

التقدير الاحصائي باستعمال بعض الدوال اللبية

Statistical Estimation Using some Kernel Functions

ا.م.د. صدى فايز محمد
 Ass. Prof. Dr. Sada Fayidh Mohammed
sada.f@uokerbala.edu.iq
 كلية الادارة والاقتصاد جامعة كربلاء , كربلاء
 College of Administration and
 Economics, University of Karbala

الاء خالد عبد الحسين
 Alaa Khalid AbdulHussein
alaa.khalid@s.uokerbala.edu.iq
 كلية الادارة والاقتصاد جامعة كربلاء , كربلاء
 College of Administration and
 Economics, University of Karbala

المستخلص: يهدف هذا البحث إلى تقديم تقدير إحصائي دقيق لدالة الكثافة الاحتمالية (PDF) لبيانات البقاء عن طريق توزيع بعض دوال كيرنل غير المتماثلة (Lindley, Gamma1, Weibull, Inverse-Gaussian) إذ تم التركيز على تطوير نموذج مرّن باستعمال دالة Weibull ككيرنل غير متماثلة، نظرًا لملاءمتها للبيانات الموجبة والمنحرفة. تم في الجانب النظري اقتراح الدالة، وتحليل خصائصها، ومقارنتها نظريًا بدوال أخرى مستخدمة في الأدبيات، من حيث الشكل والسلوك الاحتمالي وقدرة التمثيل. أما في الجانب التجريبي، فقد تم إجراء دراسة محاكاة باستعمال ثلاث أحجام عينات وأربع حالات بيانات مختلفة لتقييم أداء الدالة المقترحة مقارنة بدوال كيرنل متماثلة وغير متماثلة أخرى، باستعمال معيار الخطأ التكامل التريبي (ISE) وطريقتي Cross-Validation و Silverman في تحديد عرض الحزمة. أظهرت النتائج تفوق دالة Weibull Kernel من حيث أقل قيم للـ ISE، لاسيما عند حجم العينة 200. كما تم تطبيق النموذج على بيانات بقاء حقيقية لمرضى القسطرة في مستشفى عام، وبيّنت النتائج كفاءة النموذج المقترح من خلال تقارب دوال البقاء التقديرية مع البيانات الحقيقية، استنادًا إلى مؤشري AIC و p-value. تؤكد هذه الدراسة فعالية استعمال دوال كيرنل غير المتماثلة، لاسيما دالة Weibull، في تمثيل البيانات الموجبة المنحرفة بدقة.

الكلمات المفتاحية: دالة Weibull، الدوال اللبية غير المتماثلة، معيار الخطأ التكامل التريبي (ISE)

Abstract: In this article aims to provide accurate statistical estimation of the probability density function (PDF) for survival data using selected asymmetric kernel functions (Lindley, Gamma1, Weibull, Inverse-Gaussian). The focus was on developing a flexible model employing the Weibull function as an asymmetric kernel, given its suitability for positive and skewed data. In the theoretical part, the proposed kernel function was defined, its properties analyzed, and it was theoretically compared with existing kernels in terms of form, probabilistic behavior, and representational capability.

In the experimental part, a simulation study was conducted across three different sample sizes and data scenarios to assess the performance of the proposed kernel against other symmetric and asymmetric kernels, based on the Integrated Squared Error (ISE) criterion, using both Silverman's rule and Cross-Validation for bandwidth selection. The results showed the superiority of the Weibull kernel, particularly for a sample size of 200, where it achieved the lowest ISE values.

The model was further applied to real-world survival data from catheter patients in a general hospital. The analysis demonstrated the efficiency of the proposed model through the close agreement between the estimated and actual survival functions, evaluated via AIC and p-value metrics. This study highlights the effectiveness of asymmetric kernels especially the Weibull kernel in accurately modeling positive skewed data.

Keywords: Weibull function, non-symmetric sine functions, integral squared error criterion (ISE)

1. المقدمة (introduction)

يشكل التقدير غير المعلمي للكثافة الاحتمالية أحد المرتكزات المنهجية المهمة في الإحصاء الحديث، لما يوفره من إطار رياضي مرّن يمكن الباحث من تحليل سلوك المتغيرات العشوائية من دون الحاجة إلى افتراض شكل توزيعي مسبق. وتعد دوال كيرنل إحدى أبرز الأدوات التي اعتمدتها هذه المنهجية، نظرًا لقدرتها على تمثيل الأنماط المعقدة في البيانات، وتكيفها مع التوزيعات غير المتماثلة أو ذات الذيل الطويل التي يصعب التعامل معها عن طريق النماذج المعلمية التقليدية.

وقد شهد هذا المجال تطوراً معرفياً لافتاً خلال العقود الأخيرة، بدءاً من استعمال الدوال اللبية المتماثلة مثل Gaussian و Epanechnikov ، وصولاً إلى إدخال دوال غير متماثلة مستندة إلى توزيعات موجبة مثل Gamma و Inverse Gaussian، بهدف تحسين تمثيل الكثافة في المناطق الحدية وتقليل الانحياز في البيانات المحصورة ضمن المجال الموجب. وعلى الرغم من ما وفرته هذه الدوال من مزايا تطبيقية، إلا أن العديد منها ما زال يفتقر إلى المرونة البنيوية اللازمة للتكيف مع الأنماط الحادة أو الانحرافات الكبيرة التي تظهر في بيانات البقاء والموثوقية والبيانات المالية. وفي هذا البحث سوف يتم التطرق الى دالة غير متماثلة مقترحة من توزيع ويبل 2. منهجية البحث

1.2 مشكلة البحث (Problem of Research)

تشهد بيانات البقاء والأعمار خصائص توزيعية معقدة مثل الانحراف والذيل الطويل، ما يجعل التوزيعات المعلمية الكلاسيكية عاجزة عن تمثيلها بدقة، لاسيما عند الحواف. وعلى الرغم من استعمال دوال كيرنل غير متماثلة مثل Gamma و Lognormal، إلا أن مرونتها تظل محدودة. كما أن الأدبيات لم تُعطِ اهتماماً كافياً لتطوير دوال كيرنل مبنية على توزيع ويبل، وعلى الرغم من ملائمتها للبيانات الموجبة وغير المتماثلة، ما يبرز الحاجة إلى دالة كيرنل جديدة تتمتع بدعم موجب وبنية مرنة تقلل التحيز وتعزز دقة التقدير نتيجة قصور التوزيعات الكلاسيكية والمعلمية عن تمثيل البيانات غير المتماثلة أو المحصورة ضمن المجال الموجب. ومحدودية الدوال اللبية المنشورة في التكيف مع البيانات ذات الذيل الطويل أو السلوك الحدي وغياب نماذج كيرنلية غير متماثلة مستندة إلى توزيع ويبل في الأدبيات، وعلى الرغم من ملائمتها الطبيعية لبيانات الأعمار والبقاء والحاجة إلى بناء دالة كيرنل جديدة تتمتع بخواص رياضية ملائمة وتُستخدم لتقدير الكثافة الاحتمالية من دون تحيز في الحواف أو تشوه في التمثيل الاحتمالي وتتلخص مشكلة البحث في تقديم دالة كيرنل غير متماثلة مقترحة من توزيع ويبل لتمثل البيانات الموجبة .

2.2 هدف البحث (Aim of Research)

يهدف هذا البحث إلى تطوير نموذج تقديري مرّن للكثافة الاحتمالية يستند إلى دالة كيرنل غير متماثلة مبنية على توزيع ويبل، وذلك لتحسين تمثيل البيانات الموجبة وغير المتماثلة، كبيانات البقاء والأعمار التي تنتم غالباً بخصائص حدية أو انحرافات حادة يصعب احتواؤها باستعمال دوال كيرنل تقليدية إذا تم بناء دالة كيرنل غير متماثلة جديدة تستند إلى توزيع ويبل، وتُستخدم في التقدير غير المعلمي للكثافة الاحتمالية واشتقاق الخصائص النظرية للدالة المقترحة، مثل شرط عدم السالبية، خاصية التكامل إلى واحد، والدعم الموجب وتحليل الأداء الإحصائي للدالة من حيث التحيز والتباين باستعمال مقياس الخطأ التكامل المتوسط (MISE) وتعبيره التقريبي (AMISE) وتقييم كفاءة الدالة المقترحة عن طريق دراسات محاكاة عددية ومقارنتها مع دوال منشورة مثل Gamma kernel.

3. الجانب النظري

1.3 مفهوم التقدير غير المعلمي وأهميته في الإحصاء

يُعرف التقدير غير المعلمي لدالة الكثافة الاحتمالية بأنه أسلوب يهدف إلى تقدير دالة التوزيع الكامنة وراء مجموعة من البيانات دون افتراض شكل توزيعي محدد لها. وبدلاً من ذلك، تُستخدم أساليب رياضية تعتمد مباشرة على القيم الفعلية في العينة لتقريب شكل هذه الدالة. وتكمن أهمية هذا النوع من التقدير في كونه لا يقيد بفرضيات قد تكون خاطئة أو غير مسوغة، ما يتيح إمكانية استخلاص خصائص التوزيع كما تظهر فعلياً في البيانات دون تشويه ناتج عن اختيار نموذج غير ملائم.

ومن أكثر الأساليب استعمالاً في هذا النوع من التقدير هو التقدير باستعمال دوال كيرنل، والذي يُمثل الصيغة العامة له على النحو الآتي:

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

إذ إن:

$\hat{f}_h(x)$: هو التقدير الكيرنلي للكثافة عند النقطة x ،

$K(\cdot)$: هي دالة الكيرنل، وتُحدد شكل التأثير المحلي لكل نقطة،

h : هو معامل التنعيم (Bandwidth) ،

n : عدد مشاهدات العينة، و x_i هي عناصرها.

2.3. دالة الكثافة الاحتمالية (Probability Density Function)

تُعد دالة الكثافة الاحتمالية (PDF) إحدى الركائز الأساسية في نظرية الاحتمالات والإحصاء التطبيقي، إذ تُمثل الإطار الرياضي الذي يُستخدم لوصف سلوك المتغيرات العشوائية المستمرة. فعند التعامل مع بيانات كمية لا تأخذ إلا قيماً مستمرة كما هو الحال في قياسات الزمن، الطول، الوزن، أو العمر لا يُمكن الحديث عن احتمال وقوع المتغير في نقطة مفردة، لأن هذا الاحتمال يساوي صفراً نظرياً. ولذلك، يتم تمثيل توزيع الاحتمالات عبر ما يُعرف بدالة الكثافة الاحتمالية، والتي تُحدد مدى تركيز الاحتمالات حول القيم المختلفة على محور الأعداد الحقيقية.

