

بناء أنموذج أنحدار شبه معلمي باستخدام بيانات دائرية

Building a semi-parametric regression model using circular data

أ.م.د. أينااس عبد الحافظ محمد²

Enas Abdul Hafedh Mohammed

جامعة كربلاء كلية الإدارة والاقتصاد قسم الإحصاء

enas.albasri@uokerbala.edu.iq

م.م. صبيحة نعمه ضهد¹

Sabiha N. Dhahad

جامعة كربلاء كلية الإدارة والاقتصاد قسم الإحصاء

sabihand@stu.edu.iq

المستخلص :

يتطلب الحفاظ على الوظيفة البصرية تحسينها بفهم الخطأ الأنكساري (Refractive Error) والتغيرات المرتبطة بالعمر ، تعمل العين البشرية كنظام بصري يركز الصور المرئية على الشبكية ، إذ أن الخطأ الأنكساري هو أحد الأسباب الرئيسية لتدهور صورة الشبكية في العيون غير المصححة (Namba, H., Sugano & others , 2020) المصدر : علاقة العمر بالاستجماتزم هدف البحث لبناء أنموذج أنحدار شبه معلمي باستخدام الأنموذج شبه المعلمي الدائري-الخطي-الدائري المدمج من خلال دمج الجزء المعلمي المقدر بطريقة الأماكن الأعظم وبفرض الخطأ العشوائي بتوزيع مشترك (خطي-دائري) والجزء اللامعلمي المقدر باستخدام دوال نواة دائرية (Circular Kernel) مختلفة بواسطة معلمة الدمج (α) ، تم استعمال بيانات خاصة بالبصر وتم الحصول عليها من (مركز النور التخصصي لطب وجراحة العيون) في محافظة ذي قار وباستعمال جهاز (Auto Kerato – Refracto Tonometer TOPCON TRK. 2P) .

تضمنت عينة الدراسة (400 حالة) تمثل لبيانات قياس العين اليمنى (OD cyl axis) والعين اليسرى (OD cyl axis) وعمر الحالة (Patient Age) لأشخاص يعانون من الخطأ الأنكساري (Refractive Error) .

الكلمات المفتاحية : أنموذج الأنحدار شبه المعلمي ، دالة النواة

Abstract:

Maintaining visual function requires improving it by understanding refractive error and age-related changes. The human eye acts as an optical system that focuses visual images on the retina. Refractive error is one of the main causes of retinal image deterioration in uncorrected eyes (Namba, H., Sugano & others, 2020). Source: Age-astigmatism relationship

The aim of the research is to build a semiparametric regression model using the integrated circular-linear-circular semiparametric model by integrating the parameter part estimated by the maximum likelihood method and assuming a random error with a joint distribution (linear-circular) and the nonparametric part estimated using different circular kernel functions by the integration parameter (α). Visual data were used and obtained from (Al Noor Specialized Center for Ophthalmology and Eye Surgery) in Thi Qar Governorate using (Auto Kerato – Refracto Tonometer TOPCON TRK. 2P) device. The study sample included (400 cases) representing the data of measuring the right eye (OD cyl axis) and the left eye (OD cyl axis) and the age of the case (Patient Age) for people suffering from refractive error (Refractive Error).

Keywords: Semi-parametric regression model, Kernel function.

1- المقدمة : The Introduction

يتطلب الحفاظ على الوظيفة البصرية تحسينها بفهم الخطأ الأنكساري (Refractive Error) والتغيرات المرتبطة بالعمر ، تعمل العين البشرية كنظام بصري يركز الصور المرئية على الشبكية ، إذ أن الخطأ الأنكساري هو أحد الأسباب الرئيسية لتدهور صورة الشبكية في العيون غير المصححة (Namba, H., Sugano & others , 2020) .

من هنا جاء البحث لبناء أنموذج احصائي نستخدم تحليل الأنحدار كأداة احصائية لهذا الغرض من خلال تقدير العلاقة بين متغيرات الأنموذج . تضمن البحث أربع مباحث ، المبحث الأول يحوي المقدمة ومنهجية البحث . ركز المبحث الثاني على الدراسة النظرية لأنموذج الأنحدار شبه المعلمي باستخدام بيانات دائرية ، تطرق لبناء أنموذج الأنحدار شبه المعلمي المدمج وطرق تقدير الجزء المعلمي والجزء اللامعلمي بواسطة معلمة الدمج . المبحث الثالث تناول الجانب التجريبي والتطبيقي من خلال إجراء خطوات المحاكاة (Simulation) باستعمال برنامج (Matlab) لبناء أنموذج البحث ، في الجانب التطبيقي تم استعمال بيانات مرضى الخطأ الأنكساري (Refractive Error) في تطبيق الأنموذج شبه المعلمي الدائري-الخطي-الدائري .

2. منهجية البحث

1-2 الأهمية : Importance of thesis

تكم أهمية الأطروحة ببناء أنموذج أنحدار شبه معلمي يجمع بين المرونة والدقة وتحقيق موثوقية عالية في نتائج التقدير والتنبؤ بما يتوافق مع طبيعة البيانات الدائرية ، وتظهر أهمية هذا الأنموذج من خلال التطبيق على بيانات دائرية حول ظاهرة الخطأ الأنكساري للعين .

2-2 مشكلة البحث : Problem of thesis

تكمّن مشكلة البحث بوجود العديد من الظواهر في الواقع التي لها خاصية دورية وتكون بياناتها ضمن المدى $(0, 2\pi)$ وهي من البيانات التي لها أهمية في مجالات عديدة ويواجه تحليل هذه البيانات العديد من التحديات المرتبطة بتحليل البيانات الدائرية كأختلاف التوزيع للبيانات الدائرية والخطية التي أعتمد عليها الأنموذج أيضاً طبيعة البيانات الدائرية التي تتكرر كل 360° مما يجعل صعوبة في التفسير، كما يوجد أختلاف في طرق تقدير المتوسط الدائري والأنحراف المعياري الدائري.

3-2 هدف البحث : Aim of thesis

يهدف البحث الى وبناء انموذج انحدار شبه معلمي بأستخدام بيانات دائرية يتناسب مع طبيعة البيانات الدائرية و يفسر العلاقة بين متغيرات الدراسة ،

2- الجانب النظري

1-3 مفهوم البيانات الدائرية : The concept of circular data

البيانات الدائرية هي نوع من البيانات التي تمثل نقاطاً على محيط دائرة الوحدة المقاسة بالدرجات أو بالراديان. على مدى القرن الماضي ، تم التركيز بشكل كبير على البيانات الدائرية في أنواع مختلفة من المجالات مثل الفيزياء وعلم النفس والأرصاد الجوية والجيولوجيا وعلم الحفريات وعلم الأحياء وعلم الفلك وغيرها الكثير. يمكن تعريف البيانات الدائرية في أنواع مختلفة من التوزيعات مثل التوزيعات النقطية والقلبية والمثلثية وتوزيعات فون ميزس (VM) والملفوفة العادية (WN) وتوزيعات بواسون (WP). توزيع VM هو التوزيع الأكثر شيوعاً للبيانات الدائرية لأنه يأخذ الدور الذي يؤديه التوزيع الطبيعي في الإحصائيات الخطية القياسية. في الواقع ، يتشكل مثل التوزيع الطبيعي ، فيما عدا أن ذبوله مقطوعة. على عكس البيانات الخطية ، أصل البيانات الدائرية غير محدد لأنها موجودة على الأسطح المحيطية والكروية. يمكن أن يؤثر الاختلاف في خصائص البيانات على الوصف والاستدلال في التحليل الإحصائي. لذلك ، لم تعد طريقة الإحصاء الخطي الكلاسيكية مناسبة عند معالجة البيانات الدائرية. (Mohammad, 2021) ، يتم جمع البيانات الدائرية عندما يكون موضوع الاهتمام هو اتجاه أو وقت من اليوم. تظهر هذه البيانات الخاصة في العديد من التطبيقات: علوم الأرض (مثل اتجاهات الرياح) ، والطب (مثل إيقاع الساعة البيولوجية) ، والبيئة (مثل حركات الحيوانات) ، والطب الشرعي (وقوع الجريمة) (Claire and Thanh, 2022)

2-3 الأنحدار شبه المعلمي : the semiparametric regression

أقترح الباحث (Robinson, 1988) والباحث (Speckman Paul, 1988) دالة الأنحدار الخطي الجزئي وهي أحد دوال الأنحدار شبه المعلمي ، وعرف (Issa, Aseel Muslim. 2011) الأنحدار شبه المعلمي (semiparametric) بأنه أسلوب إحصائي يحقق خصائص عامة للأنحدار المعلمي (Parametric) واللامعلمي (Nonparametric) ويحقق نفس الهدف وهو الحصول على منحني للبيانات يطابق أو يقترب من التوافق لمنحنى المتغير المعتمد وذلك بدمج أساليب التقدير المعلمية واللامعلمية . (Tarrad, Alaa Jaber & Hassan, Raad Fadel. 2013) ويعبر عن النموذج شبه المعلمي بالآتي :

$$y_i = f(x_i, B) + g(\theta_i) + \varepsilon_i \quad \dots (1)$$

أذ :

y_i المتغير التابع ، $f(x_i, B)$ الجزء المعلمي ، $g(\theta_i)$ الجزء اللامعلمي ، ε_i الخطأ العشوائي
يجمع الأنحدار شبه المعلمي بين مكونات الأنحدار المعلمي واللامعلمي . بذلك يضمن سهولة التفسير من الجزء المعلمي ويحتفظ بالمرونة من الجزء اللامعلمي . (DERICK L. RIVERS, 2009)

