مجلة

كلية التـــراث الجامعة

مجلة علمية محكمة

متعددة التخصصات نصف سنوية

العدد السابع والثلاثون

15 حزيران 2023

ISSN 2074-5621

رقم الايداع في دار الكتب والوثائق 719 لسنة 2011

مجلة كلية التراث الجامعة معترف بها من قبل وزارة التعليم العالي والبحث العلمي بكتابها المرقم (ب 3059/4) والمؤرخ في (4/7 /2014)

مقارنة مقدر المربعات الصغرى المبتورة الدائرية الحصينة ومقدر M الدائري الحصين لانموذج الانحدار الدائري JS باستعمال المحاكاة

الباحثة هدى هديب عباس الباحث سهيل نجم عبود

جامعة بغداد – كلية الادارة والاقتصاد- قسم الاحصاء

المستخلص:

في هذا البحث تم استعمال اسلوب المحاكاة مونت- كارولوا لغرض مقارنة مقدر المربعات الصغرى المبتورة الدائرية (Robust Circular Trimmed Least square) مع طريقة M الدائرية الحصينة في حالة عدم تلوث البيانات وفي حالة وجود تلوث في البيانات من خلال اتجاهين الاول تلوث بنقاط الانعطاف العالية التي يمثل التلوث في المتغير المستقل الدائري والثاني التلوث في المتغير العمودي الذي يمثل المتغير المعتمد الدائري باستعمال ثلاثة معايير للمقارنة هي وسيط الخطأ المعياري (Median SE) ووسيط متوسط مربعات الخطأ (Median MSE) ووسيط متوسط الجتا للبواقي الدائرية (Median A(k)) . وتم التوصل الى ان طريقة المربعات الصغرى افضل من طرائق طريقة المربعات الصغرى المبتورة الدائرية الحصينة وطريقة مقدر M الدائري الحصين في حالة عدم احتواء البيانات على قيم ملوثة كونها سجلت اقل معيار وسيط متوسط مربعات خطأ (Median MSE) واقل وسيط خطأ معياري (Median SE) واكبر قيمة لمعيار وسيط متوسط جتا البواقي الدائرية A(K) ولكافة احجام العينات الافتراضية (n=20, 50, 100) . وفي حالة التلوث في البيانات العمودية اتضح عدم افضلية طريقة المربعات الصغرى الدائرية عند كافة نسب التلوث ولكافة احجام العينات. وافضلية طرائق التقدير الحصينة وانه كلما ازادت نسبة التلوث في البيانات العمودية زادت افضلية طرائق التقدير الحصينة بحيث يقل معيار وسيط متوسط مربعات خطأ (Median MSE) ومعيار وسيط خطأ معياري (Median SE) وتزداد قيمة لمعيار وسيط متوسط جتا البواقي الدائرية A(K) ولكافة احجام العينات الافتراضية. وفي حالة التلوث في نقاط الرفع العالية عدم افضلية طريقة المربعات الصغرى الدائرية بنسبة كبيرة عند كافة نسب التلوث ولكافة احجام العينات. وانه كلما ازادت نسبة التلوث في نقاط الرفع زادت افضلية طرائق التقدير الحصينة بحيث يقل معيار وسيط متوسط مربعات خطأ (Median MSE) ومعيار وسيط خطأ معياري (Median SE) وتزداد قيمة لمعيار وسيط متوسط جتا البواقي الدائرية A(K) ولكافة احجام العينات الافتراضية.

الكلمات المفتاحية : الحصانة , الانحدار الدائري , مقدر M الدائري , مقدر المربعات الصغرى, مقدر المربعات الصغرى المبتورة الدائرية الحصينة , الاحصاء الدائري, البيانات الدائرية, المشاهادت الشاذة.

:Abstract

In this research, the Monte Carrollo simulation method was used for the purpose of comparing the Robust Circular Trimmed Least square with the robust circular M method in the case of no data contamination and in the case of data contamination through two directions. The first is contamination with high inflection points that represent Pollution in the circular independent variable and the second pollution in the vertical variable that represents the circular dependent variable using three criteria for comparison: the median of the standard error (Median SE), the median of the mean squared error (Median MSE), and the median of the median cos of the circular remainder (Median A(k)). It was concluded that the method of least squares is better than the methods of the method of the impartial circular truncated least squares method and the method of the impartial circular M estimator in the event that the data did not contain contaminated values, as it recorded the least standard median mean square error (Median MSE)

and the least median standard error (Median SE) and greater A value for an intermediate standard cos of the circular residuals A(K) for all sizes of hypothetical samples (n=20, 50, 100). In the case of contamination in vertical data, it became clear that the circular least squares method was not preferred at all contamination rates and for all sample sizes. And the preference of the robust estimation methods, and that the greater the percentage of pollution in the vertical data, the greater the preference of the robust estimation methods, so that the median standard of average square error (Median MSE) and the standard of median standard error (Median SE) decrease, and the value of the median standard of the cosine of the circular residuals A(K) decreases for all sizes default samples. In the case of contamination at high elevation points, the circular least squares method is not preferred by a large percentage at all contamination rates and for all sample sizes. And that the higher the pollution percentage in the lifting points, the more the preference of the robust estimation methods increases, so that the median standard of mean square error (Median MSE) and the standard of median standard error (Median SE) decrease, and the value of the median standard cos of the circular residuals A (K) increases for all sizes of virtual samples.

Keywords: robust, circular regression, circular M estimator, least squares estimator, impervious circular truncated least squares estimator, circular statistics, circular data, anomalous observations.

1. المقدمة

إنَّ ما يشغل الباحث في حقول المعرفة كافة هو طبيعة البيانات فمن النادر جدا أن تكون هذه البيانات مهيأة مباشرة لإستخدام الأساليب الإحصائية المناسبة للتقدير , فعند حصول حالة خرق لإحدى الشروط المطلوبة للتقدير أو عدم الدقة في البيانات فانه يستوجب البحث عن أساليب مختلفة لمعالجة تلك الحالات. ومن البيانات التي تواجهنا في كثير من التطبيقات العملية هي البيانات التي تكون مقاسة بشكل زوايا بالدرجات أو الراديان مثل زوايا كسر العظم وزوايا انحناء قرنية العين وغيرها. لقد زادت التطبيقات على المتغيرات الدائرية في العقدين الماضيين, وتنوعت في العديد من المجالات بما في ذلك علم الأحياء والأرصاد الجوية والطب. وعلى الرغم من أن إنموذج الإنحدار الدائري الأول يعود إلى **(Gould, 1969)** وبعده تم اقتراح أنواع مختلفة من هذه النماذج ؛ لكن لا تزال دراسة القيم المتطرفة وحصانة نماذج الانحدار الدائري غير مدروسة بشكل جيد. اذ استعملت اساليب الكشف عن القيم المتطرفة بناءً على نموذج الانحدار الدائري البسيط من خلال توسيع الطرائق الشائعة من الانحدار الخطي. وقد يحدث في بعض الأحيان عند القيام بعملية التحليل الاحصائي لمجموعة من البيانات أن بعض المشاهدات تنحرف أو تشذ عن الكمية الأكبر من المشاهدات الموجودة معها والتي يطلق على تلك المشاهدات المنحرفة تسمية شواذ **(Outliers)** والتي في حالة وجودها ضمن البيانات فأن المقدرات التقليدية تخفق في أعطاء تقديرات دقيقة عن معلمات المجتمع الإحصائي الذي سحبت منه تلك البيانات بسبب اختراق الشروط الأساسية لها , لذا فأن إستخدام التحليلات الاحصائية لبيانات تحتوي على قيم شاذة يعد مشكلة حقيقية ويجب تجنبها. وفي نفس الوقت تواجهنا بيانات مسجلة بالدرجات او بالزوايا النصف قطرية في مجموعات مختلفة من مجالات البحث العلمي منها اتجاهات الرياح اليومية واتجاهات تيار المحيط واتجاه مستوى كسر العظام. لا يمكن استخدام نماذج الانحدار الاعتيادية معها لذلك لابد من استخدام انموذج يطبق على تلك البيانات الدائرية . فقد تم اقتراح صيغ مختلفة من هذه النماذج الدائرية؛ ولكن لا تزال مشكلة دراسة القيم الشاذة وانموذج الانحدار الدائري الحصين لم تؤخذ بعين الاعتبار لحد الان والتي تعد مشكلة حقيقية عندما تواجهنا مشكلة وجود قيم شاذة ضمن البيانات المدروسة. تنبأ الباحث **(Goud, 1969)** بالمتوسط الاتجاهي لمتغير استجابة دائري ومجموعة متغيرات مستقلة خطية وتم بعد اقتراحه العديد من نماذج الانحدار الدائري **(Alshqaq, et al, 2021,1)**. فقد تمكن الباحثان جامالامادكا وسارما **(Jammalamadaka & Sarma, 1993)** من طرح مفهوم الانحدار الدائري لبيانات تمثل اتجاهات تقاس كزوايا في مستوى مع الإشارة إلى إحساس ثابت بالدوران واتجاه ثابت للصفر , حيث ناقشا بعض الطرائق التجريبية والاختبارات المقاربة لتحديد درجة معادلة الانحدار وطورا بعض الخوارزميات العددية لايجاد معاملات الانحدار الدائري **(Jammalamadaka & Sarma, 1993)** . في عام (2008) إستعمل **(Mohamed)** وآخرون طريقة الأمكان الأعظم لتقدير معاملات إنموذج الإنحدار الخطي البسيط وإنموذج الإنحدار الخطي البسيط عندما تكون المتغيرات دائرية بخطأ عشوائي له توزيع **Von-Mises** بمتوسط صفر وتباين **k** . في عام (2013) إقترح **(A.H. Abuzaid)** وآخرون إحصاءة عددية جديدة سميت متوسط الخطأ الدائري **(Mean Circular error)**. في نفس العام (2013) استعمل

