Warith Scientific Journal

استخدام طريقة (BYC) للتقدير معلمات نموذج الانحدار الدائري اللوجستي مع تطبيق عملي
Using a Method(BYC) to Estimate the Parameters of a Circular Logistic Regression Model with a Practical Application"

ضحى جبار لفته
Duha Jabber Lafta
duhahjaber@gmail .com
كلية الادارة والاقتصاد جامعة كربلاء ,العراق

أ. د شروق عبد الرضا السباح
Ass. Prof. Dr. Sherouk Abdul Redha Al-Sabbah
Sherouk 12@gmail.com
كلية الادارة والاقتصاد جامعة كربلاء ,العراق

المستخلص:
في هذا البحث تمت مقارنه بين طريقة الامكان الاعظم الدائرية وطريقة (BYC) الدائرية الحصينة لتقدير معلمات انموذج الانحدار اللوجستي الدائري لمتمثلة بطريقة الامكان الاعظم(MaximumLikelihood Method) وفي حالة وجود القيم الشاذة في البيانات كما تم ايجاد التقديرات لحصينة المعلمات انموذج الانحدار اللوجستي الدائري , باستعمال معيار متوسط مربعات الخطأ MSE للمقارنه عن طريق اجراء محاكاة مونت-كارلوا لعينات صغيرة ومتوسطة وكبيرة الانحدار و بأختلاف نسب التلوث المختلفة لدراسة سلوك طرائق التقدير قيد الدراسة , وتم التوصل الى ان طريقة (BYC)هي الافضل من باقي طرائق التقدير كونها حققت اقل معايير مقارنة ولكن هذه الطريقة كانت مقاربة لطريقة الامكان الاعظم الاعتيادية في حال كون وكلما زاد حجم العينة زادت افضلية طرائق التقدير الحصينة الدائرية ,اما في الجانب التطبيقي فقد من هذه الدراسة فقد تم استخدام طريقة (BYC) لغرض تقدير معلمات الانموذج وتطبيقها على البيانات الحقيقية التي تمثل عينة عشوائية بسيطة مؤلفة من (120) متغيرين دائريين هما تأثير أتجاه الرياح على كمية الامطار .
الكلمات المفتاحية: طريقة الامكان الاعظم الدائرية وطريقة (BYC) الدائرية الحصينة ، انموذج الانحدار اللوجستي الدائري .

ABSTRACT : The research included a comparison between the circular maximum likelihood method and the robust circular (BYC) method to estimate the parameters of the circular logistic regression model represented by the Maximum Likelihood Method. In the case of the presence of outliers in the data, robust estimation methods were relied upon to estimate the parameters of the circular logistic regression model, namely the (BY) method, using the mean square error (MSE) criterion, conducting a Monte-Carlo simulation for small, medium and large regression samples and with different pollution ratios to study the behavior of the estimation methods under study. It was concluded that the (BYC) method is better than the rest of the estimation methods, as it achieved the lowest comparison criteria, but this method was close to the regular maximum likelihood method in the case that the larger the sample size, the greater the preference of the robust circular estimation methods. As for the applied aspect of this study, the (BYC) method was used to estimate the parameters of the circular logistic regression model. The model and its application to real data representing a simple random sample of (120) circular variables, namely the effect of wind direction on rainfall.
Keywords: **Circular Maximum Likelihood Method, Circular Boundary (BYC) Method, Circular Logistic Regression Model.**

1-المقدمة:(Introduction)

يعد انموذج الانحدار اللوجستي من نماذج الانحدار اللاخطية حيث يوضح العلاقة بين المتغير التابع ثنائي الاستجابة والمتغيرات التوضحية (التفسيرية) وقد استعمل العديد من الباحثين أنموذج الانحدار اللوجستي اذ ترجع بدايات العمل في موضوع الانحدار اللوجستي الى الباحث(1838 (Verhulst، ويعد اول من استخدم دالة اللوجستك (Logistic Function) لوصف نمو المجتمع فقد أطلق عليها اسم دالة النمو (Growth Function)، في عام 1920 قام الباحثان (Pearl and Reed) بأستخدام الدالة لحساب نمو السكان واطلق عليها فيما بعد دالة اللوجستك بدلاً من دالة النمو، حيث ان التسمية الحالية لانموذج اللوجستك ما زال يعرف به حتى يومنا هذا، كما أستخدم هذا الانموذج من قبل عدد من الباحثين .

ويمكن وصف البيانات الدائرية على انها البيانات التي تكون على شكل متجهات تجمعها نقطة اصل واحدة وتقاس بالقياس الزاوي إذ يمثل بعد كل متجه عن المتجه المحدد بالاتجاه الصفري و اتجاه الدوران مقدار الزاوية التي تدل على اتجاه ذلك المتجه , وان أقل مقدار ممكن تأخذه هذه الزاوية هو الصفر درجة واعلى مقدار هو 360 درجة , مختلفة عن البيانات الاعتيادية التي يكون مدها بين (-∞, ∞), وان مقدار هذه الزاوية لا يعبر عن مقدار الإزاحة بين المتجه الصفري والمتجه نفسه فنجد اكبر ازاحة هي عندما يكون مقدار الزاوية 180 وليس 360 التي تعتبر اقل ازاحة بينها وبين المتجه الصفري , فمقدار الزاوية يعتمد بصوره دقيقة على امرين الأول هو النقطة التي تحدد ليمر بها الاتجاه الصفري والامر الاخر هو اتجاه الدوران هل هو بتجاه دوران عقارب الساعة او عكس اتجاه دوران عقارب الساعة .

ان طـــــرق التقـــــدير التقليـدية لمعلمات أنمـــــوذج الانحـــــــــدار اللوجستـــــي عنـــــــد تحليل البيانات الثنائية (Binary Data Response) تكون ضـعيفة في معالجة المشـكلات بين البيانات، وتعد هذه الطرائق غير كفوءة في حالة عدم تحقق بعض او كل الشـروط المطلوبة، ومن بين احدى تلك المشاكل وجود القيم الشاذة والتي تعرف على انها مشاهدات تنحرف وبشكل ملحوظ عن بقية المشاهدات لمجموعة البيانات، وأن وجودها ضمن البيانات يعد غير مرغوب فيه لانها قد تخرق فرضية التوزيع الطبيعي ومن ثمَ يؤثر بشكل كبير في نتائج التحليل الاحصائي الامر الذي يستوجب اسـتخدام طرائق لمعالجة تلك المشـاكل تتمثل بكونها طرائق بديلة عن الطرائق التقليدية وتكون اكثر كفاءة في التقدير تدعى(بطرائق التقدير الحصـينة) (Robust estimation methods) لكونها قليلة الحساسية تجاه الشواذ أي لا تتأثر كثيرا بوجود القيم الشاذة، اذ يتم الحصول من خلالها على مقدرات حصينة تمتلك كفاءة عالية، وتكون ذات أداء قريب من أداء الطرائق التقليدية عند عدم وجود القيم الشاذة وذات أداء أفضل في حالة وجوده.

