n = 45\ , \alpha = 0.05$

	$\hat{\beta}^M$	$SE(\hat{\beta}^M)$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	0.555	1.137	0.090	0.474		
X1	5.213	0.156	20.672	0.284		
X2	−0.575	1.146	−0.047	0.511		6.007
X3	0.110	1.044	0.104	0.471		
X4	0.798	1.044	2.259	0.105		
X5	4.155	1.144	8.614	0.024	***	
X6	4.080	0.953	10.405	0.026	***	

من الجدول رقم (5-3) تبين ان معاملات الانحدار X_5, X_6 معاملات معنوية لان قيمة $\overline{P}.value$ اقل من (0.05) اما بقية المعاملات X_3, X_2, X_4, X_1 فهي غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05). نلاحظ ان X_2 غير معنوية رغم انها خالية من الشواذ .

الجدول(6-3).. نتيجة المحاكاة لطريقة WM-Huber حيث $n = 45\ , \alpha = 0.05$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^M)$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	0.555	0.311	−0.046	0.516		

X1	5.213	0.348	2.973	0.041	***	1.617
X2	−0.575	0.375	−0.019	0.508		
X3	0.110	0.355	0.048	0.480		
X4	0.798	0.357	2.938	0.030	***	
X5	4.155	0.373	11.199	0.000	***	
X6	4.080	0.340	12.196	0.000	***	

من النتائج المعروضة في الجدول رقم (3−5) نلاحظ ان كلا من (X_5, X_6, X_4, X_1) هي معاملات معنوية والسبب يعود الى ان قيمة $\overline{P}.value$ اقل من (0.05) اما بقية المعاملات X_3, X_2 فان قيمة $\overline{P}.value$ اكبر لهذا فإنهما غير معنوية كما نلاحظ ان قيمة الخطأ المعياري بهذه الطريقة (WM−Huber) هي اقل من طريقة (M−Huber).

الجدول(3−7). نتيجة المحاكاة لطريقة M−Huber حيث $n = 65$, $\alpha = 0.05$

	$\hat{\beta}^M$	$SE(\hat{\beta}^M)$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.850	0.850	−0.004	0.505		5.945
X1	4.742	0.143	22.706	0.254		
X2	−0.742	0.854	−0.178	0.547		
X3	0.008	0.779	−0.078	0.525		
X4	0.995	0.795	2.965	0.060	***	
X5	4.053	0.842	10.666	0.024	***	
X6	4.000	0.698	13.172	0.020	***	

من الجدول (3−7) نلاحظ ان(X_3, X_2, X_1) هي معاملات انحدار غير معنوية والسبب يعود الى ان قيمة $\overline{P}.value$ اكبر من (0.05) بينما المعاملات الأخرى(X_5, X_6, X_4)عكس ذلك . ان(X_2) هي معامل غير معنوي رغم انها خالية من القيم الشاذة.

الجدول(3−8). نتيجة المحاكاة لطريقة WM−Huber حيث $n = 65$, $\alpha = 0.05$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^{M})$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.239	0.254	−0.033	0.506		1.717
X1	4.742	0.284	3.568	0.017	***	
X2	−0.742	0.302	−0.063	0.516		
X3	0.008	0.283	−0.050	0.518		
X4	0.995	0.284	3.664	0.010	***	
X5	4.053	0.299	13.645	0.000	***	
X6	4.000	0.270	15.259	0.000	***	

من خلال نتائج الجدول (3−8) والتي تبين لنا ان قيمة الخطأ المعياري بطريقة (WM−Huber) هي اقل من طريقة (M−Huber). كما ان معاملات الانحدار(X_5, X_6, X_4, X_1) ذات معنوية وان بقية المعاملات(X_3, X_2)غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05).

الجدول(3−9). نتيجة المحاكاة لطريقة M−Huber حيث $n = 85$, $\alpha = 0.05$

	$\hat{\beta}^{M}$	$SE(\hat{\beta}^{M})$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.092	0.477	−0.058	0.508		
X1	2.487	0.086	20.879	0.257		

X2	−0.263	0.476	−0.052	0.514		3.962
X3	0.018	0.435	−0.002	0.505		
X4	0.956	0.436	3.605	0.033	***	
X5	4.048	0.469	13.327	0.006	***	
X6	4.024	0.393	16.165	0.008	***	

من الجدول (9−3) تبين قيمة $\overline{P}.value$ اكبر من (0.05) لمعاملات الانحدار (X_3, X_2, X_1) وهذا يعني انهما غير معنويتين عكس ما هو عليه بمعاملات الانحدار الأخرى (X_5, X_6, X_4). ان (X_2) رغم خلوها من الشواذ الا انها غير معنوية

الجدول(10−3). نتيجة المحاكاة لطريقة WM−Huber حيث $n = 85$, $\alpha = 0.05$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^{M})$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.092	0.221	−0.084	0.518		
X1	2.487	0.246	4.129	0.004	***	
X2	−0.263	0.257	0.081	0.478		1.773
X3	0.018	0.245	−0.030	0.506		
X4	0.956	0.245	4.180	0.003	***	
X5	4.048	0.254	16.040	0.000	***	
X6	4.024	0.233	17.451	0.000	***	

الجدول رقم (3−10) بين لنا ان كلا من (X_3, X_2) غير معنويتين بسبب ان قيمة $\overline{P}.value$ اكبر من (0.05) وان بقية المعاملات (X_5, X_6, X_4, X_1) تكونان معنويات وان قيمة الخطأ المعياري لطريقة WM−Huber) هي اقل من طريقة (M−Huber).

