AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

تقدير مَعْلَمَات انموذج انحدار kink بوجود مشكلة الأبعاد العالية مع تطبيق عملي

زهراء كريم فهد [1] **أ . د باسم شلبيه مسلم [2]**

جامعة واسط / كلية الإدارة والاقتصاد **الجامعة العراقية / كلية الإدارة والاقتصاد**

المستخلص

يُعدّ موضوع انحدار kink من المواضيع الحديثة، التي رَكّزتْ عليه البحوث في الآونة الأخيرة؛ لأهمّيته التي تبزر من ارتباطه بموضوع الاستمرارية للانحدار. ولكن هناك ندرة منها في العراق، ويتميّز هذا الانحدار بانقسام مَعْلَمَات على مَعْلَمَتين: أحدهما قبل نقطة القطع kink وبعده، ما يجعل عدد المتغيّرات التوضيحية فيه يتضاعف تبعاً لوجود نقطة القطع kink في متغيّر توضيحي واحد أو أكثر. وقد يسبّب هذا العدد المضاعف في أنْ يجعل عدد المتغيّرات التوضيحية في انموذج انحدار kink أكبر من حجم العيّنة، أي ($n < p$) ، أو أنْ يكون عدد المتغيّرات التوضيحية في أصل البيانات أكبر من حجم العيّنة الـ kink ، وفي كلتا الحالتين تظهر مشكلة الأبعاد العالية في الانموذج High Dimensions ، وفي هذه الحالة يسمّى انموذج انحدار kink بوجود مشكلة الأبعاد العالية High Dimensions Kink Regression (HDKR) . في هذا البحث تَمَّ عرض طريقة المربّعات الصغرى الجزائية تقدير انموذج انحدار kink بوجود مشكلة الأبعاد باستعمال طريقة المربّعات الصغرى الجزائية Penalized Least Squares method (PLS) باعتماد دوال جزاء Penalty functions مختلفة منها: دالّة جزاء لاسوLASSO ، ودالّة جزاء سكاد SCAD ، ودالّة Minimax Concave Penalty (MCP) فضلاً عن دالّة جزاء ايلاستك- نت Elastic- net ، التي اقترح استعمالها من قبل الباحثة إذْ تَمَّ تطبيق المربّعات الصغرى الجزائية PLS بدوال الجزاء المذكورة على بيانات حقيقية real data تخصّ القيمة السوقية والمتغيّرات المؤثّرة فيه (وهي بيانات محاسبية) لشركة بغداد للمشروبات الغازية، والتي تتضمّن 48 متغيّراً توضيحياً، وباعتماد مقياس متوسّط مربّعات الخطأ للأنموذج للمفاضلة المقدّرات، تبيّن أنّ طريقة المربّعات الصغرى الجزائية المعتمدة على دالّة الجزاء ايلاستك نت Elastic- netهي الأكفأ في عملية الاختيار والتقدير.

الكلمات المفتاحية : انموذج انحدار KINK ، المربّعات الصغرى الجزائية ، لاسو ، سكاد ، MCP، ايلاستك -نت.

Abstract

Kink regression is one of the event topics that research has focused on recently because of its importance, which emerges from its association with the subject of continuity of regression, but there is a scarcity of them in Iraq, as this regression is characterized by the division of its parameters into two parameters, one of which is before and after the kink cut-off point, making the number of explanatory variables double according to the presence of the kink cut-off point in one or more explanatory variables, and this double number may cause the number of explanatory variables in the kink regression model to be larger of the sample size (n<p) or that the number of explanatory variables in the origin of the data is greater than the sample size of the kink, and in both cases the problem of high dimensions appears in the model, and in this case the kink regression model is called High Dimensions Kink Regression (HDKR). In this research, the method Penalized Least Squares method (PLS) was presented, to estimate parameters of the (HDKR) model by adopting different penalty functions, including the LASSO penalty, the SCAD penalty, the Minimax Concave Penalty (MCP) in addition to the Elastic-net penalty function, which was suggested to be used by the researcher. By applying the methods of the PLS on real data related to the market value and the variables affecting it (accounting data) of the Baghdad Soft Drinks Company, and by adopting the scale of the mean squares error for the model for the comparison, it was found that the method of PLS based on the penalty function Elastic - net is the most efficient in the selection and estimation.

Keyword: kink regression, PLS, LASSO, SCAD, MCP, Elastic - net

1- المقدّمة:

1-1 تمهيد:

قد يواجه الباحث عند دراسته لظاهرة ما بأسلوب تحليل الانحدار حالة من عدمِ تحقّق استمرارية مَعْلَمَات الانموذج؛ بسبب امتلاك بعض المتغيّرات التوضيحية أو كلّها نقطة قطع عندها، تنقسم المَعْلَمَة الواحدة مَعْلَمَتين لمنطقتي قبل القطع وبعده Kink ، ويتنج ما يسمّى انموذج انحدار kink ممّا يجعل عدد المتغيّرات التوضيحية فيه يتضاعف تبعاً لوجود نقطة القطع kink في متغيّر توضيحي واحد أو أكثر. وقد يسبّب هذا العدد المضاعف في أنْ يجعل عدد المتغيّرات التوضيحية في انموذج انحدار kink أكبر من حجم العيّنة، أي ($n < p$) أو أنْ يكون عدد المتغيّرات التوضيحية في أصل البيانات أكبر من حجم العيّنة الـ kink وفي كلتا الحالتين تظهر مشكلة الأبعاد العالية في الانموذج High Dimension . وفي هذه الحالة يسمّى

انموذج انحدار kink بوجود مشكلة الأبعاد العالية High Dimensions Kink Regression (HDKR). وهذ النوع من الانحدار يتّصف بالحداثة على رغم من ارتباطه من بموضوع الانحدار الخطّي غير المستمرّ الشائع. وقد يحدث في الظواهر الاقتصادية والمحاسبية الصحّية والبيولوجية والاجتماعية وغيرها. ويتّصف انموذج انحدار kink بالعشوائية في ترتيب مشاهدات المتغيّر التوضيحي الذي يمتلك نقطة (kink)، وهذا يؤثّر على طرائق التقدير الاعتيادية، إذْ تصبح غير مناسبة لعملية تقدير مَعْلَمَاته وتحديد المتغيّرات التوضيحية المؤثّرة فعلاً على متغيّر الاستجابة؛ بسبب فقدان خصائصها الجيدة، ما يعطي للباحث سبباً وموضوعاً للبحث فيه.

كانت أول محاولة تطرّقت لموضوع انحدار kink عام 2009م للباحث card وآخرون [7] ، إذْ تَمَّ دراسة النماذج اللامَعْلَمَية التي توجد في متغيّراتها نقاط kink ؛ لغرض تحديد التأثيرات الهامشية في تلك النماذج، ولم يظهر مصطلح kink بشكل واضح إلى أنْ قدّم في عام 2010 الباحث Nielsen وآخرون [18] بحثاً ظهر فيه مصطلح Kink بشكل واضحٍ وصريحٍ.

وفي العام (2015) م، قام الباحثون (Card et al.) [10]بتقديم شروطًا كافية لـ Kink Regression Design (RKD) لتحديد التأثيرات السببية القابلة للتفسير في الانموذج العام غير القابل للفصل، أي إنّ الانحدار المستمرّ يكون دالّة معرفة. المتغيّر التوضيحي يمتلك (Kink). وقد تَمَّ تطبيقه لتقدير تأثير معدّلات الإعانة على فترات البطالة في النمسا بشكل تجريبي. وقد وضَّح الباحثون أنّ الانموذج (RKD) هو انموذجٌ قابلٌ للتفسير بسهولة عند المعالجة بطريقة الفترات حول منطقة الـ (Kink).عام (2018)، قدّم الباحثان Tarkhamtham& Yamaka [20]، بحثاً تضمّن تقدير مَعْلَمَات أنموذج انحدار kink الخطّي باستعمال طريقة الانتروبي العظمى العامّة بمقياسي ريني وتساليس، فضلا عن مقياس (Shannon). وبإجراء المقارنة بين التقديرات لمَعْلَمَات الانموذج من خلال المحاكاة تبيّن أفضلية الطريقة بمقياسي ريني وتساليس على الطريقة، بالمقياس الثالث على الرغم من تقارب النتائج للطريقة بالمقاييس الثلاثة. أمّا ما يخصّ بحوث مشكلة الأبعاد العالية، فكانت أول محاولة للبحث من قبل الباحثين Fran &Friedman [10]. في عام 1993 وتطرّقا إلى موضوع انحدار الحرف العام(generalization of ridge regression) والاختيار الجزئي للمتغيّرات التوضيحية من خلال إضافة دالّة جزائية penalty function هي $(\gamma \sum_j |\beta_j|^q)$ إلى مجموع مربّعات الخطأ، وهذه الدالّة مكافئة إلى للشرط $\sum_j |\beta_j|^q \leq t$). وقد سُمّي هذا الإجراء بـ (Bridge)، وتضّمنت طرائق التقدير: طريقة المربّعات الجزائية Least penalized (PLS) Squares، وطريقة المركبات الرئيسة Principal Components Regression (PC) . وقدّم الباحث Tibshirani [22] في عام (1996) بحثاً تضمّن اقتراحاً لطريقة تقدير جديدة تعتمد على دالّة جزائية على غرار أسلوب الباحثين Fran & Friedman، سمّيتْ اختصاراً LASSO (Least Absolute Shrinkage and Selection Operator) إذْ تَمّ إضافة الدالّة الجزائية إلى مجموع مربّعات الخطأ لغرض تصغيرها، وهي تعدّ طريقة تقدير واختيار في آن واحد مع خلال إيجاد التقديرات المتبعثرة (Sparse Estimators)، وهذان البحثان يُعدّان الأهمّ في مناقشة الأبعاد العالية، بعدها في عام 2011 نشر الباحث Huang وآخرون [12] بحثا، والباحث ,Brimacombe في عام 2014 [6] بحوث، وكذلك الباحث Tibshirani,في العام نفسه [23] بحثا، والباحثان Kim & Lee في عام 2017[15] بحثا، والباحث Wang وآخرون في عام 2018[24] بحثا، والباحث Zhang وآخرون في عام 2021[28] بحثا. البحوث جميعها ناقشت عملية اختيار