(**A. Ibrahim**) وآخرون طريقة المربعات الصغرى (**OLS**) لتقدير معلمات انموذج الانحدار الدائري لـ **JS** واعتمدوا على إحصاءة **COVARATIO** وطريقة حذف الصف لتحديد واكتشاف القيم الشاذة في إنموذج **JS**. في عام (2015) إستعمل (**Rambli**) وآخرون إحصاءة **COVARATIO** لتحديد واكتشاف القيم الشاذة في إنموذج الإنحدار الدائري الذي اقترحه كل من (**Downs & Mardia, 2002**) وتم تقييم اداء الطريقة عن طريق تجارب محاكاة مونت-كاروا. في عام (2017) اقترح (**A. Mahmood**) وآخرون احصاءة جديدة لتحديد القيم الشاذة المتعددة في إنموذج الإنحدار الدائري البسيط تعتمد على حساب المسافة الدائرية الحصينة بين البواقي الدائرية ومعلمة الموقع الدائرية وتم تقييم اداء الطريقة المقترحة باستعمال المحاكاة. في نفس العام (2017) اقترحا (**Jha & Biswas**) مقدر الامكان الأعظم المبتور الحصين (**Robust Maximum Trimmed estimator RMTE**) الذي إعتمد على مسافة الـ (جتا) (**Cosine Distance**) لتقدير انموذج الانحدار الدائري –الدائري (**Circular- Circular regression**) المقترح من قبل (**Kato et al., 2008**). في عام (2019) عرف (**Alkasadi**) وآخرون القيم الشاذة في إنموذج الإنحدار الدائري المتعدد (**MCRM**) الذي اقترح من قبل (**Ibrahim, 2013**) والذي يدرس العلاقة بين متغيرين دائريين أو اكثر, حيث قاموا بتوسيع احصاءة **DFFITS** الذي اقترحت من قبل (**Belsley, 1980**) واختبروا الطريقة باستعمل تجارب محاكاة مونت-كارلوا . في عام (2020) اقترح (**A. Mahmood**) وآخرون طرائق حصينة لتقدي معلمات إنموذج الإنحدار الدائري وهي طريقة الامكان الاعظم الموزونة (**Maximum Weighted Likelihood Estimation MWLE**) وطريقة الامكان الاعظم المبتورة (**Maximum Trimmed Likelihood Estimation MTLE**) وتم مقارنة الطريقتين الحصينتين باستعمال محاكاة مونت- كارلوا باستعمال المعايير الاحصائية متوسط مربعات الخطأ (**MSE**) للمقدرات والتباين (**Variance**) ومقدار التحيز (**Bias**). في عام (2021) وَسَع (**Alshqaq**) وآخرون مقدرات **M** الحصينة (**M-Estimation**) ومقدرات المربعات الصغرى المبتورة (**Least-Trimmed squares LTS**) ومقدرات مربعات اقل وسيط (**Least-median squares LMS**) من إنموذج الإنحدار الخطي التقليدي الى إنموذج الإنحدار الدائري **JS** (**Jammalamadaka & Sarma, 1993**) الذي يتصف بصفاته الجيدة وحساسيته في اكتشاف القيم الشاذة. في هذا البحث تم استعمال اسلوب المحاكاة مونت- كارلو في مقارنة طريقتي التقدير المستعملة.

2. مشكلة البحث:

قد يحدث في بعض الأحيان عند القيام بعملية التحليل الاحصائي لمجموعة من البيانات أن بعض المشاهدات تنحرف أو تشذ عن بقية المشاهدات الموجودة معها والتي يطلق على تلك المشاهدات المنحرفة تسمية شواذ (**Outliers**) والتي يتم تمثليها بانموذج احصائي متمثل بانموذج انحدار دائري فان تقدير معاملات هذا الانموذج وفق الطرائق التقليدية قد تخفق في إعطاء تقديرات دقيقة لمعلمات الانموذج الممثل للظاهرة المدروسة|.

3. اهمية البحث:

يُعد وجود قيم شاذة ضمن البيانات مشكلة حقيقية ويجب تجنبها اذ لايمكن معها استعمال نماذج الانحدار الاعتيادية لذلك لابد من استعمال انموذج يطبق على تلك البيانات الدائرية , لذلك اقترح العديد من الباحثين صيغ مختلفة من هذه النماذج الدائرية؛ ولكن لا تزال مشكلة دراسة القيم الشاذة في ظل البيانات الدائرية لم تؤخذ بعين الاعتبار لحد الان والتي تعد مشكلة حقيقية عندما تواجهنا مشكلة وجود قيم شاذة ضمن البيانات الدائرية من هنا تاتي اهمية البحث من كونه يعطي طرائق حصينة ضمن اطار البيانات الدائرية ونذجتها فيما يسمى بانموذج الانحدار الدائري الحصين.

3. الإحصاء الدائري (Circular Statistics) (Pewsey et al., 2013, 16)

يشير مصطلح الإحصاء الدائري إلى فرع معين من فرع الإحصاء الذي يتعامل مع البيانات التي يمكن تمثيلها كنقاط على محيط دائرة الوحدة (**Unit Circle**). إذ يطلق على هكذا نوع من البيانات بالبيانات الدائرية (**Circular Data**) . ومصطلح البيانات الدائرية يستخدم لغرض تمييزها عن البيانات الخطية (**Linear Data**) المعتاد عليها كثيراً في التحليلات. حيث إن الداعم (**Support**) للبيانات الدائرية هو دائرة الوحدة , بينما للبيانات الخطية يكون الداعم هو خط الأعداد الحقيقي . تدخل البيانات الدائرية في مختلف التخصصات مثل علم الأحياء ، الطب ، تحليل الصور ، علوم الأرض ، الفيزياء ، الدراسات السياسية ، وعلم الفلك. في الفضاء ثنائي الأبعاد ، ويمكن تمثيل أي نقطة إما بإحداثياتها الديكارتية كـ (**u ،v**) أو بإحداثياتها القطبية كـ (**r**، **θ**) ، حيث **r** هي المسافة من محيط الدائرة إلى نقطة الأصل. في التحليل الدائري ، يتم التركيز على الاتجاه فقط ، لذلك يتم اعتبار المتجه **r** بطول الوحدة (أي **r = 1**). لذلك ، يمكن تمثيل أي نقطة على الدائرة كـ (**Cos (θ)** ،**Sin (θ)**).ومن الأمثلة على البيانات الدائرية الاتجاهات المقاسة باستخدام أدوات مثل البوصلة أو المنقلة أو ريشة الطقس او الفرجال. فمن المعتاد تسجيل مثل هذه الاتجاهات بالزوايا المعبر عنها بالدرجات (**Degrees**) أو الراديان (**Radians**) إما في اتجاه عقارب الساعة أو عكس اتجاه عقارب الساعة من نقطة أصل الدائرة ، والتي يشار إليها باسم الاتجاه الصفري . فالمطلوب هو تحديد موقع النقطة والاتجاه وليس كما في البيانات على الخط الحقيقي ؛ فالقيم الموجودة

على يسار نقطة الأصل (0) تكون سالبة والقيم الموجودة على اليمين تكون موجبة. وبالنسبة للبيانات الدائرية ، كل زاوية تعرف كنقطة على محيط دائرة الوحدة ، تمامًا كما تحدد كل قيمة لمتغير خطي كنقطة على الخط الحقيقي. ومع زيادة القيمة المطلقة للمتغير الخطي ، نبتعد عن نقطة الأصل. لذلك ، في الخط الحقيقي ، تكون القيمة 360 قريبة نسبيًا من القيمة 355 ولكنها بعيدة نسبيًا عن نقطة الأصل. ولكن بالنسبة للمتغيرات الدائرية تقابل الزاوية 355 درجة نقطة على محيط دائرة الوحدة قريبة من تلك المقابلة لـ 360 درجة ، فإن الزاويتين 0 و 360 درجة تحددان نفس النقطة بالضبط. هذه الطبيعة الدورية للبيانات الدائرية هي التي تجبرنا على التخلي عن الأساليب الإحصائية التقليدية المصممة للبيانات الخطية وتجعلنا نبحث عن تلك التي تخدم البيانات الدائرية وتأخذ هذه الأساليب في نظر الاعتبار بنية هذه البيانات.