الجدول(3−11). نتيجة المحاكاة لطريقة M−Huber حيث $n = 100 \ , \alpha = 0.05$

	$\hat{\beta}^M$	$SE(\hat{\beta}^M)$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.063	0.411	−0.049	0.511		3.681
X1	2.306	0.078	20.484	0.254		
X2	−0.279	0.412	−0.043	0.504		
X3	−0.032	0.368	−0.018	0.509		
X4	1.013	0.375	4.130	0.018	***	
X5	3.956	0.409	15.018	0.010	***	
X6	4.019	0.335	18.249	0.005	***	

من الجدول (3−11) نلاحظ ان(X_3, X_2, X_1) هي معاملات انحدار غير معنوية والسبب يعود الى ان قيمة $\overline{P}.value$ اكبر من (0.05) بينما المعاملات الأخرى(X_5, X_6, X_4)عكس ذلك . ان(X_2) هي معامل غير معنوي رغم انها خالية من القيم الشاذة.

الجدول(3−12). نتيجة المحاكاة لطريقة WM−Huber حيث $n = 100 \ , \alpha = 0.05$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^M)$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.063	0.203	−0.016	0.505		1.786
X1	2.306	0.224	4.476	0.002	***	

X2	−0.279	0.235	0.053	0.484		
X3	−0.032	0.224	−0.033	0.509		
X4	1.013	0.224	4.515	0.001	***	
X5	3.956	0.233	17.446	0.000	***	
X6	4.019	0.215	18.806	0.000	***	

من خلال نتائج الجدول (12−3) والتي تبين لنا ان قيمة الخطأ المعياري بطريقة (WM−Huber) هي اقل من طريقة (M−Huber). كما ان معاملات الانحدار (X_5, X_6, X_4, X_1) ذات معنوية وان بقية المعاملات(X_3, X_2)غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05).

الجدول(13−3). نتيجة المحاكاة لطريقة M−Huber حيث $n = 45$, $\alpha = 0.10$

	$\hat{\beta}^M$	$SE(\hat{\beta}^M)$	T	value $\overline{P}$.	Sig	R,S.Error
Intercept	1.123	1.531	0.234	0.440		7.854
X1	7.102	0.210	18.951	0.311		
X2	−1.060	1.508	−0.082	0.523		
X3	−0.157	1.398	−0.063	0.521		
X4	0.838	1.401	2.159	0.118		
X5	4.003	1.486	8.204	0.042	***	
X6	4.313	1.272	9.858	0.031	***	

من الجدول رقم (13-5) تبين ان معاملات الانحدار X_5, X_6 معاملات معنوية لان قيمة $\overline{P}.value$ اقل من (0.05) اما بقية المعاملات X_3, X_2, X_4, X_1 فهي غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05). نلاحظ ان X_2 غير معنوية رغم انها خالية من الشواذ .

الجدول(3-14). نتيجة المحاكاة لطريقة WM-Huber حيث $n = 45 \ , \alpha = 0.10$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^{M})$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	1.123	0.310	0.212	0.477		1.609
X1	7.102	0.348	3.064	0.051	***	
X2	−1.060	0.373	−0.015	0.507		
X3	−0.157	0.345	−0.029	0.505		
X4	0.838	0.347	2.966	0.043	***	
X5	4.003	0.372	11.297	0.000	***	
X6	4.313	0.326	12.961	0.000	***	

من النتائج المعروضة في الجدول رقم (3-14) نلاحظ ان كلا من (X_5, X_6, X_4, X_1) هي معاملات معنوية والسبب يعود الى ان قيمة $\overline{P}.value$ اقل من (0.05) اما بقية المعاملات X_3, X_2 فان قيمة $\overline{P}.value$ اكبر لهذا فإنهما غير معنوية كما نلاحظ ان قيمة الخطأ المعياري بهذه الطريقة WM-(Huber) هي اقل من طريقة (M-Huber).

الجدول(3-15). نتيجة المحاكاة لطريقة M-Huber حيث $n = 65 \ , \alpha = 0.10$

	$\hat{\beta}^{M}$	$SE(\hat{\beta}^{M})$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	0.717	1.142	0.134	0.469		7.719
X1	6.472	0.181	20.697	0.275		
X2	−0.843	1.124	−0.088	0.519		

X3	0.161	1.038	−0.029	0.507		
X4	0.869	1.048	2.603	0.089	***	
X5	4.062	1.081	10.133	0.034	***	
X6	3.955	0.920	12.000	0.037	***	

من الجدول (3−15) تبين قيمة $\overline{P}.value$ اكبر من (0.05) لمعاملات الانحدار (X_3, X_2, X_1) وهذا يعني انهما غير معنويتين عكس ما هو عليه بمعاملات الانحدار الأخرى (X_5, X_6, X_4). ان (X_2) رغم خلوها من الشواذ الا انها غير معنوية

الجدول(3−16). نتيجة المحاكاة لطريقة WM−Huber حيث $n = 65$, $\alpha = 0.10$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^{M})$	t	value $\overline{P}$.	Sig	R,S.Error
Intercept	0.717	0.254	−0.012	0.512		1.699
X1	6.472	0.279	3.606	0.017	***	
X2	−0.843	0.295	−0.056	0.518		
X3	0.161	0.276	−0.021	0.512		
X4	0.869	0.278	3.619	0.011	***	
X5	4.062	0.291	14.161	0.000	***	
X6	3.955	0.263	15.751	0.000	***	

من خلال نتائج الجدول (3−16) والتي تبين لنا ان قيمة الخطأ المعياري بطريقة (WM−Huber) هي اقل من طريقة (M−Huber). كما ان معاملات الانحدار (X_5, X_6, X_4, X_1) ذات معنوية وان بقية المعاملات (X_3, X_2) غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05).