3-3 الأنحدار الدائري (Circular regression)

عند أستخدام البيانات الدائرية لأعداد الأنحدار الخطي البسيط او المتعدد يجب التمييز بين الدور الذي يلعبه المتغير الدائري وبين وقت استخدام تقنيات الأنحدار الخطي وغير الخطي ، توجد ثلاث أنواع لنماذج الأنحدار الدائري موضحة في الجدول ادناه ، في كل نوع ، يمكن استخدام كل من أشكال الأنحدار الخطي وغير الخطي. (Scott, 2002)

جدول (1) يبين أنموذجات الأنحدار الدائري

Regression category	Response Variable	Explanatory Variable
linear –circular	Linear	Circular
circular –linear	Circular	Linear
circular –circular	Circular	Circular

لبناء أنموذج الأنحدار شبه المعلمي المدمج بأستخدام بيانات دائرية تم توظيف نوعين من الأنحدار الدائري في الجدول السابق في صياغة انموذج الدراسة الحالية ، أذ تضمن الجزء المعلمي انموذج أنحدار من النوع (Circular - Linear Regression) اما الجزء اللامعلمي تضمن أنموذج أنحدار من النوع (Circular- Circular Regression) ، تم جمع الأنموذجين المعلمي واللامعلمي للحصول على الأنموذج شبه المعلمي الدائري-الخطي- الدائري بالجدول ادناه:

جدول رقم (2) يبين أنموذج الأندحار شبه المعلمي الدائري - الخطي- الدائري

Regression category	Dependent Variable	Explanatory Variable	
Circular = Linear + Circular	Circular	Linear	Circular

الأندحار شبه المعلمي الدائري - الخطي- الدائري (Semiparametric Circular -Linear- circular Regression) ، الأندحار شبه المعلمي أحد اشكال الأندحار التي من الممكن توظيف البيانات الدائرية لهذا النوع من النماذج ، يتكون من دالتين احدهما دالة (Circular - Linear Regression) هي $(y_i = \beta_0 + \beta_1 x_i)$ ومتغير معتمد دائري (y_i) ومتغير توضيحي خطي (x_i) اما الدالة الثانية (Circular- Circular Regression) هي $(y_i = g(\theta))$ بمتغير معتمد دائري وبتغير توضيحي دائري . أنموذج الأندحار شبه المعلمي من النوع (Circular = Linear + Circular) وهو من اللأنموذجات التجميعية .

4-3 تقدير أنموذج الأندحار شبه المعلمي الدائري - الخطي- الدائري المدمج

Estimating the parameters of the Combined semiparametric Circular -Linear- circular regression model

استخدمت هذه الطريقة بالتقدير من قبل الباحثان (Olkin&Spiegelman,1987) حيث تم دمج الصيغة المعلمية واللامعلمية لدالة الكثافة الاحتمالية.

قدم الباحثان (Burman&Chaudhuri,1992) تقدير لدالة الأندحار بأستخدام خليط من مقدرات معلمية ولامعلمية . ثم قدم الباحث (Wooldridge , 1992) اقتراح خاص بهذا النوع من المقدرات ، ايضاً قدم الباحثون (Rahman,Gokhale&Ullah, 1993) مقترح للتقدير المدمج ، قدمت الباحثة اسيل عام 2011 والباحث مناف يوسف عام 2012 مقترح للتقدير شبه المعلمي المدمج . تمثل هذه الطريقة الأندحار الهجين بين الأنموذج ذو الصيغة المعلومة والمعلمت المجهولة ، والأنموذج ذو دالة الأندحار غير معلومة الصيغة ، تم في هذه الدراسة تقدير دالة الأندحار المدمج وفق طريقة (Burman&Chaudhuri,1992) وحسب الصيغة :

$$y_i = (1 - \alpha)f(x_i, \beta) + \alpha g(\theta_i) + \varepsilon_i \quad \dots (2)$$

حيث $0 < \alpha < 1$ تمثل معلمة الدمج (Combin Parameter) تم التقدير وفقاً للخطوات التالية :

- 1- تقدير الجزء المعلمي بطريقة الأماكن الأعظم (Maximum Likelihood Estimator)
- 2- تقدير الجزء اللامعلمي بواسطة المقدر اللبي (Kernel Estimator)
- 3- تقدير معلمة الدمج (Combin Parameter) بطريقة المربعات الصغرى من خلال تعويض كل من المقدرين المعلمي واللامعلمي بالشكل التالي :

$$y_i = (1 - \alpha)f(x_i, \hat{B}) + \alpha \hat{g}(\theta_i) + \varepsilon_i \quad \dots (3)$$

تعد معلمة الدمج (Combin Parameter) من المعلمت المهمة في تكوين الأندحار شبه المعلمي إذ انها تعطي وزناً لكل من الجزء المعلمي والجزء اللامعلمي في الأنموذج شبه المعلمي ، وتم استخدامها في الدراسة الحالية لتكوين أنموذج اندحار شبه معلمي لبيانات دائرية ويستوجب تقديرها ومن المتعارف عليه يمكن تقدير معلمة الدمج بطريقة المربعات الصغرى:

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [y_i - (f(x_i, \hat{B}) + \alpha f(x_i, \hat{B}) - \alpha \hat{g}(\theta_i))]^2$$

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [y_i - (f(x_i, \hat{B}) + \alpha f(x_i, \hat{B}) - \alpha \hat{g}(\theta_i))]^2$$

$$= \sum_{i=1}^n [y_i - f(x_i, \hat{B}) + \alpha(\hat{g}(\theta_i) - f(x_i, \hat{B}))]^2$$

$$(y_i - f(x_i, \hat{B})) = \alpha(\hat{g}(\theta_i) - f(x_i, \hat{B})) + \varepsilon_i$$

نضرب الطرفين بالمقدار $(\hat{g}(\theta_i) - f(x_i, \hat{B}))$ ونحصل على الصيغة الآتية لمعلمة الدمج المقدره :

$$\hat{\alpha}_{BC} = \frac{\sum_{i=1}^n (y_i - f(\hat{B}, x_i)) (\hat{g}(\theta_i) - f(\hat{B}, x_i))}{\sum_{i=1}^n (\hat{g}(\theta_i) - f(\hat{B}, x_i))^2} \quad \dots (4)$$

حيث $\hat{g}(\theta_i) \neq f(\hat{B}, x_i)$

(Rahman, M., Gokhale, D. V., & Ullah, A. (1997))

$\hat{\alpha}$ تمثل مقدر المربعات الصغرى للمعلمة α ، في حالة $\hat{\alpha} = 0$ فإن الصيغة (3) تمثل مقدر دالة اندحار معلمي ، اما اذا كانت $\hat{\alpha} = 1$ فإن الصيغة (3) تمثل مقدر دالة اندحار لامعلمي . (Ullah, A., & Vinod, H. D. ,1993) . (Fan, ((Ullah, A., & Vinod, H. D. ,1993)) . (Y., & Ullah, A. 1999)

1-4-3 تقدير الجزء المعلمي بطريقة الأماكن الأعظم (Maximum Likelihood) :

تتميز طريقة الأماكن الأعظم بأنها أكثر الطرق شيوعاً وتتميز بالكفاءة ، الكفاية ، الأتساق وتتميز بخاصية الثبات .

Allahham, N. (2015)

انموذج الجزء المعلمي مشابه لأنموذج الانحدار البسيط حيث يشترط في الأنموذج الخطي البسيط ان يتوزع الخطأ توزيع طبيعي ، وبما ان توزيع فون ميزس (VM) للبيانات الدائرية مشابه للتوزيع الطبيعي للبيانات الكمية ويمكن صياغة الأنموذج بالصيغة التالية (Jawad, Ali Muhammad. 2023)

$$y_i = \beta_0 + \beta x_i + \varepsilon_i \pmod{2\pi} \quad \dots (5)$$

y_i : المتغير التابع الدائري Dependent Variable

x_i : المتغير التوضيحي المستقل inDependent Variable

β_0 : الحد الثابت للأنموذج

β : الميل الحدي للأنموذج

ε_i : الخطأ العشوائي الذي يتبع توزيع فون ميزس الشرطي (VM) بمتوسط اتجاه صفر ومعدل التركيز $(\frac{vx}{\sigma^2})$

بما ان المتغير (y_1, y_2, \dots, y_n) يتوزع توزيع فون ميزس الدائري له دالة كثافة احتمالية (pdf) لكل مشاهدة في العينة n ، يمكن تعريف دالة الكثافة الاحتمالية لتوزيع فون ميزس للخطأ العشوائي الدائري لأنموذج الانحدار كالاتي :

Abuzaid, A. H. (2010)

$$f(y/x; \beta_0, \beta_1, \sigma^2, k.) = \frac{1}{2\pi I_0(\frac{vx}{\sigma^2})} \exp\left\{\frac{vx}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i)\right\}$$

$$L(y_i, x_i; \beta_0, \beta_1, \sigma^2, v) = \prod_{i=1}^n \frac{1}{2\pi I_0(\frac{vx}{\sigma^2})} \exp\left\{\frac{vx}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i)\right\}$$

نأخذ اللوغاريتم لدالة الأماكن الأعظم :

$$\ln L(y_i, x_i; \beta_0, \beta_1, \sigma^2, v) = \sum_{i=1}^n \ln f(y/x; \beta_0, \beta_1, \sigma^2, v)$$

$$= \sum_{i=1}^n \ln \left\{ \frac{1}{2\pi I_0(\frac{vx}{\sigma^2})} \exp\left(\frac{vx}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i)\right) \right\}$$

$$= \sum_{i=1}^n \left\{ \ln \frac{1}{2\pi I_0(\frac{vx}{\sigma^2})} \exp\left(\frac{vx}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i)\right) \right\}$$

$$= \sum_{i=1}^n \left\{ -\ln(2\pi) - \ln\left(I_0\left(\frac{vx}{\sigma^2}\right)\right) + \frac{vx}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i) \right\}$$