4. البيانات الدائرية (Circular Data) (G Pramesti, 2018, 2-3)

غالباً مايتم تمثيل القياسات في اي مجال بعدداً حقيقياً , ولكن في الواقع وفي العديد من المجالات المتنوعة يمكن قياس اي مشاهدة على انها اتجاه (**Direction**) , على سبيل المثال اتجاه الرياح , اتجاه الطيور المهاجرة كبيانات دائرية يمكن قياسها بالبوصلة او الساعة, ويمكن ايضاً اعتبار عمر الكون واتجاهات الكائنات الحية واتجاه الملوثات اعتبارها مشاهدات اتجاهية ويشار الى مثل هذه البيانات بالبيانات الاتجاهية (**Direction Data**) , حيث يمكن تمثيل الاتجاه كنقاط على محيط دائرة الوحدة او متجهات وحدة تربط نقطة الأصل بهذه النقاط. وبالتالي تسمى البيانات ثنائية الابعاد بالبيانات الدائرية (**Circular Data**) وتسمى المشاهدات في ثلاثة ابعاد بالبيانات الكروية (**Spherical Data**) . ويمكن تمثيل البيانات الدائرية بالزاوية $\boldsymbol{\theta}$ والتي مداها $[0, 2\pi]$ او $[-\pi, \pi]$. والزاوية θ تكون دورية حيث ان $\theta = \boldsymbol{\theta} + 2k\pi$. وبجانب البيانات الدائرية فان المتغير العشوائي الذي له قيمة على نصف دائرة الوحدة والتي لها المجال $[-\frac{\pi}{2}, \frac{\pi}{2}]$ او $[0, \pi]$ والتي تدعى بالبيانات المحورية (**Axial**) او البيانات النصف دائرية (**Semi-Circular Data**).

5. المشاهدات الشاذة (Outlier Observations)

يُقصد بالشذوذ في البيانات وجود وتداخل آليات عشوائية مختلفة أحدهما تنتج غالبية البيانات ، والأخرى مسؤولة عن حدوث القيم الشاذة. وحتى لو تمكنا من تحديد إنموذج لغالبية البيانات ، فليس من الواضح كيفية إختيار اتجاهات الإنحرافات عن إفتراضات الإنموذج التي تسببها القيم الشاذة. بهذا المعنى ، فمن المفترض أن تكون القيم الشاذة في البيانات قيماً غير متوقعة ، ولا يمكن للتحليل الكشف عن عملية توليدها العشوائي. ولقد عرف كل من **Barnett and Lewis (1994), Davies and Gather (1993), Markatou et al. (1998), and Gather et al. (2003)** القيم الشاذة بانها قيم مأخوذة عن ظواهر والتي تتلقى احتمالاً صغيراً في ظل الإنموذج المفترض. **(Farcomeni & Gareco, 2015, 4)**. وتعرف القيمة الشاذة **(Outlier)** أو القيمة المُلَوِثة **(Contaminants)** بانها بأنها نقاط إعتباطية في الطبيعة تمثل مشاهدة او مجموعة من المشاهدات خارجة عن النمط الطبيعي لمجموعة البيانات .(هبة الله : 2005 , 4) وانها نقاط بيانات تكون مبتعدة عن غالبية نقاط البيانات الاخرى (أي أنها مشاهدات لا تنسجم مع بقية بيانات المجموعة لأي متغير من المتغيرات لظاهرة معينة أو لمجموعة من الظواهر) , فقد تكون قيمة هذه المُشاهدة كبيرة أو قد تكون صغيرة واقعة في أحد طرفي مجموعة المُشاهدات المرتبة تصاعدياً أو تنازلياً، وأن شذوذها قد يكون في حالات كثيرة مسألة طبيعية ملازمة لبعض المتغيرات (الياسري: 2007, 6). وتكون مولدة بطريقة مختلفة عن طريقة توليد المشاهدات الاصلية. وهي المشاهدات التي تظهر بصورة غير متناسقة مع ما تبقى من مجموعة البيانات **(Obikee & et.al., 2014 , 536)**. كما عرفها الباحثان **(Rousseeum & Leroy)** عام **(1987)** بأنها مشاهدات في الانحدار التي تنحرف عن الجزء الاكبر من البيانات . **(Hekimoglu & Erenoglu , 2013, 421)** في حين عرف الباحثان **(Barnett & Lewis)** في عام **(1994)** القيم الشاذة بأنها (المشاهدات التي تظهر بصورة غير متناسقة مع ما تبقى من مجموعة البيانات) .

6. أصناف المشاهدات الشاذة (Classes of Outlier Observations)

في عام **(2015)** صنف كل من الباحثون **(Lukman , et al.)** القيم الشاذة الى ثلاث فئات قيم شاذة في المحور **V** (قيم شاذة رأسية او عامودية). وقيم شاذة في كلا الاتجاهين (محور المتغير التوضيحي , محور المتغير التابع **V**). و قيم شاذة في محور المتغير التوضيحي **U** تسمى (نقاط الرفع أو نقاط الانعطاف العالية) **(HLPs)** وقد سميت بهذا الاسم لأنها تعطف او تميل خط أنموذج الانحدار نحوها وبالتالي تسبب مشاكل اكثر خطورة من القيم الشاذة العامودية في المحور **V** . (**Lukman** , et al., **2015**, 55)

6. إنموذج الإنحدار الدائري JS (Jammalamadaka & Sarma Circular regression model)

إقترح **(Jammalamadaka and Sarma, 1993)** إنموذج انحدار دائري لمتغيرين عشوائيين دائريين **U** و **V** للتنبؤ بـ **V** بواسطة **U** في سياق التوقع الشرطي للمتجه $\mathbf{e^{iv}}$ بمعلومية **u** وكالآتي:

$$E(e^{iv}|u) = \rho(u)e^{i\mu(u)} = g_1(u) + ig_2(u), \quad \dots (1)$$

إذ أن :

$$\ldots (2) e^{iv} = \cos(v) + i \sin(v)$$

$\mu(u)$ يمثل المتوسط الشرطي الإتجاهي لـ v بمعلومية u,

$\rho(u)$ معلمة الكثافة الشرطية (**Conditional concentration parameter**) للدوال الدورية $g_1(u), ig_2(u)$ والتي تكتب كالآتي :

$$\ldots (3) E(\cos(v|u)) = g_1(u)$$

$$\ldots (4) E(\sin(v|u)) = g_2(u)$$

بعد ذلك يمكن التنبؤ بـ v إذ أن:

$$\ldots (5) \mu(u) = \hat{v} = \arctan * \frac{g_2(u)}{g_1(u)} = \begin{cases} \arctan \frac{g_2(u)}{g_1(u)} & \text{if } g_1(u) > 0 \\ \pi + \arctan \frac{g_2(u)}{g_1(u)} & \text{if } g_1(u) < 0 \\ \text{undefined} & \text{if } g_1(u) = g_2(u) = 0 \end{cases}$$

وباعتبار ان $g_1(u)$ و $g_2(u)$ دوال دورية (**Periodic Function**) ضمن الفترة 2π. لذلك يتم تقريب هذه الدوال باستعمال دوال مناسبة لذلك سيكون التقريب باستعمال متعددة حدود مثلثية لدرجة مناسبة **m** وبموجب نموذجي شبيهين بانموذج الانحدار :

$$\ldots (6) V_{1j} = g_1(u) = \cos(v_j) \simeq \sum_{k=0}^{m} \left(A_k \cos(ku_j) + B_k \sin(ku_j)\right) + \varepsilon_{1j}$$

$$\ldots (7) V_{2j} = g_2(u) = \sin(v_j) \simeq \sum_{k=0}^{m} \left(C_k \cos(ku_j) + D_k \sin(ku_j)\right) + \varepsilon_{2j}$$

لـ $j = 1, \ldots, n$, $\varepsilon = (\varepsilon_1, \varepsilon_2)$ هو متجه الاخطاء العشوائية الذي له توزيع طبيعي ثنائي (**Bivariate Normal distribution**) بمتجه متوسطات يساوي صفر ومصفوفة تباين مجهولة Σ , وان (A_k, B_k, C_k, D_k) هي معلمات الانموذجين, اذ أن **k= 0,1,2,....,m** . ويمكن تقدير الاخطاء القياسية (**Standard errors**) ومصفوفة التباين المجهولة Σ بافتراض أن $B_0 = D_0 = 0$ لضمان امكانية تعريف الانموذج.