المتغيّرات المهمّة وتقدير مَعْلَمَات انموذج الانحدار بوجود مشكلة الأبعاد العالية High Dimensions. أمّا البحوث التي ناقشت انموذج انحدار Kink بوجود بيانات ذات الأبعاد العالية فهي قليلة جدا ففي عام 2016. قدّم الباحث Lee وآخرون [16] بحثاً قدّموا فيه مقدّر لاسوLASSO لمَعْلَمَات انموذج الانحدار التي تمتلك نقطة تغيّر change point والمسمّاة (نموذج العتبة) Threshold model. ومشكلة بيانات الأبعاد العالية للحصول على المقدّرات المتبعثرة. وفي عام 2021 قدّم الباحث .Yamaka [27] بحثاً تطرّق فيه بشكل واضح لأنموذج Kink، ومشكلة بيانات الأبعاد العالية High Dimensions، والتقدير بطريقة المربّعات الصغرى الجزائية PLS، باستعمال دوال الجزاء LASSO وسكاد SCAD وMCP وS.Ridge وBridge.

2-1 مشكلة البحث:

تتأثّر عملية تقدير مَعْلَمَات انموذج الانحدار بحدوث أيّ خرق، في أيّ شرط من شروط انموذج الانحدار الخطّي الواجب توفّرها ومنها($n < p$)، ممّا يؤدّي إلى حدوث مشكلة الأبعاد العالية High Dimensional ، وأيضاً خرق شرط استمرارية مَعْلَمَاته continuous ، الذي يحدث عندما توجد لأحد المتغيّرات التوضيحية بعضها أو جميعها نقاط قطع kink points ممّا يؤدّي إلى ظهور انموذج انحدار kink (Kink Regression) (KR)، وفي كلتا الحالتين تصبح مقدّرات طريقة المربّعات الصغرى LS، وطريقة الإمكان الأعظم ML غير كفوءة، وتعطي مقدّرات ذات تباينات كبيرة لا يمكن الاعتماد عليها. مشكلة هذا البحث تتوضّح في الواقع التطبيقي بشكل كبير عندما يجتمع الخرقان معاً، يواجه الباحث بوجودهما صعوبات في عملية التقدير لمَعْلَمَات انموذج انحدار Kink بوجود الأبعاد العالية (HDKR)، للحصول على مقدّرات كفوءة في عملية التقدير واختيار المتغيّرات التوضيحية المؤثّرة.

3-1 هدف البحث:

اعتماداً على ما تقدّم في مشكلة البحث فإنّ هدف البحث هو تقدير مَعْلَمَات انموذج انحدار Kink بوجود مشكلة بيانات الأبعاد العالية High Dimensional (HDKR) ($n < p$) بطريقة المربّعات الصغرى الجزائية penalized Least (PLS) Squares باستعمال دوال جزاء المختلفة، وهي دوال لاسو LASSO وسكاد SCAD وMCP فضلاً عن مقترح مقدّم من قبل الباحثة باستعمال دالّة جزاء Elastic – net ، إذْ إنّها تُعدّ من دوال الجزاء الشائعة الاستعمال مع انموذج الانحدار العادي بوجود البيانات العالية، ولكن ليس مع انموذج انحدار kink تضمن PLS الحصول على التقديرات المتبعثرة Sparse) (Estimators، والتي تضمن عملية اختيار المتغيّرات التوضيحية المؤثّرة فعلاً، التي قد يكون الجزء الموجب منها أو السالب أو كليهما على متغيّر الاستجابة response variable ؛ نتيجة امتلاكها نقطة قطع kink point. وتجدر الإشارة إلى أنّ البحوث الأحدث في هذا المجال، وهي قليلة العدد تطرّقت إلى دوال جزاء مختلفة لم تكن Elastic – net من ضمنها [27] . لذا اقترحت الباحثة استعمال دالّة جزاء أخرى هي Elastic – net لتقدير مَعْلَمَات انموذج انحدار Kink بوجود مشكلة بيانات الأبعاد العالية، والتي تستعمل عندما تشكّل المتغيّرات التوضيحية في انموذج انحدار Kink مجموعات تَضُمّ توضيحية شديدة الارتباط؛ للتعرّف على تأثيرها في عملية التقدير لمَعْلَمَات انموذج انحدار Kink، والغاية تحسين المقدّرات لمَعْلَمَاته بطريقة PLS، بما يعود بالفائدة على الاختيار الصحيح للمتغيّرات التوضيحية المؤثّرة فعلا على متغيّر الاستجابة،

والتي تكون قيمها بعد التقدير غير مساوية للصفر ومقارنة نتائجها مع بقية نتائج دوال الجزاء السابقة الذكر في الجانب التطبيقي على بيانات حقيقية تتعلّق بالواقع العراقي، تخصّ القيمة السوقية لشركة بغداد للمشروبات الغازية، والتي تتضمّن 48 متغيّراً توضيحياً، وهي أكبر من حجم العيّنة بحيث يمثّلها انموذج انحدار Kink بوجود الأبعاد العالية High Dimensional، إذْ تَمَّ تقدير مَعْلَمَاته بطريقة PLS بدوال جزاء LASSO وسكاد SCAD وMCP و Elastic – net المقترحة الاستعمال من قبل الباحثة، وباعتماد مقياس المفاضلة متوسّط مربّعات الخطأ للأنموذج MSE يَتُمّ تحديد طريقة المربّعات الصغرى الجزائية الأفضل بالتقدير والاختيار وفقاً لنوع دالّة الجزاء، التي اعتمدت عليها الطريقة في تقدير مَعْلَمَات انموذج انحدار Kink.

4-1 هيكُلّية البحث :

لتحقيق هدف البحث سَيتُمّ تقسيم البحث على أربعة مباحث: الأول منها يحتوي على مقدّمة للبحث، ومشكلة البحث، وهدف البحث، وبعض الدراسات التاريخية حول الموضوع. أمّا المبحث الثاني فتضمّن الجانب النظري للبحث متضّمناً عرضاً لأنموذج انحدار kink منفرداً، وكذلك بوجود البيانات ذات الأبعاد العالية HDKR فضلاً عن عرض طرائق التقدير المبعثرة Sparse Estimators لمَعْلَمَات انموذج انحدار kink بوجود البيانات ذات الأبعاد العالية باستعمال طريقة PLS باعتماد دوال الجزاء المختلفة مع دالّة جزاء مقترحة للاستعمال Elastic – net. أمّا المبحث الثالث فتضمّن الجانب التطبيقي لتقدير مَعْلَمَات انحدار Kink بوجود البيانات ذات الأبعاد العالية بطريقة PLS مع دوال الجزاء المعتمدة في البحث، والبيانات التي اعتمد عليها الجانب التطبيقي تخصّ القيمة السوقية لشركة بغداد الغازية، وأخيراً المبحث الرابع الذي تضمّن الاستنتاجات والتوصيات.