وتُعرف دالة الكثافة الاحتمالية رياضياً على أنها دالة $f(x)$ تحقق الشرطين الآتيين:

1. أنها غير سالبة على المجال الحقيقي:

$$f(x) \geq 0 \quad \forall x \in R$$

2. أن التكامل الكلي لها على المجال الكامل يساوي واحداً:

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

ويُستخدم التكامل المحدود لهذه الدالة لحساب احتمال أن يقع المتغير ضمن مدة معينة، كما في العلاقة الآتية:

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

إن دالة الكثافة لا تُحدد احتمال وقوع المتغير في قيمة محددة، بل تُعبر عن مدى "كثافة الاحتمال" حول كل نقطة. ولهذا السبب، يكون تمثيلها بيانياً على شكل منحنى، إذ تُعبر المساحة تحت المنحنى بين أي نقطتين عن احتمال وقوع المتغير ضمن هذا النطاق.

3.3. التقدير غير المعلمي باستعمال دوال كيرنل

يُعد التقدير غير المعلمي لدالة الكثافة الاحتمالية باستعمال دوال كيرنل (Kernel Density Estimation KDE) من أكثر الطرائق شيوعاً وفعالية عندما لا تتوفر فرضية مسبقة عن الشكل التوزيعي للبيانات. وتكمن قوة هذا الأسلوب في قدرته على توليد تمثيل سلس ومرن لدالة الكثافة الاحتمالية انطلاقاً من العينة فقط، من دون الحاجة إلى تقسيم البيانات إلى فئات كما هو الحال في المدرج التكراري.

وتقوم الفكرة الأساسية للطريقة على تركيب دالة كيرنلية $K(\cdot)$ متركزة حول كل نقطة ملاحظة x_i من العينة، ثم تجميع هذه الدوال بشكل موزون اعتماداً على معامل التنعيم h . ويُعطى التقدير الكيرنلي للكثافة بالاعتماد على العلاقة:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{X - x_i}{h}\right) \quad (2)$$

إذ إن n هو حجم العينة، و $h > 0$ هو معامل التنعيم (Bandwidth) ، الذي يتحكم بدرجة انتشار الكثافة حول كل نقطة. أما دالة الكيرنل $K(\cdot)$ فهي عادةً دالة متماثلة تحقق شرط التكامل إلى واحد وغير سالبة، ومن أمثلتها Gaussian و Epanechnikov والدالة المثلثية.

يؤدي معامل التنعيم دوراً جوهرياً في التحكم بجودة التقدير. إذ أن القيم الصغيرة جداً لـ h تُنتج تقديراً ذو تذبذبات حادة (تباين مرتفع)، في حين أن القيم الكبيرة تؤدي إلى تمثيل مفرط التنعيم يخفي تفاصيل التوزيع (انحياز مرتفع). وقد طُوّرت العديد من

الأساليب لاختيار القيمة المثلى لـ h ، من بينها: قاعدة الإبهام (Silverman's Rule of Thumb) ، التحقق المتقاطع (Cross-Validation) ، والتقليل العددي لمقاييس الخطأ مثل MISE و AMISE.

4.3. أنواع دوال كيرنل

تُستخدم دوال كيرنل في التقدير غير المعلمي للكثافة الاحتمالية كوظائف موزونة توزع التأثير الإحصائي لنقاط العينة حول نقطة معينة. ويمكن تصنيف هذه الدوال إلى نوعين رئيسيين بناءً على خصائصها الشكلية ودعمها الرياضي: دوال كيرنل متماثلة (Symmetric Kernels) ودوال كيرنل غير متماثلة (Asymmetric Kernels). ينعكس هذا التصنيف على سلوك التقدير الناتج ومدى ملاءمته لطبيعة البيانات .

4.3.1 الدوال اللبية غير المتماثلة

تُصمم دوال الكيرنل غير المتماثلة للتعامل مع بيانات غير سالبة أو ذات دعم أحادي الجانب، وتتميز بأن دعمها يقع غالباً ضمن المجال $[0, \infty)$ وهي لا تحقق شرط التماثل السابق، بل تأخذ شكلاً مائلاً أو منحزاً، بما يتوافق مع التوزيعات المنحرفة أو المحدودة .

أولاً : دالة كيرنل Gamma

تُستخدم دالة Gamma كنواة غير متماثلة (Asymmetric Kernel) لتقدير الكثافة الاحتمالية، خاصة عند التعامل مع البيانات الموجبة والانحراف الإيجابي. يعتمد هذا النوع من الأنوية على التوزيع الاحتمالي Gamma الذي يتمتع بدرجة عالية من المرونة في تمثيل أشكال مختلفة للبيانات.

الصيغة الرياضية

تعطى دالة الكثافة الاحتمالية لتوزيع Gamma بالشكل الآتي:

$$f(x; \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3)$$

إذن: $\alpha > 0$ هو معامل الشكل (shape parameter) . $\beta > 0$ هو معامل المعدل (rate parameter) $\Gamma(\alpha)$ هي دالة غاما القياسية.

وعند استعمال Gamma كدالة كيرنل، تكون النواة:

$$K_h(x, X_i) = \frac{1}{h} f\left(\frac{x}{h}; \alpha, \beta\right) \quad (4)$$

ويكون تقدير الكثافة كما يأتي:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x, X_i) \quad (5)$$

ثانياً : دالة كيرنل Inverse-Gaussian

تُعد دالة Inverse-Gaussian من الدوال اللبية غير المتماثلة التي تُستخدم في تقدير الكثافة الاحتمالية للبيانات التي تقتصر على المجال الموجب فقط. وقد تم اقتراحها كبديل فعال للتقليل من التحيز الناتج عن الحدود، لاسيما في حالة البيانات التي تتركز قرب الصفر.

الصيغة الرياضية

تعطى دالة كيرنل Inverse-Gaussian بالصيغة:

$$K_h(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right), x > 0 \quad (6)$$

إذ إن:

x تمثل نقطة التقدير.

μ متوسط التوزيع، يُؤخذ عادة مساوياً للقيمة التي نقدر الكثافة عندها.

λ معلمة الشكل، تُعبر عن عرض الحزمة (Bandwidth).

ثالثاً : دالة كيرنل ليندلي (Lindley Kernel)

تُعد دالة Lindley Kernel من الدوال غير المتماثلة وتستند إلى توزيع Lindley الذي قدمه Ghitany وآخرون (2008). وتُستخدم عادة لتقدير الكثافة الاحتمالية عندما تكون البيانات موجبة وتحرف إلى اليمين، مثل بيانات البقاء والوقت حتى الفشل.

الصيغة الرياضية:

الصيغة العامة لدالة Lindley Kernel تُعطى كما يأتي :

$$K(x; \theta) = \frac{\theta^2}{1 + \theta} (1 + x) e^{-\theta x} \quad \text{for } x > 0, \theta > 0 \quad (7)$$

حيث θ : هو معلمة الشكل ، x يمثل القيم الموجبة للمتغير. هذه الدالة موجبة الدعم وتناسب مع توزيع Lindley الاحتمالي.

5.3. أثر اختيار معامل التنعيم (Bandwidth)

يُعد معامل التنعيم h العنصر الحاسم في عملية التقدير الكيرنلي للكثافة الاحتمالية، إذ يتحكم بشكل مباشر بدرجة سلاسة التقدير الناتج. وهو يمثل عرض النافذة التي يتم عن طريقها تجميع تأثير كل نقطة ملاحظة حول نقطة التقدير، ومن ثم فإن الاختيار المناسب لهذا المعامل يُعد شرطاً أساسياً للحصول على تقدير دقيق وواقعي للكثافة.

الصيغة العامة لتقدير الكثافة اللبية هي:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K \frac{(x - X_i)}{h} \quad (8)$$

ونلاحظ من هذه المعادلة أن h يظهر في المقام، ما يعني أن له تأثيراً في شكل وحجم الكثافة المقدرة. فإذا كانت قيمة h صغيرة جداً، فإن التقدير سيكون حساساً بشكل مفرط لتقلبات البيانات، ما يؤدي إلى انخفاض التحيز وارتفاع التباين (Overfitting). أما إذا كانت قيمة h كبيرة جداً، فإن التقدير سيكون مفرط التنعيم، ما يؤدي إلى ارتفاع التحيز وانخفاض التباين (Oversmoothing)، ويخفي التفاصيل البنوية الحقيقية للتوزيع. تحقيق التوازن بين التحيز والتباين يُعد الهدف المركزي لاختيار معامل التنعيم. ومن أجل ذلك، طُورت عدة طرائق لاختيار القيمة المثلى لـ h ، من أهمها:

قاعدة الإبهام (Silverman's Rule of Thumb) :

إذ $\hat{\sigma}$ هو الانحراف المعياري للعينة، وتُستخدم هذه القاعدة عندما يُفترض أن البيانات تتبع توزيعاً طبيعياً.

يُعد اختيار معامل التنعيم h خطوة مركزية في طريقة التقدير غير المعلمي لدالة الكثافة الاحتمالية باستعمال دوال كيرنل، إذ يتحكم هذا المعامل بدرجة التنعيم المطبقة على البيانات، وبالتالي يؤثر بشكل مباشر على شكل المقدّر النهائي. من أبرز الطرق المستعملة لتحديد هذا المعامل هي ما يُعرف بقاعدة الإبهام (Rule of Thumb) التي وضعها Silverman (1986)، والتي تستعمل عند افتراض أن البيانات تتبع توزيعاً طبيعياً وأن دالة الكيرنل المختارة هي دالة غاوسية.

تستند هذه الطريقة إلى تقليل الخطأ التكامل التربيعي الوسيط المعروف

بـ AMISE (Asymptotic Mean Integrated Squared Error)، وهو معيار يستعمل لقياس الفرق المتوقع بين دالة

الكثافة الحقيقية $f(x)$ ومقدّر الكثافة $\hat{f}_h(x)$ على مدى مجال x . وعند افتراض أن البيانات تتبع التوزيع الطبيعي القياسي $N(\mu, \sigma^2)$ ، فإن القيمة المثلى التقريبية لمعامل التنعيم تُعطى من خلال المعادلة التالية:

$$h = 1.06 \hat{\sigma} n h^{-1/5} \quad (9)$$

إذ إن :

h : معامل التنعيم (bandwidth)،

$\hat{\sigma}$: الانحراف المعياري للعينة،

n : حجم العينة.

التحقق المتقاطع (Cross-Validation) :

تُعد طريقة التحقق المتقاطع (Cross-Validation) من أكثر الأساليب استعمالاً لتحديد القيمة المثلى لمعامل التنعيم h في تقدير دالة الكثافة الاحتمالية. وتقوم الفكرة الأساسية لهذه الطريقة على تقليل خطأ التقدير عن طريق محاكاة الأداء التنبؤي للنموذج على

بيانات لم تستعمل في عملية التقدير، ما يُسهم في اختيار h الذي يُحقق أفضل توازن بين التحيز (Bias) والتباين (Variance)

6.3. الدالة المقترحة

أولاً: الصيغة الرياضية للدالة المقترحة

تتطلب هذه الدراسة من فكرة إعادة تشكيل دالة كيرنل غير متماثلة تستند إلى خصائص توزيع ويبل، لما يتمتع به هذا التوزيع من ملاءمة طبيعية لتمثيل البيانات الموجبة المنحرفة، لاسيما في مجالات تحليل البقاء والموثوقية. تُعرف دالة الكثافة الاحتمالية لتوزيع ويبل بالصيغة:

$$f(x; \lambda, \kappa) = \frac{\kappa}{\lambda} \left(\frac{x}{\lambda}\right)^{\kappa-1} \exp\left(-\left(\frac{x}{\lambda}\right)^{\kappa}\right), \quad x > 0 \quad (10)$$

إذ إن:

$\lambda > 0$ معلمة القياس (scale parameter).