(Jammalamadaka and Sarma, 1993) S. IBRAHIM et al., 2013)) (Alshqaq2, et al., 2021, 2(

7. تقدير المربعات الصغرى الدائرية (Circular Least Squares Estimation)

ليكن $(u_1, v_1), \ldots, (u_n, v_n)$, عينة عشوائية دائرية بحجم **n** , فالمعادلة (1) يمكن ان تلخص كالآتي:

$$\ldots (V^{(1)} = (V_{11}, \ldots, V_{1n})' \quad)18$$

$$\ldots (V^{(2)} = (V_{21}, \ldots, V_{2n})' \quad)19$$

$$\ldots (\varepsilon^{(1)} = (\varepsilon_{11}, \ldots, \varepsilon_{1n})' \quad)20$$

$$\ldots (\varepsilon^{(2)} = (\varepsilon_{21}, \ldots, \varepsilon_{2n})' \quad)21$$

$$\ldots (U_{nx(2m+1)} = \begin{bmatrix} 1 & \cos u_1 & \sin u_1 & \cdots & \sin mu_1 \\ 1 & \cos u_2 & \sin u_2 & \cdots & \sin mu_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cos u_n & \sin u_n & \cdots & \sin mu_n \end{bmatrix} \quad)22$$

ولغرض التبسيط سيتم تسمية متجه المعلمات للانموذجين (6) و (7) كالآتي:

$$\ldots (\lambda^{(1)} = (A_0, A_1, \ldots, A_m, B_1, \ldots \ldots B_m)' \quad)23$$

$$\ldots (\lambda^{(2)} = (C_0, C_1, \ldots, C_m, D_1, \ldots \ldots D_m)' \quad)24$$

لذلك فان المعادلتين (23) و (24) يمكن ان تكتب بصيغة المصفوفات كالآتي:

$$\ldots (V^{(1)} = U\lambda^{(1)} + \varepsilon^{(1)} \quad)25$$

$$\ldots (V^{(2)} = U\lambda^{(2)} + \varepsilon^{(2)} \quad)26$$

فان تقديرات المربعات الصغرى يكون كالآتي:

$$\ldots (\hat{\lambda}^{(1)} = \min \sum_{i=1}^{n} (V_j^{(1)} - U\lambda^{(1)})^2 \quad)27$$

$$\ldots (\hat{\lambda}^{(2)} = \min \sum_{j=1}^{n} (V_j^{(2)} - U\lambda^{(2)})^2 \quad)28$$

ومن المعادلتين (27) و (28) نحصل على تقديرات المربعات الصغرى الدائرية باشتقاق الصيغتين نسبة للمعلمات المراد تقديرها ومساواة المشتقة الاولى بالصفر ينتج الآتي:

$$\dots (\hat{\lambda}^{(1)} = (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\,\mathbf{V}^{(1)} \quad)29$$

$$\dots (\hat{\lambda}^{(2)} = (\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\mathbf{V}^{(2)} \quad)30$$

8. الحصانة في انموذج انحدار JS (Robustness in JS model)

تعد مشكلة وجود القيم الشاذة من المشاكل الشائعة في التحليل الاحصائي, والتي تعرف بانها المشاهدات التي تختلف كثيراً عن باقي المشاهدات في مجموعة البيانات. اذ وجد **(Ibrahim et al. , 2013)** الحصانة في انموذج انحدار **JS** واستنتج بان انموذج **JS** حساس لوجود الشواذ ومن المحتمل ان يؤدي الى تاثيرات فعالى على تقديرات المربعات الصغرى لانموذج **JS** . **(Ibrahim et al. , 2013, 2275)**

فلذلك يمكن دراسة الحصانة في إنموذج الانحدار الدائري **JS** من اتجاهين :

1. القيم الشاذة العمودية الدائرية **(Circular Vertical Outliers)**

وهي القيم الشاذة في المتغير المعتمد **V** , فاذا تم استبدال $\mathbf{V_{1j}}$ بـ $\mathbf{V^*_{1j} = Z_1 V_{1j}}$ و $\mathbf{V_{2j}}$ بـ $\mathbf{V^*_{2j} = Z_2 V_{2j}}$ والذي يؤدي الى:

$$\dots (\mathbf{V_{1j} = Z_1{}^{-1}V^*{}_{1j}} \quad)31$$

$$\dots (\mathbf{V_{2j} = Z_2{}^{-1}V^*{}_{2j}} \quad)32$$

فإن الإنحدار الدائري في المعادلة (1) يمكن ان تكتب كالآتي:

$$\dots (\mathbf{Z_1{}^{-1}V^*{}_{1j} = \cos(v_j) = \textstyle\sum_{k=0}^{m}\left(A_k\cos\,(ku_j) + B_k\sin\,(ku_j)\right) + \varepsilon_{1j}} \quad)33$$

$$\dots (\mathbf{Z_2{}^{-1}V^*{}_{2j} = \sin(v_j) = \textstyle\sum_{k=0}^{m}\left(D_k\cos\,(ku_j) + C_k\sin\,(ku_j)\right) + \varepsilon_{2j}} \quad)34$$

فان مقدرات المربعات في ظل وجود القيم الشاذة يمكن الحصول عليها كالآتي:

$$\dots (\quad \mathbf{\hat{\lambda}^{(1)}(U, V^*{}_{1j}) = Z_1{}^{-1}\hat{\lambda}^{(1)}(U, V_{1j})} \quad)35$$

$$\dots (\mathbf{\hat{\lambda}^{(2)}(U, V^*{}_{2j}) = Z_2{}^{-1}\hat{\lambda}^{(2)}(U, V_{2j})} \quad)36$$

2. نقاط الانعطاف الدائرية **(Circular Leverage Points)**

وهي القيم الشاذة في المتغير المستقل **U** , فاذا تم استبدال **U** ب $\mathbf{U^* = ZU}$ فان:

$$\dots (\mathbf{\hat{\lambda}^{(1)}(U^*, V_{1j}) = Z_1{}^{-1}\hat{\lambda}^{(1)}(U, V_{1j})} \quad)37$$

$$\mathbf{\hat{\lambda}^{(2)}(U^*, V_{2j}) = Z_1{}^{-1}\hat{\lambda}^{(1)}(U, V_{2j})} \quad)38\dots ($$

(Alshqaq, 2021, 2-3))Alshqaq2, et al., 2021,3(

9. تقدير M الحصين الدائري (Circular Robust M-estimation)

قدم الباحث (Huber, 1964) مقدرات أطلق عيها مقدرات الامكان الأعظم النوعية (Maximum Likelihood Type estimators) ويرمز لها بـ (M) , تعتمد فكرتها على تصغير دالة في البواقي بدلاً من تصغير البواقي نفسه.