2-مشكلة البحث: (Problem of Research)

تعد طرائق التقدير الاعتيادية طفيفه في تقدير معلمات الانموذجات عند وجود قيم شاذه في البيانات

3-هدف البحث: (Aim of Research)

يهدف البحث الى ايجاد أفضل الطرائق التقدير معلمات انموذج الانحدار اللوجسيتي الدائري في حاله تلوث البيانات في حاله عدم تلوث البيانات

4-البيانات الدائرية Circular Data

تشـير البيانات الدائرية إلى البيانات التي يتم قياسـها على مقياس دائري ، مما يعني أن القيم تمثل بدلاً من الكميات الخطية تتضـمن أمثلة البيانات الدائرية والاتجاهات أو الزوايا أو الخطوط غالباً البوصلة واتجاهات الرياح الأمثلة عليها اتجاه الطائر أو اتجاه حيوان أو اتجاه الرياح وغيرها من الظواهر هذه الأمثلة ذات بعدين يمكن أن تمثل بشـكل زاوية يمكن قياسـها ، و قد تكون الاتجاهات في ثلاثة أبعاد ممثلة بزاويتين (على غرار تمثيل النقاط على الأرض حسـب خطوط الطول والعرض الخاصـة بهم) ، كمتجهات وحدة في ثلاثة أبعاد وهذا يسـمى بالبيانات الكروية والوقت من اليوم والظواهر البيولوجية أو السـلوكية الدورية, ولكن في الواقع وفي العديد من المجالات المتنوعة يمكنً حقيقياً تمثيل قياس اي مشـاهدة على انها اتجاه (Direction)إذ يمكن تمثيل الاتجاه كنقاط على محيط دائرة الوحدة أو متجهات وحدة تربط نقطة الأصـل بهذه النقاط وبالتالي تسـمى البيانات ثنائية الأبعاد بالبيانات الدائرية (Data Circular)وتسمى المشاهدات في ثلاثة أبعاد بالبيانات الكروية(Data Spherical) ويمكن تمثيل البيانات الدائرية بالزاوية θ والتي مداها $(0, 2\pi)$ او $(-\pi, \pi)$] والزاوية θ تكون دورية.

أن مدى البيانات الدائرية يختلف عن نظيرتها في البيانات الاعتيادية إذ إن مدى البيانات الاعتيادية ممكن أن يكون بين السـالب ما لانهاية إلى ما لانهاية (-∞ , ∞) بينما البيانات الدائرية يكون مداها محصـوراً بين الصـفر و 360 درجة (0^{o} , 360^{o}) . كما هناك فرق أخر وهو أن مقدار الزاوية ليس فريداً وإنما تتغير من قيمة إلى قيمة أخرى لأن قيمة الزاوية تعتمد على اختيار ما يسـمى بالاتجاه الصـفري واتجاه بالدوران فما يعد 30 درجة عندما يأخذ الشرق الحقيقي باعتباره الاتجاه الصفري والدوران عكس اتجاه عقارب الساعة يمكن أن يكون هو 60 درجة بالنسبة للذي يأخذ الشمال الحقيقي على أنه الصفر والدوران مع اتجاه عقارب الساعة.

5-نموذج الانحدار الزاوي البسيط Simple Angular Regression Model

عندما تكون لدينا ظاهرة ما في الطبيعة او في الحياة تعتمد على متغير مستقل تم قياسه بمقاييس دائرية (على شكل زاوية محصورة قيمها بين الصفر درجة و 360 درجة)ومتغير معتمد بياناته غير دائرية انما هي بيانات كمية اعتيادية فيسـتعمل أنموذج الانحدار الزاوي البسـيط الذي صـيغته كما مبينه ادناه : [5]

$$y_i = \alpha + \beta \cos(x_i - a) + e_i \quad \ldots\ldots (1)$$

اذ أن

y_i تمثل المتغير المعتمد (التابع) (Dependent variable) المراد تحديد قيمته

x_i تمثل المتغير المستقل دائري (independent variable)

α تمثل الحد الثابت للأنموذج

β تمثل الميل الحدي للأنموذج

Aتمثل معدل الزاوي (المتوسط الدائري للزوايا) (mean circular)

أما لو كان لدينا ظاهره مقاسه بقياسات دائرية وتعتمد على متغير مستقل مقاس أيضا بالقياسات الدائرية وان عدد هذه المتغيرات هو اثنان فقط احدهما متغير مستقل (Independent variable) و الأخر متغير تابع (Dependent variable) يعتمد في قيمه على المتغير المستقل أو يتأثر بتغير قيم المتغير المستقل , ولمعرفة مقدار هذا التأثير و اتجاهه و التنبؤ بقيم المتغير التابع المستقبلية سيتم دراسة نوعين من النماذج التي تلائم هذا النوع من البيانات اعتمادا على نوع التوزيع الذي يتبعه الخطأ العشوائي واقترح أنموذج ثالث بتوزيع للخطأ مختلف .

4-أنموذج الانحدار اللوجستي الدائري Logistic regression model for circular data

يعد أنموذج الانحدار اللوجسـتي الدائري من النماذج المهمة التي تسـتعمل في التطبيقات الاحصـائية وتحليل البيانات ذات الطابع الاتجاهي ويهدف نموذج الانحدار اللوجستي للبيانات الدائرية إلى وصف العلاقة بين الاستجابة الثنائية والمتنبئ الدائري. ويستعمل إنموذج الانحدار اللوجستي بصوره عامه لتفسير العلاقة بين متغير توضـيحي واحد او (اكثر) حيث يمكن أن تكون الاسـتجابة الثنائية هي هطول الأمطار (نعم أو لا) والمتنبئ الدائري هو اتجاه الرياح. ويستعمل على نطاق واسع في العديد من العلوم الاقتصادية والاجتماعية والطبية.

في نموذج الانحدار اللوجستي الدائري تكون العلاقة بين المتغير التابع Y والمتغيرات التوضيحية $(\theta_1, \theta_2, \theta_3 \dots \theta_p)$ غير خطية حيث يكون المتغير التابع Y ثنائي الاستجابة يأخذ إحدى القيمتين (اما صفر او واحد) أما النجاح success حدوث الاستجابة(وقوع الحدث) باحتمال $pi = \frac{y_i}{n_i}$ أو الفشل failure عدم حدوث الاستجابة باحتمال (1-pi) و علية نموذج الانحدار اللوجستي الدائري يكتب من خلال الصيغة التالية :[1] [3]

$$\text{Logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + A\cos(\theta_i + \theta_i) \quad \dots(2)$$

β_0:هي قيمة المعامل اللوجستي (logodds) عندما تكون90= $(\theta_i + \theta_0)$.
θ : الزاوية التي يصل عندها اللوجت (logodds) إلى أعلى قيمة له.
A:يمثل المسافة من المحور إلى أعلى نقطة في المنحنى.
P_i : احتمال وقوع الحدث.
بافتراض ان :

$$A = \sqrt{{\beta_1}^2 + {\beta_2}^2}$$

$$\theta_0 = \arctan\frac{\beta_2}{\beta_1}$$

فان نموذج الانحدار اللوجستي يمكن اعادة كتابة بالصيغة التالية :

$$\text{Logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1\cos\theta_i + \beta_2\sin\theta_i \quad \dots(3)$$

ويمكن تحويل نموذج الانحدار اللوجستي الى الشكل الخطي بأستخدام تحويل اللوجت Logit حيث يتم اخذ اللوغاريتم الطبيعي لنسبة الترجيح وكما يلي:
$logit\ (Oi)$: لوغاريتم الطبيعي لنسبة الترجيح .

$$logit\ (p_i) = \log(\frac{p_i}{1-p_i}) = \beta_0 + \beta_1\cos\theta_i + \beta_2\sin\theta_i \quad \dots(4)$$

اذ أن :
$logit\ (p_i)$: لوغاريتم الطبيعي لنسبة الترجيح .