2- الجانب النظري:

2-1 انموذج انحدار kink بوجود البيانات العالية High dimensions:

يُعرّف انموذج انحدار Kink (kink Regression) كالآتي [14] [20] :

$$Y_i = \beta_1^-(X_{1i} - c_1)^- + \beta_1^+(X_{1i} - c_1)^+ + \cdots + \beta_j^-(X_{ji} - c_j)^- + \beta_j^+(X_{ji} - c_j)^+ + \cdots +$$

$$\beta_k^-(X_{ki} - c_k)^- + \beta_k^+(X_{ki} - c_k)^+ + \beta_0 X_{0i} + u_i \qquad , i = 1,2,..n \qquad (1)$$

إذْ إنّ :

Y_i : يمثّل مشاهدات متغيّر الاستجابة.

$X_{1i}, X_{2i}, \dots X_{ki},$: تُمثّل مشاهدات المتغيّرات التوضيحية، والتي تمتلك نقاط قطع.

$(c_1, \cdots, c_j, \cdots, c_k)$: تُمثّل نقاط (Kink) التي تمتلكها المتغيّرات التوضيحية تبعاً لتسلسلها.

$\beta_1^-, \beta_2^-, \dots \beta_k^-, \beta_1^+, \beta_2^+, \dots, \beta_k^+$: مَعْلَمَات انموذج انحدار قبل القطع وبعده.

$\beta_0 X_{0i}$: الحدّ الثابت ومتغيّره التوضيحي المساوي للواحد، وغالباً ما يَتُمّ حذفه عند عملية التقدير. .

u_i: يمثّل حدّ الخطأ العشوائي i، والذي هو مستقلّ independent عن حدود الأخطاء العشوائية الأخرى، أيْ إنّها (independent and identically distributed) (i.i.d) بوسط حسابي مساوٍ إلى الصفر وتباين مساوٍ σ^2 ويتّضح أنّ الأخطاء العشوائية لا تتوزّع توزيعاً طبيعياً في انموذج انحدار kink [27] .

ويرتبط مفهوم انحدار kink بمفهوم الانحدار غير المستمرّ ((Regression Discontinuity (RD)، الذي تَمَّ عرضه من قبل الباحثين (Thistlethwaite, D. L., & Campbell, D. T.) [11] في عام 1960، إذْ إنّ الأخير يمتلك نقطة عتبة (Threshold) لأحد المتغيّرات التوضيحيّة أو أكثر تُحدث قطعاً في خطّ الانحدار تبعاً لمشاهدات المتغيّر التوضيحي قبل نقطة العتبة وبعدها، ممّا يسبّب ظهور خطّين لأنموذج الانحدار (سفلي وعلوي) بدلاً من خطّ واحد للانحدار؛ لأنّهما منفصلات عند تلك النقطة بحيث يكون خطّ انحدار قبل نقطة العتبة وخطّ انحدار، بعدها وهذا الانفصال بخطّ الانحدار يؤدّي إلى أنْ يكون $E(Y|X)$ دالّة غير مستمرّة ايضاً، بمعنى أنّ تكون المَعْلَمَات(الميول الحدّية) المكوّنة للدالّة غير مستمرّة Discontinuity إذْ تكون لمَعْلَمَات (الميول الحدّية) في الانموذج قيمتان قبل نقط العتبة وبعدها، ونتيجة لذلك تَمَّ استعمال آلية حسابية لمعرفة سلوك خطّ الانحدار من خلال تقدير المسافة بين الخطّين المنفصلين باحتساب متوسّط المشاهدات، التي تكون قبل نقطة العتبة وبعدها، ولكن تَمَّ الاستغناء عن هذه الآلية بظهور انموذج انحدار kink ،الذي يعتمد على خطّ واحد لتحقيق صفة الاستمرارية continuity لـ $E(Y|X)$ ثُمَّ الاستمرارية للمَعْلَمَات انموذج انحدار Kink .ونعيد كتابة الانموذج (1) كالآتي:

$$Y_i = \sum_{j=1}^{k}\left[\beta_j^{-}(X_{ji} - c_j)^{-} + \beta_j^{+}(X_{ji} - c_j)^{+}\right] + \beta_0 X_{0i} + u_i \qquad (2)$$

ووفقًا لصيغة انموذج انحدار kink (2) يمكن كتابته بدلالة المصفوفات كالآتي:

$$Y_i = \beta' X_{(c)} + u_i \qquad i = 1,2,\dots,n \qquad (3)$$

ويمكن تعريف متوسّط مربّعات الخطأ كالآتي:

$$S(\beta, \underline{c}) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \beta' X_{i(\underline{c})}\right)^2 \qquad (4)$$

وللحصول على مقدّرات المربّعات الصغرى للمَعْلَمَات (β, c) ، يَتُمّ من خلال الاشتقاق الجزئي لتقليل الصيغة (4) كالآتي:

$$(\hat{\beta}, \hat{\underline{c}}) = \operatorname{argmin}_{\beta,\underline{c}} S_n(\beta, \underline{c}) \qquad (5)$$

لقد ذكر الباحث (Hansen, B. E.)[11] أنّ الصيغة (4) ، هي صيغة تربيعية بالمَعْلَمَات (β) ، ولكنّها ليست مُحدّبة (convex) بوجود مَعْلَمَات (Kink) (التركيبة تكون مُحدّبة لمجموعة من النقاط، تكون معاملاتها جميعها موجبة (وسالبة)، ومجموعها مساوي للواحد الصحيح).

وتَتُمّ عملية التقدير أولاً لموجّه المَعْلَمَات β من خلال الآتي:

$$(\hat{\beta}, \underline{c}) = \min_{\beta} S(\beta, \underline{c}) = S(\hat{\beta}_{(\underline{C})}, \underline{c}) \qquad (6)$$

إذْ إنّ:

$\hat{\beta}(\underline{c})$:مقدّرات طريقة المربّعات الصغرى لأنموذج انحدار kink لمتغيّر الاستجابة Y_i على المتغيّرات $X_{i(\underline{c})}$ بثبوت قيم نقاط الـ kink $(\underline{c})$ المتغيّرات التوضيحية كُلّها.

أمّا مقدّرات المربّعات الصغرى لنقاط kink فتكون كالآتي:

$$(\hat{\underline{c}}) = \operatorname{argmin}_{\underline{c}} S(\hat{\beta}_{(\underline{C})}, \underline{c}) = \operatorname{argmin}_{\underline{c}} S^*_{(\underline{C})} \qquad (7)$$

إذْ إنّ:

$$S^*_{(\underline{C})} = S(\hat{\beta}_{(\underline{C})}, \underline{c}) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{\beta}_{(\underline{C})} X_{i_{(\underline{C})}}\right)^2 \qquad (8)$$

ويمكن الحصول على $(\hat{\underline{c}})$ المعرّف بالصيغة (7) عددياً باستعمال البحث الشبكي (grid search) لكُل متغيّر توضيحي وبتعويض $(\hat{\underline{c}})$ بالصيغة (6) نحصل على $\hat{\beta}_{(\underline{C})}$ ، الذي باعتماده يمكن الحصول على البواقي (residuals) لانموذج انحدار kink كالآتي:

$$\hat{u}_i = \widehat{\mathbf{Y}}_{\boldsymbol{i}} - \hat{\beta}_{(\hat{\underline{C}})} X_{i_{(\hat{\underline{c}})}} \qquad i = 1 \dots n \qquad (9)$$

والبواقي في الصيغة (9) هي غير خطّية التي باستعمالها يمكن الحصول على تقدير تباين الخطأ لانموذج انحدار kink كالآتي [14] :

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{\beta}_{(\hat{\underline{C}})} X_{i_{(\hat{\underline{c}})}}\right)^2 = S^*_{(\hat{\underline{C}})} \qquad (10)$$

$$\widehat{Y}_i = \widehat{\beta}' X_{i(\hat{r})} + \hat{u} \qquad (11)$$

إنّ بواقي طريقة المربّعات الصغرى لا خطّية (nonlinear) في الانموذج (11). ويمكن تقدير تباين الخطأ إذْ إنّ $\sigma^2 = E\varepsilon_i^2$ كالآتي:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i^2 = S_n^*(\hat{c}) \qquad (12)$$

كُلّ ما تقدّم من خطوات لتقدير مَعْلَمَات انموذج انحدار kink استند إلى افتراض أنْ يكون حجم العيّنة simple size أكبر من عدد المَعْلَمَات parameters # المراد تقديرها، أيْ إنّ $(n > 2k + 1)$، وهو شرط رئيس لإجراء طريقة المربّعات الصغرى (LS)، ومن ثمّ فإنّ المقدّر سيكون وحيد (الحلّ وحيد)، يمكن من خلال ذلك استعمال الطرائق التقليدية لاختيار المتغيّرات التوضيحية المؤثّرة فعلا في انموذج انحدار kink. ومنها على سبيل المثال طريقة الانحدار المتسلسل (stepwise Regression) للحصول على أفضل معادلة تقديرية لإجراء التنبُّؤ المستقبلي للظاهرة تحت الدراسة، وهو يمثّل الهدف الرئيس لدراسة أيّ نوع من نماذج الانحدار، ومنها انموذج انحدار kink. وتجدر الإشارة إلى يوجد طرائق تقدير أخرى تعتمد على أساليب أخرى ومنها أسلوب الانتروبي [2] و [20] .