$\kappa > 0$ معلمة الشكل (shape parameter).

لإنشاء دالة كيرنل غير متماثلة من توزيع ويبل، يتم إعادة تشكيل دالة الكثافة عن طريق استبدال معلمة الشكل κ بمعامل التنعيم h المستخدم في التقدير الكيرنلي. وعليه، نحصل على الدالة اللبية المقترحة:

$$f(x; \lambda, \kappa) = \frac{h}{\lambda} \left(\frac{x}{\lambda}\right)^{h-1} \exp\left(-\left(\frac{x}{\lambda}\right)^h\right), \quad x > 0 \quad (11)$$

هذه الدالة تمثل كيرنل ويبل غير متماثل، متمركز عند الصفر ودعومه من 0 إلى ∞ .

و لغرض التحقق من خصائص دالة الكيرنل و التي هي تتمثل بعدم السالبية: لأن $\kappa > 0$ و $x > 0$ و $h > 0$ ، فإن $K_h(x) \geq 0$.

اما الخاصية الثانية و التي تتمثل بالتكامل إلى واحد، لإثبات أن الدالة متكاملة إلى واحد نغير المتغير بالتكامل: نأخذ بعين الاعتبار دالة الكيرنل المبينة على توزيع ويبل كما في المعادلة (11):

ونهدف إلى إثبات أن هذه الدالة تتكامل إلى واحد، أي:

$$\int_0^{\infty} K_h(x) dx = 1 \quad (12)$$

نقوم بتغيير المتغير كما يلي:

$$u = \left(\frac{x}{\lambda}\right)^h \Rightarrow x = \lambda \cdot u^{\frac{1}{h}} \quad (13)$$

ثم حساب dx بالنسبة إلى du

$$dx = \lambda \cdot \frac{1}{h} \cdot u^{\left(\frac{1}{h}\right)-1} du \quad (14)$$

نُعوّض في التكامل الأصلي:

$$\int_0^{\infty} \frac{h}{\lambda} \left(\frac{x}{\lambda}\right)^{h-1} \exp\left(-\left(\frac{x}{\lambda}\right)^h\right) dx \quad (15)$$

ليصبح:

$$\int_0^{\infty} \frac{h}{\lambda} \cdot u^{(h-1)/h} \cdot e^{-u} \cdot \lambda \cdot \frac{1}{h} \cdot u^{(1/h-1)} du \quad (16)$$

تبسيط الثوابت والأسس

$$\frac{h}{\lambda} \cdot \lambda \cdot \frac{1}{h} = 1$$

يصبح التكامل بالشكل التالي

$$\int_0^{\infty} e^{-u} du = 1 \quad (17)$$

تُظهر النتيجة أن دالة كيرنل ويبل مُطبقة بشكل صحيح ومُقاسة بشكل سليم، لأنها تتكامل إلى واحد على الفترة $(0, \infty)$.
الشكل العام لتقدير الكثافة اللبية باستعمال هذه الدالة، عند استعمال هذه الدالة في تقدير الكثافة اللبية غير المعلمية، نحصل على:

$$\hat{f}h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i) \quad (18)$$

لكن بما أن الدالة غير متماثلة ودعما $x > 0$ ، يُعاد بناءها كدالة تعتمد على المسافة النسبية بين x و X_i ، غالباً بصيغة:

$$\hat{f}h(x) = \frac{1}{n} \sum_{i=1}^n \frac{h}{\lambda} \left(\frac{x}{\lambda}\right)^{h-1} \exp\left(-\left(\frac{x}{\lambda}\right)^h\right) \quad (19)$$

مع ضمان أن $x > 0$.

ثانياً: مميزات الدالة المقترحة

تتميز الدالة اللبية المقترحة، المستندة إلى توزيع ويبل، بعدد من الخصائص النظرية والعملية التي تجعلها مناسبة جداً لتقدير الكثافة الاحتمالية للبيانات الإيجابية وغير المتماثلة، لاسيما في سياق بيانات البقاء والانحرافات الحادة. ويمكن عرض هذه المميزات كما يأتي:

- 1- موجبة الدعم (Positive Support)
- 2- غير متماثلة (Asymmetric Shape)
- 3- قابلة للتكامل إلى واحد (Normalization)
- 4- السيطرة على شكل الكثافة (Shape Flexibility)
- 5- التحيز عند الحواف منخفض (Low Boundary Bias)
- 6- القابلية للتطبيق العملي

الجانب التجريبي

1.4. مفهوم المحاكاة

تُعد المحاكاة أحد الأساليب الإحصائية والتجريبية الفعالة التي تُستخدم لدراسة سلوك النماذج الإحصائية في ظل ظروف متنوعة ومعقدة يصعب تحليلها نظرياً أو تجريبياً بشكل مباشر. وتهدف المحاكاة إلى توليد بيانات صناعية (مصطنعة) من توزيعات محددة ومعروفة، ثم تطبيق نماذج التقدير أو التحليل عليها لتقييم أدائها تحت سيناريوهات مختلفة من أهم مزايا أسلوب المحاكاة:

- القدرة على مقارنة أداء طرائق متعددة في ظروف ماثلة.
 - التحكم الكامل في خصائص البيانات مثل حجم العينة، نوع التوزيع، وجود التحيز أو القيم المتطرفة.
 - توفير بيئة مثالية لاختبار الفرضيات النظرية أو التحقق من مدى ملائمة الطرائق المقترحة.
- وتكمن أهمية المحاكاة في كونها تُعدّ جسراً بين النظرية والتطبيق، إذ تُساعد الباحث في التحقق من صلاحية النماذج الإحصائية المقترحة واختبارها قبل استعمالها في تحليل بيانات حقيقية. كما أنها تُعدّ أداة مركزية في الأبحاث الحديثة لتقييم مرونة الطرائق الجديدة واستجابتها لمختلف التحديات الإحصائية.

4.2 مراحل المحاكاة

تم تنفيذ المحاكاة في هذا البحث على وفق مراحل منهجية متسلسلة تهدف إلى تقييم أداء مجموعة من دوال كيرنل المتماثلة وغير المتماثلة في تقدير دالة الكثافة الاحتمالية. وقد جاءت مراحل المحاكاة على النحو الآتي:

1- تحديد التوزيع للبيانات

تم اختيار توزيع مستمر يمثل الحالة المرجعية التي تُقاس نسبة إليها دقة التقديرات. إذ تم توليد البيانات العشوائية على وفق هذا التوزيع الحقيقي لضمان وجود مرجع واضح للمقارنة.

- 2- اختيار دوال كيرنل المستخدمة في التقدير شملت المقارنة بين دوال كيرنل متماثلة هي Gaussian، Triangular، Biweight، Epanechnikov، و دوال غير متماثلة هي Lindley، Weibull، Gamma1، و Inverse-Gaussian، بهدف فحص تأثير شكل الدالة في دقة التقدير.
- 3- تحديد أحجام العينات تم اعتماد حجوم (200,100,50) حيث يمثل اربع حالات متنوعة تعكس سيناريوهات توزيع مختلفة من حيث التماثل والانحراف والتشتت.
- 4- اختيار طرائق تقدير عرض الحزمة (Bandwidth) تم اعتماد طريقتين لحساب عرض الحزمة: h طريقة Silverman : التي تعتمد على قاعدة تقليدية لحساب h باستعمال التباين وحجم العينة. طريقة Cross-Validation : التي تعتمد على تقليل الخطأ التقديري الكلي للوصول إلى القيمة المثلى لـ h.
- 5- توليد البيانات وتقدير الكثافة لكل سيناريو، تم توليد البيانات وحسب حجم العينة المحدد، وتم استعمال كل دالة كيرنل مع كل طريقة لحساب h لتقدير دالة الكثافة الاحتمالية على البيانات المولدة.
- 6- حساب معيار الأداء (ISE) تم استعمال معيار الخطأ التكامل التريبي (ISE) كمقياس رئيس لتقييم دقة كل تقدير، وذلك عن حساب الفارق التراكمي المربع بين الكثافة التقديرية والكثافة الحقيقية.
- 7- توثيق النتائج وتحليلها تم تلخيص النتائج في جداول مفصلة لكل حالة، فقد أظهرت قيم ISE و h لكل دالة كيرنل وطريقة تقدير، ثم تم رسم أشكال بيانية توضّح تقارب دوال الكثافة التقديرية من الدالة الحقيقية.
- 8- مقارنة الأداء واستنتاج النتائج في نهاية المحاكاة، تم تحليل الجداول والأشكال لتحديد أفضل دالة كيرنل من حيث دقة التقدير في كل سيناريو. وقد تبين أن دالة Weibull غير المتماثلة حققت أدنى ISE في معظم الحالات، بينما كانت دالة Biweight الأفضل بين الدوال المتماثلة.
- 3.4 التجربة عند حجوم (200,100,50)**
- يتم من خلال التجربة تثبيت طريقة تحديد قيمة عرض الحزمة بالطريقتين Silverman و Cross-Validation مع اختيار التوزيع للمتجة المولد وهو التوزيع الاسي، مع تغيير حجوم العينات .
- التجربة : في حالة العينة n=50 الحالة الاولى .**
- جدول (1) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 50 (الحالة الاولى)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.024515468	0.24066516	1.218592527	0.454489796
Weibull	0.016237190	0.123159065	1.218592527	0.454489796
Gamma1	0.217724415	0.144451656	1.218592527	0.454489796
Inverse-Gaussian	0.179102580	0.191510262	1.218592527	0.454489796
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.18292346	0.118186593	1.218592527	0.454489796
Epanechnikov	0.199888417	0.121556958	1.218592527	0.454489796
Biweight	0.120636818	0.09995659	1.218592527	0.454489796
Triangular	0.120919984	0.100810408	1.218592527	0.454489796

- من جدول (1) يتبين لنا الآتي :
- 1- أن الدالة Weibull حققت أقل قيمة لمعيار الخطأ التكامل التريبي ISE عند استعمال طريقة سيلفرمان، إذ سجلت القيمة (0.0162)، ما يدل على دقة تقدير عالية. تليها دالة Lindley بقيمة (0.0245)، ثم دالة Inverse-Gaussian بقيمة (0.1791)، وأخيراً دالة Gamma1 التي سجلت أعلى خطأ (0.2177). وعند مقارنة نفس الدوال باستعمال طريقة التحقق المتقاطع (ISE_CV)، يتبين أن دالة Weibull حافظت أيضاً على أدائها الأفضل مسجلة أقل قيمة (0.1232)، تليها Gamma1 بقيمة (0.1445)، ثم Inverse-Gaussian بقيمة (0.1915)، وأخيراً Lindley التي سجلت أعلى خطأ

(0.2407). وعليه، يمكن القول إن دالة Weibull أظهرت الأداء الأفضل بين الدوال غير المتماثلة على وفق المعيارين ISE_Silverman و ISE_CV معاً.