فمن المعروف أن فكرة طريقة المربعات الصغرى **(OLS)** تعتمد على تصغير مجموع مربعات الخطأ (البواقي) اقل ما يُمكن أي:

$$\dots (\min_{\hat{\lambda}^{(p)}} \textstyle\sum_{i=1}^{n}\left(V_j - U\lambda^{(p)}\right)^2 \quad ; p = 1, 2 \quad)39$$

ولكن طريقة **M-estimate** الحصينة تقوم بتصغير دالة الهدف بدلاً من تقليل مجموعة مربعات الخطأ لدالة الهدف حيث انها تعتمد على تصغير دالة الهدف الآتية:

$$\dots (\min_{\hat{\lambda}^{(p)}} \textstyle\sum_{i=1}^{n} P\left(V_j - U\lambda^{(p)}\right)^2 \quad ; p = 1, 2 \quad)40$$

إذ أن P دالة محدبة متماثلة (Symmetric Convex Function) غير متناقصة حدودها $[0, \infty)$ مستمرة وقابلة للأشتقاق نحصل منها على مقدرات حصينة اذ تكون اقل حساسية للقيم الشاذة من مجموع المربعات . إذ ان P المنطقية والمعقولة يجب ان تمتلك الخصائص الآتية :

$p(e) \geq 0$
$p(0) = 0$

$p(e) = p(-e)$

$p(e) = p(e')$ لكل $|e| \leq |e'|$

اذ أنه يمكن الحصول على مقدرات معلمات إنموذج الانحار الدائري بتصغير المعادلتين:

$$\hat{\boldsymbol{\lambda}}^{(\mathbf{1})}(U^*, V_{1j}) = min \sum_{i=1}^{n} P\left({V^{(1)}}_j - U\lambda^{(1)}\right)^2 \quad \ldots (41)$$

$$\hat{\boldsymbol{\lambda}}^{(\mathbf{2})}(U^*, V_{2j}) = min \sum_{i=1}^{n} P\left({V^{(2)}}_j - U\lambda^{(2)}\right)^2 \quad \ldots (42)$$

بأخذ المشتقة الجزئية بالنسبة للمعلمات ومساواتها للصفر وكما يأتي:

$$\ldots (43) \begin{cases} \sum_{i=1}^{n} \psi\left(V_j - U\lambda^{(p)}\right)^2 = 0 \\ \sum_{i=1}^{n} \psi\left(V_j - U\lambda^{(p)}\right)^2 = 0 \end{cases}$$

إذ أن ψ تمثل المشتقة الجزئية للدالة $\psi = P'$ بالنسبة للمعلمات. وأن المعادلة (43) يمكن ان نحصل منها على مقدرات المربعات الصغرى .

وتعد دالة Huber او Biweight من الدوال المستخدمة بشكل واسع في طريقة M الحصينة والتي يمكن ان تعرف من خلال مصفوفة الوزن $W = diag(w)$, اذ أن $w = \frac{\psi(\varepsilon)}{\varepsilon}$, فان المعادلة (43) يمكن اعادة صياغتها كالآتي:

$$\ldots (44) \begin{cases} W^{(1)}\left(V^{(1)} - U\lambda^{(1)}\right) = 0 \\ W^{(2)}\left(V^{(2)} - U\lambda^{(2)}\right) = 0 \end{cases}$$

اذ ان :

$$\ldots (45) \begin{cases} W^{(1)} = \dfrac{\psi\left(V^{(1)} - U\lambda^{(1)}\right)}{\left(V^{(1)} - U\lambda^{(1)}\right)} \\[2ex] W^{(2)} = \dfrac{\psi\left(V^{(2)} - U\lambda^{(2)}\right)}{\left(V^{(2)} - U\lambda^{(2)}\right)} \end{cases}$$

فيمكن كتابة المعادلة (45) كالآتي:

$$\ldots (46) \begin{cases} \sum_{i=1}^{n} W^{(1)}(V^{(1)} - U\lambda^{(1)})U \\ \sum_{i=1}^{n} W^{(2)}(V^{(2)} - U\lambda^{(2)})U \end{cases}$$

ويمكن دمج المعادلات في (47) في المصفوفة المفردة الآتية:

$$U^T W^{(1)} U \lambda^{(1)} = U^T W^{(1)} U V^{(1)} \quad \ldots (48)$$

$$\ldots (49) \quad U^T W^{(2)} U \lambda^{(2)} = U^T W^{(2)} U V^{(2)}$$

هذا يعني ان المقدرات الحصينة لانموذج الانحدار الدائري JS يمكن الحصول عليها كالآتي:

$$\ldots (50) \hat{\boldsymbol{\lambda}}^{(\mathbf{1})} = \left(U^T W^{(1)} U\right)^{-1} U^T W^{(1)} V^{(1)}$$

$$\ldots (51) \hat{\boldsymbol{\lambda}}^{(\mathbf{2})} = \left(U^T W^{(2)} U\right)^{-1} U^T W^{(2)} V^{(2)}$$

اذ ان $W^{(1)}$ و $W^{(2)}$ هي مصفوفة الاوزان ببعد nxn .

اقترح الباحث Huber عند العدد الصحيح M=1.345 دالة الهدف الآتية:

$$\ldots (52)\ P(\varepsilon) = \begin{cases} \varepsilon^2 & for\ |\varepsilon| \leq M \\ 2M|\varepsilon| - M^2 & for\ |\varepsilon| > M \end{cases}$$

اذن فان ψ تعطى كالآتي:

$$\ldots (53)\ \psi(\varepsilon) = \begin{cases} 2\varepsilon & for\ |\varepsilon| \leq M \\ 2M sgn(\varepsilon) & for\ |\varepsilon| > M \end{cases}$$

فان دالة الوزن تعطى كالآتي:

$$\ldots (54)\ W(\varepsilon) = \begin{cases} 1 & for\ |\varepsilon| \leq M \\ \frac{M}{|\varepsilon|} & for\ |\varepsilon| > M \end{cases}$$

وبسبب تزايد الاهتمام بطريقة مقدرات M خلال العقود الماضية دفع الكثير من الباحثين والإحصائيين بالاهتمام بأسلوب التقديرات الحصينة أمثال (Hample, Hinich, Talwar, and Rews) إلى اقتراح كثير من الدوال $\psi_i(.)$. بحيث تعطي مقدر غير شديد الحساسية Non Sensitivity وغير متأثر بالشواذ .

10. مقدر المربعات الصغرى المبتورة الحصينة الدائرية (Circular Robust Least Trimmed Squares estimator)

اقترحت هذه الطريقة من الباحث (Rousseew) عام (1984 م) , أن مقدر (RLTS) يحقق نقطة انهيار عالية (BP =0.5) , كونه لا يتأثر بوجود القيم الشاذة , أن المقدر الناتج من هذه الطريقة يدعى بمقدر المربعات الصغرى المبتورة ويرمز له بالرمز(LTS) ويتم حسابه وفقاً للصيغة الآتية :

$$\hat{\lambda}_{(p)LTS} = \underset{\lambda}{\arg\min}\ Q_{LTS}(\lambda) \quad \dots (55)$$

وان :

$$\dots (56)\ Q_{LTS}\left(\lambda_{(p)}\right) = \sum_{i=1}^{h} r_j^{\,2}$$

h : ثابت وشروطه ($\frac{n}{2} < h < n$) . وان $r^2_{(1)} \le r^2_{(2)} \le \dots \le r^2_{(n)}$ وان:

$$r = 1,2,\dots,p\ ; \quad p = 1,2$$

تمثل مربعات البواقي الاتجاهية المرتبة من اقل قيمة الى اعلى قيمة اذ أن:

$$r_i^2 = \frac{\left(V^{(p)}{}_j - U\lambda^{(p)}\right)^2}{1+\hat{\lambda}_{(p)}{}'\hat{\lambda}_{(p)}} \quad , \quad i = 1,2,\dots,n \quad \dots (57)$$

(A. S. Shvedov **, 2016, 11)**

ان h هو عدد النقاط البيانات الجيدة (الغير مشذبه) والتي تكون مستبعدة في المجموع , وأن المقدر يعطي تقدير حصين من خلال تعريف $n-h$ من النقاط والتي يكون لها أكبر البواقي كالشواذ , مما يسمح باستبعاد هذه النقاط من مجموعة البيانات بشكل كامل اعتماد على قيمة h التي تكون قريبة جدًا من نقاط البيانات الجيدة وذلك لان أعلى عدد من النقاط الجيدة يستعمل في التقدير, وفي هذه الحالة فأن مقدر (RLTS) سوف يعطي افضل تقدير ممكن .