الجدول(3-17). نتيجة المحاكاة لطريقة M-Huber حيث $n = 85$, $\alpha = 0.10$

	$\hat{\beta}^M$	$SE(\hat{\beta}^M)$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	0.347	0.678	0.115	0.468		5.524
X1	3.841	0.121	20.442	0.259		
X2	−0.542	0.655	−0.102	0.522		
X3	−0.013	0.618	−0.029	0.507		
X4	1.113	0.620	3.303	0.049	***	
X5	3.918	0.653	12.323	0.019	***	
X6	3.897	0.554	14.473	0.020	***	

من الجدول (3-17) نلاحظ ان(X_3, X_2, X_1) هي معاملات انحدار غير معنوية والسبب يعود الى ان قيمة $\bar{P}.value$ اكبر من (0.05) بينما المعاملات الأخرى(X_5, X_6, X_4)عكس ذلك . ان(X_2) هي معامل غير معنوي رغم انها خالية من القيم الشاذة.

الجدول(3-18). نتيجة المحاكاة لطريقة WM-Huber حيث $n = 85$, $\alpha = 0.10$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^M)$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	0.347	0.222	0.048	0.488		1.781
X1	3.841	0.245	4.219	0.004	***	
X2	−0.542	0.255	0.017	0.493		
X3	−0.013	0.239	−0.036	0.510		

X4	1.113	0.240	−0.290	0.003	***	
X5	3.918	0.254	15.931	0.000	***	
X6	3.897	0.228	17.865	0.000	***	

من خلال نتائج الجدول (18−3) والتي تبين لنا ان قيمة الخطأ المعياري بطريقة (WM−Huber) هي اقل من طريقة (M−Huber). كما ان معاملات الانحدار (X_5, X_6, X_4, X_1) ذات معنوية وان بقية المعاملات(X_3, X_2)غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05).

الجدول(19−3). نتيجة المحاكاة لطريقة M−Huber حيث $n = 100$, $\alpha = 0.10$

	$\hat{\beta}^M$	$SE(\hat{\beta}^M)$	t	value$\overline{P}$.	Sig	R,S.Error
Intercept	0.203	0.483	−0.003	0.494		4.416
X1	2.759	0.087	20.538	0.252		
X2	−0.364	0.468	−0.107	0.530		
X3	0.055	0.438	0.029	0.490		
X4	0.985	0.435	3.825	0.026	***	
X5	4.030	0.465	14.426	0.010	***	
X6	3.956	0.396	17.059	0.012	***	

من الجدول (19−3) نلاحظ ان(X_3, X_2, X_1) هي معاملات انحدار غير معنوية والسبب يعود الى ان قيمة $\overline{P}.value$ اكبر من (0.05) بينما المعاملات الأخرى(X_5, X_6, X_4)عكس ذلك . ان(X_2) هي معامل غير معنوي رغم انها خالية من القيم الشاذة.

الجدول(20−3). نتيجة المحاكاة لطريقة WM−Huber حيث $n = 100$, $\alpha = 0.10$

	$\hat{\beta}^{WM}$	$SE(\hat{\beta}^{M})$	t	value $\bar{P}$.	Sig	R,S.Error
Intercept	0.203	0.205	−0.046	0.509		1.811
X1	2.759	0.228	4.344	0.003	***	
X2	−0.364	0.236	−0.055	0.514		
X3	0.055	0.221	0.012	0.495		
X4	0.985	0.221	4.613	0.001	***	
X5	4.030	0.234	17.370	0.000	***	
X6	3.956	0.209	19.367	0.000	***	

من خلال نتائج الجدول (20−3) والتي تبين لنا ان قيمة الخطأ المعياري بطريقة (WM−Huber) هي اقل من طريقة (M−Huber). كما ان معاملات الانحدار(X_5, X_6, X_4, X_1) ذات معنوية وان بقية المعاملات(X_3, X_2)غير معنوية لان قيمة $\overline{P}.value$ اكبر من (0.05).

3.4.2 دراسة محاكاة لطريقة WFBP.WM

أجريت دراسات المحاكاة لمعرفة أداء خوارزمية WFBP.WM مقارنة مع.WFBP.M . ان خوارزمية (WBP.M) تشبه (WFBP.WM) ماعدا ان (WBP) تم ادراجها مع M−Huber. سيتم إجراء مقارنة أخرى داخل خوارزمية WFBP.WM لمعرفة أي اختبار ستكون هذه الخوارزمية فيه أكثر استقراريه. يتم إنشاء مصفوفة التصميم $X_{n\times 6}$ للمتغيرات الستة المستقلة بشكل عشوائي من التوزيع الطبيعي متعدد المتغيرات بمتوسط صفر وتباين $\rho^{|i-j|}$ ومصفوفة التباين والتباين المشترك $\rho = 020$. يتم استبدال قيم لـ X_1 بقيم من توزيع $\chi^2_{(0.05,50)}$ لإنشاء نقاط جذب و $m = \alpha \times n$ من المشاهدات , وكذلك $\{, X_4, \}$ ملوثة باستخدام آلية التلوث السابقة لإنشاء نقاط جذب أخرى حيث α النسبة المئوية للتلوث. بعد