نقوم بأجراء الأشتقاق الجزئي لكل من المعلمات $(\beta_0, \beta_1, \sigma^2, v)$ المطلوب تقديرها :

$$\frac{\partial \ln L}{\partial \beta_0} = \sum_{i=1}^n \frac{vx_i}{\sigma^2} \sin(y_i - \beta_0 - \beta_1 x_i) \quad \dots (6)$$

$$\frac{\partial \ln L}{\partial \beta_1} = \sum_{i=1}^n \left[\frac{vx_i}{\sigma^2} \sin(y_i - \beta_0 - \beta_1 x_i) x_i \right] \quad \dots (7)$$

$$\frac{\partial \ln L}{\partial v} = \sum_{i=1}^n \left[\frac{-x_i I_1\left(\frac{vx_i}{\sigma^2}\right)}{\sigma^2 I_0\left(\frac{vx_i}{\sigma^2}\right)} + \frac{x_i}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i) \right] \quad \dots (8)$$

$I_1\left(\frac{vx_i}{\sigma^2}\right)$ هي المشتقة الأولى ل $I_0\left(\frac{vx_i}{\sigma^2}\right)$

$$\frac{\partial \ln L}{\partial \sigma^2} = \sum_{i=1}^n \left[\frac{vx_i^2 I_1\left(\frac{vx_i}{\sigma^2}\right)}{(\sigma^2)^2 I_0\left(\frac{vx_i}{\sigma^2}\right)} - \frac{vx_i^2}{2(\sigma^2)^2} \cos(y_i - \beta_0 - \beta_1 x_i) \right] \dots (9)$$

نجعل المعادلات (6) و(7) و(8) و(9) مساوية للصفر نحصل :

$$\sum_{i=1}^n \frac{vx_i}{\sigma^2} \sin(y_i - \beta_0 - \beta_1 x_i) = 0 \quad \dots (10)$$

$$\sum_{i=1}^n \left[\frac{vx_i}{\sigma^2} \sin(y_1 - \beta_0 - \beta_1 x_i) x_i \right] = 0 \quad \dots (11)$$

$$\sum_{i=1}^n \left[\frac{-x_i I_1 \left(\frac{vx_i}{\sigma^2} \right)}{\sigma^2 I_0 \left(\frac{vx_i}{\sigma^2} \right)} + \frac{x_i}{\sigma^2} \cos(y_i - \beta_0 - \beta_1 x_i) \right] = 0 \quad \dots (12)$$

$$\sum_{i=1}^n \left[\frac{vx_i^2 I_1 \left(\frac{vx_i}{\sigma^2} \right)}{(\sigma^2)^2 I_0 \left(\frac{vx_i}{\sigma^2} \right)} - \frac{vx_i^2}{2(\sigma^2)^2} \cos(y_i - \beta_0 - \beta_1 x_i) \right] = 0 \quad \dots (13)$$

وبتطبيق خاصية الوسط الحسابي المذكورة في المعادلة رقم () على المعادلة رقم (2-14) نحصل على :

$$\sum_{i=1}^n [\sin(y_i - \beta_1 x_i) \cos \beta_0 - \cos(y_i - \beta_1 x_i) \sin \beta_0] = 0$$

$$\cos \beta_0 \sum_{i=1}^n \sin(y_i - \beta_1 x_i) - \sin \beta_0 \sum_{i=1}^n \cos(y_i - \beta_1 x_i) = 0$$

بالقسمة على $\cos \beta_0$

$$\sum_{i=1}^n \sin(y_i - \beta_1 x_i) - \frac{\sin \beta_0}{\cos \beta_0} \sum_{i=1}^n \cos(y_i - \beta_1 x_i) = 0$$

$$\tan \beta_0 = \frac{\sin \beta_0}{\cos \beta_0}$$

اذن

$$\sum_{i=1}^n \sin(y_i - \beta_1 x_i) - \tan \beta_0 \sum_{i=1}^n \cos(y_i - \beta_1 x_i) = 0$$

$$\tan \beta_0 = \frac{\sum_{i=1}^n \sin(y_i - \beta_1 x_i)}{\sum_{i=1}^n \cos(y_i - \beta_1 x_i)}$$

$$\beta_0 = \tan^{-1} \frac{\sum_{i=1}^n \sin(y_i - \beta_1 x_i)}{\sum_{i=1}^n \cos(y_i - \beta_1 x_i)}$$

$$\hat{\beta}_0 = \begin{cases} \tan^{-1} \left(\frac{S}{C} \right) & \text{if } S \geq 0, C > 0 \\ \frac{\pi}{2} & \text{if } S \geq 0, C = 0 \\ \tan^{-1} \left(\frac{S}{C} \right) + \pi & \text{if } C < 0 \\ \tan^{-1} \left(\frac{S}{C} \right) + 2\pi & \text{if } S < 0, C \geq 0 \\ \text{undefined} & \text{if } S = 0, C = 0 \end{cases} \quad \dots (14)$$

حيث :

$$S = \sum_{i=1}^n \sin(y_i - \beta^0 x_i)$$

$$C = \sum_{i=1}^n \cos(y_i - \beta^0 x_i)$$

حيث: β^0 تمثل القيم الأولية للمعاملات المقدره
 اما $\hat{\beta}$ يمكن الحصول عليها من الصيغة التالية :

$$\hat{\beta} \approx \beta^0 + \frac{\sum_{i=1}^n \sin(y_i - \hat{\beta}_0 - \beta^0 x_i)}{\sum_{i=1}^n \cos(y_i - \hat{\beta}_0 - \beta^0 x_i)} \quad \dots (15)$$

اما $\hat{\nu}$ يمكن الحصول عليها حسب القانون التالي :

$$\hat{\nu} = A^{-1} \left(\frac{1}{n} \sum_{i=1}^n \cos(y_i - \hat{\beta}_0 - \hat{\beta}x_i) \right) \dots (16)$$

حيث : A^{-1} هي نسبة دالة بيسل (modified Bessel) من النوع الأول من الترتيب الأول ومن الترتيب الصفري لمعلمة التركيز ν في توزيع فون ميزس وقد اعطى (Dodson 1978) تقديرات تقريبية بسيطة للحصول على A^{-1} وهي Hassan, S. F. (2015).

$$A^{-1}(R) = \begin{cases} 2R + R^2 + 0.833R^5 & R > 0.053 \\ -0.4 + 1.39R + 0.43(1 - R)^{-1} & 0.53 \leq R < 0.85 \\ (R^3 - 4R^2 + 3R)^{-1} & R \geq 0.85 \end{cases} \dots (17)$$

2-4-3 تقدير الجزء اللامعلمي بطريقة المقدر اللبي (Kernel Estimator)

يعد المقدر اللبي (Kernel Estimator) أسلوب لاعملي يستخدم لتقدير الدوال الأحصائية بالأعتماد على البيانات ، يتميز بالبساطة وسهولة البرمجة ورسم القيم المقدره بيانيا وتقدير دالة الأندحار اللامعلمية ، تم اقتراح هذا المقدر من قبل الباحثين (Parzen,1962) و (Rosenblat,1956)

بفرض (y_i, θ_i) حيث $i = 1, 2, \dots, n$ تمثل عينة عشوائية لمتغيرين دائريين (y, θ) يمكن تمثيل العلاقة بين المتغيريين من خلال أنموذج الأندحار التالي :

$$y_i = [g(\theta_i) + \varepsilon_i](\text{mod } 2\pi) \quad i = 1.2. \dots n \quad \dots (18)$$

حيث ε_i تتوزع بشكل زوايا عشوائية بمتوسط اتجاهي صفر وتركيز محدود ومستقل عن θ_i .
تم تقدير الجزء اللامعلمي بواسطة النواة (Kernel) .

. بفرض ان المتغير العشوائي $(X=x_1. x_2. \dots x_n)$ له دالة كثافة احتمالية غير معلومة $f_x(x)$ فإن المقدر اللبي لهذه الدالة يوضح بالشكل التالي :

$$\hat{f}_x(x) = \frac{1}{nh} \sum_{i=1}^n K \left(\frac{x-x_i}{h} \right) \quad \dots (19)$$

حيث $h > 0$

$K(.)$ هي دالة لبية تسمى دالة النافذة .