5- دالة اللوجت (logit function)

إن المتغير التابع يأخذ قيمتين وهيَ (y=1) ظهور الصفة و (y=0) عدم ظهور الصفة واذا جمعنا عدد الحالات التي يكون فيها (y=1) وقسمنا على العدد الكلي للحالات سوف نحصل على متوسط قيمة المتغير التابع y والتي تقابل تماماً احتمال ان تكون قيمة المتغير التابع يساوي واحد P(y=1) وهنا سوف نتعامل مع قيم احتمالية للمتغير التابع الذي يأخذ القيمة واحد P(y=1) بدلاً من المتغير التابع y نفسهُ وهذا الشيء يفتح الطريق للتعامل مع المتغير التابع P(y=1) كمتغير متصل ولكن هذه الطريقة سوف تواجه مشكلة عند توفيق البيانات بأنموذج خطي لان سوف تظهر قيم متوقعة للمتغير التابع تتجاوز الواحد الصحيح او تقل عن الصفر وهذا الشيء يتناقض مع مفهوم الاحتمالات كون المتغير التابع مقيداً بقيم محددة بين الصفر والواحد ولحل هذه المشكلة هيَ اجراء تحويل بسيط وهوَ استخدام نسبة الترجيح Odds بدلاً من الاحتمالات Probabilities. [7] [9]

6-نسبة الترجيح Odds

هوَ عبارة عن نسبة احتمال حدوث حدث ما الى احتمال عدم حدوثه وغالباً ما يتم التعبير عنهُ على شكل نسبة بين عددين ، فمثلاً توقع حدوث ثلاثة ولادات مشوهة من خمسة ولادات حية فهذا يعني ان احتمال التشوه هوَ $(\frac{3}{5} = 0.6)$ وعند التعبير عن الاحتمال السابق باستعمال بنسب الترجيح فيقال بأن نسبة ترجيح تشوه ولادي هوَ 3 الى 2 هذا يعني ان التشوه يحدث في ثلاثة ولادات ولا يحدث في ولادتين وبذلك يظهر الفرق بين نسبة الترجيح والاحتمالات.
وعادةً ما يتم التعبير عن نسبة الترجيح بعدد كسري فمثلاً نسبة الترجيح 7 الى 3 يمكن ان يتم التعبير عنهَ على شكل قيمة 2.33 وهذا يعني :

- اذا كان قيمة نسبة الترجيح = 1 فهذا يعني احتمال وقوع الحدث = احتمال عدم وقوع الحدث
- اذا كان قيمة نسبة الترجيح > 1 فهذا يعني احتمال وقوع الحدث > احتمال عدم وقوع الحدث
- اذا كان قيمة نسبة الترجيح < 1 فهذا يعني احتمال وقوع الحدث < احتمال عدم وقوع الحدث

و يمكن التعبير عن نسبة الترجيح بالصيغة التالية [7]

$$O_i = \frac{P_i}{1-P_i} = e^{\alpha+Bxi} \quad \dots(5)$$

اذ أن
O_i : نسبة ترجيح وقوع الحدث .
P_i : احتمال وقوع الحدث

α : ثابت المعادلة .
B : المعامل اللوجستي .
من المعروف انه قيمة p محصورة دائماً بين الصفر والواحد $0 \leq p \leq 1$ وبناء على المعادلات السابقة فان نسبة الترجيح O_i عبارة عن مقدار موجب محصور بين $(0, \infty)$ أي أن: $0 \leq O_i < \infty$
يلاحظ ان نسبة الترجيح قد حل مشكلة الحدود العليا للاحتمال P واصبح يأخذ اي قيمة من الصفر وحتى ما لانهاية وبهذا فان نسبة الترجيح قد ساهم في حل نصف المشكلة وهي ازالة الحد الاعلى للاحتمالات ويمكن تحويل نموذج الانحدار اللوجستي الى الشكل الخطي بأستخدام تحويل اللوجت Logit حيث يتم اخذ اللوغاريتم الطبيعي لنسبة الترجيح وكما يلي:

$$logit\ (Oi) = \log(\frac{p_i}{1-p_i}) = \alpha + BX \qquad ...(6)$$

من المعادلة (2-5) نحصل على :

$$p_i = e^{\alpha+BX}(1 - p_i)$$
$$p_i = e^{\alpha+BX} - p_i\, e^{\alpha+BX}$$
$$p_i\,(1 + e^{\alpha+BX}) = e^{\alpha+BX}$$

$$p_i = \frac{e^{\alpha+BX}}{1+e^{\alpha+BX}} = \frac{1}{1+e^{-(\alpha+BX)}} \qquad ...(7)$$

اذ أن
logit (Oi): لوغاريتم الطبيعي لنسبة الترجيح .

7-طرائق التقدير Methods of Estimation

نستعرض في هذا البحث بعض طرائق تقدير معلمات أنموذج الانحدار اللوجستي الدائري بهدف ايجاد مقدرات لمعلمات النموذج المدروس بحيث تكون هذه المقدرات اقرب مايمكن للقيمة الحقيقية للمعلمات المراد تقديرها، ومن ثَمَّ فأن هذه المقدرات تتصف بصفات جيدة تؤهلها لتكوين نموذج تقديري يتم الاعتماد عليه في أغراض مختلفة والوصول الى نتائج اكثر دقة، ويكون ذلك من خلال اختيار الطريقة المناسبة للتقدير.

7-1طريقة الإمكان الأعظم لتقدير الدائرية لتقدير معلمات أنموذج الانحدار اللوجستي الدائرية

تُعد طريقة الامكان الاعظم من أهم الطرائق المستعملة في عملية التقدير التي اقترحت من قبل العالم ا (Fisher) عام (1920) والتي تهدف إلى جعل دالة الامكان للمتغيرات العشوائية (random variables) أعظم ما يمكن والاكثر شيوعاً واستعمالاً، كونها تتميز بعدة خصائص منها:-[10]

1- الكفاية (Sufficient) .
2- أقل تباين (Minimum Variance) .
3- الثبات (Invariance) .
4- عدم التحيز(Unbiased) بازدياد حجم العينة.
5- الاتساق (consistency) .
6- الكفاءة (Efficiency) .