ذلك يتم تلوث متجهه الخطا العشوائي $e_{m\times1}$ بتوزيع مربع كاي بدرجة حرية 50 والجزء المتبقي من متجهه الخطأ العشوائي$e_{[(n-m)\times1]}$ بالتوزيع الطبيعي $N\{0,2\}$. افترض أن معاملات الانحدار للمجتمع يشار اليها $\beta_{7\times1}$ وهي معرفه $\beta_{7\times1} = (1,1,1,1,0,0,0)$ لمتغير الاستجابة $y_{n\times1}$ على النحو الاتي

$$y = X_{n\times7}\,\beta_{7\times1} + e_{n\times1} \qquad (3.3)$$

تم أخذ سيناريو المحاكاة هذا في الاعتبار عندما $\alpha = 0.05$ لـ $n = \{45,65,85,100\}$ حيث ان n عدد العينات . تم تصميم دراسة المحاكاة هذه بحيث تحتوي على ثلاثة معاملات غير صفرية وثلاثة معاملات صفرية. أفضل طريقة هي تلك التي تقوم بتشخيص المعاملات المعنوية وغير معنوية بصورة صحيحه . لهذا الغرض ، تم استخدام كلتا الطريقتين للحصول على نتائج 1000 دوره للبيانات وتكررت كل واحدة منها 200 مرة. ان متوسط معاملات $\hat{\beta}^{MM}$ WFBPP، وهو متوسط WFBP للحدود الدنيا والعليا من فترات الثقة Ave.L.C.I. و Ave.U.C.I. ، على التوالي , $p(\hat{\tau}_j)$ متوسط , $\overline{P}.value$ ليتم حسابهم لكلا للطريقتين. اتخاذ القرار بشأن رفض أو قبول الفرضية الصفرية التي تفترض أن معاملات الانحدار تساوي الصفر ستعتمد على فترات الثقة ، $\overline{P}.value$. . إذا كان الصفر يقع خارج فترة الثقة بين Ave.L.C.I. و Ave.U.C.I. ، يجب رفض فرضية العدم ، كما يجب رفض فرضية العدم أيضًا عندما ($\overline{P}.value < 0.05$) . عندما يتم قبول الفرضية البديلة لمعامل الانحدار المحدد ، فهذا يعني أنها مختلفة عن الصفر. سوف نميز المعاملات المعنوية بثلاث نجوم (***). من ناحية أخرى ، أفضل نموذج هو ذلك الذي يحتوي على اقل خطأ معياري من الآخر.

الجدول(3-21). نتيجة المحاكاة لطريقة WFBP.M حيث $\alpha = 0.05$, $n = 45$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$P$	Sig
Intercept	2.136	51.81	−10.334	25.729		0.393	
X1	1.813	−2049.26	−11.126	27.345		0.320	
X2	1.820	108.79	−10.485	26.209		0.560	
X3	1.236	−75.75	−9.573	23.715		0.461	

X4	0.009	8.29	−10.513	22.920		0.489	
X5	0.156	187.12	−11.243	24.649		0.489	
X6	0.669	77.32	−9.424	20.590		0.488	

من نتائج جدول (3−21) تبين ان معاملات الانحدار جميعها غير معنويه لذلك تم قبول فرضية العدم ,

كما نلاحظ أيضا ان الصفر يقع داخل الفتره لـ (Ave.LCI و Ave.UCI)

الجدول(3−22). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 45$, $\alpha = 0.05$

Variable	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	2.136	1.020	−0.349	4.484		0.024	**
X1	1.813	1.068	−0.414	4.660		0.032	**
X2	1.820	1.080	−0.369	4.548		0.035	**
X3	1.236	1.147	−0.248	4.211		0.033	**
X4	0.009	0.001	−1.256	3.222		0.261	
X5	0.156	−0.003	−1.345	3.487		0.279	
X6	0.669	0.039	−1.127	2.898		0.276	

من الجدول رقم (3−22) نلاحظ ان المعاملات($X_1, X_2, X_3,$) معنويات اما بقية المعاملات (X_4, X_6, X_5) فهي غير معنوية وذلك لان قيمة $\overline{P}.value$ اكبر من (0.05).

الجدول(3-23). نتيجة المحاكاة لطريقة WFBP.M حيث $n = 65 \ , \alpha = 0.05$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	2.457	7.07	−6.113	17.633		0.393	
X1	2.798	35.82	−6.639	18.541		0.297	
X2	1.597	−4.74	−6.285	17.701		0.576	
X3	2.226	6.76	−5.617	16.440		0.410	
X4	0.427	−10.96	−6.637	15.454		0.480	
X5	−0.296	−11.38	−7.097	16.670		0.457	
X6	−0.174	4.95	−5.951	13.786		0.501	

من الجدول (3-23) وعنده زيادة حجم العينة الى (65) فأن أداء الطريقة (WFBP.M) لم يتغير فقد بقت جميع المعاملات $(X_5, X_6, X_4, , X_3, X_2, X_1)$ غير معنوية بسبب كبر قيمة $\overline{P}.value$. ان كلا من $(X_3, X_2 X_5, X_6)$ غير معنويات على الرغم من خلوهم من الشواذ

الجدول(3-24). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 65 \ , \alpha = 0.05$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	2.457	1.048	0.131	3.710	**	0.005	**
X1	2.798	1.084	0.072	3.827	**	0.013	**
X2	1.597	1.076	0.104	3.720	**	0.014	**
X3	2.226	1.095	0.175	3.494	**	0.010	**
X4	0.427	0.011	−0.846	2.479		0.267	

	0.264		2.694	−0.883	0.001	−0.296	X5
	0.298		2.236	−0.736	0.037	−0.174	X6

من نتائج الجدول (24−3) تبين ان زيادة في حجم العينة أدى الى ان (Ave.L.C.I.) لمعاملات الانحدار ($X_1, X_2, X_3,$) اكبر من الصفر واقل من نسبة التلوث ($\alpha = 0.05$) لذا تعد معنوية اما بقية المعاملات (X_4, X_6, X_5) فهي اصغر من الصفر وذات قيم سالبة .