ويمكن كتابة انموذج انحدار kink بوجود البيانات ذات الأبعاد العالية HDKR ، عندما يكون عدد المتغيّرات التوضيحية أكبر من حجم العيّنة، أيْ إنّ $(p > n)$ في الانموذج (1) كالآتي:

$$Y_i = \beta_1^-(X_{1i} - c_1)^- + \beta_1^+(X_{1i} - c_1)^+ + \cdots + \beta_j^-(X_{ji} - c_j)^- + \beta_j^+(X_{ji} - c_j)^+ + \cdots + \beta_k^-(X_{ki} - c_k)^- + \beta_k^+(X_{ki} - c_k)^+ + \beta_0 X_{0i} + \varepsilon_i \quad , (p \geq n) \tag{13}$$

إنّ الصيغة (13) تُمثّل انموذج انحدار kink بوجود مشكلة البيانات ذات الأبعاد العالية HDKR سواء أكان حدثت المشكلة بسبب أنّ عدد المتغيّرات التوضيحية في انموذج انحدار kink أكبر من حجم العيّنة في الظاهرة المدروسة في الأصل أم نتيجة لتضاعف عددها لامتلاكها نقاط قطع Kink points، ممّا يجعل عددها المتضاعف أكبر من حجم العيّنة، وفي كلتا الحالتين يبرز سوءًا في نتائج التقدير أو استحالة في إيجاد مقدّرات مَعْلَمَات الانموذج بطرائق التقدير المستعملة جميعها، ومنها طريقة المربّعات الصغرى التي يَتُمّ الحصول عليها من خلال الصيغة (4)، إذا لم يوخذ بالحسبان أنّ عدد المتغيّرات التوضيحية أكبر من حجم العيّنة، وهذا الأمر يوجب البحث عن طرائق تقدير جديدة لمَعْلَمَات أنموذج انحدار Kink بوجود البيانات ذات الأبعاد العالية High Dimensions؛ كونه يتعلّق بشيئين مهمّين: الأول هو تقدير مَعْلَمَات الانموذج التي تضاعفت، والشيء الثاني هو اختيار المتغيّرات التوضيحية المهمّة Selection Variables وابقاؤها في انموذج انحدار kink وإهمال المتبقّي، وهذا الأمر يَتُمّ باتّجاهين:

الاتّجاه الأول: هو استعمال خوارزميات لاختيار المتغيّرات التوضيحية المهمّة؛ للتغلّب على مشكلة الأبعاد العالية، ثُمَّ إجراء التقدير للمَعْلَمَات المرتبطة بها بطرائق التقدير المتعارف عليها في انموذج انحدار kink.

الاتّجاه الثاني: وهو الاتّجاه الذي يذهب معه البحث، إذْ يَتُمّ استعمال طرائق تقدير تعمل على اختيار المتغيّرات التوضيحية المهمّة والتقدير في آنٍ واحد، من خلال اختيار المتغيّرات التوضيحية المهمّة في الانموذج بتقدير مَعْلَمَاتها المرتبطة بها، والحصول على قيمها، وإهمال المتغيّرات التوضيحية غير المؤثّرة على متغيّر الاستجابة بوضع أصفار للمَعْلَمَات المرتبطة بها. وتُعدّ طريقة المربّعات الصغرى الجزائية (PLS) (Penalized Least Squares method) من طرائق التقدير

الشائعة الاستعمال لتقدير مَعْلَمَات نماذج الانحدار العادية بوجود البيانات ذات الأبعاد العالية، ولكنّها حديثة الاستعمال لأنموذج الانحدار kink ؛ بسبب خصوصية الأنموذج لفقدانه صفة الاستمرارية لمَعْلَمَاته؛ للحصول على التقديرات أو المقدّرات المتبعثرة (sparse estimators).

وتقوم طريقة (PLS) Penalized Least Squares method على تقليل مجموع مربّعات الخطأ لأنموذج انحدار kink والمعرّف بالصيغة (4) بوجود بعض القيود على مَعْلَمَات الانموذج والتي يتُمّ صياغتها كدالّة جزاء penalty function ، إذ يَتُمّ تطبيق أسلوب مضاعف لاكرانج Lagrange Multiplier على دالّة هدف (objective function) ، التي تتكوّن من حاصل جمع مجموع مربّعات الخطأ لأنموذج انحدار kink ، والمعرّف بالصيغة (4) مع دالّة جزاء (penalty function) و كالآتي [1][24] :

$$Z(\beta, \underline{c}) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \beta' X_{i(\underline{c})}\right)^2 + \sum_{j=1}^{2k+1} pen(\beta_i|\alpha) \qquad (14)$$

ومن الصيغة (14) يلحظ الآتي :

1- $Z(\beta, \underline{c})$ تُمثّل دالّة الخسارة الجزائية (penalized loss function) .
2- يمثّل الحدّ الأول من الصيغة (14) دالّة الخسارة (loss function) ، والحدّ الثاني منها يمثّل دالّة الجزاء (penalty function)، التي وجودها مهمٌ للحصول على المقدّرات المتبعثرة (sparse estimators) .
3- عندما تصبح دالّة الخسارة الجزائية $Z(\beta, \underline{c})$ بدون دالّة الجزاء (penalty function) تكون مساوية للدالّة (4) لأنموذج انحدار kink فقط؛ ولذا نحصل على مقدّرات المربّعات الصغرى بسهولة؛ لأنّ البيانات لا تكون ذات أبعاد عالية.
4- α: تُمثّل مَعْلَمَة الجزاء (penalty parameter)، وتسمّى ايضاً بمَعْلَمَة الضبط(Regularized parameter) .
5- عندما تتّجه دالّة الجزاء (penalty function)إلى ما لانهاية ($\sum_{j=1}^{2k+1} pen(\beta_i|\alpha) \to \infty$) يصبح الانموذج فارغاً من المقدّرات؛ لأنّ جميعها ستصبح أصفاراً.

وإشارة إلى الملاحظة التي أوردها الباحث (Hansen, B. E.)[11] تعدّ الصيغة (14) هي صيغة تربيعية بالمَعْلَمَات (β)، وأنّها غير قابلة للاشتقاق بالنسبة لنقاط القطع (kink) ، ولذا يمكن إيجاد مقدّرات (حلّ) مربّعات الصغرى الجزائية لمَعْلَمَات أنموذج انحدار kink ذي الأبعاد العالية المعرّف في الصيغة (13) بوجود شرط $(n \leq 2k+1)$وفق الصيغة الآتية [25] :

$$Est.(\beta) = \arg\min_{\beta}\left\{\sum_{i=1}^{n}\left(Y_i - \beta' X_{i(\underline{c})}\right)^2 + \sum_{j=1}^{2k+1} pen(\beta_i|\alpha)\right\} \qquad (15)$$

والصيغة (15) تُمثّل مقدّرات المربّعات الصغرى الجزائية (PLS) لمَعْلَمَات انموذج انحدارkink ذي الأبعاد العالية (HDKR) ، وهذه المقدّرات الجزائية لابُدّ من تحقّق الشروط الآتية:

1- أنْ تكون قيم نقاط القطع الـ kink ($\underline{c}$) معلومة، التي تمتلكها المتغيّرات التوضيحية في انموذج انحدار kink ذي الأبعاد العالية (HDKR).

2- إنّ دالّة الجزاء (penalty function) التي على وفقها يتُمّ تقدير مَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية (HDKR) معرّفة، وتجدر الإشارة هنا أنّ هذا الشرط متحقّق دائما ، إذ توجد العديد من أنواع دوال الجزاء، ومعرّفة وبتنوّعها تتنوّع مقدّرات المَعْلَمَات كما سيمرّ بنا لاحقاً .

3- أنْ تكون مَعْلَمَة الجزاء (penalty parameter) التي تتضمّنها دالّة الجزاء (penalty function) معلومة.

ووفقًا تحقق الشروط المذكورة آنفًا، سيَتُمّ تقديم مقدّرات المربّعات الصغرى الجزائية (PLS) وفق دوال الجزاء المحدّدة في البحث للحصول على المقدّرات المتبعثرة(Sparse Estimators) لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية.