وان ايجاد قيم المعلمات التي تعظم دالة الامكان لمشاهدة واحدة وان هذه الطريقة لا تتطلب اي اشتراطات تخص المتغيرات المستقلة بمعنى انها طريقة صالحة للتقدير سواء كانت المتغيرات المستقلة اسمية أو رتبية او فئوية لذا استعملنا تقديرات الامكان الاعظم مع النماذج اللوجستية الدائرية وحسب الصيغ الاتية :

$$L_i = pin(Y_i = y_i)$$

اذ ان :

$$L_i = \binom{n_i}{v_i} p^{v_i}{}_i (1 - p_i)^{n_i - v_i} \qquad \dots(8)$$

اما الصيغة لتقدير الامكان الاعظم لـ(n) من المشاهدات فتكون :

$$L(L_i) = \prod_{i=1}^{n} \left(\binom{n_i}{v_i} p^{v_i}{}_i (1 - p_i)^{n_i - v_i} \right) \qquad \dots(9)$$

ومن خلال استخدام الدالة الاسية نحصل على التالي :

$$\frac{p_i}{1 - p_i} = e^{\beta_0 + \beta_1 cosu_i + \beta_2 sinu_i} \Rightarrow p_i = \frac{e^{\beta_0 + \beta_1 cosu_i + \beta_2 sinu_i}}{1 + e^{\beta_0 + \beta_1 cosu_i + \beta_2 sinu_i}}$$

$$L(\beta)=\prod_{i=1}^{n}\left(\begin{array}{c}\binom{n_i}{v_i}\left(\frac{e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}{1+e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}\right)^{v_i}\\ \left(1-\frac{e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}{1+e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}\right)^{n_i-v_i}\end{array}\right)\ldots(10)$$

وبأخذ اللوغاريتم للطرفين ينتج:

$$\log\ L(\beta)=\left(\begin{array}{c}\sum_{i=1}^{k}\ \log\binom{n_i}{v_i}+v_i\log\left(\frac{e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}{1+e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}\right)\\ +(n_i-v_i)\log\left(\frac{e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}{1+e^{\beta_0+\beta_1 cosu_i+\beta_2 sinu_i}}\right)\end{array}\right)\quad\ldots(11)$$

وباشتقاق الدالة آنفاً لجميع بالنسبة للمعلمات $(\beta_0,\beta_1,\beta_2)$ على الترتيب ومساواتها للصفر:

$$\frac{\log\ L(\beta)}{\partial\beta_0}=\sum_{i=1}^{k}v_i-\sum_{i=1}^{k}\frac{n_i}{1+e^{-(\beta_0+\beta_1 cosu_i+\beta_2 sinu_i)}}=0\ \ldots(12)$$

$$\frac{\log\ L(\beta)}{\partial\beta_1}=\sum_{i=1}^{k}v_i cosu_i-\sum_{i=1}^{k}\frac{n_i cosu_i}{1+e^{-(\beta_0+\beta_1 cosu_i+\beta_2 sinu_i)}}=0\qquad(12)$$

$$\frac{\log\ L(\beta)}{\partial\beta_2}=\sum_{i=1}^{k}v_i sinu_i-\sum_{i=1}^{k}\frac{n_i sinu_i}{1+e^{-(\beta_0+\beta_1 cosu_i+\beta_2 sinu_i)}}=0\qquad(14)$$

المعادلات السابقة(12)و(13)و(14) لا يمكن حلها بالطرائق التحليلية الاعتيادية لانها معادلات غير خطية ولذلك تم حلها باستعمال الطريقة العددي باستخدام طريقة نيوتن- رافسون للحصول على مقدرات طريقة الإمكان الأعظم، $(\hat{\beta}_0,\hat{\beta}_1,\hat{\beta}_2)$.

2-6-2- طريقة BYC الدائرية الحصينة: Circular BY Estimators estimator [8] [11]

تعد طريقة (BYC)احدى طرق التقدير الحصينة لأنموذج الانحدار الدائري اللوجستي وهي امتداد لطريقة الانحدار الحصين لنموذج الانحدار اللوجستي التي اقترحها الباحثان (Bianco and Yohai) في عام (1996)، حيث تستخدم طريقة(BYC) في تقدير المعلمات نموذج الانحدار اللوجستي الخطي في حالة وجود شواذ او نسبة من التلوث في البيانات إن فكرة طريقة التقدير (BYC) نشأت من خلال تحوير الأسلوب المتبع في طريقة الاعظم الصغرى إذ يتم استبدال مجموع الانحرافات للمقدار الامكان الاعظم وكالتالي:

$$\hat{\beta}_{MLE}=\arg\min_{\beta\in R^p}\sum_{i=1}^{k}D_i(\beta)\qquad\ldots(15)$$

حيث ان يبدأ التقدير باستخدام مقدر الامكان الاعظم MLE ويرمز للانحرافات (D_i) وتساوي المقدار الآتي:

$$D_i(\beta)=\{-v_i LnF(\eta_i)-(1-v_i)Ln[1-F(\eta_i)]\}\qquad\ldots(16)$$

ويمكن الحصول على المقدراتβ بطريقة (BYC) وذلك باستبدال دالة الانحرافات(D_i) بالدالة ρ كالتالي :

$$\hat{\beta}_{BYC}=\arg\min_{\beta\in R^p}\sum_{i=1}^{k}\rho\left[D\left(F(\eta_i),v\right)+G\left(1-F(\eta_i)\right)\right]\qquad\ldots(17)$$

بالنسبة للدالة ρ تستعمل Tukeys biasquare دالة الهدف

بالنسبة للدالة ρ تستعمل Tukeys biasquare دالة الهدف وهي دالة معرفة وقابلة للاشتقاق وغير متناقصة تم تعريفها من قبل الباحثين Bianco and Yoha) وتكون على النحو التالي :

$$\rho(ui)=\begin{cases}u-\left(\frac{u^2}{2m}\right), if\ u\le m\\ \frac{m}{2}, \text{other wise}\end{cases}\qquad\ldots(18)$$

حيث ان :

m:تمثل عدد المتغيرات التوضيحية في نموذج الانحدار عدد موجب .

يتم الحصول على المشتقة الجزئية للدالة $\rho(ui)$وتكون كالتالي :

$$\bar{\rho}(ui) = \psi(ui) \qquad \dots(19)$$

وان $\boldsymbol{\psi}$: هي دالة مشتقة من P .
كون الدالة التجميعية للبواقي(F) غير معلومة يتم الاعتماد على المعادلة (18) للبواقي لايجاد الدالة التجميعية الآتية:

$$G(u) = \int_0^u \psi(-Lng)dg \qquad \dots(\mathbf{20})$$

13-الجانب التجريبي: **(Simulation)**

يؤدي اسلوب المحاكاة في تمثيل الجوانب العملية دوراً مهماً في معالجة المشكلات وتنفيذها ولا سيما بعد التطور الواسع والكبير في مجال الحاسبات الإلكترونية ما دفع الكثير من الباحثين الى إعتماد اسلوب المحاكاة في الكثير من البحوث التي تهدف الى دراسة سلوك أية مقدرات أو احصاءات اختبار أو أنموذج أو توزيع احصائي نظراً لصعوبة معرفة ذلك نظرياً . [13] [4]

3 مراحل تطبيق تجارب المحاكاة ***Stages of the application of simulation experiments*** [3]

لقد تضمنت تجارب المحاكاة لهذه الدراسة كتابة عدد من البرامج بلغة (*MATLAB*) والمثبتة في الملحق رقم (*1*) في توليد البيانات لغرض المقارنة بين الطرائق بأختلاف احجام العينات ونسب تلوث مختلفة، مع الاخذ بنظر الاعتبار أن الانموذج الذي تم الاعتماد عليه في هذا البحث يكون وفق الصيغة (20) الواردة في الجانب النظري.