2- أما بالنسبة للدوال اللبية المتماثلة، فقد أظهرت دالة Biweight تفوقاً واضحاً حسب معيار ISE عند استعمال قاعدة سيلفرمان، إذ حققت أقل قيمة بلغت (0.1206)، تليها دالة Triangular بقيمة (0.1209)، ثم Gaussian بقيمة (0.1829)، وأخيراً Epanechnikov التي سجلت (0.1999). أما عند استعمال طريقة التحقق المتقاطع، فقد حافظت دالة Biweight على مركزها الأول بأقل خطأ (0.1000)، تلتها دالة Triangular (0.1008)، ثم Gaussian (0.1182)، وأخيراً Epanechnikov (0.1216). وبناءً على نتائج كلا المعيارين، تُعد دالة Biweight الأفضل أداءً بين الدوال اللبية المتماثلة في هذه الحالة.

التجربة: في حالة العينة $n=50$ الحالة الثانية.

جدول (2) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 50 (الحالة الثانية)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.024389747	0.92912990	1.247119971	0.313061224
Weibull	0.025155181	0.269326893	1.247119971	0.313061224
Gamma1	0.266504669	0.189362838	1.247119971	0.313061224
Inverse-Gaussian	0.239094310	0.258684421	1.247119971	0.313061224
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.225369737	0.166287111	1.247119971	0.313061224
Epanechnikov	0.236876162	0.166794735	1.247119971	0.313061224
Biweight	0.172424033	0.175669639	1.247119971	0.313061224
Triangular	0.174890921	0.174686441	1.247119971	0.313061224

من جدول (2) يتبين لنا الآتي :

1- أن دالة Lindley حققت أقل قيمة لمعيار الخطأ التكاملية التربيعي (ISE) عند استعمال قاعدة سيلفرمان، إذ سجلت (0.0244)، تليها Weibull بقيمة (0.0252)، ثم Inverse-Gaussian بقيمة (0.2391)، وأخيراً Gamma1 التي سجلت أعلى خطأ بقيمة (0.2665). أما عند استعمال طريقة التحقق المتقاطع (ISE_CV)، فقد جاءت دالة Gamma1 في المرتبة الأولى بأقل خطأ (0.1894)، تليها Inverse-Gaussian (0.2587)، ثم Weibull (0.2693)، وأخيراً Lindley التي سجلت أعلى قيمة (0.9291). وعليه، فإن دالة Weibull تُعد الأفضل أداءً بين الدوال غير المتماثلة، كونها سجلت أداءً جيداً ومستقرًا ضمن المعيارين مقارنةً بباقي الدوال.

2- أما بالنسبة للدوال اللبية المتماثلة، فقد أظهرت دالة Biweight تفوقاً حسب معيار ISE باستعمال قاعدة سيلفرمان، إذ سجلت أقل قيمة بلغت (0.1724)، تليها Triangular (0.1749)، ثم Gaussian (0.2254)، وأخيراً Epanechnikov التي سجلت أعلى خطأ بقيمة (0.2369). بينما عند استعمال معيار ISE_CV، حققت Gaussian أفضل نتيجة (0.1663)، تليها Epanechnikov (0.1668)، ثم Triangular (0.1747)، وأخيراً Biweight (0.1757). وعند النظر إلى كلا المعيارين معاً، نجد أن دالة Triangular حققت توازناً جيداً في الأداء، ما يجعلها الخيار الأنسب من بين الدوال اللبية المتماثلة في هذه الحالة.

التجربة: في حالة العينة $n=50$ الحالة الثالثة.

جدول (3) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 50 (الحالة الثالثة)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.022636453	0.089314058	1.784347439	1
Weibull	0.026407906	0.053190575	1.784347439	1
Gamma1	0.300812061	0.255524585	1.784347439	1
Inverse-Gaussian	0.258909171	0.272700063	1.784347439	1
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.282556499	0.234296409	1.784347439	1

Epanechnikov	0.297006306	0.242977469	1.784347439	1
Biweight	0.212167123	0.182306129	1.784347439	1
Triangular	0.214701167	0.184715595	1.784347439	1

من جدول (3) يتبين لنا الآتي :

1- أن دالة Lindley سجلت أقل قيمة للخطأ التكاملي التربيعي باستعمال قاعدة سيلفرمان (ISE_Silverman) وقد بلغت (0.0226)، تليها Weibull بقيمة (0.0264)، ثم Inverse-Gaussian (0.2589)، وأخيراً Gamma1 التي حققت أعلى خطأ بلغ (0.3008). أما بالنسبة لطريقة التحقق المتقاطع (ISE_CV)، فقد جاءت Weibull في المرتبة الأولى بأقل خطأ (0.0532)، تليها Lindley (0.0893)، ثم Gamma1 (0.2555)، وأخيراً Inverse-Gaussian (0.2727). وبمقارنة الأداء عبر كلا المعيارين، نجد أن دالة Weibull أظهرت أفضل أداء إجمالي من إذ الاستقرار وانخفاض الخطأ، ما يجعلها الدالة الأنسب ضمن الدوال غير المتماثلة في هذه الحالة.

2- أما بالنسبة للدوال اللبية المتماثلة، فقد أظهرت Biweight أفضل أداء يحسب معيار ISE_Silverman، إذ سجلت (0.2122)، تليها Triangular (0.2147)، ثم Gaussian (0.2826)، وأخيراً Epanechnikov (0.2970). أما عند استعمال طريقة التحقق المتقاطع (ISE_CV)، فقد حافظت Biweight أيضاً على المركز الأول بأقل خطأ (0.1823)، تليها Triangular (0.1847)، ثم Gaussian (0.2343)، وأخيراً Epanechnikov (0.2430). وبذلك، فإن دالة Biweight تُعد الأفضل أداءً ضمن الدوال المتماثلة وفقاً لكلا المعيارين، ما يعزز كفاءتها في تقدير الكثافة الاحتمالية في هذه الحالة التجريبية.

التجربة : في حالة العينة $n=50$ الحالة الرابعة .

جدول (4) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 50 (الحالة الرابعة)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.01713042	0.15594716	1.995065963	1
Weibull	0.0086319	0.108992853	1.995065963	1
Gamma1	0.347392486	0.311307114	1.995065963	1
Inverse-Gaussian	0.297443261	0.319267687	1.995065963	1
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.321784227	0.289959651	1.995065963	1
Epanechnikov	0.329648666	0.295600384	1.995065963	1
Biweight	0.283509897	0.251633882	1.995065963	1
Triangular	0.284387227	0.254863171	1.995065963	1

من جدول (4) يتبين لنا الآتي :

1- أن دالة Weibull سجلت أقل قيمة للخطأ التكاملي التربيعي باستعمال قاعدة سيلفرمان (ISE_Silverman) وقد بلغت (0.0086)، تليها Lindley (0.0171)، ثم Inverse-Gaussian (0.2974)، وأخيراً Gamma1 التي سجلت أعلى قيمة خطأ (0.3474). أما عند استعمال طريقة التحقق المتقاطع (ISE_CV)، فقد واصلت Weibull تحقيق أفضل أداء بأقل خطأ بلغ (0.1090)، تلتها Lindley (0.1559)، ثم Gamma1 (0.3113)، وأخيراً Inverse-Gaussian (0.3193). وبناءً على كلا المعيارين، تنصدر دالة Weibull ترتيب الدوال غير المتماثلة في هذه الحالة، ما يؤكد كفاءتها العالية في تقدير الكثافة الاحتمالية مقارنةً ببقية الدوال.

2- أما بالنسبة لدوال كيرنل المتماثلة، فقد أظهرت Biweight أفضل أداء عند استعمال قاعدة سيلفرمان (ISE_Silverman) بتسجيلها أقل خطأ (0.2835)، تليها Triangular (0.2844)، ثم Gaussian (0.3218)، وأخيراً Epanechnikov (0.3296). وباستعمال طريقة التحقق المتقاطع (ISE_CV)، استمرت Biweight في التفوق محققة أقل خطأ (0.2516)، تليها Triangular (0.2549)، ثم Gaussian (0.2900)، وأخيراً Epanechnikov (0.2956). لذلك، تعد دالة Biweight الأفضل أداءً من بين الدوال المتماثلة في هذه الحالة، ما يعكس قدرتها على تحقيق تقدير دقيق وثابت للكثافة الاحتمالية.

وبشكل عام من نتائج الجداول الأربعة (1-3) إلى (4-3)، والتي استعرضت أداء الدوال غير المتماثلة Weibull، Lindley، Gamma1 و Inverse-Gaussian وفقاً لمعيار ISE_Silverman و ISE_CV، يمكن استخلاص الآتي:

1- بالنسبة لمعيار ISE_Silverman، أظهرت دالة Weibull تفوقاً واضحاً في ثلاث حالات من أصل أربع، فقد سجلت أقل قيمة للخطأ التكاملي التربيعي في الجداول (3-1)، و(3-4)، وكانت ثاني أفضل قيمة في الجدول (3-3)، ما يبين استقرارها وأداءها العالي. أما في الجدول (3-2)، فقد جاءت دالة Lindley في الصدارة بفارق طفيف، إلا أن الأداء العام عبر الحالات يُرجح كفة دالة Weibull كأفضل دالة على وفق هذا المعيار. أما على وفق معيار ISE_CV، فقد واصلت دالة Weibull تقديم الأداء الأفضل، إذ حققت أقل قيمة للخطأ في الجداول (3-1)، و(3-3)، و(3-4)، بينما تفوقت دالة Gamma1 فقط في الجدول (3-2). ونلاحظ أن دالة Weibull لم تقتصر على تحقيق أفضل النتائج في معظم الحالات فحسب، بل كانت الفروق لصالحها واضحة مقارنةً ببقية الدوال. عند النظر إلى أداء كل دالة على وفق كلا المعيارين معاً، نجد أن دالة Weibull تميزت بانخفاض مستمر وثابت في قيم الخطأ، سواء باستعمال قاعدة سيلفرمان أو طريقة التحقق المتقاطع، ما يجعلها أفضل خيار بين الدوال غير المتماثلة في تقدير الكثافة الاحتمالية عند حجم العينة 50 عبر الحالات الأربع.