الجدول(25−3). نتيجة المحاكاة لطريقة WFBP.M حيث $n = 85$, $\alpha = 0.05$

Sig	$\bar{P}$. value	Sig	Ave.U.C.I.	Ave.L.C.I.	$\hat{\beta}^{MM}$	Obs.t	
	0.300		12.982	−3.520	5.73	3.961	Intercept
	0.221		13.478	−3.696	36.30	3.794	X1
	0.451		13.064	−3.621	−7.07	2.254	X2
	0.318		11.892	−3.174	2.16	2.992	X3
	0.422		10.939	−4.132	3.61	0.854	X4
	0.431		11.820	−4.502	1.26	0.210	X5
	0.411		9.957	−3.742	−4.85	−0.278	X6

من الجدول(25−3) ان أداء الطريقة (WFBP.M) لم يتغير فقد بقت جميع المعاملات($X_1, X_2, X_3, , X_4, X_6, X_5$) غير معنوية بسبب كبر قيمة $\bar{P}.value$ حتى مع زيادة حجم العينة الى (85) ان كل من ($X_6, X_5 X_2, X_3$) غير معنويات على الرغم من خلوهم من الشواذ

الجدول(3-26). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 85$, $\alpha = 0.05$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	3.961	1.031	0.329	3.316	**	0.002	**
X1	3.794	1.093	0.309	3.395	**	0.003	**
X2	2.254	1.039	0.282	3.296	**	0.003	**
X3	2.992	1.067	0.356	3.081	**	0.002	**
X4	0.854	0.013	−0.631	2.105		0.273	
X5	0.210	−0.018	−0.693	2.267		0.258	
X6	−0.278	0.028	−0.563	1.917		0.303	

من نتائج الجدول (3-26) تبين ان زيادة في حجم العينة الى(85) أدى الى ان (Ave.L.C.I.) لمعاملات الانحدار(X_3, X_2, X_1 ,) اكبر من الصفر واقل من نسبة التلوث ($\alpha = 0.05$) لذا تعد معنوية اما بقية المعاملات ($X_5, X_6, X_4,$) فهي اصغر من الصفر وذات قيم سالبة .

الجدول(3-27). نتيجة المحاكاة لطريقة WFBP.M حيث $n = 100$, $\alpha = 0.05$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	3.340	1.38	−1.716	11.141		0.194	
X1	2.741	64.70	−1.911	11.559		0.127	
X2	2.869	−17.16	−1.793	11.312		0.330	
X3	5.028	4.70	−1.520	10.317		0.209	
X4	−0.359	−3.65	−2.504	9.291		0.361	

X5	−1.134	−6.26	−2.685	10.131		0.375	
X6	−0.963	0.06	−2.301	8.501		0.395	

من الجدول (3−27) وعنده زيادة حجم العينة الى (100) فأن أداء الطريقة (WFBP.M) لم يتغير فقد بقت جميع المعاملات $(X_5, X_6, X_4, , X_3, X_2, X_1)$ غير معنوية بسبب كبر قيمة $\overline{P}.value$. ان كلا من $(X_3, X_2 X_5, X_6)$ غير معنويات على الرغم من خلوهم من الشواذ

الجدول(3−28). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 100$, $\alpha = 0.05$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	3.340	1.010	0.392	3.127	**	0.001	**
X1	2.741	1.053	0.356	3.191	**	0.001	**
X2	2.869	1.087	0.387	3.170	**	0.001	**
X3	5.028	1.087	0.451	2.968	**	0.000	**
X4	−0.359	0.003	−0.564	1.952		0.275	
X5	−1.134	−0.004	−0.615	2.123		0.252	
X6	−0.963	−0.010	−0.524	1.779		0.289	

من نتائج الجدول (3−28) تبين ان زيادة في حجم العينة أدى الى ان (Ave.L.C.I.) لمعاملات الانحدار(X_3, X_2, X_1 ,) اكبر من الصفر واقل من نسبة التلوث ($\alpha = 0.05$) لذا تعد معنوية اما بقية المعاملات $(X_5, X_6, X_4,)$ فهي اصغر من الصفر وذات قيم سالبة .

الجدول (3-29). نتيجة المحاكاة لطريقة WFBP.M حيث $n = 45$, $\alpha = 0.10$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$P$	Sig
Intercept	9.73	2.312	−8.561	28.450		0.393	
X1	203.35	3.213	−9.118	29.533		2.262	
X2	−33.66	2.602	−8.513	28.004		0.583	
X3	10.63	1.047	−7.824	26.077		0.461	
X4	−12.30	1.696	−8.870	25.042		0.472	
X5	7.34	−0.281	−9.358	26.688		0.460	
X6	3.06	−0.662	−7.840	22.684		0.473	

من نتائج جدول (3-29) تبين ان معاملات الانحدار جميعها غير معنويه لذلك تم قبول فرضية العدم ,

كما نلاحظ أيضا ان الصفر يقع داخل الفتره لـ (Ave.LCI و Ave.UCI)

الجدول (3-30). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 45$, $\alpha = 0.10$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\bar{P}$	Sig
Intercept	3.082	2.312	−0.373	4.304		0.029	**
X1	1.124	2.213	−0.451	4.365		0.225	
X2	0.404	2.602	−0.406	4.215		0.303	
X3	3.884	1.047	−0.279	4.003		0.027	**
X4	2.906	1.696	−1.322	2.965		0.435	

X5	−1.388	−0.281	−1.387	3.172		0.378	
X6	3.320	−0.662	−1.182	2.683		0.412	

من الجدول رقم (3−30) نلاحظ ان المعاملات(X_3, X_2, X_1 ,) معنويات اما بقية المعاملات ($X_5, X_6, X_4,$) غير معنوية وذلك لان قيمة $\overline{P}.value$ اكبر من (0.10).