2-2 المقدّرات المتبعثرة (Sparse Estimators)

يمكن الحصول على المقدّرات المتبعثرة لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية عند تحقّق الشروط للصيغة (14)، ولذا سيتُمّ تعريف بعض دوال الجزاء (penalty functions) التي اعتمدها البحث، ومن ثَمّ تُعرف المقدّرات، التي تُسمّى بأسماء دوال الجزاء؛ للتميز بينها وكالآتي:

2-2-1 مقدّرات لاسو (LASSO Estimators)

يتُمُّ الحصول مقدّرات لاسو (LASSO estimator) لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية عند تعويض دالّة جزاء لاسو (LASSO penalty) المقترحة من قبل الباحث (Tibshirani) في الصيغة (15) مع تحقّق بقية الشروط ، ومختصر(LASSO) يعود على (Least Absolute Shrikage and selection Operator). وتعرفُ دالّة جزاء لاسو كالآتي [23][26][1][27]:

$$\mathrm{pen}^{\mathrm{LAS}}(\beta_j|\alpha) = \alpha|\beta_j| \quad , j = 1, \dots, 2k+1 \qquad (16)$$

وتقوم دالّة جزاء لاسو (LASSO penalty) بتقليص بعض مَعْلَمَات انموذج انحدارkink ذي الأبعاد العالية إلى الصفر، ثُمَّ يحذفها من الأنموذج كُلّه، وسيكون مقدّر لاسو كالآتي:

$$\mathrm{Est}(\beta_{\mathrm{LAS}}) = \arg\min_{\beta}\left\{\sum_{i=1}^{n}\left(Y_i - \beta' X_{i(\underline{c})}\right)^2 + \alpha\sum_{j=1}^{2k+1}|\beta_j|\right\} \qquad (17)$$

ولتحقيق الشرط الأول من شروط الصيغة (14) اقترح الباحث Yamaka[27] أسلوبا للحصول على مقدّرات المربّعات الصغرى الجزائية (PLS) لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية (HDKR)، وحلّ المعادلة المعرفة بالصيغة (14) من خلال اتّباع أسلوب الدمج بين طريقة التركيز والبحث الشبكي (Concentration method and grid search) لذا تمّ وضع قيم أولية لنقاط القطع kink ، بحيث يكون لدينا $(\tilde{c} = [\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_k]$ ومن ثَمَّ يمكن إعادة كتابة الصيغة (14) بعد تعويض القيم الأولية لنقاط kink كخطوة أولى كالآتي:

$$\tilde{\beta}_{LAS} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} \left(Y_i - \beta' X_{i(\underline{\tilde{c}})}\right)^2 + \alpha \sum_{j=1}^{2k+1} |\beta_j| \right\} \quad (18)$$

ويلاحظ من الصيغة (18) أن الجزاء يكون على المَعْلَمَات β فقط، ولا يمتدُّ إلى نقاط القطع $\underline{c}$ ؛ لأنّ المَعْلَمَات عند تقليص لا يصبح لها أهمّية نتيجة حذفها من الأنموذج. ولكي يتُمّ تقدير نقاط القطع $\underline{c}$ نستعمل الصيغة الآتية:

$$\underline{\hat{c}} = \arg\min_{c} \left\{ \sum_{i=1}^{n} \left(Y_i - \tilde{\beta}'_{LAS} X_{i(\underline{\tilde{c}})}\right)^2 + \alpha \sum_{j=1}^{2k+1} \left|\tilde{\beta}_{LAS_{(j)}}\right| \right\} \quad (19)$$

$$\hat{\beta}_{LAS} = \arg\min_{\tilde{\beta}_{LAS}} \left\{ \sum_{i=1}^{n} \left(Y_i - \tilde{\beta}'_{LAS} X_{i(\underline{\hat{c}})}\right)^2 + \alpha \sum_{j=1}^{2k+1} \left|\tilde{\beta}_{LAS_{(j)}}\right| \right\} \quad (20)$$

والصيغة (20) تُمثّل مقدّرات طريقة المربّعات الصغرى الجزائية لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية. ولتطبيق هذه الصيغة يستوجب تحديد قيمة مَعْلَمَة الجزاء α ، وهو ما تضمّنه الشرط الثالث من الشروط [27] [1]:

$$\hat{\beta}_{LAS(j)} = \left(|\hat{\beta}_j| - \alpha\right)_+ \text{sign}(\hat{\beta}_j) \quad (21)$$

، وتجدر الإشارة أنّه عندما تكون $(\alpha = 0)$ فإنّ مقدّر لاسو سيعود على مقدّر المربّعات الصغرى (LS) .

2-2-2 مقدّرات اسكاد (SCAD Estimators)

يتمُّ الحصول مقدّرات اسكاد (SCAD estimators) لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية عند تعويض دالّة جزاء اسكاد (SCAD penalty) المقترحة من قبل الباحثين (Fan & Li) [3][1] في الصيغة (14) مع تحقّق بقية الشروط، ومختصر(SCAD) يعود على (Smoothly Clipped Absolute Deviation) وتعرّف دالّة جزاء اسكاد كالآتي :

$$\text{pen}^{\text{SCAD}}(\beta_j|\alpha, a) = \begin{cases} \alpha|\beta_j| & if\ |\beta_j| \le \alpha \\ -\left(\dfrac{(|\beta_j|^2 - 2a\alpha|\beta_j| + \lambda^2)}{2(a-1)}\right) & if\ \alpha < |\beta_j| \le \alpha a \\ \dfrac{(a+1)\alpha^2}{2} & if\ |\beta_j| > \alpha a \end{cases} \quad (22)$$

إذْ إنّ :

a : مَعْلَمَة إضافية في دالّة الجزاء يتُمّ تقديرها بمعيار العبور الشرعي (cross-validation) ولكن سيتُمّ استعمال (a=3.7) كقيمة مثلى وفق مقترح (Fan & Li) [3][1] .

ولتجنّب التكرار بخطوات تحقّق الشرط الأول، سيتُمُّ كتابة صيغة مقدّرات اسكاد (SCAD) مباشرة كالآتي :

$$\hat{\beta}_{SCAD} = \arg\min_{\tilde{\beta}_{SCAD}} \left\{ \sum_{i=1}^{n} \left(Y_i - \tilde{\beta}'_{SCAD} X_{i(\underline{\hat{c}})} \right)^2 + \alpha \sum_{j=1}^{2k+1} pen^{SCAD}(\beta_j | \alpha, a) \right\} \qquad (23)$$

الصيغة (23) تُمثّل مقدّرات اسكاد (SCAD estimators) لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية، والحلّ للصيغة، والحصول على المقدّرات يكون كالآتي [1]:

$$\hat{\beta}_{SCAD(j)} = \begin{cases} (|\hat{\beta}_j| - \alpha)_+ \ sign(\hat{\beta}_j) & \text{if } |\beta_j| \le \alpha \\ [(a-1)\hat{\beta}_j - sign(\hat{\beta}_j) a\alpha]/(a-2) & \text{if2 } \alpha < |\hat{\beta}_j| \le \alpha a \\ \hat{\beta}_j & \text{if } |\hat{\beta}_j| > \alpha a \end{cases} \qquad (24)$$

والتي نحتاج إلى تطبيقها توفّر قيمة a ، والتي تَمَّ تثبيتها بقيمة 3.7 ، وقيمة مَعْلَمَة الجزاء α . إنّ دالّة الجزاء سكاد (SCAD) تحقّق صفات دالّة الجزاء الجيدة، وهي: عدم التحيز،(unbiasedness)، والتبعثر(sparsity)، والاستمرارية (continuity).

3-2-2 مقدّرات MCP (Minimax Concave Penalty)

وهي المقدّرات التي تعتمد على دالّة الجزاء MCP ، غير المُحدّبة، والتي اقترحت من قبل الباحثين Fan and Li [17]. وقد ذكر الباحثان أنّها من دوال الجزاء الجيدة؛ لأنّ مقدّراتها تتّصِفُ بخصائص عدم التحيّز Unbiasedness ، والتبعثر Sparsity ، والاستمرارية Continuity فضلاً عن تمتّعها بخصائص الاوراكل Oracle Properties ، والتي تتضمّن الاتّساق والمتعلّقة بخاصّية التبعثر وامتلاك المقدّر توزيعاً طبيعياً محاذياً لتقدير مَعْلَمَات أنموذج الانحدار الخطّي بوجود بيانات الأبعاد العالية. ويعرّف مقدّر MCP كالآتي[3] [22] :

$$\hat{\beta}_{MCP} = \arg\min_{\beta_{MCP}} \left\{ \frac{1}{2n} \sum_{i=1}^{n} \left(Y_i - \tilde{\beta}'_{En} X_{i(\underline{\hat{c}})} \right)^2 + \sum_{j=1}^{2k+1} pen^{MCP}(\beta_j | a) \right\} \qquad (25)$$

إذْ إنّ :

$$pen^{MCP}(\beta_j | a) = \begin{cases} \alpha\beta - \frac{\beta^2}{2a} & , \ \beta \le a\alpha \\ \frac{1}{2} a \alpha^2 & , \ \beta > a\alpha \end{cases} \quad , \ a > 1 \qquad (26)$$

4-2-2 مقدّرات ايلاستك –نت (Elastic- net Estimators) (المقترحة الاستعمال مع PLS).