اذ يتم وصف مراحل تجربة المحاكاة من خلال الخطوات الآتية:

الخطوة الاولى

تعيين القيم الافتراضية لمعلمات الانموذج *(*β 0,β 1,β 2)، وتعد هذه الخطوة من الخطوات المهمة والاساسية التي تعتمد عليها الخطوات اللاحقة، اذ تم فيها اختيار قيم المعلمات والنماذج المفترضة، وكما مبين في الجدول(1) وكتالي :-

جدول (3-1)

القيم الافتراضية للمعلمات والنماذج المفترضة

CASE	β_0	β_1	β_2	T
CASE 1	**-0.5**	**2**	**-0.3**	**5%**
CASE 2	**-1.5**	**0.2**	**-0.3**	**10%**

الخطوة الثانية

توليد قيم المتغيرات التوضيحية وهي من المراحل المهمة التي تعتمد عليها المراحل الاخرى من خلال الاعتماد على اسلوب مونت- كارلو في المحاكاة حيث يتم توليد خمسة متغيرات توضيحية وبالاعتماد على الجانب التطبيقي ويكون ذلك على مرحلتين وكالآتي:

المرحلة الاولى:عندما تكون نسبة التلوث تساوي صفراً اي في حالة كون البيانات لا تحتوي على قيم شاذة، حيث يتم توليد متغيرات عشوائية على عدد المتغيرات التوضيحية.

المرحلة الثانية: في هذه الخطوة سيتم تلويث البيانات اي اضافة قيم شاذة للبيانات وبنسب تلوث مختلفة.

الخطوة الثالثة

توليد قيم المتغير التابع (y) (ثنائي الاستجابة) ، اذ ان هناك أكثر من طريقة لتوليد متغيرات تتبع توزيع فون مايسز *(von Mises distribution)* إذ يلعب هذا التوزيع دور مهم في البيانات الدائرية مشابه للدور الذي يلعبه التوزيع الطبيعي للبيانات الاعتيادية ودالة الكثافة الاحتمالية له معرفة بالشكل الآتي : [5] [9]

$$f(x) = \begin{cases} \dfrac{1}{2\pi\, I_{0(k)}} e^{\{k\cos(x-\mu)\}} & 0 \le x < 2\pi \ , \ 0 \le k < \infty \\ 0 & other\ way \end{cases} \qquad \dots(\mathbf{21})$$

اذأن μ تمثل المتوسط الدائري او متوسط الاتجاه
k معلمة التركيز

الذي يسمى تعديل دالة بيتسل (*Bessel function*) من النوع الأول و الرتبة الصفرية وهو مساوي للمقدار $I_{0(k)} = \frac{1}{2\pi}\int_0^{2\pi} e^{k\,cos(x)}dx$ أما تعديل بينسل من النوع الأول و رتبه أعلى من الصفر و لتكن الرتبة (p) فيمكن أن يكتب بالصور ألأتيه : $\sum_{r=0}^{\infty}(\frac{k}{2})^{2r}(\frac{1}{r!})^2$

$$I_{p(k)} = \int_0^{2\pi} cos(px)e^{k\,cos(x)}dx$$

أما الدالة التراكمية للتوزيع *(c.d.f)* فيمكن ان تكتب بالشكل التالي

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{1}{2\pi\, I_{0(k)}}\displaystyle\int_o^x e^{\{k\,cos(t-\mu)\}}dt & 0 \le x < 2\pi \\ 1 & x \ge 2\pi \end{cases} \qquad \dots(\mathbf{22})$$

وذلك بمتوسط $\mu = 60^{\circ}$ و قيم مختلفة لمعلمة التركيز k حيث كما يأتي :-

k	1	2	5	9	12

الخطوة الرابعة

من أهم العوامل المؤثرة والتي يتم اختيارها هي كالآتي:

*1.*اختيار احجام مختلفة للعينات بشكل يتناسب مع معرفة مدى تاثير حجم العينة على دقة وكفائة النتائج المستحصلة من طرائق التقدير المستخدمة في هذه الدراسة ,اذ تم الاعتماد على خمس احجام مختلفة للعينات المفترضة وهي $n=(10,30, 50, 70, 100)$.

*2.*اعتماد نسب مختلفة لتلوث البيانات وذلك بافتراض ثلاثة مستويات لهذه النسب وهي $(\boldsymbol{T = 5\%, 10\%})$ ، تمثل نسبة التلوث المفترضة للدراسة ل أداء طرائق التقدير تحت ظروف بيانات غير مثالية، تم إدراج نسب مختلفة من التلوث ضمن تجارب المحاكاة. وقد جرى ذلك من خلال استبدال جزء من المشاهدات الحقيقية ببيانات شاذة يتم توليدها من توزيع مختلف عن التوزيع الحقيقي للنموذج، بحيث تم تمثيل التلوث بنسب ϵ تراوحت بين مستويات منخفضة ومتوسطة وعالية. وتعكس هذه الآلية نموذج البيانات الملوّثة (*Contaminated Data Model*)، والذي يُعبَّر عنه بأنّ العينة تتكوّن من خليط من التوزيع الحقيقي وتوزيع آخر يشوّه البيانات. وتهدف هذه الخطوة إلى تقييم متانة طرائق التقدير وقدرتها على الحفاظ على خصائصها الإحصائية في ظل وجود الانحرافات والقيم المتطرفة التي يمثلها التلوث.

*3.*تكرار التجربة مساوياً الى *(1000)* مرة لكل تجربة .

الخطوة الخامسة

يتم في هذه المرحلة تقدير معلمات انموذج الانحدار اللوجستي وفق طرائق التقدير التي تم عرضها في الجانب النظري من البحث وهي كالآتي:

1- الامكان الاعظم الدائرية (*MaximumLikelihood Method*)

2- طريقة (*BYC*) الدائرية الحصينة.

جدول رقم (2)

T=5% جميع التشكيلات المدروسة عندما تكون نسبة تلويث البيانات (Bias MSE) و يوضح قيم متوسط مربعات الخطاء

n			100		150		200		250		300		Best
μ	k	method	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	
$6^{\circ}0$	1	**MLE**	0.068718	0.90596	0.05472	0.59185	0.041512	0.50841	0.039556	0.49696	0.026992	0.48568	**BYC**
		BYC	0.04792	0.70178	0.040219	0.31175	0.025726	0.50059	0.021992	0.41998	0.01304	0.39265	
	2	**MLE**	0.09804	0.70203	0.07918	0.70087	0.07712	0.7006	0.050345	0.50025	0.03817	0.47787	
		BYC	0.0556	0.6134	0.030866	0.60579	0.01709	0.50403	0.008396	0.40168	0.003697	0.48279	

	5	**MLE**	0.050546	0.8693	0.048473	0.84551	0.023563	0.74027	0.034151	0.73295	0.011408	0.72251	
		BYC	0.032591	0.8635	0.03234	0.83279	0.023267	0.82619	0.040828	0.81694	0.003387	0.40979	
	9	**MLE**	0.06245	0.85092	0.06188	0.74862	0.01877	0.74828	0.017228	0.718	0.05161	0.69562	
		BYC	0.0786	0.83027	0.077127	0.8173	0.058983	0.81523	0.056273	0.81354	0.03184	0.5943	
	12	**MLE**	0.06245	0.85092	0.066188	0.74862	0.01877	0.74828	0.027228	0.818	0.02161	0.69562	
		BYC	0.0786	0.83027	0.067127	0.8173	0.058983	0.81523	0.016273	0.71354	0.01184	0.5943	

من الجدول رقم (2)

عند القيم الافتراضية $(T = 5\%, \mu = 60^{\circ})$ **نلحظ ما يأتي:**

- تفوق طريقة طريقة(**BYC**) الحصينة على باقي طرائق التقدير اذا سجلت هذه الطريقة اقل معدل تحيز (Bias) واقل متوسط مربعات خطأ (MSE) عند كافة احجام العينات المفترضة ثم تليها طريقة (**MLE**) بالمرتبة الثانية.
- نلحظ وجود دالة متناقصة بين حجم العينة ومعدل التحيز ومتوسط مربعات الخطأ قيمة MSE تكون أفضل كلما زاد حجم العينة في أغلب الحالات بالنسبة للطريقة الحصينة .