(Rousseeuw & Driessen, 2006, 31-33) (Chen, 2015, 5-6)

11. المحاكاة (Simulation)

تم إعتماد إسلوب المحاكاة مونت-كارلو **(Monte-Carlo Simulation)** لتقدير معلمات انموذج الانحدار الدائري لـ **JS** باستعمال مقدر **S** الدائري الحصين ومقارنتها مع طريقة المربعات الصغرى الدائرية (**Circular Least squares method**) باستعمال ثلاثة معايير للمقارنة هي وسيط الخطأ المعياري **(Median SE)** ووسيط متوسط مربعات الخطأ **(Median MSE)** ووسيط متوسط الجتا للبواقي الدائرية **(Median A(k))** للوصل الى افضل التقديرات. ولغرض اختبار طرائق التقدير تم اجراء الخطوات الآتية:

11.1 تحديد الإنموذج (Model identification)

لغرض التبسيط وتوضيح اكثر لإنموذج الانحدار الدائري سيتم إعتبار الحالة للانموذج عندما تكون k=1 لذلك سوف نحصل على الانموذج الآتي:

$$V_{1j} = g_1(u) = \cos v_j \simeq A_1\cos 0u_j + B_0\sin 0u_j + A_1\cos 1u_j + B_1\sin 1u_j + \varepsilon_{1j}$$

$$\simeq A_0 + A_1\cos u_j + B_1\sin u_j + \varepsilon_{1j} \quad \dots (58)$$

$$V_{2j} = g_2(u) = \sin v_j \simeq C_0\cos 0u_j + D_0\sin 0u_j + C_1\cos 1u_j + D_1\sin 1u_j + \varepsilon_{2j}$$

$$\simeq C_0 + C_1\cos u_j + D_1\sin u_j + \varepsilon_{2j} \quad \dots (59)$$

$$A_0, \quad A_1, \quad B_1, \quad C_0, \quad C_1, \quad D_1$$

ولغرض التبسيط سيتم اعتبار ان $A_0,\ C_0$ مساوية الى الصفر بينما القيم الافتراضية لـ A_1, B_1, C_1, D_1 يتم الحصول عليها من خلال معادلات كثيرة الحدود المثلثية المضـــــــــــــــــافة للقيمة a $\cos(a+v)$ و $\sin(a+u)$ فعندما a=2. بحيث يكون :

$$\cos(2+u) = -0.0416\,cosu - 0.9093\,sinu$$
$$\sin(2+u) = 0.9093\,cosu - 0.0416\,sinu$$

وبالمقارنة مع المعادلات :

$$w^{(1)} = \left(v^{(1)} - u\lambda^{(1)}\right) = 0$$

$$w^{(2)} = \left(v^{(2)} - u\lambda^{(2)}\right) = 0$$

اذ أن:

$$w^{(1)} = \frac{\psi\left(v^{(1)} - u\lambda^{(1)}\right)}{\left(v^{(1)} - u\lambda^{(1)}\right)}$$

$$w^{(2)} = \frac{\psi\left(v^{(2)} - u\lambda^{(2)}\right)}{\left(v^{(2)} - u\lambda^{(2)}\right)}$$

فان متجه القيم الافتراضية للمعلمات A_1, B_1, C_1, D_1 يكون كالآتي:

-0.04161, -0.09093, 0.09093, -0.04161

11.2 توليد الأخطاء العشوائية (random errors generating)

ستم الأخذ بنظر الاعتبار الاخطاء العشوائية تكون غير مترابطة ε_{1j} , ε_{2j} وتتوزيع توزيع طبيعي ثنائي المتغيرات بمتوسط $\boldsymbol{\mu}$**=0** وتباينات $\sigma^2{}_1$ و $\sigma^2{}_2$ هي **0.03, 0.03** على التوالي .

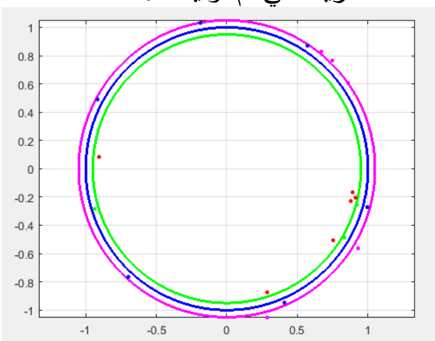
11.3 توليد المتغيرات العشوائية الدائرية (Circular random variables generating)

تم توليد المتغيرات العشوائية الدائرية في انموذج الانحدار الدائري لـ **JS** وهي المتغير العشوائي المستقل **u** والمتغير العشوائي الدائري المعتمد **v** من توزيع **Von-misses** بمتوسط **π** ومعلمة كثافة مساوية الى **2** أي أن: **(Alshqaq2, et al., 2021, 4)**

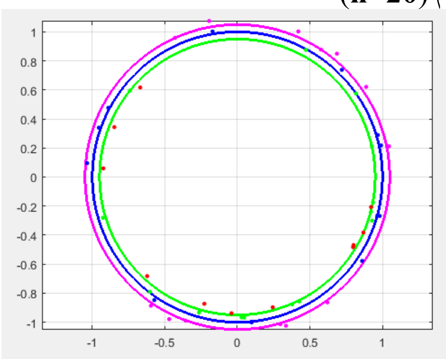
$$\boldsymbol{v \sim vM(\pi, 2)}$$

$$\boldsymbol{u \sim vM(\pi, 2)}$$

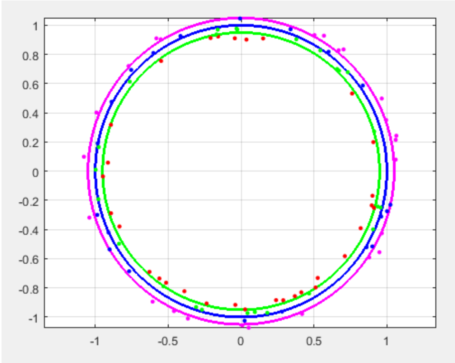
والأشكال (1) و (2) و (**3**) توضح البيانات الدائرية التي تم توليدها .



شكل (1) البيانات الدائرية المولدة بحجم **(n=20)**



شكل (2) البيانات الدائرية المولدة بحجم **(n=50)**



شكل (3) البيانات الدائرية المولدة بحجم **(n=100)**

11.4 توليد احجام العينات (samples generating)

سيتم توليد عينات بحجم **n=10, 50 , 100** للتعرف على سلوك طرائق التقدير عند احجام عينات مختلفة.

11.5 تلويث البيانات (Contaminant data)

لغرض التحقق من حصانة المقدرات مقابل التلوث في قيم المتغيرات الدائرية **u** و **v** سيتم اختيار ثلاث نسب للتلوث وهي **(%5, %10, %20, %30, %40, %50)** وحسب السيناريوهات الآتية:

1. استعمال بيانات المتغيرين بدون تلويث وتطبيق طرائق التقدير مع البيانات غير الملوثة.
2. تلويث قيم المتغير العشوائي الدائري المعتمد (البيانات العمودية) **v** فقط عن طريق موقع المشاهدة **d** وليكن $_d$ ويتم التلويث كالآتي:

$$v_d^* = v_d + \delta\pi \quad \dots (60)$$

اذ أن v_d^* هي قيم المتغير المعتمد **v** بعد التلويث, δ نسبة التلويث اذ ان $0 \le \delta \le 1$

3. تلويث قيم المتغير العشوائي الدائري المستقل (نقاط الانعطات) **u** فقط عن طريق موقع المشاهدة **d** وليكن u_d ويتم التلويث باستحصال نسب مختلفة من البيانات الاصلية عند الموقع **d** باستعمال توزيع **Von-misses** بمتوسط 2π ومعلمة كثافة 6 بدلاً من البيانات الاصلية المولدة في الفقرة **(12.3)** اي ان :

$$u_d \sim vM(2\pi, 6)$$

لغرض التعرف على افضلية طرائق التقدير التي تم استعمالها في تقدير معلمات انموذج الانحدار الدائري لـ **JS** تم استعمال ثلاثة معايير وكما يأتي :

1. **وسيط الخطأ المعياري (Median of standard error)**

تم استعمال معيار وسيط الخطأ المعياري لمعلمات الانموذج بموجب الصيغة الآتية:

$$M(SE)(\hat{\lambda}_j) = \sqrt{\frac{\sum_{j=1}^{s}(\hat{\lambda}_{ij}-\bar{\lambda})^2}{s}} \quad \dots (61) \quad ; j = 1,2,\dots,6$$

اذ ان $\bar{\lambda} = \frac{\sum_{j=1}^{s}\hat{\lambda}_{ij}}{s}$ ويمثل متوسط التقديرات المستحصل عليها بموجب كل طريقة.