إنّ استعمال طريقة PLS بدوال جزاء penalty functions مختلفة يؤدّي إلى الحصول على مقدّرات مختلفة، ولكُلّ منها خصائصها. لذا تقترح الباحثة استعمال دالّة الجزاء Elastic- net penalty) التي تستعمل عندما تُشكّلُ المتغيّرات التوضيحية مجموعات تحتوي على متغيّرات مستقلّة شديدة الارتباط، الناتج من البيانات العالية الأبعاد أو بوجود التعدّد

الخطّي. وغاية الاستعمال هو معرفة مدى التأثير الحاصل من المتغيّرات التوضيحية، التي تكون بشكل مجموعة في عملية التقدير لمَعْلَمَات انموذج انحدار Kink، ومن ثَمّ اختيار المتغيّرات المؤثّرة التي تكون تقدير مَعْلَمَاتها غير مساوية للصفر.

يتُمُّ الحصول مقدّرات ايلاستك –نت (Elastic- net estimators) لمَعْلَمَات انموذج انحدار kink ذي الأبعاد العالية عند تعويض دالّة جزاء (Elastic- net penalty) المقترحة من قبل الباحثين (Zou & Hastie) [1] في الصيغة (14) مع تحقّق بقية الشروط ، وهي خليطٌ من داليتي جزاء: الأولى منهما (L_2 quadratic penalty of ridge regression) ، والثانية هي (L_1 penalty of Lasso) باستعمال تركيبة مُحدّبة (convex combination) . ويمكن تعريف الدالّة كالآتي [1]:

$$\text{pen}^{\text{Elastic}-\text{net}}(\beta_j|\alpha) = \alpha_1 \sum_{j=1}^{2k+1}|\beta_j| + \alpha_2 \sum_{j=1}^{2k+1} {\beta_j}^2 \quad , j = 1, \dots, 2k+1 \qquad (27)$$

ولتلافي تكرار الخطوات في تحقّق الشرط الأول، سيتُمُّ كتابة صيغة مقدّرات ايلاستك-نت (Elastic- net estimators) مباشرة كالآتي :

$$\hat{\beta}_{En} = \arg\min_{\tilde{\beta}_{En}} \left\{ \sum_{i=1}^{n} \left(Y_i - \tilde{\beta}'_{En} X_{i(\hat{c})} \right)^2 + \alpha \sum_{j=1}^{2k+1} \text{pen}^{\text{Elastic}-\text{net}}(\beta_j|\alpha) \right\} \qquad (28)$$

ويمكن عرض الحلّ للصيغة (28) كالآتي:

$$\hat{\beta}_{En(j)} = \left(|\hat{\beta}_j| - 0.5\alpha\right)_+ \text{sign}(\hat{\beta}_j) \qquad (29)$$

إنّ المقدّرات جميعها، التي تَمَّ ذكرها تعتمد على قيمة مَعْلَمَة الجزاء penalty parameter ، التي يستوجب تحديدها، وهذا ما ينطبق مع الشرط الثالث.

وتجدر الإشارة إلى أنّ دوال الجزاء التي تمَّ ذكرها للتعامل مع البيانات العالية الأبعاد، قد تَمّ توفير حزمها البرمجية ضمن برنامج R ، وقد تمَّ توظيفها مع حزم برمجية للتعامل مع انحدار Kink والمقدّمة من قبل الباحث Yamaka [2] لتقدير مَعْلَمَات انموذج انحدار Kink بوجود البيانات العالية High Dimensions ، ومنها الحزمة لإجراء اختبار والد (Wald Test).

5-2-2 اختبار والـــد (Wald Test)

من الاختبارات التي سيتُمُّ توظيفها لاختبار وجود Kink في المتغيّرات التوضيحيّة في أنموذج الانحدار بوجود الأبعاد العالية، هو اختبار والـد (Wald Test)، والذي أُقترِحَ من قبل الباحث ابراهام والد (Abraham wald) في عام (1943) ، وفرضيته كالآتي[2]:

H_0 : المتغيّر التوضيحي لا يمتلك (Kink).

H_1 : المتغيّر التوضيحي يمتلك (Kink).

أمّا صيغة اختبار والد wald test فهي كالآتي:

$$w = \frac{(\hat{\beta}-\beta_0)^2}{var(\hat{\beta})} \sim \chi_1^2 \qquad (30)$$

وقد ذكر الباحث (Yamaka) أنّ هناك صعوبة في التعامل مع الاختبار نظرياً، لذا وضع الباحث حزمة بلغة (R) باسم ((Kinktest)، والتي تُعدّ هي الوحيدة في هذا المجال [2].

6-2-2 اختبار العلاقة بين متغيّر الاستجابة **والمتغيّر التوضيحي من كونها (خطّية ام غير خطّية) في HDRM .**

بشكل عام تكون العلاقة بين متغيّر الاستجابة والمتغيّرات التوضيحية في انموذج انحدار Kink هي غير خطّية Non liner relation ، ولكن في ظلّ وجود البيانات ذات الأبعاد العالية قد يكون لبعضها علاقة خطّية، ويتُمّ تحديدها بالاعتماد على قيم المقدّرات التي يتُمّ الحصول عليها بالمربّعات الصغرى الجزائية PLS وفقاً لدوال الجزاء، فالمتغيّرات التي مَعْلَمَاتها المرتبطة بها أصفار علاقتها خطّية، والتي لها قيم غير الأصفار علاقتها غير خطّية [26] .

3- الجانب التطبيقي

في هذا المبحث سيتُمّ تحليل بيانات حقيقية real data (بيانات محاسبية) عن القيمة السوقية لشركة بغداد للمشروبات الغازية، التي تُعدُّ من الشركات التجارية المدرجة في سوق الأوراق المالية في العراق، والذي تَمّ الحصول عليها من منشوراته للمدة الزمنية المعتمدة في البحث، وكما وردت في المصدر [4] ، وتُمثّل هذه البيانات متغيّرات توضيحية عددها 48 متغيّرا تؤثّر على القيمة السوقية للشركة، والتي تُمثّل متغيّر الاستجابة، وتَمَّ استعمال برامج بلغة R للحصول على نتائج التحليل. والمتغيّرات التوضيحية هي، إذْ تمَّ الحصول على مشاهدات بحجم 32 عن القيمة السوقية الموسمية للشركة، كمتغيّر استجابة والمتغيّرات التوضيحية المؤثّرة عليها بقدر 48 متغيّر للمدّة الزمنية من 2012-2019 وهي: (نسبة السيولة (X_1)، حجم الشركة (X_2)، معدّل النموّ في المبيعات (X_3) ، العائد على الأصول 1 (X_4)، الرافعة المالية (X_5)، توزيع الأرباح (X_6)، التدفّق النقدي التشغيلية (X_7)، الموجودات المتداولة (X_8)، المطلوبات (X_9)، الموجودات (X_{10})، الملكية المؤسّسية (X_{11})، كبار المساهمين (X_{12})، عمر الشركة (X_{13})، نموّ الموجودات (X_{14}) ، إصدار أسهم جديدة (X_{15})، حجم الإنفاق الاستثماري (X_{16}) ، معدّل دوران المدنيين (X_{17})، نسبة النقدية (X_{18})، نسبة التداول السريعة (X_{19})، نسبة التداول (X_{20})، راس المال العامل (X_{21})، راس المال (X_{22})، صافي الربح (X_{23}) ، المبيعات (X_{24}) ، المطلوبات المتداولة (X_{25}) ، معدّل دوران الأصول (X_{26}) ، معدّل دوران الأصول الثابتة (X_{27}) ، طول مدّة الدفع (X_{28}) ، معدّل دوران الدائنين (X_{29}) طول دورة النشاط (X_{30}) ، طول مدّة التخزين (X_{31}) ، معدّل دوران المخزون (X_{32})، طول مدّة التحصيل (X_{33})، معدّل تغطية الفوائد (X_{34})، نسبة الخصوم إلى حقوق الملكية (X_{35})، نسبة إجمالي الأصول على حقوق الملكية (X_{36})، نسبة إجمالي الخصوم إلى إجمالي الأصول (X_{37})، نسبة المديونية (X_{38})، العائد على المبيعات (X_{39})، إجمالي الربح إلى المبيعات (X_{40})، العائد على المال المستثمر (X_{41})، العائد على الأصول 2 (X_{42}) ، العائد على حقوق الملكية (X_{43})، القيمة الدفترية

للسهم (X_{44}) ، نسبة التوزيع النقدية (X_{45})، العائد النقدي للسهم (X_{46})، مكرّر الربحية (X_{47})، ربح السهم (X_{48}). ووفقا لذلك يكون أنموذج الانحدار لعيّنة الدراسة كالآتي :

$$Y_i = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{48} X_{48} + \varepsilon_i \qquad , n=32 \ , p=48 \qquad (31)$$

إذْ إنّ : (ε) يمثّل حدّ الخطأ العشوائي الذي يتوزّعُ توزيعا طبيعياً.