عندما تكون (k=1)

- عند حجم عينة (100و150و250و300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE).

عندما تكون (k=2)

- عند حجم عينة (100و150و250و300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE).

عندما تكون (k=5)

- عند حجم عينة (100و150و300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE)
- عند حجم عينة (250و300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)

عندما تكون (k=9)

- عند حجم عينة (100و250و200) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)
- عند حجم عينة (150و 300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE)

عندما تكون (k=12)

- عند حجم عينة (100و1500و200) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة
- (BYC)
- عند حجم عينة (250و 300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE)

جدول رقم (3)
يوضح قيم متوسط مربعات الخطاء و Bias MSE جميع التشكيلات المدروسة عندما تكون نسبة تلويث البيانات (T=10%)

n			100		150		200		250		300	
μ	k	method	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$6^{\circ}0$	1	**MLE**	0.00328	0.050426	3.224E-03	0.050425	2.293E-	0.040425	2.157E-03	0.0401214	2.043E-03	0.0294211
		BYC	0.00143	0.043892	0.001225	0.043880	0.00166	0.043879	0.000170	0.0688554	0.0001285	0.0356937
	2	**MLE**	0.00828	0.03556	0.000258	0.01555	0.00019	0.01073	0.00014	0.0013571	0.000129	0.005446
		BYC	0.01209	0.091052	0.00741	0.04095	0.00568	0.03961	0.00187	0.028273	0.0015899	0.021338
	5	**MLE**	0.002367	0.06389	0.002174	0.06174	0.00160	0.04574	0.00164	0.044711	0.00118246	0.03066
		BYC	0.002121	0.049335	0.00204	0.04393	0.00222	0.04384	0.00204	0.035724	0.00150893	0.03163
	9	**MLE**	0.002682	0.037916	0.0013909	0.03756	0.00140	0.03289	0.0013927	0.022196	0.001045782	0.031428
		BYC	0.008508	0.053899	0.0064850	0.05372	0.00348	0.04370	0.0031850	0.043179	0.002485082	0.022651
	12	**MLE**	0.007682	0.015001	0.0033431	0.01498	0.00314	0.01497	0.0029539	0.014978	0.002349209	0.014742
		BYC	0.004508	0.038618	0.0035202	0.03856	0.00310	0.03855	0.0026766	0.03855	0.001594281	0.017873

من الجدول رقم (3)

عند القيم الافتراضية $(T = 10\% , \boldsymbol{\mu = 60^{\circ}})$ **نلحظ ما يأتي:**

- تفوق طريقة طريقة(**BYC**) الحصينة على باقي طرائق التقدير اذا سجلت هذه الطريقة اقل معدل تحيز (Bias) واقل متوسط مربعات خطأ (MSE) عند كافة احجام العينات المفترضة ثم تليها طريقة (**MLE**) بالمرتبة الثانية.
- نلحظ وجود دالة متناقصة بين حجم العينة ومعدل التحيز ومتوسط مربعات الخطأ قيمة MSE تكون أفضل كلما زاد حجم العينة في أغلب الحالات بالنسبة للطريقة الحصينة .

عندما تكون (k=1)

- عند حجم عينة (250و300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE).
- عند حجم عينة (100و200و150=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)

عندما تكون (k=2)

- عند حجم عينة (100و200و300و150و150,=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)

عندما تكون (k=5)

- عند حجم عينة (100و150و200=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE)
- عند حجم عينة (250و300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)

عندما تكون (k=9)

- عند حجم عينة (100و250و200) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)
- عند حجم عينة (150و 300=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE)

عندما تكون (k=12)

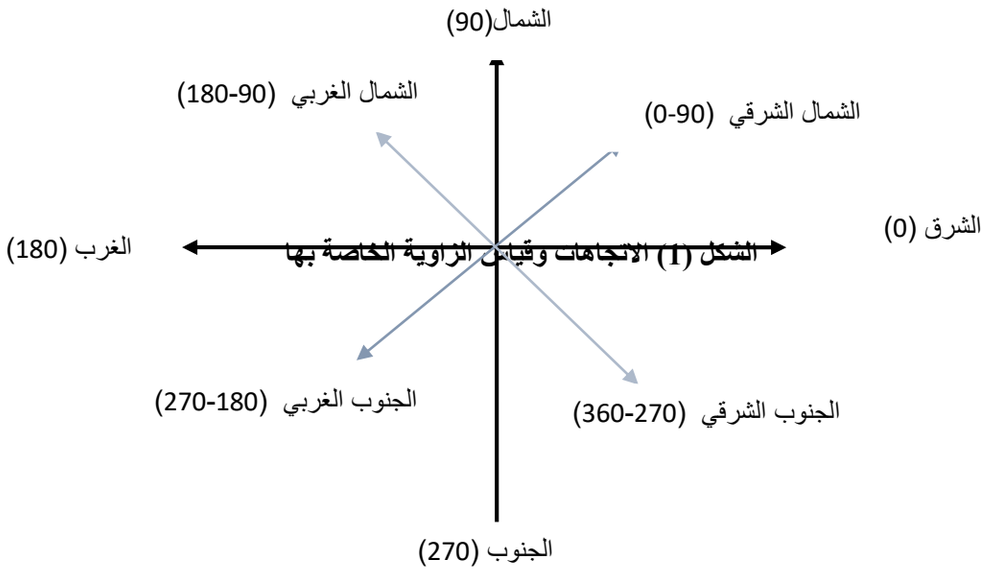
- عند حجم عينة (100و150و300) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(MLE) حيث ظهرت باقل MSE مقارنة بطريقة (BYC)
- عند حجم عينة (250و 200=n) كانت افضل طريقة لتقدير نموذج الانحدار اللوجستي هي طريقة(BYC) حيث ظهرت باقل MSE مقارنة بطريقة (MLE)