2- أما بالنسبة للدوال المتماثلة فإن لمعيار ISE_Silverman، سجلت دالة Biweight أقل قيمة للخطأ التكاملي التربيعي في الجداول (1)، (3)، و(4)، بينما كانت في المركز الثاني في الجدول (2)، ما يعكس تفوقها الواضح واستقرار أدائها. أما دالة Triangular، فقد حافظت على موقع متقدم نسبياً، لكنها لم تتفوق على Biweight في أي حالة، في حين سجلت كل من Gaussian وEpanechnikov قيمة أعلى نسبياً في معظم الحالات، ما يضعف أداءهما تحت هذا المعيار. أما عند استعمال معيار ISE_CV، فقد واصلت دالة Biweight تفوقها في الجداول (1)، (3)، و(4)، وسجلت المركز الرابع فقط في الجدول (2). وعلى الرغم من أن Gaussian حققت أفضل قيمة في الجدول (2)، إلا أن أداءها لم يكن ثابتاً عبر باقي الحالات. وتُظهر Triangular استقراراً نسبياً، إذ جاءت في المرتبة الثانية في أكثر من حالة.

عند الدمج بين المعيارين، نلاحظ أن دالة Biweight هي الأكثر استقراراً وكفاءةً بين الدوال المتماثلة، إذ احتلت المرتبة الأولى في معظم الحالات، على وفق كلا المعيارين. لذا، يمكن عدها أفضل دالة كيرنل متماثلة لأغراض تقدير الكثافة الاحتمالية عند حجم عينة 50، عبر الحالات التجريبية الأربعة.

التجربة : في حالة العينة $n=100$ الحالة الأولى .

جدول (5) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 100 (الحالة الأولى)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.019612397	0.0641654	1.384800372	0.979795918
Weibull	0.018467851	0.054067783	1.384800372	0.979795918
Gamma1	0.285943569	0.260187129	1.384800372	0.979795918
Inverse-Gaussian	0.262725286	0.269691136	1.384800372	0.979795918
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.265523214	0.241777699	1.384800372	0.979795918
Epanechnikov	0.275574742	0.25248501	1.384800372	0.979795918
Biweight	0.202566773	0.181752512	1.384800372	0.979795918
Triangular	0.206220664	0.184355734	1.384800372	0.979795918

من جدول (5) يتبين لنا الآتي:

1- بالنسبة للدوال غير المتماثلة، سجلت دالة Weibull أقل قيمة للخطأ التكاملي التربيعي على وفق طريقة سيلفرمان (ISE_Silverman) وبلغت (0.0185)، تليها Lindley (0.0196)، ثم Inverse-Gaussian (0.2627)، وأخيراً Gamma1 التي سجلت أعلى خطأ (0.2859). أما عند استعمال طريقة التحقق المتقاطع (ISE_CV)، فقد حافظت Weibull على موقعها الأول بأقل قيمة خطأ (0.0541)، تليها Lindley (0.0642)، ثم Gamma1 (0.2602)، وأخيراً Inverse-Gaussian (0.2697). يتضح من ذلك أن دالة Weibull قدمت الأداء الأفضل والأكثر استقراراً بين الدوال غير المتماثلة في هذه الحالة، وفقاً لكلا المعيارين.

2- أما بالنسبة للدوال اللبية المتماثلة، فقد سجلت دالة Biweight أقل قيمة للخطأ على وفق طريقة سيلفرمان (0.2026)، تليها Triangular (0.2062)، ثم Gaussian (0.2655)، وأخيراً Epanechnikov (0.2756). وباستعمال طريقة التحقق

المتقاطع، استمرت Biweight في تحقيق أفضل أداء بأقل قيمة (0.1818)، تلتها (0.1844) Triangular، ثم Gaussian (0.2418)، وأخيراً Epanechnikov (0.2525) وعليه، فإن دالة Biweight تُعد الخيار الأمثل من بين الدوال المتماثلة في هذه الحالة، نتيجة لأدائها المتفوق ضمن كلا المعيارين.

التجربة : في حالة العينة $n=100$ الحالة الثانية .

جدول (6) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختبار عرض الحزمة على وفق معيار ISE عند حجم عينة 100 (الحالة الثانية)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.020366966	0.091178705	1.602799913	0.898979592
Weibull	0.013674719	0.055936415	1.602799913	0.898979592
Gamma1	0.288093202	0.241544527	1.602799913	0.898979592
Inverse-Gaussian	0.256048765	0.268066463	1.602799913	0.898979592
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.269252259	0.224109073	1.602799913	0.898979592
Epanechnikov	0.282060801	0.23371209	1.602799913	0.898979592
Biweight	0.200991607	0.166282646	1.602799913	0.898979592
Triangular	0.204302877	0.168876121	1.602799913	0.898979592

من جدول (6) يتبين لنا ما يأتي:

1- بالنسبة للدوال غير المتماثلة، سجلت دالة Weibull أقل قيمة للخطأ التكاملي التريبيعي باستعمال قاعدة سيلفرمان وبلغت (0.0137)، تلتها Lindley بقيمة (0.0204)، ثم Inverse-Gaussian بقيمة (0.2560)، وأخيراً Gamma1 التي حققت أعلى خطأ بقيمة (0.2881). أما وفقاً لطريقة التحقق المتقاطع، فقد واصلت Weibull تحقيق أفضل أداء بأقل خطأ (0.0559)، تلتها Lindley بقيمة (0.0912)، ثم Gamma1 بقيمة (0.2415)، وأخيراً Inverse-Gaussian بقيمة (0.2681). يتضح من ذلك أن دالة Weibull تُعد الأفضل أداءً ضمن هذه المجموعة في هذه الحالة التجريبية، وفقاً لكلا المعيارين.

2- أما بالنسبة للدوال المتماثلة، فقد سجلت دالة Biweight أقل قيمة لمعيار الخطأ التكاملي التريبيعي باستعمال قاعدة سيلفرمان بقيمة (0.2010)، تلتها Triangular (0.2043)، ثم Gaussian (0.2693)، وأخيراً Epanechnikov (0.2821) وعند استعمال معيار ISE_CV، واصلت Biweight أيضاً تفوقها محققة أقل قيمة (0.1663)، تلتها Triangular (0.1689)، ثم Gaussian (0.2241)، وأخيراً Epanechnikov (0.2337). يشير ذلك إلى أن دالة Biweight أظهرت أداءً متميزاً وثابتاً ضمن الدوال المتماثلة وفق المعيارين معاً.

التجربة : في حالة العينة $n=100$ الحالة الثالثة .

جدول (7) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختبار عرض الحزمة على وفق معيار ISE عند حجم عينة 100 (الحالة الثالثة)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.021056838	0.096796219	1.308011199	0.898979592
Weibull	0.027044276	0.076981043	1.308011199	0.898979592
Gamma1	0.287492586	0.262444602	1.308011199	0.898979592
Inverse-Gaussian	0.27173814	0.280007786	1.308011199	0.898979592
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.25752621	0.241211117	1.308011199	0.898979592
Epanechnikov	0.263317597	0.24490187	1.308011199	0.898979592
Biweight	0.224130221	0.209878745	1.308011199	0.898979592
Triangular	0.225045941	0.211106478	1.308011199	0.898979592

من جدول (7) يتبين لنا ما يأتي:

1- بالنسبة للدوال غير المتماثلة، أظهرت دالة Lindley أفضل أداء باستعمال قاعدة سيلفرمان، إذ سجلت أقل قيمة للخطأ التكاملي التريبيعي (0.0211)، تلتها Weibull (0.0270)، ثم Inverse-Gaussian (0.2717)، وأخيراً Gamma1 التي حققت أعلى

خطأ (0.2875). أما عند استعمال طريقة التحقق المتقاطع، فقد حافظت Weibull على أدائها الجيد بأقل قيمة خطأ (0.0770)، تلتها Lindley (0.0968)، ثم Gamma1 (0.2624)، وأخيراً Inverse-Gaussian (0.2800) وعموماً، يتضح أن كل من Weibull وLindley أظهرتا نتائج مقاربة، إلا أن Weibull تتفوق قليلاً عند استعمال طريقة التحقق المتقاطع، بينما تنصدر Lindley عند استعمال قاعدة سيلفرمان.

2- أما بالنسبة للدوال المتماثلة، فقد سجلت دالة Biweight أقل قيمة لمعيار الخطأ التكاملي التريبيعي باستعمال قاعدة سيلفرمان وبلغت (0.2241)، تلتها Triangular (0.2250)، ثم Gaussian (0.2575)، وأخيراً Epanechnikov (0.2633) وعند استعمال طريقة التحقق المتقاطع، واصلت Biweight تفوقها بقيمة (0.2099)، تلتها Triangular (0.2111)، ثم Gaussian (0.2412)، وأخيراً Epanechnikov (0.2449). تشير هذه النتائج إلى أن دالة Biweight حافظت على أدائها الأفضل بين الدوال المتماثلة وفق المعيارين معاً.

التجربة : في حالة العينة $n=100$ الحالة الرابعة .

جدول (8) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 100 (الحالة الرابعة)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.0222578	0.194908646	1.65584773	0.75755102
Weibull	0.009187768	0.110700337	1.65584773	0.75755102
Gamma1	0.30777549	0.254897759	1.65584773	0.75755102
Inverse-Gaussian	0.274656714	0.28909973	1.65584773	0.75755102
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.279620927	0.233541334	1.65584773	0.75755102
Epanechnikov	0.28975204	0.241209362	1.65584773	0.75755102
Biweight	0.228463044	0.189372141	1.65584773	0.75755102
Triangular	0.230327628	0.191714542	1.65584773	0.75755102

من جدول (8) يتبين ما يأتي:

1- بالنسبة للدوال غير المتماثلة، جاءت دالة Weibull في المرتبة الأولى من إذ الأداء الأفضل باستعمال قاعدة سيلفرمان، إذ سجلت أقل قيمة لمعيار الخطأ التكاملي التريبيعي بلغت (0.0092)، تلتها Lindley (0.0223)، ثم Inverse-Gaussian (0.2747)، وأخيراً Gamma1 التي سجلت أعلى خطأ (0.3078). وعند استعمال طريقة التحقق المتقاطع، حافظت Weibull أيضاً على المركز الأول بأقل خطأ (0.1107)، تلتها Lindley (0.1949)، ثم Gamma1 (0.2549)، وأخيراً Inverse-Gaussian (0.2891). توّضح هذه النتائج أن دالة Weibull أظهرت أداءً متفوقاً ومستقراً على وفق المعيارين، ما يؤكد كفاءتها في تقدير الكثافة الاحتمالية عند هذا الحجم من العينة.

2- أما بالنسبة للدوال المتماثلة، فقد سجلت دالة Biweight أقل قيمة لمعيار ISE عند استعمال قاعدة سيلفرمان وبلغت (0.2285)، تلتها Triangular (0.2303)، ثم Gaussian (0.2796)، وأخيراً Epanechnikov (0.2898) وباستعمال طريقة التحقق المتقاطع، واصلت Biweight تفوقها أيضاً بأقل خطأ (0.1894)، تلتها Triangular (0.1917)، ثم Gaussian (0.2335)، وأخيراً Epanechnikov (0.2412). وبذلك، يتضح أن دالة Biweight حققت أفضل أداء بين الدوال المتماثلة في هذه الحالة وفقاً لكلا المعيارين.