وهي دالة رياضية حقيقية محددة تعمل على تنقية المقدر من التعرجات ويجب ان تحقق الشرطين : Ban Ahmed Matras, & Alaa Mahmoud Mohamed. (2013).

$$1) K(x) \geq 0$$

$$2) \int_{-\infty}^{\infty} K(x) dx = 1$$

اما **Circular Kernel** (النواة الدائرية) يرمز لها (K_ν) مع معلمة التركيز $\nu > 0$ هي دالة حقيقية ، عرف (Marzio et.al.2009) النواة الدائرية هي نواة الجيب (sin) ذات الرتبة r ومعلمة تركيز (تمهيد) $\nu > 0$ بأنها دالة حقيقية أي $R \rightarrow (0.2\pi) \in K_\nu$

ليبيانات المتغير التوضيحي الدائري والمتغير التابع الدائري ممكن تقدير دالة الأندحار اللامعلمية من خلال تقدير دالة كثافة النواة الدائرية (Kernel circular density estimator) ، يمكن التعبير عن مقدر دالة النواة الدائرية بالشكل التالي : (Marzio, 2013)

$$\hat{f}(\theta; \nu) = \frac{1}{n} \sum_{i=1}^n K_\nu(\theta - \theta_i), \quad 0 \leq \theta \leq 2\pi \quad \dots (20)$$

حيث $K_\nu(\theta - \theta_i)$ تمثل دالة النواة الدائرية (Kernel circular) $\nu > 0$ تمثل معلمة التركيز (معامل التجانس)

اما بالنسبة لأنواع النواة الدائرية من الممكن توظيف دوال الكثافة الدائرية ، وفي هذه الأطروحة تم استخدام دالتين :
1- **فون ميزس (Von Mesis)** : فهو يستخدم على نطاق واسع بسبب احاديته وشكله المتمائل ومعادلته التالية المذكورة سابقا . (Alonso-Pena, M., Ameijeiras-Alonso, J., & Crujeiras, R. M. (2021)).

$$f(\theta; \mu, \nu) = \frac{1}{2\pi I_0(\nu)} \exp\{\nu \cos(\theta - \mu)\} \quad \dots (21) \quad K_\nu =$$

حيث $0 \leq \theta \leq 2\pi$

مع هذه النواة المحددة يتم إعطاء مقدر دالة الكثافة في المعادلة (22) ، وهي عبارة عن مجموعة من توزيعات فون ميزس (Von Mesis) المتمركزة في θ_i ومع معلمة التركيز ν . لمعلمة التركيز تأثير في تقليل بعض معايير الخطأ بين الدالة الحقيقية اللامعلمية والدالة التقديرية . (Oliveira&Rodriguez , 201) ، (Sikaroudi, A. E. (2017))

$$\hat{f}(\theta; v) = \frac{1}{n2\pi I_0(v)} \sum_{i=1}^n \exp\{v \cos(\theta - \theta_i)\}, \quad 0 \leq \theta \leq 2\pi \quad \dots (22)$$

2- كوشي المغلف (Wrapped Cauchy) :

$$K_{wc} = f(\theta; \mu, p) = \frac{1}{2\pi} \frac{1 - p^2}{1 + p^2 - 2pcos(\theta)} \quad 0 < p < 1 \quad \dots (23)$$

حيث p هي معلمة التركيز

يمكن كتابة المقدر اللامعلمي لدالة الأنحدار $f(\theta_i)$ عند النقطة الدائرية (δ) بالشكل التالي : المصدر : Di Marzio, M., Fensore, S., & Taylor, C. C. (2023)

$$\hat{f}(\delta) = atan^2[\hat{g}_1(\delta) . \hat{g}_2(\delta)] \quad \dots (24)$$

حيث تمثل الدالة $[atan^2[\hat{g}_1(\delta) . \hat{g}_2(\delta)]$ بأنحدار الزاوية بين المحور والمتجه من الأصل يمكن حسابها بالشكل التالي :

$$atan^2[\hat{g}_1(\delta) . \hat{g}_2(\delta)] = \begin{cases} atan\left(\frac{\hat{g}_2(\delta)}{\hat{g}_1(\delta)}\right) & \text{if } \theta > 0 \\ \pi + atan\left(\frac{\hat{g}_2(\delta)}{\hat{g}_1(\delta)}\right) & \text{if } \hat{g}_2(\delta) \geq 0 \quad \hat{g}_1(\delta) < 0 \\ -\pi + atan\left(\frac{\hat{g}_2(\delta)}{\hat{g}_1(\delta)}\right) & \text{if } \hat{g}_2(\delta) < 0 \quad \hat{g}_1(\delta) < 0 \\ \frac{\pi}{2} & \text{if } \hat{g}_2(\delta) > 0 \quad \hat{g}_1(\delta) = 0 \\ \frac{-\pi}{2} & \text{if } \hat{g}_2(\delta) < 0 \quad \hat{g}_1(\delta) = 0 \\ \text{undefined} & \text{if } \hat{g}_2(\delta) = 0 \quad \hat{g}_1(\delta) = 0 \end{cases}$$

$$\hat{g}_1(\delta) = \frac{1}{n} \sum_{i=1}^n \sin(y_i) W(\theta_i - \delta) \quad \dots (25)$$

$$\hat{g}_2(\delta) = \frac{1}{n} \sum_{i=1}^n \cos(y_i) W(\theta_i - \delta) \quad \dots (26)$$

W تمثل الوزن .

يمكن إعطاء أوزان النواة عن طريق أخذ نواة دائرية كدالة للوزن :

$$W(\theta_i - \delta) = n^{-1} K_v(\theta_i - \delta) \left\{ \sum_{i=1}^n K_v(\theta_i - \delta) \sin^2(\theta_i - \delta) - \sin(\theta_i - \delta) \sum_{i=1}^n K_v(\theta_i - \delta) \sin(\theta_i - \delta) \right\} \quad \dots (27)$$

تشير K_v الى النواة الدائرية و v معلمة التمهيد

(Oliveira, M., Crujeiras, 2014) . (Meilán-Vila, A. (2020) .

3- الجانب التجريبي و التطبيقي

3-1-3 الجانب التجريبي

تم تقديم انموذجين للأنحدار شبه المعلمي الدائري -الخطي الدائري، يختلف كل أنموذج عن الآخر باختلاف **Circular Kernel** (النواة الدائرية) واختلاف معلمة الدمج .

3-1-3 المحاكاة Simulation المصدر (حسام أطروحة + أطروحة علي جواد)

وهي عملية تصميم أنموذج مشابه للنظام الحقيقي في المعطيات والفرضيات لغرض معرفة سلوك النظام الحقيقي والحصول على نتائج تساعد في اتخاذ القرار . يعرف أيضاً بأنها أحد الأساليب التي تستعمل لتوليد بيانات لغرض تحليلها دون الحاجة الى إجراء تجربة حقيقية ومراقبتها وتسجيل البيانات . وهي أيضاً أسلوب يستعمل أنموذج رياضي نظري شبيه او أنموذج بديل للأنموذج الحقيقي دون المحاولة للحصول على الأنموذج الحقيقي .

3-1-3 توليد البيانات :

1- تم استعمال اربعة أحجام للعيينة (400 , 300 , 150 , 100 = n) ولكل تجربة محاكاة تكرارات (1500) . الجانب التجريبي للجزء المعلمي :

1- اختيار أنموذج الأنحدار البسيط :

$$y_i = \beta_0 + \beta x_i + \varepsilon_i \pmod{2\pi} \quad \dots (28)$$

وأختيار قيم افتراضية للمعلمات كما في الجدول ادناه :

رقم التجربة	B_0	B_1
1	1.5	1
2	0.9	1.9
3	0.05	2.5

ويكون التقدير بطريقة الأماكن الأعظم (MLE) تم توليد البيانات للمتغير الدائري والمتغير الخطي في الأنموذج المعلمي باستخدام الدوال الشرطية ويفضل ذلك عند الأهتمام بدراسة المتغير الدائري بناءً على قيم المتغير الخطي .

المصدر (Imoto, T., Shimizu, K., & Abe, T. (2019). A cylindrical distribution with heavy-tailed (linear part). *Japanese Journal of Statistics and Data Science*, 2, 129-154.

1- توليد بيانات المتغير التابع (y_i) حسب التوزيع الخطي - الدائري وتم أستعمال التوزيع الشرطي وهو توزيع فون ميزس ($\varepsilon_i \sim VN(\mu, \frac{vx}{\sigma^2})$) حسب الدالة :

$$f(y/x) = \frac{1}{2\pi I_0\left(\frac{vx}{\sigma^2}\right)} \cdot \exp\left\{\frac{vx}{\sigma^2} \cos(y - \mu)\right\}$$

بفرض : $\mu=0$ و $v, \sigma^2 = 1$

2- المتغير المستقل (x_i) يتوزع التوزيع الشرطي الطبيعي حسب الدالة :

$$f(x/y) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2\sigma^2}(x - \lambda + v \cos(y - \mu))^2\right\}$$

بفرض : $\sigma^2 = 1$ و $\lambda, v = 0$

3- توليد الأخطاء العشوائية بفرض انها تتوزع توزيع دائري وتم افتراض توزيع فون ميزس (VM) .
بفرض : $\varepsilon_i \sim VN(\mu, v)$ و $\mu=0$ و $v = 1$

الجانب التجريبي للجزء اللامعلمي :

1- اختيار أنموذج الأنحدار اللامعلمي :

$$y_i = g(\theta_i) + \varepsilon_i \pmod{2\pi}$$

يكون تقدير الأنحدار اللامعلمي من خلال تقدير دالة كثافة النواة الدائرية (Kernel Density Estimation) وتم استعمال نوعين من دوال النواة الدائرية :

1- فون ميزس (Von Mesis) :

بفرض معلمة تركيز تأخذ ثلاث قيم ($v = 0.25, 0.50, 0.75$)

توليد بيانات المتغير الدائري (θ_i) يتوزع توزيع فون ميزس (VM) حسب الدالة :

$$f(\theta) = \frac{1}{2\pi I_0(k)} \exp\{v \cos(\theta - \mu)\}$$

2- كوشي المغلف (Wrapped Cauchy) :

بفرض معلمة تركيز تأخذ ثلاث قيم ($p = 0.25, 0.50, 0.75$)

توليد بيانات المتغير الدائري (θ_i) يتوزع توزيع كوشي الملفوف (WC) حسب الدالة :

$$f(\theta) = \frac{1}{2\pi} \frac{1 - p^2}{1 + p^2 - 2p \cos(\theta - \mu)}$$

الجانب التجريبي لأنموذج الأنحدار شبه المعلمي الدائري -الخطي-الدائري المدمج :

$$y_i = (1 - \alpha)f(x_i, \beta) + \alpha g(\theta_i) + \varepsilon_i$$

وتم اختيار معلمة دمج بثلاث قيم لكل تجربة :