والأنموذج (31) وفقاً للبيانات المتوفّرة عن القيمة السوقية لشركة بغداد للمشروبات الغازية هو أنموذج انحدار بوجود مشكلة الأبعاد العالية high dimensions. ولاختبار وجود kink في المتغيّرات التوضيحية من عدمه يتُمُّ استعمال اختبار والد، وبقدر 48 مرّة. وقد أوضحتُ النتائج امتلاك المتغيّرات التوضيحية جميعها نقاط قطع kink ، التي وفقها تَمّ مضاعفة مَعْلَمَات الأنموذج (31) لينتج أنموذج انحدار kink ، ويكتب كالآتي :

$$\mathbf{Y_i = \beta_1^-(x_{1i} - c_1)^- + \beta_1^+(x_{1i} - c_1)^+ + \cdots + \beta_{48}^-(x_{48i} - c_{48})^- + \beta_{48}^+(x_{48i} - c_{48})^+ + \beta_0 + \varepsilon_i} \quad (32)$$

ووفقاً لأنموذج انحدار kink المعرّف بالصيغة (32)، تَمَّ تطبيق طرائق التقدير الجزائية المعتمدة في البحث، ويتُمّ تقدير المَعْلَمَات المرتبطة بالمتغيّرات المؤثّرة بمتغيّر الاستجابة، أمّا غير المهمّة فتُهمل بوضع أصفارٍ لمَعْلَمَاتها، وبذلك يتُمُّ اختزال عددها، وهذه هي الفكرة الأساس لاستعمال مقدّرات المربّعات الصغرى الجزائية، ولأيٍّ عددٍ من المتغيّرات التوضيحية، وقدْ تَمّ الحصول على النتائج الآتية:

الجدول (1)

يوضّحُ اختيار المتغيّرات التوضيحية المؤثّرة بالقيمة السوقية لشركة بغداد للمشروبات الغازية ومقدّرات المتبعثرة لمَعْلَمَاتها وفق أنموذج انحدار Kink ، وبوجود الأبعاد العالية بحسب طرائق التقدير المعتمدة في البحث

رمز المتغيّر	**B.LASSO**	**B.SCAD**	**B.MCP**	**B. Elastic-net**
$\mathbf{X_1^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_1^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_2^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_2^+}$	0.09923497	0.00000000	0.00000000	0.00000000

$\mathbf{X_3^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_3^+}$	0.01849150	0.00000000	0.00000000	0.00000000
$\mathbf{X_4^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_4^+}$	0.00944329	0.00000000	0.00000000	0.00000000
$\mathbf{X_5^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_5^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_6^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_6^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_7^-}$	0.18921829	0.00000000	0.00000000	0.28582608
$\mathbf{X_7^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_8^-}$	0.47196982	0.00000000	0.00000000	0.00000000
$\mathbf{X_8^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_9^-}$	-0.61365068	0.00000000	0.00000000	-0.25733131
$\mathbf{X_9^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{10}^-}$	-0.30060551	0.00000000	0.00000000	-0.03297005
$\mathbf{X_{10}^+}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{11}^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{11}^+}$	0.26642667	0.00000000	0.00000000	0.00000000
$\mathbf{X_{12}^-}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{12}^+}$	-0.03449220	0.00000000	0.00000000	-0.07414811

$\mathbf{X_{13}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{13}^{+}}$	-0.32043850	0.00000000	0.00000000	-0.20312748
$\mathbf{X_{14}^{-}}$	-0.09020662	0.00000000	0.00000000	-0.05182135
$\mathbf{X_{14}^{+}}$	0.03004589	0.00000000	0.00000000	0.00000000
$\mathbf{X_{15}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{15}^{+}}$	0.20877143	0.00000000	0.00000000	0.14635680
$\mathbf{X_{16}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{16}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{17}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{17}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{18}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{18}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{19}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{19}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{20}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{20}^{+}}$	-0.04491346	0.00000000	0.00000000	0.00000000
$\mathbf{X_{21}^{-}}$	0.00412599	0.00000000	0.00000000	0.00000000
$\mathbf{X_{21}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{22}^{-}}$	-0.01769235	0.00000000	0.00000000	0.00000000
$\mathbf{X_{22}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000

$\mathbf{X_{23}^{-}}$	-0.74795438	-1.26682334	-1.33222921	-0.73404106
$\mathbf{X_{23}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{24}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{24}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{25}^{-}}$	-0.26204421	0.00000000	0.00000000	-0.10706052
$\mathbf{X_{25}^{+}}$	0.12787504	0.00000000	0.00000000	0.00000000
$\mathbf{X_{26}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{26}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{27}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{27}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{28}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{28}^{+}}$	0.02561247	0.00000000	0.00000000	0.00000000
$\mathbf{X_{29}^{-}}$	-0.02444059	0.00000000	0.00000000	0.00000000
$\mathbf{X_{29}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{30}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{30}^{+}}$	-0.07458572	0.00000000	0.00000000	0.00000000
$\mathbf{X_{31}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{31}^{+}}$	-0.04824234	0.00000000	0.00000000	0.00000000
$\mathbf{X_{32}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{32}^{+}}$	-0.00905124	0.00000000	0.00000000	0.00000000

$\mathbf{X_{33}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{33}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{34}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{34}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{35}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{35}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{36}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{36}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{37}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{37}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{38}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{38}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{39}^{-}}$	-0.02222485	0.00000000	0.00000000	-0.03242193
$\mathbf{X_{39}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{40}^{-}}$	0.00000000	0.00000000	0.00000000	-0.04048649
$\mathbf{X_{40}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{41}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{41}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{42}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{42}^{+}}$	-0.00398878	0.00000000	0.00000000	0.00000000

$\mathbf{X_{43}^{-}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{43}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{44}^{-}}$	0.00000000	0.00000000	0.00000000	-0.04847817
$\mathbf{X_{44}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{45}^{-}}$	0.00376925	0.00000000	0.00000000	0.00000000
$\mathbf{X_{45}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{46}^{-}}$	0.00000006	0.00000000	0.00000000	0.00000000
$\mathbf{X_{46}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{47}^{-}}$	-0.00116432	0.00000000	0.00000000	-0.04924142
$\mathbf{X_{47}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
$\mathbf{X_{48}^{-}}$	-0.00000057	0.00000000	0.00000000	-0.04197665
$\mathbf{X_{48}^{+}}$	0.00000000	0.00000000	0.00000000	0.00000000
عدد المتغيّرات المختارة	30	1	1	14

ويلاحظ من الجدول (1) وجود 96 متغيّراً ، وكُلّ متغيّرين متتابعين يخصّان متغيّرًا واحدا من متغيّرات البيانات المحاسبية للشركة؛ بسبب وجود القطع kink. ويوضّحُ أنّ طريقة LASSO قد اختارت (اختزلت) 28 متغيّراً مهمّاً ومؤثّراً على القيمة السوقية لشركة بغداد للمشروبات الغازية من مجموع المتغيّرات التوضيحية في الأنموذج Kink، وقدّرتْ المَعْلَمَات المرتبطة بها، واستبعدت (أُهملَتْ) 68 متغيّراً بجعل مقدّراتها ومَعْلَمَاتها المرتبطة مساوية إلى الصفر، في حين أنّ طريقة Elastic- net قد اختارت (اختزلت) 14 متغيّراً مهمّاً ومؤثّراً على القيمة السوقية لشركة بغداد للمشروبات الغازية من مجموع المتغيّرات التوضيحية في الأنموذج Kink. وقدّرتْ المَعْلَمَات المرتبطة بها، واستبعدت (أُهملتْ) 82 متغيّراً بجعل

مقدّراتها مَعْلَمَاتها المرتبطة مساوية إلى الصفر. أمّا الطريقتان MCP و SCAD فقد تَمّ اختيار متغيّرٍ واحدٍ فقط ولكليهما وهو صافي الربح (X_{23}) قبل القطع، والذي اختارته أصلاً الطريقتان الأخريتان. ووفقًا لما تقدّمَ نحدّد المتغيّرات المهمّة وفقاً للطريقتين LASSO و Elastic- net وقبل القطع Kink ، وبعده؛ لأنّهما الأفضل في عملية الاختزال لمتغيّرات أنموذج HDKR و كالآتي:

الجدول (2)

يبيّنُ المتغيّرات التوضيحية المؤثّرة فعلاً على القيمة السوقية لشركة بغداد للمشروبات الغازية التي تمَّ اختيارها من الطريقتين LASSO و Elastic- net قبل القطع Kink وبعده.