استناداً إلى نتائج الجداول الأربعة (5) إلى (8) اللاسما بحجم العينة 100، والتي تناولت مقارنة أداء دوال كيرنل غير متماثلة (Lindley)، (Weibull)، (Gamma1)، (Inverse-Gaussian) ومتماثلة (Gaussian)، (Epanechnikov)، (Biweight)، (Triangular) باستعمال طريقتي Silverman و Cross-Validation لاختيار عرض الحزمة على وفق معيار ISE، يمكن تلخيص النتائج كما يأتي:

1- بالنسبة للدوال غير المتماثلة، يتبين أن دالة Weibull قدّمت الأداء الأفضل في ثلاث حالات من أصل أربع على وفق معيار ISE_Silverman، إذ حققت أقل قيمة للخطأ في الجداول (5)، (6)، و(8)، وكانت قريبة من الأفضل في الجدول (7). أما على وفق معيار ISE_CV، فقد حافظت Weibull على صدارتها في ثلاث جداول أيضاً، ما يدل على ثباتها وموثوقيتها العالية في تقدير الكثافة. في المقابل، أظهرت دالة Lindley أداءً جيداً في بعض الحالات، لكنها لم تكن الأفضل عموماً. أما Gamma1 و Inverse-Gaussian فقد سجلتا أداءً ضعيفاً نسبياً عبر معظم الحالات على وفق كلا المعيارين.

2- بالنسبة للدوال المتماثلة، تُظهر النتائج أن دالة Biweight قدّمت أفضل أداء إجمالي بين نظيراتها، إذ سجلت أقل خطأ على وفق معيار ISE_Silverman في جميع الجداول الأربعة، واحتفظت كذلك بالمركز الأول أو الثاني ضمن معيار ISE_CV،

ما يبرز كفاءتها العالية. في المقابل، جاءت دالة Triangular في المرتبة الثانية في أغلب الحالات، في حين سجلت Gaussian و Epanechnikov قيمًا أعلى للخطأ، ما يشير إلى أداء أقل استقرارًا مقارنةً بـ Biweight. بناءً على ذلك، يمكن اعتماد دالتي Weibull و Biweight كأفضل الخيارات في تقدير الكثافة الاحتمالية عند حجم عينة 100، لما أظهرته من تفوق واضح وثبات في الأداء عبر التجارب الأربع، وعلى وفق كلا المعيارين.

التجربة : في حالة العينة $n=200$ الحالة الاولى .

جدول (9) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 200 (الحالة الاولى)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.118603337	0.501300072	0.059385412	0.030204082
Weibull	0.059448143	0.08047379	0.059385412	0.030204082
Gamma1	1.608724328	2.167267082	0.059385412	0.030204082
Inverse-Gaussian	1.826987922	1.916637945	0.059385412	0.030204082
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	2.291823083	2.690966239	0.059385412	0.030204082
Epanechnikov	2.271330106	2.685837691	0.059385412	0.030204082
Biweight	2.787030896	2.920125022	0.059385412	0.030204082
Triangular	2.763574749	2.909700061	0.059385412	0.030204082

من جدول (9) يتبين الآتي:

1- بالنسبة للدوال غير المتماثلة، سجلت دالة Weibull أقل قيمة لمعيار ISE باستعمال قاعدة سيلفرمان، إذ بلغت (0.0594)، تليها دالة Lindley (0.1186) التي سجلت (1.6087)، وأخيرًا Inverse-Gaussian التي سجلت أعلى قيمة بلغت (1.8270). أما وفق معيار ISE_CV، فقد حافظت دالة Weibull على تفوقها بأقل قيمة (0.0805)، تليها Lindley (0.5013)، ثم Inverse-Gaussian (1.9166)، وأخيرًا Gamma1 (2.1673). وهذا يشير بوضوح إلى أن دالة Weibull هي الأفضل أداءً بين الدوال غير المتماثلة في هذه الحالة من حيث دقة تقدير الكثافة باستعمال كلا المعيارين.

2- بالنسبة للدوال المتماثلة، سجلت جميع الدوال قيمًا مرتفعة نسبيًا لمعيار ISE، ما يدل على ضعف الأداء العام لها في هذه الحالة. فقد كانت أدنى القيم ضمن معيار ISE_Silverman تعود إلى دالة (2.2713) Epanechnikov و Gaussian (2.2918)، بينما كانت أعلى القيم من نصيب Biweight (2.7870) و Triangular (2.7636). أما وفق معيار ISE_CV، فقد كانت Gaussian الأفضل أداءً نسبيًا بقيمة (2.6910)، تليها Epanechnikov (2.6858)، بينما سجلت Biweight و Triangular أعلى القيم، ما يشير إلى أن دوال كيرنل المتماثلة جميعها قدمت أداءً ضعيفًا نسبيًا مقارنة بالدوال غير المتماثلة في هذه الحالة، ولم تُظهر أية دالة منها تميزًا واضحًا.

التجربة : في حالة العينة $n=200$ الحالة الثانية .

جدول (10) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 200 (الحالة الثانية)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.082232122	0.380990234	0.060945138	0.030204082
Weibull	0.031624464	0.054152947	0.060945138	0.030204082
Gamma1	1.757145252	2.484924156	0.060945138	0.030204082
Inverse-Gaussian	1.952049044	2.030783363	0.060945138	0.030204082
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	2.433636144	3.102902505	0.060945138	0.030204082
Epanechnikov	2.390210671	3.075629035	0.060945138	0.030204082
Biweight	3.288381614	3.609210703	0.060945138	0.030204082

Triangular	3.235309074	3.585097951	0.060945138	0.030204082
------------	-------------	-------------	-------------	-------------

من خلال جدول (10) يتضح ما يأتي:

1- بالنسبة للدوال غير المتماثلة، حققت دالة Weibull الأداء الأفضل بوضوح عند استعمال قاعدة سيلفرمان، حيث سجلت أقل قيمة لمعيار ISE بلغت (0.0316)، تلتها دالة Lindley التي سجلت (0.0822)، ثم Gamma1 بقيمة مرتفعة (1.7571)، وأخيراً Inverse-Gaussian بأعلى خطأ بلغ (1.9520). وبالنسبة لطريقة التحقق المتقاطع، واصلت دالة Weibull تفوقها بتسجيلها أقل قيمة ISE CV بلغت (0.0542)، تلتها Lindley (0.3810)، ثم Inverse-Gaussian (2.0308)، وأخيراً Gamma1 التي سجلت أعلى خطأ (2.4849). يدل هذا الأداء الثابت على أن دالة Weibull ما تزال تحتفظ بمركزها كأفضل دالة غير متماثلة من حيث دقة التقدير.

2- أما فيما يخص دوال كيرنل المتماثلة، فقد سجلت جميعها قيمة مرتفعة للخطأ التكاملي التريبيعي، سواء باستعمال سيلفرمان أو التحقق المتقاطع. فوفق معيار ISE Silverman، جاءت Gaussian و Epanechnikov بأفضل القيم نسبياً (2.4336) و (2.3902)، بينما سجلت Biweight و Triangular أعلى القيم. وضمن معيار ISE CV، كانت Gaussian الأفضل مرة أخرى نسبياً (3.1029)، تلتها Epanechnikov (3.0756)، في حين حافظت Biweight و Triangular على أدائهما الأضعف. يشير ذلك إلى أن أداء الدوال المتماثلة بقي ضعيفاً في هذه الحالة، من دون تقديم أي دالة منها لنتائج مميزة بالمقارنة مع نظيراتها غير المتماثلة.

التجربة : في حالة العينة $n=200$ الحالة الثالثة .

جدول (11) مقارنة أداء دوال كيرنل المتماثلة و غير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختيار عرض الحزمة على وفق معيار ISE عند حجم عينة 200 (الحالة الثالثة)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.116338472	0.386284867	0.054452669	0.030204082
Weibull	0.086314716	0.060227944	0.054452669	0.030204082
Gamma1	2.063108917	2.74067567	0.054452669	0.030204082
Inverse-Gaussian	2.165680714	2.255804307	0.054452669	0.030204082
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	2.767783983	3.29425515	0.054452669	0.030204082
Epanechnikov	2.732874931	3.280343623	0.054452669	0.030204082
Biweight	3.520416715	3.882818036	0.054452669	0.030204082
Triangular	3.477840543	3.834404228	0.054452669	0.030204082

من خلال جدول (11) يتضح ما يأتي:

1- بالنسبة للدوال غير المتماثلة، حققت دالة Weibull أقل قيمة للخطأ التكاملي التريبيعي باستعمال قاعدة سيلفرمان، إذ بلغت (0.0863)، تلتها Lindley بقيمة (0.1163)، في حين جاءت Gamma1 و Inverse-Gaussian بنتائج مرتفعة جداً تجاوزت (2.0)، ما يدل على ضعف أدائهما في هذه الحالة. أما عند استعمال طريقة التحقق المتقاطع، فقد حافظت Weibull على تفوقها مسجلة أقل خطأ (0.0602)، تلتها Lindley (0.3863)، ثم Inverse-Gaussian (2.2558)، وأخيراً Gamma1 التي سجلت أعلى خطأ (2.7407). يتضح من ذلك أن دالة Weibull قدّمت الأداء الأفضل من حيث دقة التقدير وفق كلا المعيارين.

2- أما فيما يخص الدوال اللبية المتماثلة، فقد أظهرت جميعها أداءً ضعيفاً في هذه الحالة، فقد سجلت Gaussian و Epanechnikov أقل القيم ضمن هذه المجموعة لكنها ما زالت مرتفعة (2.7678 و 2.7329) على الترتيب وفق (ISE Silverman)، بينما جاءت Biweight و Triangular بأعلى الأخطاء. وتكررت النتيجة ذاتها مع معيار ISE CV، إذ حققت Gaussian و Epanechnikov أفضل القيم نسبياً (3.2943 و 3.2803 على الترتيب)، بينما سجلت Biweight و Triangular أعلى الخطأ. يشير هذا إلى أن أداء الدوال المتماثلة بقي ضعيفاً ولم تتمكن من مجاراة دقة الدوال غير المتماثلة، لا سيما دالة Weibull.

التجربة : في حالة العينة $n=200$ الحالة الرابعة .