الجدول (3−31). نتيجة المحاكاة لطريقة WFBP.M حيث $n = 65$, $\alpha = 0.10$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$P$	Sig
Intercept	2.97	3.801	−4.817	20.510		0.370	
X1	63.32	3.612	−4.931	20.693		0.234	
X2	−11.62	3.066	−4.697	20.007		0.615	
X3	1.55	3.928	−4.383	18.853		0.420	
X4	0.42	1.218	−5.395	17.893		0.473	
X5	0.72	1.630	−5.634	18.648		0.478	
X6	−0.70	−1.597	−4.884	16.059		0.448	

من الجدول (3−31) وعنده زيادة حجم العينة الى (65) فأن أداء الطريقة (WFBP.M) لم يتغير إذ بقت جميع المعاملات ($X_5, X_6, X_4, , X_3, X_2, X_1$) غير معنوية بسبب كبر قيمة $\overline{P}.value$. ان كلا من ($X_3, X_2 X_5, X_6$) غير معنويات على الرغم من خلوهم من الشواذ

الجدول (3−32). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 65$, $\alpha = 0.10$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\overline{P}$	Sig
Intercept	1.728	3.801	0.073	3.758	**	0.005	**

X1	1.091	3.612	0.023	3.701	**	0.069	**
X2	0.900	3.066	0.049	3.645	**	0.099	**
X3	1.888	3.928	0.098	3.476	**	0.013	**
X4	0.868	1.218	−0.882	2.511		0.313	
X5	−0.396	1.630	−0.925	2.615		0.305	
X6	0.978	−1.597	−0.828	2.229		0.333	

من نتائج الجدول (32−3) تبين ان زيادة في حجم العينة أدى الى ان (Ave.L.C.I.) لمعاملات الانحدار ($X_1, X_2, X_3,$) اكبر من الصفر واقل من نسبة التلوث ($\alpha = 0.10$) لذا تعد معنوية اما بقية المعاملات ($X_4, X_6, X_5,$) فهي اصغر من الصفر وذات قيم سالبة .

الجدول (33−3). نتيجة المحاكاة لطريقة WFBP.M حيث $\alpha = 0.10$, $n = 85$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$P$	Sig
Intercept	63.48	4.282	−3.071	14.778		0.301	
X1	1175.57	5.109	−3.140	14.920		0.193	
X2	−347.38	4.159	−3.59	14.631		0.553	
X3	−167.99	4.013	−2.781	13.429		0.369	
X4	69.15	2.396	−3.782	12.529		0.449	
X5	−47.08	0.838	−3.981	13.217		0.429	
X6	98.19	−1.551	−3.394	11.392		0.415	

من الجدول(33-3) ان أداء الطريقة (WFBP.M) لم يتغير حيث بقت جميع المعاملات$(X_5, X_6, X_4, , X_3, X_2, X_1)$ غير معنوية بسبب كبر قيمة $\overline{P}.value$ حتى مع زيادة حجم العينة الى (85) ان كلاً من $(X_3, X_2 X_5, X_6)$ غير معنويات على الرغم من خلوهم من الشواذ

الجدول (34-3). نتيجة المحاكاة لطريقة WFBP.WM حيث $n = 85 \ , \alpha = 0.10$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\overline{P}$	Sig
Intercept	1.287	4.282	0.292	3.309	**	0.001	**
X1	1.087	5.109	0.289	3.314	**	0.022	**
X2	0.970	4.159	0.256	3.240	**	0.043	**
X3	1.399	4.013	0.317	3.053	**	0.005	**
X4	0.273	2.396	−0.681	2.070		0.305	
X5	−0.159	0.838	−0.702	2.200		0.287	
X6	0.404	−1.551	−0.591	1.909		0.292	

من نتائج الجدول (34-3) تبين ان زيادة في حجم العينة الى(85) أدى الى ان (Ave.L.C.I.) لمعاملات الانحدار($, X_3, X_2, X_1$) اكبر من الصفر واقل من نسبة التلوث ($\alpha = 0.10$) لذا تعد معنوية اما بقية المعاملات $(X_5, X_6, X_4,)$ فهي اصغر من الصفر وذات قيم سالبة .

الجدول (35-3). نتيجة المحاكاة لطريقة WFBP.M حيث $n = 100 \ , \alpha = 0.10$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$P$	Sig
Intercept	6.75	2.651	−1.413	11.483		0.175	
X1	86.26	4.859	−1.505	11.552		0.114	

X2	−17.22	4.078	−1.437	11.347		0.339	
X3	4.35	1.872	−1.219	10.550		0.221	
X4	2.33	0.834	−2.207	9.572		0.362	
X5	1.19	−1.075	−2.309	10.074		0.378	
X6	−5.62	1.296	−1.999	8.633		3.56	

من الجدول (3−35) وعند زيادة حجم العينة الى (100) فأن أداء الطريقة (WFBP.M) لم يتغير فقد بقت جميع المعاملات $(X_5, X_6, X_4, , X_3, X_2, X_1)$ غير معنوية بسبب كبر قيمة $\overline{P}.value$. ان كلا من $(X_3, X_2 X_5, X_6)$ غير معنويات على الرغم من خلوهم من الشواذ

الجدول (3−36). نتيجة المحاكاة لطريقة WFBP.WM حيث $\alpha = 0.10$, $n = 100$

	Obs.t	$\hat{\beta}^{MM}$	Ave.L.C.I.	Ave.U.C.I.	Sig	. value$\overline{P}$	Sig
Intercept	1.085	2.651	0.375	3.120	**	0.001	**
X1	1.052	4.859	0.354	3.105	**	0.005	**
X2	1.037	4.078	0.351	3.069	**	0.011	**
X3	1.122	1.872	0.407	2.910	**	0.002	**
X4	0.050	0.834	−0.605	1.906		0.267	
X5	−0.024	−1.075	−0.610	2.034		0.266	
X6	0.084	1.296	−0.522	1.735		0.307	

من نتائج الجدول (3−36) تبين ان زيادة في حجم العينة أدى الى ان (Ave.L.C.I.) لمعاملات الانحدار ($X_1, X_2, X_3,$) اكبر من الصفر واقل من نسبة التلوث ($\alpha = 0.10$) لذا تعد معنوية اما بقية المعاملات ($X_4, X_6, X_5,$) فهي اصغر من الصفر وذات قيم سالبة .