رقم التجربة	قيم معلمة الدمج α		
1	0.05	0.01	0.1
2	0.2	0.5	0.7
3	0.3	0.6	0.9

2-1-3 الجانب التجريبي لبناء انموذج الانحدار شبه المعلمي الدائري - الخطي - الدائري المدمج بأستعمال دوال نواة مختلفة

تم اختيار الأنموذج شبه المعلمي الدائري-الخطي-الدائري المدمج والأفضل وحسب الصيغة :

$$y_i = (1 - \alpha)f(x_i, \beta) + \alpha g(\theta_i) + \varepsilon_i$$

من خلال بناء العديد من الأنموذجات شبه المعلمي الدائري -الخطي الدائري بأجراء العديد من تجارب المحاكاة المذكور نتائجها في الجداول الملحق () يتضمن كل جدول (case 27) أي (27 أنموذج) وبأستعمال طريقة الأماكن الأعظم في تقدير الجزء المعلمي واستخدام طريقة تقدير دالة النواة الدائرية (Kernel Density Estimation) في تقدير الجزء اللامعلمي وبواسطة دالة فون ميزس (Von Mesis) ودالة نواة كوشي المغلف (Wrapped Cauchy) . أذ تم تغيير القيم الافتراضية لمعاملات كل أنموذج واختيار الأنموذج الأفضل على أساس متوسط مربع الخطأ (MSE) ومعامل التحديد (R^2) . وتم إعادة بناء كل أنموذج بأستعمال حجوم عينات مختلفة (50,200,300,400) . للحصول على أفضل أنموذج ، تمت المقارنة بين الأنموذجات الأفضل التي تم بناءها من كل تجربة وحسب نوع دالة النواة المختارة كما يأتي :

1- بأستعمال دالة نواة فون ميزس (Von Mesis)

الجدول (3) مقارنة أنموذجات الانحدار شبه المعلمي الدائري-الخطي-الدائري المدمج المقدر بأستعمال دالة النواة فون ميزس (VM) لتقدير الجزء اللامعلمي لحجوم عينات (50,200,300,400)

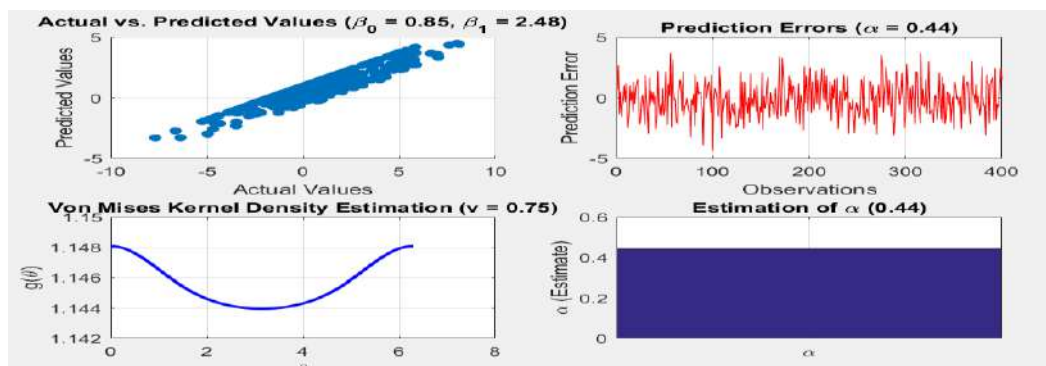
حجم العينة	الأنموذج شبه المعلمي الدائري-الخطي-الدائري المدمج	معلمة التركيز	MSE	R^2
50	$\bar{y}_1 = 0.98 (6.34 + 1.28x_i) + 0.02\hat{g}(\theta_i)$	0.25	0.8208	0.7514
	$\bar{y}_2 = 0.77 (1.32 + 1.86x_i) + 0.23\hat{g}(\theta_i)$	0.50	0.7640	0.8528
	$\bar{y}_3 = 0.84 (0.72 + 2.69x_i) + 0.16\hat{g}(\theta_i)$	0.25	0.8817	0.8966
200	$\bar{y}_1 = 0.98 (2.38 + 1.03x_i) + 0.02\hat{g}(\theta_i)$	0.50	0.8188	0.7568
	$\bar{y}_2 = 0.90 (1.21 + 1.95x_i) + 0.10\hat{g}(\theta_i)$	0.50	0.7653	0.8574
	$\bar{y}_3 = 0.92 (0.89 + 2.49x_i) + 0.08\hat{g}(\theta_i)$	0.25	0.8543	0.8978
300	$\bar{y}_1 = 0.99 (1.18 + 1.45x_i) + 0.01\hat{g}(\theta_i)$	0.25	0.7121	0.7651
	$\bar{y}_2 = 0.65 (1.71 + 1.88x_i) + 0.35\hat{g}(\theta_i)$	0.50	0.7422	0.8581
	$\bar{y}_3 = 0.82 (0.85 + 2.57x_i) + 0.18\hat{g}(\theta_i)$	0.25	0.8337	0.8991
400	$\bar{y}_1 = 0.98 (2.24 + 1.07x_i) + 0.02\hat{g}(\theta_i)$	0.75	0.7381	0.7971
	$\bar{y}_2 = 0.84 (1.72 + 1.90x_i) + 0.16\hat{g}(\theta_i)$	0.75	0.7063	0.8964
	$\bar{y}_3 = 0.66 (0.85 + 2.48x_i) + 0.44\hat{g}(\theta_i)$	0.75	0.7047	0.8997

يتبين من الجدول (3-1) العديد من أنموذجات الانحدار شبه المعلمي الدائري - الخطي - الدائري المدمج المقدر وفقاً لحجوم عينات مختلفة (50,200,300,400) تم الحصول على ثلاث أنموذجات من كل حجم وتحديد الأنموذج الأفضل من كل تجربة وبأستعمال دالة نواة فون ميزس (VM) . تم الحصول على كل أنموذج في الجدول أعلاه من خلال اختيار أنموذج واحد من كل تجربة موضحة بجدول يتضمن (case 27) . حققت التجربة الثالثة عند الحالة (case 6) وحجم 400 أفضل أنموذج كما في الجدول () من الملحق () .

- تأثير الجزء المعلمي (المتغير الخطي) قوي جداً وضعف تأثير الجزء اللامعلمي (المتغير الدائري) لجميع الأنموذجات المختارة من الجداول الممثلة لكل تجربة .
- الأنموذج الأفضل الذي حقق اقل قيمة لمتوسط مربع الخطأ البالغة (MSE = 7047) ، حقق أعلى قوة تفسيرية من خلال قيمة معامل التحديد البالغة ($R^2 = 8997$) التي تفسر حجم تأثير المتغيرات المستقلة على المتغير التابع يعني ذلك ان أداء أنموذج الانحدار شبه المعلمي الدائري الخطي - الدائري المدمج جيد جداً وبمعلمة دمج ($\alpha = 0.44$) وبقائمة تباين ($\sigma^2 = 1$)
- انخفاض متوسط مربع الخطأ (MSE) عند الحجم 400 مما يدل على انخفاضه مع حجوم العينات الكبيرة .
- معلمة التمهيد المقدر ($\hat{\rho}$) اخذت اعلى القيم عند الحجم 400 مما ساهم في زيادة مرونة الجزء اللامعلمي في الأنموذج كما يتضح من الشكل (1) :

- المخطط في الجهة العليا الأيسر (Actual vs. Predicted Values) المحور الأفقي يمثل القيم الفعلية (Actual Values) للبيانات المستهدفة ، والمحور الرأسي يمثل المتنبأ بها (Predicted Values) يتضح أن هناك علاقة خطية بين القيم الفعلية والقيم المتنبأ بها حيث كانت النقاط تتبع خطاً قطعياً من اليسار الى اليمين (زاوية 45 درجة تقريباً) هذا يعني ان الأنموذج يحقق تنبؤاً جيد مع وجود بعض الانحرافات الصغيرة في التنبؤات عند المقدرات ($b_0 = 0.85$, $b_1 = 2.48$).
- المخطط العلوي في الجهة اليمنى يشير الى أخطاء التنبؤ (Prediction Error) عند قيمة ($\hat{\alpha} = 0.44$). مثل المحور الأفقي المشاهدات (Observations) والمحور الرأسي يمثل أخطاء التنبؤ لكل ملاحظة تتراوح بين قيم سالبة وقيم موجبة فهي موزعة بشكل عشوائي حول الصفر مما يعني وجود بعض الانحرافات عن القيم الفعلية.
- المخطط السفلي الأيسر يعرض تقدير كثافة النواة (Kernel Density Estimation) باستخدام دالة نواة Von Mises (VM) للجزء اللامعلمي في الأنموذج حيث مثل المحور الأفقي قيم المتغير الدائري (θ) والمحور الرأسي مثل قيمة دالة الكثافة ($g(\theta)$) وهي تأخذ شكل منحنى وبأستعمال معلمة التركيز المقدر ($\hat{\nu} = 0.75$) تشير الى التشتت في تقدير الكثافة وهي تتحكم بشكل المنحنى حيث تمثل مدى دقة النواة في تقدير القيم حول θ .
- المخطط السفلي الأيمن (Estimation of α) يعرض قيمة ($\hat{\alpha} = 0.44$) المقدره وهي تمثل معلمة الدمج الجزء المعلمي واللامعلمي في النموذج ، المحور الرأسي يعرض قيمة (α) المقدره مما يشير الى مقدار مساهمة الجزء اللامعلمي في الأنموذج التنبؤي . يمكن صياغة الأنموذج بالمعادلة والشكل ادناه :

$$\hat{y}_3 = 0.66 (0.85 + 2.48x_i) + 0.44\hat{g}(\theta_i) \quad \dots \quad (29)$$