جدول (12) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و طريقة Cross-Validation في اختبار عرض الحزمة على وفق معيار ISE عند حجم عينة 200 (الحالة الرابعة)

AsKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Lindley	0.030303188	0.154063057	0.233561408	0.151428571
Weibull	0.008047204	0.044088595	0.233561408	0.151428571
Gamma1	0.053292372	0.094833566	0.233561408	0.151428571
Inverse-Gaussian	0.044908406	0.054125373	0.233561408	0.151428571
SKernel	ISE_Silverman	ISE_CV	h_Silverman	h_CV
Gaussian	0.118872684	0.168350907	0.233561408	0.151428571
Epanechnikov	0.117985904	0.166946647	0.233561408	0.151428571
Biweight	0.218572121	0.246543051	0.233561408	0.151428571
Triangular	0.212216584	0.242021352	0.233561408	0.151428571

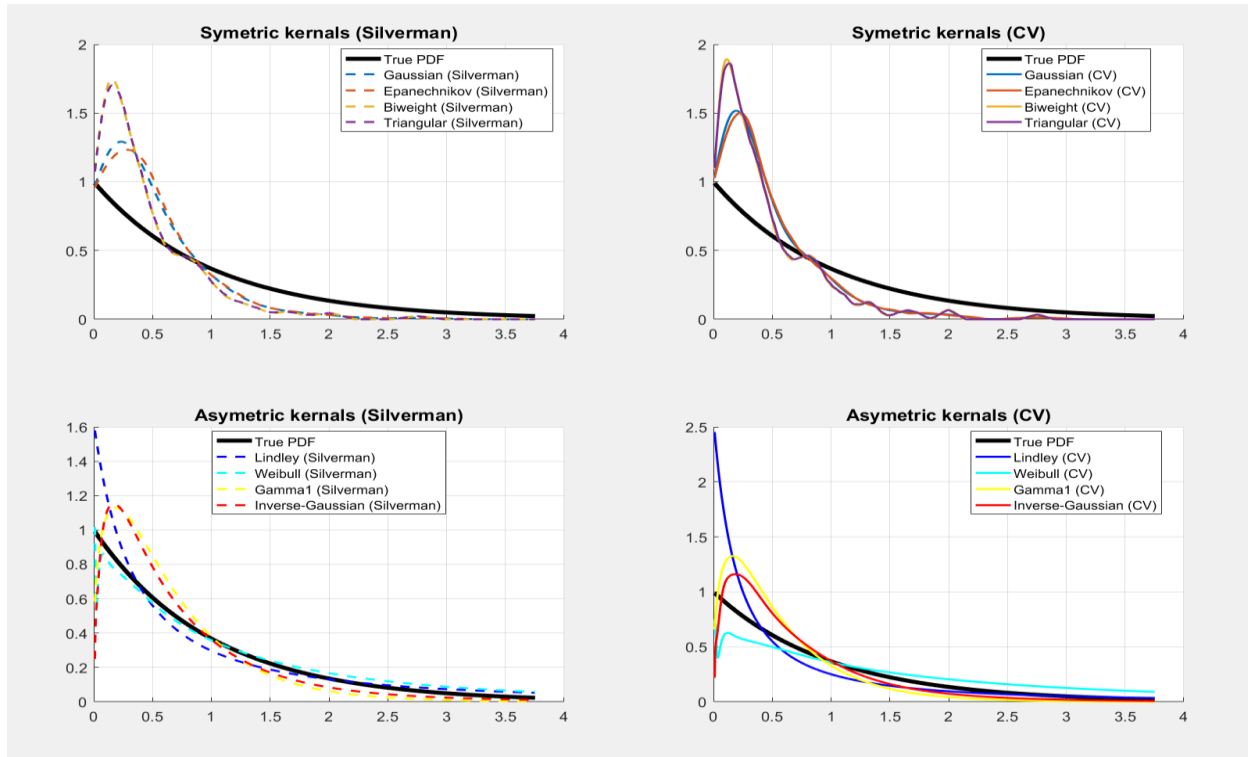
جدول (12) يتبين :

أن دالة Weibull قد سجلت أقل قيمة للخطأ التكاملي التربيعي باستعمال قاعدة Silverman ، إذ بلغت (0.0080)، تلتها دالة Inverse-Gaussian بقيمة (0.0449)، ثم Gamma1 (0.0533)، وأخيراً Lindley (0.0303). أما عند استعمال طريقة Cross-Validation ، فقد استمرت دالة Weibull بتحقيق الأداء الأفضل عن طريق تسجيلها أقل خطأ بقيمة (0.0441)، تليها Inverse-Gaussian (0.0541)، ثم Gamma1 (0.0948)، وأخيراً Lindley (0.1541) بناءً على ذلك، فإن دالة Weibull هي الأفضل من بين الدوال غير المتماثلة في هذه الحالة وفق كلا المعيارين. أما بالنسبة للدوال اللبية المتماثلة، فقد حققت دالة Epanechnikov أفضل أداء على وفق معيار Silverman مسجلة خطأ قدره (0.1180)، تليها Gaussian (0.1189)، بينما جاءت Biweight (0.2186) و Triangular (0.2122) بترتيب أكبر. وعند استعمال Cross-Validation ، حافظت Epanechnikov على تفوقها بتسجيلها (0.1669)، تليها Gaussian (0.1684)، ثم Triangular و Biweight. وبذلك تُعد دالة Epanechnikov الفضلى بين الدوال المتماثلة في هذه الحالة على وفق المعيارين معاً.

من الجداول الأربعة الخاصة بحجم العينة 200، يتبين أن دالة Weibull كانت الأكثر استقراراً وتفقاً من بين الدوال غير المتماثلة، حيث سجلت أقل قيمة لمعيار الخطأ التكاملي التربيعي (ISE_Silverman) في جميع الحالات الأربعة، وهي: (0.0594)، (0.0316)، (0.0863)، و (0.0080) على الترتيب. ما عند استعمال طريقة Cross-Validation ، فقد واصلت دالة Weibull تحقيق أقل القيم كذلك، وهي: (0.0541)، (0.0542)، (0.0602)، و (0.0441)، ما يدل على دقتها وثباتها العالي في التقدير.

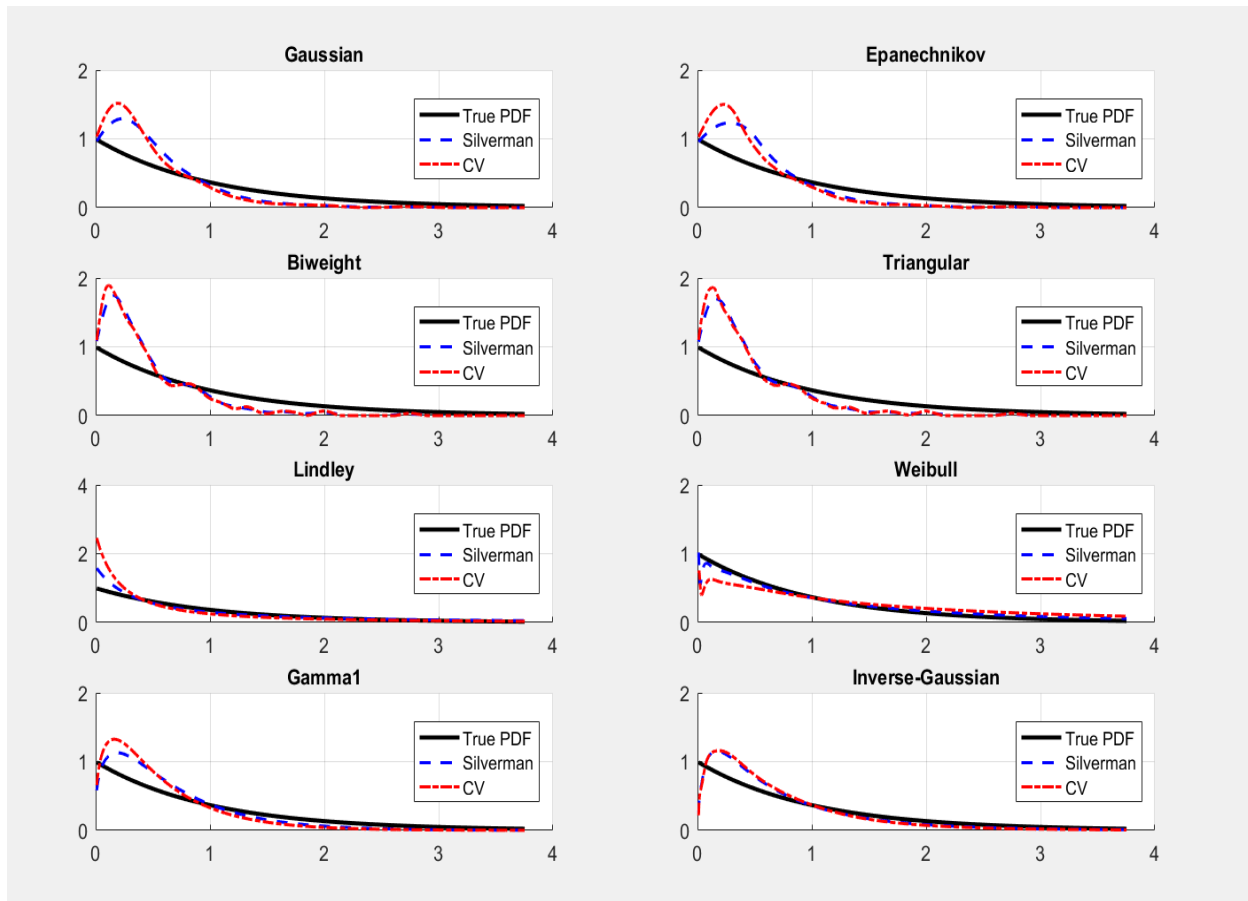
أما بالنسبة للدوال اللبية المتماثلة، فقد أظهرت دالة Epanechnikov أداءً أفضل نسبياً وفق معيار Silverman في الحالات الثانية والرابعة، بينما سجلت دالة Biweight أقل القيم في الحالة الثالثة، وكانت Triangular الأفضل في الحالة الأولى. وعند استعمال معيار Cross-Validation ، برزت Epanechnikov أيضاً بأداء متميز في الحالتين الثانية والرابعة، في حين تفوقت Biweight في الحالة الثالثة، و Triangular في الحالة الأولى. وبمقارنة الأداء الكلي، يمكن عد دالة Epanechnikov الأكثر تفوقاً من بين الدوال المتماثلة في هذا الحجم من العينة.

وبمقارنة الأداء العام للدوال غير المتماثلة مقابل المتماثلة، يتضح أن الدوال غير المتماثلة، وعلى وجه الخصوص دالة Weibull ، قدّمت أداءً متفوقاً في تقدير الكثافة الاحتمالية، إذ حققت مستويات منخفضة من الخطأ التكاملي التربيعي عبر مختلف أحجام العينات والحالات. وقد تم تسجيل أقل قيمة مطلقة لمعيار ISE_Silverman بين جميع الدوال عند استعمال دالة Weibull ، حيث بلغت 0.008047204 وذلك في الحالة الرابعة عند حجم العينة $n = 200$ باستعمال طريقة Silverman ، مع عرض حزمة $h = 0.233561408$. هذا الأداء يُبرز مدى كفاءة الدوال غير المتماثلة في تمثيل البيانات المنحرفة مقارنةً بالدوال المتماثلة، التي أظهرت أداءً أضعف نسبياً في هذا السياق.