الـفصـــل الرابـــــع

الاستنتاجات والتوصيات

1.4. الاستنتاجات:-

1- ان مقدر M-Huber عنده وجود القيم الشاذة (leverage point) في مصفوفة المتغيرات التوضيحية يكون غير كفؤ لان يكون مقاوم للقيم الشاذة في بواقي الانحدار فقط

2- ان استخدام تقنية البوتستراب الموزون احتماليا يعطينا عينات خالية من القيم الشاذة لكي يتم تطبيق مقدر M-Huber الحصين الذي يكون مقاوم للقيم الشاذة التي تصل الى 50%

3- ان اختبار حصانة معاملات الانحدار البوتسترابي كانت اكثر موثوقية

2.4. التوصيات:-

بناء على ما تقدم يوصي الباحث بما يأتي

1. عند تشخيص وجود قيم شاذة ونقاط جذب في المصفوفة X يوصي الباحث باستخدام طريقة WM-Huber في بيانات عينة الدراسة لتقدير معالم النموذج بطريقة حصينة.

2. يوصي الباحث باستخدام طريقة(WFBP.WM) عندما تكون احجام العينات صغيرة وملوثة بالقيم الشاذة والنقاط الجذب لاختبار الفرضيات حول معلمات الانحدار البوتسترابي لمعرفة معنوية المتغيرات.

المصـــــــــادر

اولاً- المصادر العربية

1- الراوي، خاشع محمود (1987) المدخل الى تحليل الانحدار ، جامعه الموصل، وزاره التعليم العالي والبحث العلمي ،العراق.

2- كاظم، اموري هادي و مسلم ،باسم شليبة (2002) "القياس الاقتصادي المتقدم النظرية والتطبيق" ، مطبعة دنيا الامل ،بغداد ،العراق.

3- محمود مهدي حسن & ,دلير صليوا دغا. (2008). تحديد القيم الشاذة باستخدام الطرق الاستكشافية ومقارنتها مع الطرق المعلمية.

ثانياً- المصادر الانكليزية

Alma, Ö. G. (2011). Comparison of robust regression methods in linear regression. Int. J. Contemp. Math. Sciences, 6(9), 409-421.

Arslan, O. (2012). Weighted LAD-LASSO method for robust parameter estimation and variable selection in regression. Computational Statistics & Data Analysis, 56(6), 1952-1965.

Barnett, V. & Lewis, T. (1978), "Outliers In Statistical Data", Jon Wiley And Sons, New York.

Barrera, M. S. M. (2003) . Fast And Stable Bootstrap Methods For Robust Estimates. Computing Science And Statistics, 34, 346-359

Barrera, M.S. , M.& Zamar, R. H. (2002) Botstrapping Robust Estimates Of Regression, J. Ann. Statist., 30, Pp.556-582.

Box, G. E. (1953). Non-normality and tests on variances. Biometrika, 40(3/4), 318-335.

Bross, I.D.J. (1961), "Outliers In Patternend Experiments: Strategic Re-Appraisal". Technometics . 3,19-102.

Edgeworth, F. Y. (1887). On Observations Relating To Several Quantities. Hermathena, 6(13), 279-285.

Efron, B. (1979). "Bootstrap Meth- Ods: Another Look At The Jackknife", Ann. Statist., P. 7, 1-26.

Ellis, S. P., & Morgenthaler, S. (1992). Leverage and breakdown in L 1 regression. Journal of the American Statistical Association, 87(417), 143-148.

Fox, M. (2002). Optical properties of solids.

Giloni, A., Sengupta, B., & Simonoff, J. S. (2006). A mathematical programming approach for improving the robustness of least sum of absolute deviations regression. Naval Research Logistics (NRL), 53(4), 261-271.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. Journal of the american statistical association, 69(346), 383-393.

Hampel, F. R. (1992). Introduction to Huber (1964) robust estimation of a location parameter. In Breakthroughs in statistics (pp. 479-491). Springer, New York, NY.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. And Stahel, W. A. (1986). Robust Statistics. New York: Wiley.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (2011). Robust statistics: the approach based on influence functions (Vol. 196). John Wiley & Sons.

Hampel, F.R. 1971., A General Qualitative Definition Of Robustness, The Annals Mathematical Of Statistics., 42,1887-1896.

Huber, P. J. (1965). A robust version of the probability ratio test. The Annals of Mathematical Statistics, 1753-1758.

Huber, P. J. (1972). The 1972 wald lecture robust statistics: A review. The Annals of Mathematical Statistics, 43(4), 1041-1067.

Huber, P. J. (1973). Robust Regression: Asymptotic, Conjectures, And Monte Carlo. The Annals Of Statistics, 1: 799-821.

Huber, P. J., & Ronchetti, E. M. (1981). Robust statistics john wiley & sons. New York, 1(1).

Huber, P. J., & Ronchetti, E. M. (1981). Robust statistics, ser. Wiley Series in Probability and Mathematical Statistics. New York, NY, USA, Wiley-IEEE, 52, 54.

Hubert, M., & Rousseeuw, P. J. (1997). Robust regression with both continuous and binary regressors. Journal of Statistical Planning and Inference, 57(1), 153-163.