14-الجانب التطبيقي

يتضمن هذا الجانب وصف البيانات المتمثلة بأتجاه الرياح وتأثيرها على سقوط الامطار , ونظراً لأهمية اتجاه الرياح عند التخطيط لأنشاء مصدات الرياح لتفادي خطر الرياح القوية او التقليل من الاتربة او الاوساخ التي تحملها اضافة إلى أن اتجاه الرياح يساعد معرفة اوقات سقوط الامطار و وكذلك اتجاه الرياح يساعد في توجيه تدفق الهواء داخل المنازل او المباني واختيار اماكن النوافذ التي يتم فتحها ولان بيانات اتجاه الرياح هي بيانات دائرية تتطابق مع طبية البيانات الازمة للدراسة لذا تم اختيار التطبيق في هذا الجانب ,إذ سيتم التطبيق طرائق التقدير التي تم ذكرها انفا على اتجاه الرياح في مدينة كربلاء (مركز المحافظة) واتجاه الضغط الجوي الاعلى وتم اخذ البيانات من دائرة الاحوال الجوية في مدينة بغداد لشهر كانون الثاني من سنة 2024 كما في الملحق. [7]

14-1 وصف البيانات الحقيقية *Data description*

جمعت البيانات الحقيقية الخاصة وبعينة حجمها (120) الخاصة بالارصاد الجوي الخاصة بهطول الامطار في محافظة بغداد في شهر ايلول ولغرض تسهيل مهمة تحليل هذه البيانات فقد تم اعتبار متغير الاستجابة (x) هو ($x=0$) عدم وجود تساقط للامطار ،($x=1$) وجود تساقط للامطار (خط العرض 32° شمالًا وخط الطول 35 غربًا) وعلى ارتفاع 1 5 أمتار فوق سطح البحر إذ تم تقسيم اليوم إلى اربع ارباع كل ربع له فتره الزمنية ستة ساعات , الربع الأول يبدأ من الساعة (12:00 ليلا) إلى الساعة (6:00 صباحا) والربع الثاني من الساعة (6:00 صباحا) الى الساعة (12:00 ظهرا) والربع الثالث يبدأ من الساعة (12: ظهرا) إلى الساعة (6:00 مساءا أما الربع الرابع فيبدأ من الساعة (6:00 مساءا) إلى الساعة (12:00 ليلا) .وتم اعتبار اتجاه الشرق هو الاتجاه الصفري والدوران عكس اتجاه عقرب الساعة أي عندما يكون اتجاه الرياح باتجاه الشرق يأخذ الزاوية صفر او 360 وعنما يكون اتجاه الرياح بتجاه الشمال يأخذ الزاوية (90) وعنما يكون بتجاه الغرب يأخذ الزاوية (180) بينما اذا كان اتجاه الرياح نحو الجنوب تكون الزاوية (270) , أما اذا كان اتجاه الرياح بين أي من الاتجاهات فأنه يأخذ قياس زاوية بينهما أيضا , أي لو كان اتجاه الرياح شمالي شرقي فتكون قيمة الزاوية بين (الصفر و 90) , واذا كان اتجاه الرياح شمال غربي فان قياس الزاوية يكون بين (90 إلى 180) واذا كان اتجاه الرياح جنوب غربي فيكون قياس الزاوية بين (180 إلى 270) واما اذا كان اتجاه الرياح جنوب شرقي فان قياس الزاوية يكون بين (270 إلى 360) وكما موضح بالشكل ادناه . [1] [15]



الشكل (1) الاتجاهات وقياس الزاوية الخاصة بها

الجدول (4) يبين البيانات الحقيقية الخاصة بأتجاه الرياح (بالراديان) حيث ان القيم السالبة تمثل الرياح من اتجاهات غربية إلى جنوبية غربية القيم الموجبة تمثل الرياح من اتجاهات شرقية إلى شمالية شرقية والقيمة القيمة −3.14 و +3.14 تمثلان نفس الاتجاه.

y	θ	i	y	θ	i	y	θ	I
1	2.24	77	0	0.24	39	0	-1.77	1
1	2.3	78	0	0.29	40	0	-1.44	2
1	2.35	79	0	0.34	41	0	-1.66	3
1	2.4	80	0	0.1	42	0	-1.61	4
1	2.46	81	0	0.45	43	0	-1.56	5
1	2.53	82	0	0.5	44	0	-1.5	6
1	2.56	83	1	0.55	45	0	-1.45	7
1	2.6	84	0	0.91	46	0	-1.4	8
1	2.67	85	1	0.66	47	0	-1.35	9
1	2.72	86	1	0.71	48	0	-1.39	10
1	2.77	87	1	0.7	49	0	-1.24	11
1	2.82	88	1	0.82	50	0	-1.19	12
1	2.88	89	1	0.87	51	0	-1.14	13
1	2.93	90	1	0.92	52	0	-1.08	14
1	2.98	91	1	0.98	53	1	-1.03	15
1	3.04	92	1	1.03	54	0	-0.98	16
1	-3	93	1	1.08	55	0	-0.92	17
1	3.14	94	0	1.14	56	0	-0.87	18
1	-3.14	95	1	1.19	57	1	-0.82	19
1	-3.09	96	1	1.24	58	0	-0.77	20

1	-3.04	97	1	1.29	59	0	-0.71	21
1	-2.98	98	1	1.35	60	0	-0.66	22
1	-2.93	99	1	1.4	61	0	-0.63	23
1	-2.88	100	1	1.45	62	0	-0.55	24
1	-2.82	101	1	1.5	63	0	-0.5	25
1	-2.77	102	1	1.56	64	0	-0.45	26
1	-2.72	103	1	1.61	65	0	-0.4	27
1	-2.67	104	1	1.66	66	0	-0.34	28
1	-2.61	105	1	1.72	67	0	-0.29	29
1	-2.56	106	0	1.77	68	0	-0.53	30
0	-2.51	107	1	1.82	69	0	-0.18	31
1	-2.46	108	1	1.87	70	0	-0.13	32
1	-2.4	109	1	1.93	71	1	-0.08	33
1	-2.35	110	1	2.03	72	0	-0.03	34
0	-2.3	111	1	2.03	73	0	0.03	35
1	-2.24	112	1	2.09	74	1	0.09	36
0	-2.19	113	1	2.14	75	0	0.13	37
1	-2.14	114	1	2.19	76	0	0.21	38
0	-1.98	117	1	-2.03	116	1	-2.09	115
0	-1.82	120	0	-1.87	119	0	-1.93	118

15-الطرق الاستكشافية لبيانات المتغيرات التوضيحية Exploratory Methods for data of Explanatory variables

هناك عدة طرق يتم الاعتماد عليها لفحص البيانات في حالة احتوائها على قيم شاذة(Outlier Value) ومن هذه الطرق:

16- التحليل باستخدام Cook's Distance للانحدار اللوجستي الدائري [13] [12]

تم تطبيق تحليل Cook's Distance على بيانات الحقيقية قيد الدراسة للتحقق من وجود مشاهدات شاذة على النموذج

ويوضح الجدول (5) أعلى خمس مشاهدات من حيث التأثير على تقديرات النموذج:

i	قيمة المتغير المستقل	الاستجابة	Cook's D
11	-1.24	1	0.15
30	-0.53	0	0.12
52	0.92	1	0.10
60	1.35	1	0.08
68	1.72	0	0.07

من خلال الجدول (5)

- نتائج اختبار Cook's Distance حيث يتبين أن المشاهدات رقم (11) و(30) سجلت أكبر القيم لاختبار Cook's Distance اذ سجلت (0.15),(0.12) على التوالي مما يدل على ان لهما تأثيرًا قوي وملحوظ على النموذج.
- أن المشاهدات رقم (52) و(60)و(68) سجلت قيم متوسطة لاختبار Cook's Distance اذ سجلت (0.10),(0.08),(0.07) على التوالي مما يدل على ان لهما تأثيرًا اقل على النموذج.