الشكل (1) يشير الى المقدرات المعلمية واللامعلمية للنموذج الثالث بواسطة (VM) ومعلمة الدمج لحجم (400)

2- بأستعمال دالة نواة Wrapped Cauchy (WC)

الجدول (4) مقارنة أنموذجات الأندثار شبه المعلمي الدائري-الخطي-الدائري المدمج المقدره بأستعمال دالة النواة كوشي المغلف (WC) لتقدير الجزء اللامعلمي لحجوم عينات (50,200,300,400)

حجم العينة	الأنموذج شبه المعلمي الدائري-الخطي-الدائري المدمج	معلمة التركيز (ρ)	MSE	R^2
50	$\hat{y}_1 = 0.94 (6.22 + 1.01x_i) + 0.06\hat{g}(\theta_i)$	0.05	0.7204	0.8704
	$\hat{y}_2 = 0.98 (1.71 + 2.14x_i) + 0.02\hat{g}(\theta_i)$	0.20	0.6995	0.8709
	$\hat{y}_3 = 0.85 (0.68 + 2.45x_i) + 0.15\hat{g}(\theta_i)$	0.05	0.6525	0.9155
200	$\hat{y}_1 = 0.96 (1.24 + 0.93x_i) + 0.04\hat{g}(\theta_i)$	0.80	0.9650	0.8584
	$\hat{y}_2 = 0.58 (1.64 + 1.93x_i) + 0.42\hat{g}(\theta_i)$	0.20	0.5738	0.9621
	$\hat{y}_3 = 0.93 (0.79 + 2.53x_i) + 0.07\hat{g}(\theta_i)$	0.80	0.7253	0.9729
300	$\hat{y}_1 = 0.99 (2.88 + 0.91x_i) + 0.01\hat{g}(\theta_i)$	0.80	0.9139	0.9201
	$\hat{y}_2 = 0.97 (2.76 + 1.01x_i) + 0.03\hat{g}(\theta_i)$	0.80	0.5752	0.9548
	$\hat{y}_3 = 0.72 (0.89 + 1.17x_i) + 0.28\hat{g}(\theta_i)$	0.80	0.7243	0.9723
400	$\hat{y}_1 = 0.46 (2.28 + 1x_i) + 0.54\hat{g}(\theta_i)$	0.80	0.9041	0.9511
	$\hat{y}_2 = 0.98 (1.66 + 2x_i) + 0.02\hat{g}(\theta_i)$	0.80	0.5647	0.9517
	$\hat{y}_3 = 0.69 (1.72 + 3.01x_i) + 0.31\hat{g}(\theta_i)$	0.20	0.5379	0.9764

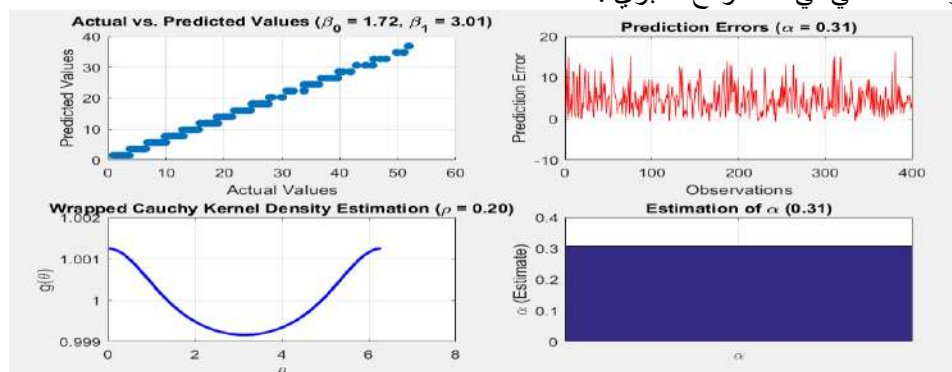
يتبين من الجدول (3-3) العديد من أنموذجات الأندثار شبه المعلمي الدائري - الخطي - الدائري المدمج المقدره وفقاً لحجوم عينات مختلفة وبأستعمال دالة نواة كوشي المغلف (WC) تم الحصول على كل انموذج في الجدول أعلاه من خلال اختيار انموذج واحد من كل تجربة موضحة بجدول في الملاحق () يتضمن كل جدول (case 27) تتفاوت كل حالة عن الأخرى بالقيمة الافتراضية لكل من (التباين (σ^2) ، ومعلمة التركيز (ρ) وقيم معلمات الجزء المعلمي الافتراضية (b_0 و b_1) و معلمة الدمج α) . وعند مقارنة الأنموذجات نلاحظ التالي :

- يتبين من الجدول ان تأثير الجزء المعلمي (المتغير الخطي) قوي جداً وضعف تأثير الجزء اللامعلمي (المتغير الدائري) لجميع الأنموذجات المختارة من الجداول الممثلة لكل تجربة .
- واضح من النتائج تذبذب في قيم (MSE) ، وقيم معامل التحديد (R^2) أيضاً فيها تذبذب واضح ويعطي قيم جيدة لجميع الأنموذجات مما يدل على قوة تفسير المتغيرات المستقلة للمتغير التابع
- من خلال النتائج في الجدول نلاحظ ان الأنموذج الأفضل الذي حقق اقل قيمة لمتوسط مربع الخطأ البالغة (MSE = 0.5379) ، حقق أعلى قوة تفسيرية من خلال قيمة معامل التحديد البالغة ($R^2 = 0.9764$) التي تفسر حجم تأثير المتغيرات المستقلة على المتغير التابع يعني ذلك ان أداء انموذج الأنحدار شبه المعلمي الدائري الخطي – الدائري المدمج جيد جداً وبمعلمة دمج ($\alpha = 0.31$) وبلغت قيمة معلمة التركيز ($\hat{\rho} = 0.20$) صياغة الأنموذج بالمعادلة والشكل ادناه :

$$\hat{y}_3 = 0.69 (1.72 + 3.01x_i) + 0.31\hat{g}(\theta_i) \quad \dots \quad (30)$$

يشمل الشكل (3-3) :

- المخطط في الجهة العليا الأيسر (Actual vs. Predicted Values) المحور الأفقي يمثل القيم الفعلية (Actual Values) للبيانات المستهدفة ، والمحور الرأسي يمثل المتنبأ بها (Predicted Values) يتضح أن هناك علاقة خطية بين القيم الفعلية والقيم المتنبأ بها حيث كانت النقاط تتبع خطاً قطرياً من اليسار الى اليمين (زاوية 45 درجة تقريباً) هذا يعني ان الأنموذج يحقق تنبؤاً جيد مع وجود بعض الانحرافات الصغيرة في التنبؤات عند المقدرات
- المخطط العلوي في الجهة اليمنى يشير الى أخطاء التنبؤ (Prediction Error) عند قيمة ($\hat{\alpha} = 0.31$). مثل المحور الأفقي المشاهدات (Observations) والمحور الرأسي يمثل أخطاء التنبؤ لكل ملاحظة تتراوح بين قيم سالبة وقيم موجبة فهي موزعة بشكل عشوائي حول الصفر مما يعني وجود بعض الانحرافات عن القيم الفعلية.
- المخطط السفلي الأيسر يعرض تقدير كثافة النواة (Kernel Density Estimation) باستخدام دالة نواة Wrapped Cauchy (WC) للجزء اللامعلمي في الأنموذج حيث مثل المحور الأفقي قيم المتغير الدائري (θ) والمحور الرأسي مثل قيمة دالة الكثافة $g(\theta)$ وهي تأخذ شكل منحنى وبأستعمال معلمة التركيز ($\rho = 0.20$) تشير الى التشتت في تقدير الكثافة وهي تتحكم بشكل المنحنى حيث تمثل مدى دقة النواة في تقدير القيم حول θ .
- المخطط السفلي الأيمن (Estimation of α) يعرض قيمة ($\alpha = 0.31$) المقدره وهي تمثل معلمة دمج الجزء المعلمي واللامعلمي في النموذج ، المحور الرأسي يعرض قيمة (α) المقدره مما يشير الى مقدار مساهمة الجزء اللامعلمي في الأنموذج التنبؤي .