2. **وسيط متوسط مربعات الخطأ (Median of mean square error)**

تم استعمال معيار وسيط الخطأ المعياري لمعلمات الانموذج بموجب الصيغة الآتية:

$$M(MSE) = \frac{\varepsilon_{ij}}{s} \quad \dots (62) \quad ; j = 1,2,\dots,6$$

3. **وسيط متوسط جتا البواقي الدائرية**

(Median of mean cosines of circular residuals)

تم استعمال معيار وسيط متوسط جتا البواقي الدائرية بموجب الصيغة الآتية:

$$A(K) = \frac{\sum_{j=1}^{s}Cos(\varepsilon_{ij})}{s} \quad \dots (61) \quad ; j = 1,2,\dots,6$$

(Alshqaq2, et al., 2021, 5)

جدول (1) نتائج تحليل المقدرات المستعملة في حالة عدم وجود تلوث في البيانات عندا **a=3**

n	Criteria	Estimates			Best
		CLS	CRM	CRLTS	
20	Median MSE	2.45333	5.34551	6.45365	CLS
	Median SE	3.94744	8.87848	8.88778	CLS
	Median A(K)	0.99595	0.88554	0.85548	CLS
50	Median MSE	2.43323	5.23453	6.23113	CLS
	Median SE	3.91463	8.76755	8.77655	CLS
	Median A(K)	0.99646	0.89554	0.88133	CLS
100	Median MSE	2.33213	5.21124	6.13333	CLS
	Median SE	3.88543	8.73345	8.89865	CLS
	Median A(K)	0.99813	0.89994	0.89453	CLS

جدول (2) نتائج تحليل المقدرات المستعملة في حالة وجود تلوث في البيانات العمودية عندما **a=3**

n	Criteria	Estimates			Best
		CLS	CRM	CRLTS	
5% Vertical					
20	Median MSE	6.89755	5.52767	5.78655	CRM
	Median SE	8.99655	5.67555	5.79877	CRM
	Median A(K)	0.87344	0.98651	0.95577	CRM
50	Median MSE	6.77787	5.51644	5.67676	CRM
	Median SE	8.98655	5.63543	5.67676	CRM
	Median A(K)	0.88644	0.98699	0.96771	CRM
100	Median MSE	6.65567	5.44235	5.65776	CRM
	Median SE	8.97887	5.52123	5.65656	CRM
	Median A(K)	0.88921	0.98998	0.97961	CRM
10% Vertical					
20	Median MSE	6.85343	5.52134	5.64354	CRM
	Median SE	8.97343	5.91134	5.75545	CRM
	Median A(K)	0.81123	0.98833	0.96678	CRM
50	Median MSE	6.84566	5.51334	5.33334	CRM
	Median SE	8.96112	5.14771	5.73343	CRM
	Median A(K)	0.82788	0.98921	0.96955	CRM
100	Median MSE	6.83356	5.11272	5.32235	CRM
	Median SE	8.94782	5.11153	5.71656	CRM
	Median A(K)	0.83235	0.98978	0.96999	CRM
10 % Vertical					
20	Median MSE	7.56655	3.91766	4.96113	CRM
	Median SE	9.88976	3.27676	4.68796	CRM

	Median A(K)	0.76411	0.98886	0.96832	CRM
50	Median MSE	7.54545	3.82198	4.95466	CRM
	Median SE	9.78998	3.11676	4.61445	CRM
	Median A(K)	0.79234	0.99164	0.96987	CRM
100	Median MSE	7.45553	3.71888	4.94113	CRM
	Median SE	9.67877	3.10443	4.58086	CRM
	Median A(K)	0.79899	0.99651	0.97786	CRM
30% Vertical					
20	Median MSE	8.78666	3.49876	3.78909	CRM
	Median SE	9.99987	3.55655	3.63544	CRM
	Median A(K)	0.67655	0.98575	0.95456	CRM
50	Median MSE	8.55655	3.39565	3.66776	CRM
	Median SE	9.76766	3.54355	3.62214	CRM
	Median A(K)	0.70466	0.98934	0.96411	CRM
100	Median MSE	8.53355	3.38675	3.64544	CRM
	Median SE	9.68977	3.51355	3.61123	CRM
	Median A(K)	0.72343	0.99134	0.97348	CRM
40% Vertical					
20	Median MSE	8.52343	3.33664	3.44544	CRM
	Median SE	9.66766	3.51123	3.49665	CRM
	Median A(K)	0.53145	0.98366	0.95775	CRM
50	Median MSE	8.51123	3.31242	3.42245	CRM
	Median SE	9.56555	3.48656	3.41552	CRM
	Median A(K)	0.53111	0.98444	0.96534	CRM
100	Median MSE	8.49865	3.11233	3.31765	CRM
	Median SE	9.45655	3.43331	3.34555	CRM
	Median A(K)	0.56233	0.98833	0.97633	CRM
50 % Vertical					
20	Median MSE	9.89556	2.77756	2.89765	CRM
	Median SE	10.67889	2.57866	2.69888	CRM
	Median A(K)	0.94765	0.98775	0.97887	CRM
50	Median MSE	9.78788	2.76767	2.86767	CRM
	Median SE	10.55656	2.56565	2.68978	CRM
	Median A(K)	0.94656	0.98876	0.97886	CRM
100	Median MSE	9.66767	2.74545	2.76767	CRM
	Median SE	10.44454	2.44566	2.66777	CRM
	Median A(K)	0.93343	0.98965	0.98156	CRM

جدول (3) نتائج تحليل المقدرات المستعملة في حالة وجود تلوث في نقاط الرقع عندما **a=3**

n	Criteria	Estimates			Best
		CLS	CRM	CRLTS	
5% Leverage					
20	Median MSE	5.78544	4.46433	4.47553	CRM
	Median SE	7.66754	6.66866	6.78654	CRM
	Median A(K)	0.54533	0.96723	0.96446	CRM
50	Median MSE	5.67655	4.45656	4.46343	CRM
	Median SE	7.57767	6.64121	6.74465	CRM
	Median A(K)	0.55923	0.97834	0.96766	CRM
100	Median MSE	5.64576	4.44433	4.45675	CRM
	Median SE	7.55657	6.61335	6.70655	CRM
	Median A(K)	0.56896	0.98164	0.97671	CRM
10% Leverage					
20	Median MSE	7.56754	3.65644	3.66235	CRM
	Median SE	9.64444	5.54333	5.56755	CRM
	Median A(K)	0.62211	0.98745	0.98275	CRM
50	Median MSE	7.55444	3.65345	3.65776	CRM
	Median SE	9.63345	5.54113	5.55211	CRM
	Median A(K)	0.66755	0.98813	0.98544	CRM
100	Median MSE	7.54343	3.74544	3.63424	CRLTS
	Median SE	9.61211	5.59454	5.53113	CRLTS
	Median A(K)	0.68567	0.97174	0.98993	CRLTS
20% Leverage					
20	Median MSE	7.53334	3.63575	3.62343	CRLTS
	Median SE	9.61125	5.52354	5.52175	CRLTS
	Median A(K)	0.71798	0.96765	0.97867	CRLTS
50	Median MSE	7.50311	3.58837	3.63123	CRLTS
	Median SE	9.59977	5.51454	5.56436	CRLTS
	Median A(K)	0.74565	0.96887	0.96164	CRLTS
100	Median MSE	7.49865	3.57854	3.65766	CRLTS
	Median SE	9.53324	5.50966	5.58897	CRLTS
	Median A(K)	0.75181	0.97546	0.97043	CRLTS
30% Leverage					
20	Median MSE	8.67666	2.79777	2.75444	CRLTS
	Median SE	9.78777	3.68989	3.68888	CRLTS
	Median A(K)	0.78877	0.97954	0.97898	CRLTS
50	Median MSE	8.66544	2.79666	2.74333	CRLTS
	Median SE	9.76444	3.66766	3.65332	CRLTS

	Median A(K)	0.79555	0.97999	0.97995	CRLTS
100	Median MSE	8.57866	2.78664	2.66875	CRLTS
	Median SE	9.66787	3.65665	3.64565	CRLTS
	Median A(K)	0.80565	0.98676	0.98787	CRLTS
40% Leverage					
20	Median MSE	8.98998	2.73323	2.66444	CRLTS
	Median SE	9.97777	3.68644	3.61355	CRLTS
	Median A(K)	0.71556	0.98896	0.98355	CRLTS
50	Median MSE	8.97988	2.62112	2.55443	CRLTS
	Median SE	9.96755	3.65454	3.58977	CRLTS
	Median A(K)	0.72865	0.98788	0.98444	CRLTS
100	Median MSE	8.95676	2.61113	2.53444	CRLTS
	Median SE	9.95655	3.61425	3.55656	CRLTS
	Median A(K)	0.73351	0.98565	0.98551	CRLTS
50% Leverage					
20	Median MSE	8.99898	2.21997	2.43232	CRLTS
	Median SE	9.99866	3.31233	3.34433	CRLTS
	Median A(K)	0.97653	0.98775	0.98655	CRLTS
50	Median MSE	8.97866	2.20554	2.33443	CRLTS
	Median SE	9.97788	3.30665	3.31344	CRLTS
	Median A(K)	0.98843	0.98854	0.98776	CRLTS
100	Median MSE	8.86664	2.03453	2.23133	CRLTS
	Median SE	9.85564	3.27966	3.24786	CRLTS
	Median A(K)	0.98886	0.98995	0.98895	CRLTS