17الطريقة المستخدمة في تقدير المعلمات **Method used in Estimate parameter**

سيتم في هذا الجزء تقدير معلمات الانموذج بأستعمال طريقة الامكان الاعظم التي ظهرت افضليتها في الجانب التجريبي (MLE)وذلك على أساس انها تعطي اقل(MSE) للانموذج من الطرائق الأخرى اذ تم استخراج قيم وكما في الجدول الآتي:

Param. / Method	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
MLE	0.8	2.6	-2.5

18-الاختبارات المعنوية للنماذج المقدرة Goodness-of-fit test for estimated models

لاختبار العلاقات المقدرة في النماذج الثلاثة التي تم تقديرها سنختبر الفرضة التي ينص فرض العدم منها على أن العلاقة المقدرة غير معنوية أما الفرض البديل فينص على أن العلاقة المقدرة معنوية وكما يلي :

الجدول(6) اختبارات النماذج بالنسبة للبيانات الحقيقية

Method	Nagelkerke_R2	pValue_Model
MLE	0.734	.000
BYC	0.730	.000

نلاحظ من الجدول(6)في أعلاه أن قيمة اختبار Nagelkerke_R2 المحتسبة في طريقة MLE هي أعلى من القيمة الجدولية بمستوى ثقة مقداره (0.95) اذ سجلت ومن ثم يكون القرار هو رفض فرضية العدم أي أن العلاقة المقدرة هي علاقة معنوية وأن الأنموذج هو مثال للبيانات بصورة صحيحة، أما بالنسبة في طريقة (BYC) نلاحظ أن قيمة اختبار Nagelkerke_R2 المحتسبة اذ سجلت0.730 أعلى من القيمة الجدولية بمستوى ثقة مقداره (0.95) ي أن العلاقة المقدرة هي علاقة معنوية وأن الأنموذج هو مثال للبيانات بصورة صحيحة وكذلك بالنسبة لبقية طرائق التقدير.

19-الاستنتاجات

1- اظهرت طريقة الامكان الاعظم*BYC* أداءً أفضل مقارنةً بطرائق قيد الدراسة ، حيث حققت أقل معايير للمقارنة وكانت فعالة بشكل ملحوظ بناءً عليه، تُعد طريقة الامكان الاعظم اقل كفاءة.
وقد اخت طريقة الامكان الاعظم بالمرتبة الثانية من حيث دقة التقديرات بناءً على معيار *MSE* بالنسبة لطرائق التقدير كافة.
2- قدرة الطرائق المدروسة على التعامل مع نموذج الانحدار الدائري وتقديم مقدرات جيدة على العموم لهذا النوع من النماذج .
3- على الرغم من وجود تذبذب في قيم متوسط مربعات الخطأ (*MSE*) بين طرائق التقدير المختلفة، فإنّ معظم القيم تقع ضمن نطاق مقبول إحصائياً، إذ إنها قريبة من الصفر، مما يدل على كفاءة معتبرة في أداء طرائق التقدير قيد الدراسة.
4- تنادأ إلى الجانب التطبيقي، أظهرت طريقة *BYC* أداءً متميزاً؛ إذ حقق معامل التحديد *R2* قيمة بلغت 0.734، وهي قيمة تتجاوز القيمة الجدولية المقابلة لمستوى الثقة 0.95. ويعكس هذا التفوق أن النموذج المقدَّر يتمتع بدلالة إحصائية عالية، فضلاً عن امتلاكه قدرة تفسيرية ملائمة لتمثيل البيانات الحقيقية قيد الدراسة بشكل موثوق.

20-التوصيات :-

بناءً على الاستنتاجات السابقة يمكن بالتوصيات الاتية :

1. يوصي الباحث باستعمال طريقة ***(BYC)*** في تقدير معلمات انموذج الانحدار الدائري في حالة وجود قيم شاذة في مقدرات المتغير التابع وحسب نتائج التطبيق وكذلك في حال وجود قيم ذات تأثير عال في المتغيرات التوضيحية .
2. يمكن التوصية بمعرفة مدى تأثر زيادة نسبة التلوث (زيادة عدد الشوارد وزيادة مستوى شرودها) على كفاءة الطريقة الحصينة وطرائق عمق الانحدار.
3. توسيع استعمال الانحدار الدائري في الدراسات الاقتصادية والاجتماعية والطبية.
4. ضرورة اعتماد توثيق لحالات الارصاد الجوي وتجاه الرياح وذلك من اجل توفير قاعدة بيانات لتسهيل عمل الباحثين .

المصادر الاجنبية:

1. Ismail, M. A. (2001). Linear Regression Analysis. Institute of Public Administration, Riyadh, Kingdom of Saudi Arabia.

2. Al-Sabbagh, A. A. S. (2009). Building a Regression Model for Children's Malnutrition in Karbala Governorate (Doctoral dissertation, University of Baghdad, Iraq).
3. Alaa, E. S. (2003). Comparison Between SIMEX and LLS Methods for Estimating the Regression Function Using Simulated Data (Master's thesis, College of Administration and Economics, University of Baghdad).
4. Al-Muqdad, R. S. M. (2013). Robust and Modified Methods in Analyzing Iraqi Data with an Application on the Circular Model (Doctoral dissertation, College of Administration and Economics, University of Baghdad).
5. Fahmi, M. B. (2005). SPSS Statistics Without Pain: A Practical Guide with Applications on SPSS Program. Institute of Public Administration, Riyadh, Kingdom of Saudi Arabia.
6. Qasim, M. A. A. (2019). Analytical and Numerical Solutions for Some Fractional Differential Equations (Doctoral dissertation, College of Graduate Studies, University of Baghdad).
7. Abuzaid, A. H., & Ahmed, H. I. E. (2021). On outliers detection in circular logistic regression. J. Appl. Probab. Stat, 16, 95-110.
8. Abuzaid, A. H., & Allahham, N. R. (2015). SIMPLE CIRCULAR REGRESSION MODEL ASSUMING WRAPPED CAUCHY ERROR. Pakistan Journal of Statistics, 31(4).
9. Abuzaid, A. H., Mohamed, I. B., & Hussin, A. G. (2012). Boxplot for circular variables. Computational Statistics, 27, 381-392.
10. Abuzaid, A., Mohamed, I., Hussin, A. G., & Rambli, A. (2011). COVRATIO statistic for simple circular regression model. Chiang Mai Journal of Science, 38(3), 321-330.
11. Akhtaruzzaman, M. D., Shafie, A. A., & Khan, M. R. (2016). Gait analysis: Systems, technologies, and importance. Journal of Mechanics in Medicine and Biology, 16(07), 1630003.
12. **Al-Daffaie, K., & Khan, S. (2017, May). Logistic regression for circular data. In AIP conference proceedings (Vol. 1842, No. 1, p. 030022). AIP Publishing LLC.**
13. Al-Daffaie, K., & Khan, S. (2018). Performance enhancement of AlGaN-based UV-LEDs inserted with a pin-doped AlGaN layer between the active region and electron-blocking layer. Materials Science in Semiconductor Processing, 83, 133-138.