الشكل (2) يشير الى المقدرات المعلمية واللامعلمية بواسطة دالة النواة كوشي (WC) ومعلمة الدمج لحجم (400)

2-3 الجانب التطبيقي

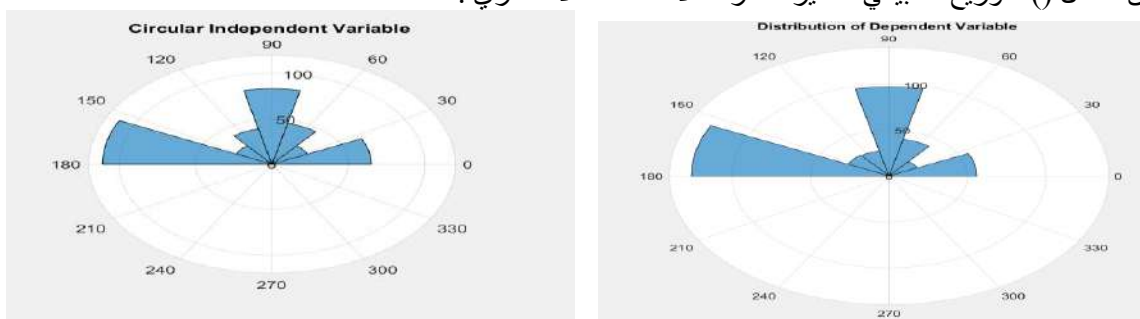
1-2-3 مقدمة Introduction

الهدف من هذا الفصل هو الحصول على أنموذج أن أنموذج الأنحدار شبه المعلمي الدائري - الخطي – الدائري المدمج يعطي صورة واضحة لتأثير الأصابة بمرض الخطأ الأنكساري (Refractive Error) وتأثير كل من العين اليسرى المصابة وعمر المصاب على العين اليمنى . يتطلب الحفاظ على الوظيفة البصرية تحسينها بفهم الخطأ الأنكساري (Refractive Error) والتغيرات المرتبطة بالعمر ، تعمل العين البشرية كنظام بصري يركز الصور المرئية على الشبكية ، أذ ان الخطأ الأنكساري هو أحد الأسباب الرئيسية لتدهور صورة الشبكية في العيون غير المصححة (Namba, H., Sugano & others , 2020) المصدر : علاقة العمر بالاستجماتزم

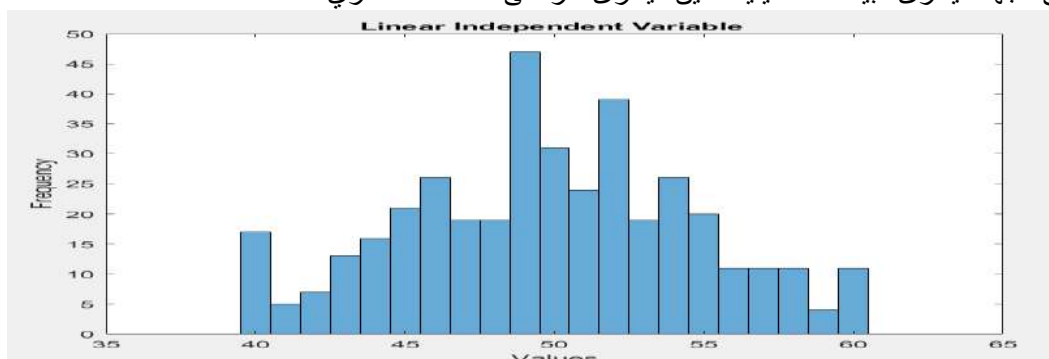
2-2-3 البيانات الحقيقية

تكونت البيانات الحقيقية للخطأ الأنكساري (Refractive Error) للعين اليمنى (OD) مثل المتغير المعتمد (y) وبيانات العين اليسرى (OS) مثل المتغير التوضيحي (θ) والمتغير التوضيحي (X) مثل العمر لمرضى الخطأ الأنكساري لعينة مكونة من (400) حالة كما في الجدول () وتم تمثيل البيانات بالأشكال أدناه :

يتضح من الشكل () توزيع البيانات الحقيقية ضمن دائرة الوحدة للمتغير لبيانات الخطأ الأنكساري للعين اليمنى والذي يمثل المتغير المعتمد. أيضا نلاحظ من الشكل () البيانات الحقيقية لمتغير الخطأ الأنكساري للعين اليسرى وهي موزعة ضمن دائرة الوحدة وتمثل المتغير التوضيحي للجزء اللامعلمي للأنموذج الأندحار شبه المعلمي الدائري - الخطي - الدائري المدمج ويبين الشكل () التوزيع الطبيعي لمتغير العمر لحالات الخطأ الأنكساري .



شكل (3) الشكل في الجهة اليمنى البيانات الحقيقية للعين اليمنى لمرضى الخطأ الأنكساري الشكل في الجهة اليسرى البيانات الحقيقية للعين اليسرى لمرضى الخطأ الأنكساري



شكل (4) البيانات العمرية الحقيقية لمرضى الخطأ الأنكساري

3-2-3 تقدير معاملات النموذج : Estimation of model Parameters

تم تقدير معاملات الجزء المعلمي للأنموذج الأندحار شبه المعلمي الدائري - الخطي - الدائري المدمج أذ ممثل المتغير المعتمد (y) ببيانات مرضى الخطأ الأنكساري للعين اليمنى ، ومثل المتغير (x) البيانات العمرية لمرضى الخطأ الأنكساري النتائج موضحة كما في الجدول أدناه :

جدول (5) تقدير معاملات الجزء المعلمي للأنموذج الأندحار شبه المعلمي الدائري - الخطي - الدائري المدمج

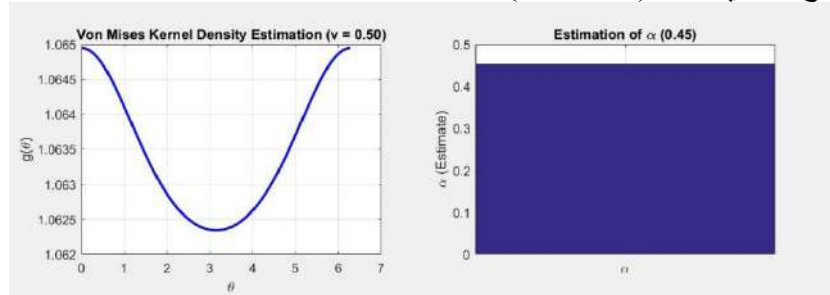
Parameter Method	\hat{B}_0	\hat{B}_1
MLE الأماكن الأعظم	0.9632	2.6412

نلاحظ من الجدول () قيم المعلمات التي تم تقديرها للجزء المعلمي هي متقاربة مع القيم المقدرة بأستعمال البيانات المولدة في الجانب التجريبي (المحاكاة)

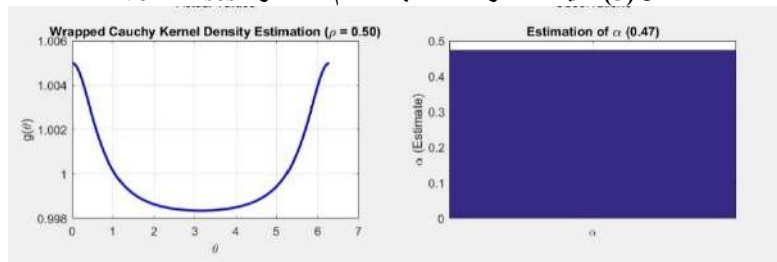
- الشكل () المخطط الأيسر يعرض تقدير كثافة النواة (Kernel Density Estimation) بأستخدام دالة نواة Von Mises للجزء اللامعلمي في الأنموذج حيث مثل المحور الأفقي قيم المتغير الدائري (θ) الممثلة لبيانات الخطأ الأنكساري للعين اليسرى والمحور الرأسي مثل قيمة دالة الكثافة المقدرة للبيانات الحقيقية ($g(\theta)$) وهي تأخذ شكل منحنى وبأستعمال معلمة التركيز ($\nu = 0.50$) تشير الى التشتت في تقدير الكثافة وهي تتحكم بشكل المنحنى حيث تمثل مدى دقة النواة في تقدير القيم حول θ .
- المخطط الأيمن يمثل قيمة معلمة الدمج المقدرة للبيانات الحقيقية (Estimation of α) ، المحور الرأسي يعرض قيمة (α) المقدرة مما يشير الى مقدار مساهمة الجزء اللامعلمي في الأنموذج التنبؤي ، أذ بلغت ($\alpha = 0.45$) .
- الشكل () المخطط الأيسر يعرض تقدير كثافة النواة (Kernel Density Estimation) بأستخدام دالة نواة Wrapped Cauchy (WC) للجزء اللامعلمي في الأنموذج حيث مثل المحور الأفقي قيم المتغير الدائري (θ) والمحور الرأسي مثل قيمة دالة الكثافة ($g(\theta)$) وهي تأخذ شكل منحنى وبأستعمال معلمة التركيز المقدرة (ρ)

0.50) تشير الى التشتت في تقدير الكثافة وهي تتحكم بشكل المنحنى حيث تمثل مدى دقة النواة في تقدير القيم حول θ .

- المخطط الأيمن (Estimation of α) يعرض قيمة معلمة الدمج المقدرة وهي تمثل معلمة دمج الجزء المعلمي واللامعلمي في النموذج ، المحور الرأسي يعرض قيمة (α) المقدرة مما يشير الى مقدار مساهمة الجزء اللامعلمي في اللانموذج التنبؤي بلغت ($\alpha = 0.47$).