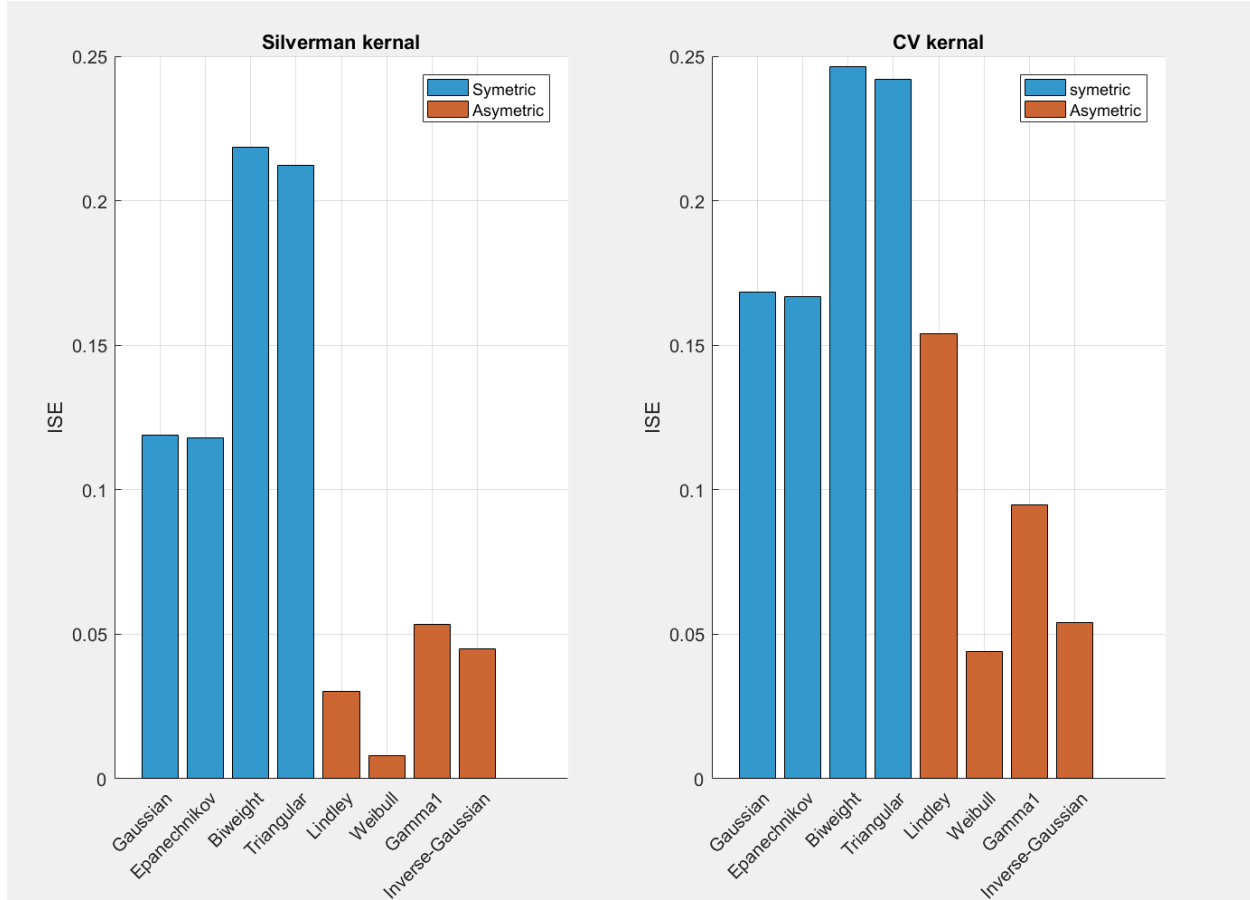
شكل (1) مقارنة أداء دوال كيرنل المتماثلة وغير المتماثلة في تقدير دالة الكثافة الاحتمالية مقارنة بالقيمة الحقيقية في التجربة

من الشكل (1) أنفاً، يتبين لنا أن دالة Weibull (باللون السماوي) تُعد الأكثر تطابقاً مع الكثافة الحقيقية (الخط الأسود العريض) ضمن جميع دوال كيرنل، سواء المتماثلة أو غير المتماثلة، ولا سيما في المنطقتين القريبة من الذروة والذيل الأيسر للتوزيع. هذا التوافق البصري الواضح يعكس قدرة Weibull العالية على تمثيل البيانات غير المتماثلة بدقة ومرونة، ويُعزز من فاعليتها في تقليل الخطأ التقديري مقارنة بباقي الدوال. كما يُلاحظ أن بقية الدوال، ولا سيما المتماثلة مثل Gaussian و Triangular، فشلت في التقاط الشكل الفعلي للتوزيع، ما يدل على محدودية أدائها عند وجود انحراف. وعليه، يُمكن الاستنتاج أن الطريقة المقترحة باستعمال دالة Weibull تُوفر تقديرًا أكثر واقعية وثباتًا في مثل هذا النوع من البيانات.



شكل (2) مقارنة أداء دوال كيرنل المختلفة باستعمال طريقتي Silverman و Cross-Validation في تقدير دالة الكثافة الاحتمالية مقارنة بالدالة الحقيقية في التجربة

من الشكل (2) أنفاً، يتبين لنا أن الدالة المقترحة Weibull قد أظهرت توافقاً عالياً مع الدالة الحقيقية للكثافة الاحتمالية، سواء عند استعمال طريقة Silverman أو Cross-Validation. فقد تميز منحناها بالانسيابية والدقة في تمثيل شكل التوزيع الحقيقي، دون مبالغة في القيم أو اهتزازات حادة على الأطراف، بخلاف ما لوحظ في بعض الدوال الأخرى. هذا يعكس قدرة Weibull على التعامل بمرونة مع البيانات الموجبة وغير المتماثلة، ويؤكد تفوقها في توفير تقدير أكثر دقة واستقراراً للكثافة الاحتمالية مقارنة بباقي دوال كيرنل المستعملة.



شكل (3) مقارنة قيم الخطأ التكاملية التربيعي (ISE) بين دوال كيرنل المتماثلة وغير المتماثلة باستعمال طريقتي Silverman و Cross-Validation في التجربة

من الشكل (3) أنفاً ، يتبين لنا أن الدالة المقترحة Weibull من بين الدوال غير المتماثلة قد سجلت أقل قيمة للخطأ التكاملية التربيعي ISE باستعمال كل من طريقتي Silverman و Cross-Validation ، مقارنة بجميع دوال كيرنل الأخرى، سواء المتماثلة أو غير المتماثلة. هذا يشير إلى تفوق أدائها في تقدير دالة الكثافة الاحتمالية بشكل أكثر دقة واستقرار، ويؤكد ملاءمتها العالية للبيانات الموجبة غير المتماثلة، على خلاف دوال كيرنل المتماثلة التي أظهرت قيم ISE أعلى في كلا الطريقتين.

5 الاستنتاجات والتوصيات

1.5 الاستنتاجات

1. أظهرت النتائج التجريبية أن الدوال غير المتماثلة، وعلى وجه الخصوص دالة Weibull ، قدّمت أداءً أكثر دقة مقارنةً بالدوال الكيرنل المتماثلة في تقدير دالة الكثافة الاحتمالية لبيانات موجبة غير متماثلة، كما هو الحال في بيانات البقاء قيد الدراسة.
2. تم تحقيق أدنى قيمة للخطأ التكاملية التربيعي (ISE) عند استعمال دالة Weibull مع طريقة Silverman ، إذ بلغت القيمة 0.008، ما يعكس تفوقاً ملحوظاً في تمثيل طبيعة البيانات.
3. من خلال اختبار Kolmogorov–Smirnov ومعيار AIC ، تبين أن نموذج التقدير باستعمال Weibull Kernel مع طريقة CV هو الأكثر كفاءة من بين النماذج المدروسة، حيث حقق أعلى قيمة P-value وأقل قيمة AIC.

4. أثبتت طريقة المحاكاة فعاليتها في اختبار أداء دوال التقدير المختلفة تحت سيناريوهات متعددة، وبيّنت بشكل واضح مدى حساسية النتائج لطبيعة الدالة المختارة وحجم العينة

2.5 التوصيات

1. يُوصى باعتماد دوال كيرنل غير متماثلة، لاسيما دالة Weibull ، عند التعامل مع بيانات موجبة وغير متماثلة مثل بيانات البقاء، نظرًا لما توفره من دقة أعلى في التقدير.
2. يُوصى استعمال احجام مختلفة في تقدير الدوال المدروسة .
3. يُفضل استعمال طريقة التحقق المتقاطع (CV) لاختبار عرض الحزمة لما أثبتته من أداء أفضل مقارنة بطريقة Silverman التقليدية.
4. في التطبيقات الواقعية، يُوصى بتطبيق اختبار K-S ومعيار AIC معًا لتقييم ملائمة الانموذج المختار.
5. ينبغي إجراء دراسات مستقبلية لاختبار كفاءة دوال كيرنل غير متماثلة أخرى وتوسيع التجارب لتشمل توزيعات بقاء مختلفة وسيناريوهات متنوعة.
6. يُنصح بتطبيق المنهجية المقترحة على بيانات طبية أخرى، مثل بيانات مدة المكوث في المستشفيات أو أزمّة الشفاء، لتقييم عمومية الانموذج في سياقات تطبيقية مختلفة

6.المصادر

- [1] Bareche, A., & Aïssani, D. (2008). Kernel density in the study of the strong stability of the M/M/1 queueing system. *Operations Research Letters*, 36(5), 535-538.
- [2] Belzile, L. R., Desgagné, A., Genest, C., & Ouimet, F. (2022). Normal approximations for the multivariate inverse Gaussian distribution and asymmetric kernel smoothing on \mathbb{R}^d -dimensional half-spaces. *arXiv preprint arXiv:2209.04757*.
- [3] Bouezmarni, T., & Scaillet, O. (2005). Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data. *Econometric Theory*, 21(2), 390–412.
- [4] Ghitany, M. E., Atieh, B., & Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4), 493-506.
- [5] Hanif, M. (2016). A nonparametric approach to the estimation of jump-diffusion models with asymmetric kernels. *Cogent Mathematics*, 3(1), 1179247.
- [6] Jones, M. C., Marron, J. S., & Sheather, S. J. (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, 91(433), 401–407.
- [7] Maiti, S. S., & Mukherjee, I. (2016). Some estimators of the PDF and CDF of the Lindley Distribution. *arXiv preprint arXiv:1604.06308*.
- [8] Markovich, L. A. (2015). Gamma kernel estimation of the density derivative on the positive semi-axis by dependent data. *arXiv preprint arXiv:1502.02373*.
- [9] Maswadah, M. (2007, May). Kernel inference on the Weibull distribution. In *Proc. Of the Third National Statistical Conference. Lahore, Pakistan (Vol. 14, pp. 77-86)*.
- [10] Samiuddin, M., & El-Sayyad, G. M. (1990). On nonparametric kernel density estimates. *Biometrika*, 77(4), 865-874.
- [11] Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels. *Nonparametric statistics*, 16(1-2), 217-226.
- [12] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.
- [13] Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.
- [14] Yang, F., & Yue, Z. (2014). Kernel density estimation of three-parameter Weibull distribution with neural network and genetic algorithm. *Applied Mathematics and Computation*, 247, 803-814.