Hubert, M., & Rousseeuw, P. J. (1997). Robust regression with both continuous and binary regressors. Journal of Statistical Planning and Inference, 57(1), 153-163.

Iglewicz, B., & Hoaglin, D. C. (1993). How to detect and handle outliers (Vol. 16). Asq Press.

Imon, A.H.M.R, And Ali.M.M, 2005. “Bootstrapping Regression Residuals”J.Korean Data Inform .Sci.Soc.,16,665-682.

Keller, G., & Warkack, B. (2000). Statisticsjor Management and Economics 5th Ed. Pacific Groove.

Midi, H., Uraibi, H. S., & Talib, B. A. (2009a). Dynamic robust bootstrap method based on LTS estimators. European Journal of Scientific Research, 32(3), 277-287.

Midi,H., Ramli N.M. & Imon ,A.H.M.R. (2009b). The performance of diagnostic-robust generaliozed potential approach for the identification of multiple high leverage points in linear regression, Journal of Applied Statistics. 36(5):1-15.

Rousseeuw P. J. And Leroy. M.A ., 2003. "Robust Regression And Outlier Detection", Illustrated Edn., New York:Wiley-IEEE.

Rousseeuw P.J. And Van Driessen, K. (1999). A Fast Algorithm For The Minimum Covariance Determinant Estimator. Technometrics. 41:212–223

Rousseeuw, P. And Van Zomeren, B. (1990).Unmasking Multivariate Outliers And Leverage Points. Journal Of American Statistical Associations. 85: 633-639

Rousseeuw, P. J. (1984). Least Median Of Squares Regression. Journal Of Theamerican Statistical Association. 79: 871–880.

Rousseeuw, P. J. and Zomeren B. Van, (1990). Unmasking multivariate outliers and leverage points, J. AM. STAT. ASSOC., 85, 633-639

Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. Technometrics, 41(3), 212-223.

Rousseeuw, P.J. and Leroy, A.M. (1987). Robust Regression and Outlier Detection. New York: Wiley

Susanti, Y., & Pratiwi, H. (2014). M estimation, S estimation, and MM estimation in robust regression. International Journal of Pure and Applied Mathematics, 91(3), 349-360.

Tillé, Y. (2006). Sampling algorithms (pp. 31-39). Springer New York.

Uraibi, H. S., & Midi, H. (2019). ON ROBUST BIVARIATE AND MULTIVARIATE CORRELATION COEFFICIENT. Economic Computation & Economic Cybernetics Studies & Research, 53(2).

Uraibi, H. S., Midi, H., & Rana, S. (2015). Robust stability best subset selection for autocorrelated data based on robust location and dispersion estimator. Journal of Probability and Statistics, 2015.

Uraibi, H. S., Midi, H., & Rana, S. (2017). Robust multivariate least angle regression. J. Scienceasia, 43(1), 56-60.

Uraibi, H. S., Midi, H., Talib, B. A., & Yousif, J. H. (2009). Linear regression model selection based on robust bootstrapping technique. American Journal of Applied Sciences, 6(6), 1191.

Venables, W. N., & Ripley, B. D. (1999). Tree-based methods. In Modern applied statistics with S-Plus (pp. 303-327). Springer, New York, NY.

Willems G. & Aelst S., (2005) Fast and Robust Bootstrap for LTS, Computational Statistics & Data Analysis., 48,703-715.

Yohai, V.J. (1987). High Breakdown Point And High Efficiency Robust Estimates For Regression. The Annals Of Statistics. 15: 642-656.

Abstract

When the regression data contains an outlier, there is a serious problem that leads to a breakdown least squares estimator, and robust regression methods should be recommended. It is well known that the M-Huber method has a 0.50 high breakdown point to resistant outliers, but it breaks down when a single leverage point is present in the design matrix of X. The first target of this thesis is to suggest the weighted M-Huber method by assigning a specific weight for each leverage point. The performance of the proposed method has been tested with the original M-Huber by using real data and simulation. the results show our weighted M-Huber method is more efficient and reliable than M-Huber.

Bootstrap is one of the random sampling methods with replacement, that was proposed to address the problem of small samples whose distributions are difficult to derive. The distribution of bootstrap samples is empirical or free and due to its random sampling with replacement, the probability of choosing a specific observation may be equal to one. Unfortunately, when the original sample data contains an outlier, there is a serious problem even with the M-Huber method which has a high breakdown point is not more than 0.50, so the M-Huber method would break down when the percentage of outliers in the bootstrap sample exceeds 0.50.

It is well known that fixed-x bootstrap is resampled the residuals which probably are having outliers. Moreover, the leverage point(s) is an outlier that occurs in X-direction, so the effects of it on fixed-x bootstrap samples would be existence. However, the decision-making about the null hypothesis of bootstrap regression coefficients could not be reliable. The second objective of this thesis is to propose weighted fixed-x bootstrap with a probability approach to guarantee the

percentage of outliers in the bootstrap samples will be very low. And then weighted M−estimate should be to tackle the problem of outliers and leverage points and take a more reliable decision about bootstrap regression coefficients hypothesis test. the performance of the suggested method has been tested with others by using real data and simulation. the results show our proposed method is more efficient and reliable than others.

Ministry of Higher Education
Scientific Research
University of AI-Qadisiyah
College of Administration & Economics

Robust Test of linear Regression Bootstrapping Coefficients in The Presence of outliers with application

A thesis submitted to the Council of College of Administration Economics University of Al- Qadisiyahln Partial Fulfillment of the Requirement for the Degree of Master of Science in Statistics

M.Sc. Thesis of Statistics

Introduced by

Ali Kwesh Haroon

Supervisor

Assist. Prof.Hassan S.Uraibi

2022 A.D. **1443 A.H.**