شكل (5) القيمة المقدرة لمعلمة باستخدام دالة النواة Von Mises



شكل (6) القيمة المقدرة لمعلمة باستخدام دالة النواة Wrapped Cauchy

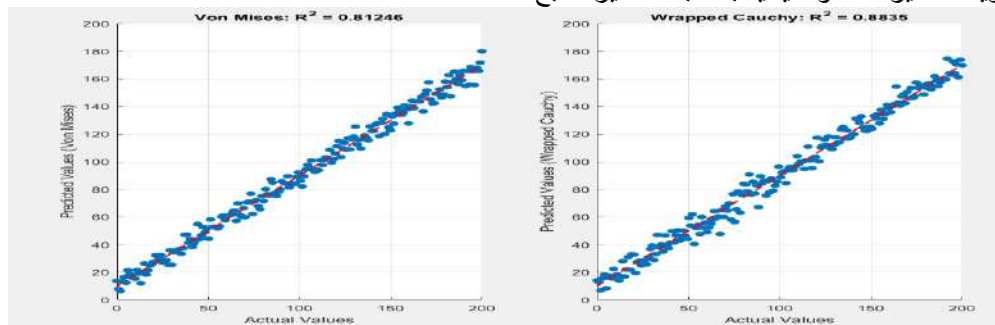
3-4-2-3 المقاضلة بين الأنموذجات المقدرة للبيانات الحقيقية:

لمعرفة أي الأنموذجين يتفوق في الكفاءة على الأنموذج الآخر سنستعمل معايير المعلومات للمفاضلة بينها وهما معيار المعلومات أكايكي (AIC) ومعيار المعلومات (BIC) إذ تعتمد هذه المعايير في حسابها على متوسط مربع الخطأ (MSE) ، يمكن تلخيص نتائج هذه المعايير مع متوسط مربع الخطأ بالجدول أدناه :

جدول (6) يبين معايير المقاضلة لمقدرات أنموذج الانحدار شبه المعلمي الدائري-الخطي – الدائري المدمج

الأنموذج شبه المعلمي الدائري -خطي – دائري المدمج	MSE	المعايير	
		AIC	BIC
دالة كثافة نواة VM	0.8937	319.0389	324.7749
دالة كثافة نواة WC	0.78322	292.4276	298.1637

نلاحظ من الجدول (6) أن أنموذج الانحدار شبه المعلمي الدائري -الخطي – الدائري المدمج المقدر باستخدام دالة نواة كوشي المغلف (WC) بمعلمة الدمج المقدرة () تفوق على أنموذج الانحدار شبه المعلمي الدائري -الخطي – الدائري المدمج المقدر باستخدام دالة نواة فون ميزس (VM) إذ حقق أقل قيم لكل من متوسط مربع الخطأ (MSE) ومعيار المعلومات (AIC) و(BIC) لذا يعد أنموذج الانحدار شبه المعلمي الدائري -الخطي – الدائري المدمج . في حين حقق أنموذج الانحدار شبه المعلمي الدائري -الخطي – الدائري المدمج المقدر باستخدام دالة نواة فون ميزس (VM) أعلى القيم بالنسبة لمتوسط مربع الخطأ (MSE) ومعيار المعلومات (AIC) ومعيار المعلومات (BIC) . جاءت هذه النتائج متفقة مع نتائج الجانب التجريبي (المحاكاة) في الفصل الثالث عند الحجم (400) حيث كانت الأفضلية لأنموذج الانحدار شبه المعلمي الدائري -الخطي – الدائري باستخدام دالة نواة كوشي المغلف (WC) . ايضاً حقق الأنموذج شبه المعلمي الدائري -الخطي – الدائري المدمج المقدر باستخدام دالة نواة كوشي المغلف (WC) أعلى معامل تحديد إذ بلغ (0.8835) يدل ذلك على القوة التفسيرية للمتغيرات التوضيحية بالنسبة للمتغير التابع



جدول (7) يبين (MSE) لمقدرات الجزء المعلمي لأنموذج الانحدار شبه المعلمي الدائري – الخطي - الدائري

- تُظهر من الشكل () العلاقة بين القيم الحقيقية (True Values) والقيم المتوقعة (Fitted Values) من النموذج. الخط الأحمر المنقط يمثل الخط المرجعي (Reference Line) ، حيث تشير النقاط القريبة من هذا الخط إلى أن النموذج يتنبأ بدقة عالية.
- يمكن ملاحظة أن معظم النقاط تقع بالقرب من الخط المرجعي، مما يشير إلى أن النموذج لديه قدرة جيدة على التنبؤ بالقيم الحقيقية. ومع ذلك، هناك بعض النقاط التي تنحرف عن الخط، مما قد يشير إلى وجود قيم شاذة أو مناطق يحتاج النموذج إلى تحسينها.
- الأداء العام للنموذج يبدو جيداً، حيث أن النقاط تتبع بشكل عام اتجاه الخط المرجعي. هذا يشير إلى أن النموذج قادر على تفسير جزء كبير من التباين في البيانات.
- هناك بعض النقاط التي تبعد بشكل ملحوظ عن الخط المرجعي، مما قد يشير إلى وجود قيم شاذة في البيانات أو أن النموذج يحتاج إلى تحسين في هذه المناطق".
- دقة النموذج تبدو معقولة، حيث أن معظم النقاط تقع بالقرب من الخط المرجعي. ومع ذلك، يمكن تحسين الدقة من خلال تحسين معاملات النموذج أو إضافة ميزات إضافية
- بعد إجراء المقارنة تبين أن أنموذج الأنحدار شبه المعلمي الدائري -الخطي-الدائري المدمج المقدر باستخدام دالة النواة Wrapped Cauchy هو الأنموذج الأفضل وبمعلمة تركيز ذات قيمة معتدلة بلغت ($\rho = 0.50$) ويمكن كتابة معادلة الأنموذج بالصيغة التالية :

$$\hat{y} = 0.63 (0.9632 + 2.6412x_i) + 0.47\hat{g}(\theta_i) \quad \dots \quad (31)$$

4- الأستنتاجات والتوصيات

1-4 الأستنتاجات

- بالنسبة للأنموذج شبه المعلمي الدائري -الخطي-الدائري المدمج :
- يجمع أنموذج الأنحدار شبه المعلمي الدائري -الخطي-الدائري المدمج يجمع بين المرونة والدقة وتحقيق موثوقية في التقدير والتنبؤ بما يتلائم وطبيعة البيانات الدائرية.
- عدم أستقرار نتائج الأنموذج في العينات الصغيرة أذ يلاحظ تذبذب قيمة معامل التحديد (R^2) وكذلك متوسط مربع الخطأ (MSE) بين الأرتفاع والأخفاض ، قد يكون ذلك بسبب تغيير معلمة التركيز الخاصة بدالة النواة أو بسبب تغيير التباين .
- أستخدام دالة النواة كوشي المغلف (Wrapped Cauchy) في تقدير الجزء اللامعلمي للأنموذج الأنحدار شبه المعلمي الدائري -الخطي-الدائري أفضل من استخدام فون ميزس (Von Mesis)

2-4 التوصيات

- نوصي بتعميم الأنموذج على بيانات لظواهر أخرى .
- إجراء دراسات حول أنموذج الأنحدار شبه المعلمي الدائري -الخطي-الدائري باعتماد عينات كبيرة لتحسين أستقرار النتائج .
- دراسة أنواع أخرى من أنموذجات الأنحدار الدائري .

المصادر:

1. Namba, H., Sugano, A., Murakami, T., Utsunomiya, H., Nishitsuka, K., Ishizawa, K., ... & Yamashita, H. (2020). Age-related changes in astigmatism and potential causes. *Cornea*, 39, S34-S38.
2. Scott, A. (2002). Circular Data: An Overview with Discussion of One-Sample Tests.
3. Fan, Y., & Ullah, A. (1999). Asymptotic normality of a combined regression estimator. *Journal of Multivariate Analysis*, 71(2), 191-240.
4. Abuzaid, A. H. (2010). *Some problems of outliers in circular data*. University of Malaya (Malaysia).
5. Allahham, N. (2015). *ON THE SIMPLE ANGULAR REGRESSION MODEL WITH WRAPPED CAUCHY ERROR* (Doctoral dissertation, Al Azhar University-Gaza).
6. Rahmani, S., Mohaghegh, S., & Moradnejad, S. (2024). Refractive error and vision related quality of life. *BMC ophthalmology*, 24(1), 83.
7. Oliveira, M., Crujeiras, R. M., & Rodríguez-Casal, A. (2014). CircSiZer: an exploratory tool for circular data. *Environmental and ecological statistics*, 21, 143-159.
8. Meilán-Vila, A. (2020). Nonparametric inference for regression models with spatially correlated errors.
9. Hussen, Namarq Qassem. Mohammed, Laith Ali & Hasan , Rammah Oday . (2022) " Using a semi-parametric regression model to study the most important factors affecting the gross domestic product of oil prices for the period" (1980-2020) . Journal of Algebraic ststistic . Volume 13,No.2 , 2022 , P . 1024-1036 .https://publishoa.com , ISSN:1309-3452

10. Rahman, M., Gokhale, D. V., & Ullah, A. (1997). A note on combining parametric and non-parametric regression. *Communications in Statistics-Simulation and Computation*, 26(2), 519-529.
11. Ullah, A., & Vinod, H. D. (1993). 4 General nonparametric regression estimation and testing in econometrics.
12. Di Marzio, M., Fensore, S., & Taylor, C. C. (2023). Kernel regression for errors-in-variables problems in the circular domain. *Statistical Methods & Applications*, 32(4), 1217-1237.
13. Ban Ahmed Matras, & Alaa Mahmoud Mohamed. (2013). Using the Albi estimator and k-mean clustering method for hand gesture recognition. *AL-Rafidain Journal of Computer Sciences and Mathematics*, 10(1).
14. Jawad, Ali Muhammad. 2023 AD “Building and Estimating an Angular Regression Model for Circular Data with a Practical Application” PhD Thesis in Statistics, College of Administration and Economics, University of Karbala.
15. Tarrad, Alaa Jaber & Hassan, Raad Fadel. 2013 AD “Parametric and semi-parametric regression models (comparative study)”. Master’s thesis in statistics, Al-Mustansiriya University, College of Administration and Economics.
16. Issa, Aseel Muslim. 2011. “Comparison of some semi-parametric estimators for estimating the electrical energy consumption function for the city of Baghdad.” Master’s thesis in statistics. College of Administration and Economics, University of Baghdad.
17. Alonso-Pena, M., Ameijeiras-Alonso, J., & Crujeiras, R. M. (2021). Nonparametric tests for circular regression. *Journal of Statistical Computation and Simulation*, 91(3), 477-500.