12. مناقشة النتائج:

من جدول (1) يتضح ان طريقة المربعات الصغرى كانت افضل من مقدر المربعات الصغرى المبتورة الدائرية (Robust Circular Trimmed Least square) وطريقة M الدائرية الحصينة في حالة عدم احتواء البيانات على قيم ملوثة كونها سجلت اقل معيار وسيط متوسط مربعات خطأ **(Median MSE)** واقل وسيط خطأ معياري **(Median SE)** واكبر قيمة لمعيار وسيط متوسط جتا البواقي الدائرية **A(K)** ولكافة احجام العينات الافتراضية **(n=20, 50, 100)** وعندما تكون قيم **a=3**. من جدول (2) وفي حالة التلوث في البيانات العمودية وعندما تكون قيمة **a=3** يتضح عدم افضلية طريقة المربعات الصغرى الدائرية عند كافة نسب التلوث ولكافة احجام العينات. وانه كلما ازادت نسبة التلوث في البيانات العمودية زادت افضلية طرائق التقدير الحصينة بحيث يقل معيار وسيط متوسط مربعات خطأ **(Median MSE)** ومعيار وسيط خطأ معياري **(Median SE)** وتزداد قيمة لمعيار وسيط متوسط جتا البواقي الدائرية **A(K)** ولكافة احجام العينات الافتراضية. وتفوق طريقة مقدر **M** على مقدر المربعات الصغرى المبتورة الدائرية (Robust Circular Trimmed Least square) . ومن جدول (3) وفي حالة التلوث في نقاط الرفع عدم افضلية طريقة المربعات الصغرى الدائرية بنسبة كبيرة عند كافة نسب التلوث ولكافة احجام العينات. وانه كلما ازادت نسبة التلوث في نقاط الرفع زادت افضلية طرائق التقدير الحصينة بحيث يقل معيار وسيط متوسط مربعات خطأ **(Median MSE)** ومعيار وسيط خطأ معياري **(Median SE)** وتزداد قيمة لمعيار وسيط متوسط جتا البواقي الدائرية **A(K)** ولكافة احجام العينات الافتراضية. وتفوق مقدر طريقة M الدائرية الحصينة على مقدر المربعات الصغرى المبتورة الدائرية (Robust Circular Trimmed Least square) عند نسبة التلوث ولكن كلما زادت نسبة التلوث تتفوق طريقة المربعات الصغرى المبتورة الدائرية (Robust Circular

(Trimmed Least square على طريقة M الدائرية الحصينة. ومن خلال ما تم التوصل اليه من نتائج اوصى الباحثان استعمال مبدأ الضبابية مع الاخذ بنظر الاعتبار الشواذ في البيانات الدائرية مما يجعل طرائق التقدير تتمتع بمرونة ودقة اكبر في التعامل مع البيانات الدائرية في حالة كون البيانات غير دقيقة. وكذلك ادخال الدوال اللبية (Kernel functions) مع طرائق التقدير الحصينة في حالة البيانات الدائرية لايجاد مقدرات ذات كفاءة عالية في تقدير الظواهر المروسة. وتوسعة طرائق التقدير المستعملة في هذه الرسالة في حالة البيانات الكروية التي لها ابعاد متعددة (Multivariate). وتقدير انموذج الانحدار الدائري لـ JS في حالة وجود اكثر من متغير مستقل.

المصادر:

1. الياسري , تهاني مهدي عباس , (2007) , " مقارنة مقدرات بيز الحصين مع مقدرات أخرى لتقدير دالة المعولية التقريبية لتوزيع ويبل" , اطروحة دكتوراه , كلية الادارة والاقتصاد- جامعة بغداد .
2. الصراف, مصطفى نزار , الراوي , اسماء غالب, اسماعيل, غرفان كمال , التقدير الاحصائي , مكتب الجزيرة للطباعة والنشر , بغداد, رقم الايداع في دار الكتب والوثائق ببغداد(1588) لسنة 2016, الطبعة الأولى.
3. Ahmed , Dheyab Ahmed; Abdulwahhab , Baydaa Ismael; Abdulah, Ebtisam Karim, (2020), " A comparison among Different Methods for Estimating Regression Parameters with Autocorrelation Problem under Exponentially Distributed Error",[3] Computer Center, College of Administrations & Economics, University of Baghdad, Iraq *Corresponding author:ahmedthieb@yahoo.com
4. ALMETWALLY , EHAB MOHAMED, ALMONGY, HISHAM MOHAMED, (2018), "COMPARISON BETWEEN M-ESTIMATION, S-ESTIMATION, AND MM ESTIMATION METHODS OF ROBUST ESTIMATION WITH APPLICATION AND SIMULATION ", Journal of Mathematical Archive-9(11), 2018, 55-63 Available online through www.ijma.info ISSN 2229 – 5046.
5. Alshqaq , Shokrya, Ahmadini, Abdullah , Ali Abuzaid , (2021) , ' Robust estimators for circular regression models ', Journal of King Saud University – Science 33 , 101576
6. Alshqaq, Shokrya Saleh, (2021), " On the least trimmed squares estimators for *JS* circular regression model ", Kuwait J.Sci., Vol.48 (3), July.2021, pp(1-13)
7. Arthur Pewsey, Markus Neuhäuser, and Graeme D. Ruxton. (2013), "Circular Statistics in R ", Oxford and New York: Oxford University Press. xiv + 183 p.; ill.; index. ISBN: 978-0-19-967113-7.
8. Farcomeni, A.; Greco, L. (2013), Robust methods for data reduction, Boca Raton, FL: Chapman & Hall/CRC Press, ISBN 978-1-4665-9062-5
9. Ferreira, Sergio L.C.; Caires , Adriana O. ; da S. Borges, Thaise; M.D.S. Lima Ariana, O.B. Silva, Laiana and N.L. dos Santos Walter ,(2017), " Robustness evaluation in analytical methods optimized using experimental designs", Microchemical Journal 131 (2017) 163–169, http://dx.doi.org/10.1016/j.microc.2016.12.004 0026-265X/© 2016 Elsevier B.V. All rights reserved.
10. Hekimoglu, S., and Erenoglu,R. Z. (2013) “A new GM-estimate with high breakdown point ” Acta Geod Geophys,NO. 48 ,PP 419-437.
11. Ibrahim, S., Rambli, A., Hussin, A.G., Mohamed, I., (2013). Outlier detection in a circular regression model using COVRATIO statistic. Commun. Stat.-Simul. Comput. 42 (10), 2272–2280.
12. Jammalamadaka, Sarma, Y., , S., 1993. Circular regression. In: Matusita, K. (Ed.), Statistical Theory and Data Analysis. VSP, Utrecht, pp. 109128109–109128128.
13. Lukman, A. F., Osowole , O. I. & Ayinde, K.(2015) “Two Stage Robust Ridge Method in a Linear Regression Model” Journal of Modern Applied Statistical Methods, NO. 14(2), PP 53-67.

14. Rusiecki m Andrzej , (2009), " Robust Learning Algorithm with LTS Error Function", Encyclopedia of Artificial Intelligence, DOI: 10.4018/978-1-59904-849-9.ch204
15. S. Shvedov, Simple proof of robustness for the least trimed squares estimator in linear regression models, Probl. Upr. , 2016, Issue 5, 10–13
16. Susanti, Yuliana, Hasih Pratiwi,Sri Sulistijowati H., Twenty Liana, (2014), "M ESTIMATION, S ESTIMATION, AND MM ESTIMATION IN ROBUST REGRESSION", International Journal of Pure and Applied Mathematics Volume 91 No. 3 2014, 349-360 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version) url: http://www.ijpam.eu doi: http://dx.doi.org/10.12732/ijpam.v91i3.