المتغيّر	اسم المتغيّر	قبل القطع kink		بعد القطع kink	
		LASSO	Elastic- net	LASSO	Elastic- net
X_2	حجم الشركة	X	X	✓	X
X_3	معدّل النموّ في المبيعات	X	X	✓	X
X_4	العائد على الأصول 1	X	X	✓	X
X_7	التدفّق النقدي التشغيلية	✓	✓	X	X
X_8	الموجودات المتداولة	✓	X	X	X
X_9	المطلوبات	✓	✓	X	X
X_{10}	الموجودات	✓	✓	X	X
X_{11}	الملكية المؤسّسية	X	X	✓	X
X_{12}	كبار المساهمين	X	X	✓	✓
X_{13}	عمر الشركة	X	X	✓	✓

X14	نمو الموجودات	✓	✓	✓	X
X15	إصدار أسهم جديدة	X	X	✓	✓
X_{20}	نسبة التداول	X	X	✓	X
X_{21}	رأس المال العامل	✓	X	X	X
X_{22}	رأس المال	✓	X	X	X
X_{23}	صافي الربح	✓	✓	X	X
X_{25}	المطلوبات المتداولة	✓	✓	✓	X
X_{28}	طول مدّة الدفع	X	X	✓	X
X_{29}	معدّل دوران الدائنين	✓	X	✓	X
X30	دورة النشاط	X	X	✓	X
X31	طول مدّة التخزين	X	X	✓	X
X32	معدّل دوران المخزون	X	X	✓	X

X 39	العائد على المبيعات	✓	✓	X	X
X 40	إجمالي الربح إلى المبيعات	X	X	X	✓
X 42	العائد على الأصول 2	X	X	✓	X
X 44	القيمة الدفترية للسهم	X	✓	X	X
X 45	نسبة التوزيع النقدية	✓	X	X	X
X 46	العائد النقدي للسهم	✓	X	X	X
X 47	مكرّر الربحية	✓	✓	X	X
X 48	ربح السهم	✓	✓	X	X

ويلاحظ من الجدول (2) أنّ وجود القطع kink في بيانات المتغيّرات التوضيحية يؤثّر على عملية الاختيار لتلك المتغيّرات، هذا من ناحية، وإنّ الطرائق تتعامل مع المتغيّرات التوضيحية قبل القطع وبعده، كأنّها متغيّرات مستقلّة عن بعضها؛ ولذا عند عملية الاختيار تثبّتُ الطرائق متغيّرات المؤثّرة فعلاً على القيمة السوقية. وقد أظهرتْ عملية الاختيار selection variables إنّ الطريقتين توافقتا في عملية الاختيار للمتغيّرات معا، وافتراقا في بعض المتغيّرات الأخرى كما موضّح في الجدول (2) لذا لتحديد الطريقة المثلى في عملية الاختيار فضلاً عن تقدير المَعْلَمَات المرتبطة بها، نعتمد مقياس متوسّط مربّعات الخطأ لنموذج DHKR بالاعتماد على مقدّرات الجدول (1) للطرائق جميعها وكانت النتائج كالآتي :

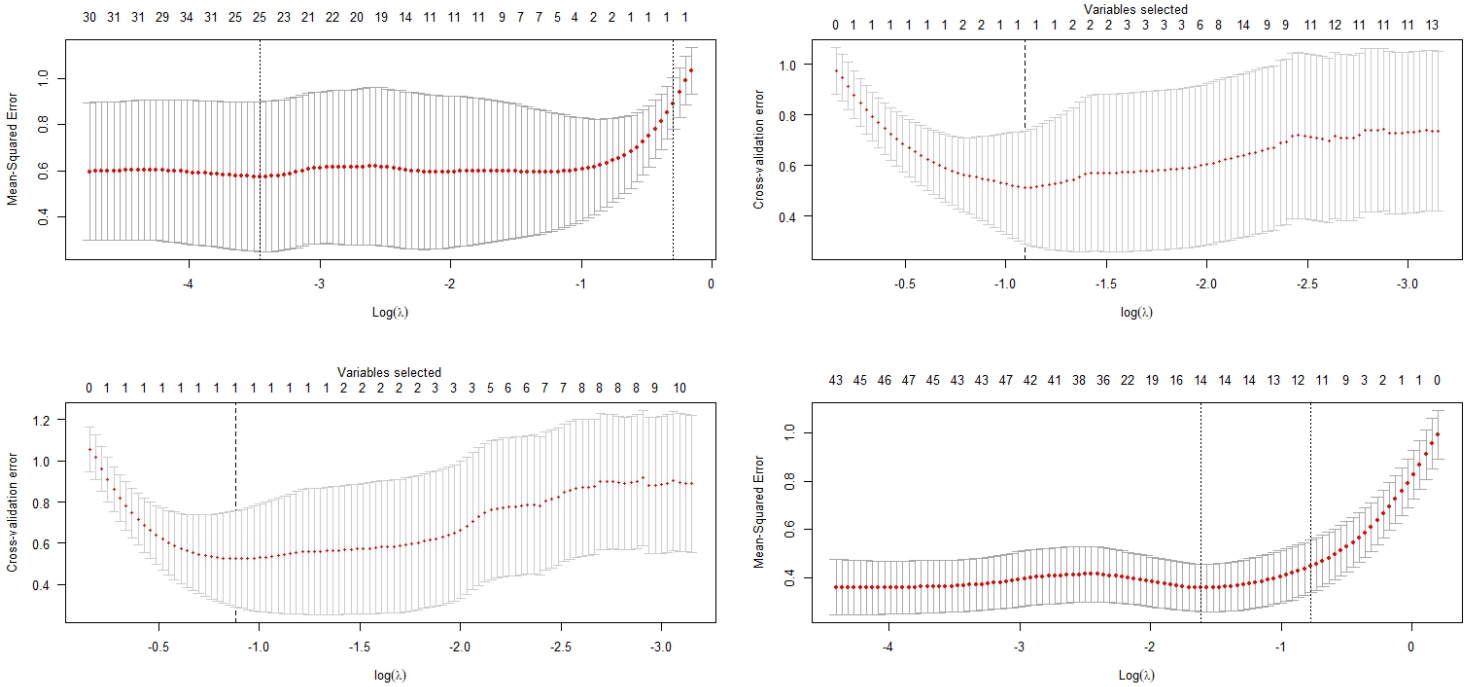
الجدول (3)

يبيّنُ قيم متوسّط مربّعات الخطأ لأنموذج انحدار Kink القيمة السوقية لشركة بغداد للمشروبات الغازية

بحسب الطرائق LASSO وSCAD و MCP و Elastic- net

الطريقة	متوسّط مربّعات الخطأ mse لأنموذج انحدار kink
LASSO	0.426532
SCAD	0.590091
MCP	0.608553
Elastic- net	0.272311

من الجدول (3) يلاحظ أنّ طريقة Elastic- net هي الطريقة الأمثل لعملية الاختيار والتقدير لأنموذج انحدار Kink القيمة السوقية لشركة بغداد للمشروبات بغداد إذْ إنّها تمتلك أقلّ mse من بقية الطرائق، وتأتي بعدها طريقة LASSO وSCAD و MPC على التوالي، وهذا ما تؤكّده الأشكال البيانية للطرائق جميعها وكالآتي :



الشكل (1) يوضّحُ الرسوم البيانية mse بحسب طرائق التقدير LASOO ، SCAD و MCP وElastic- net

ويلاحظ من الشكل (1) أنّ الرسم الرابع في الأسفل يمثّل mse طريقة Elastic- net، والذي يظهر انخفاضاً واضحا عن بقية الطرائق الأخرى، ما يدُلّ على أفضلية الطريقة في الاختيار والتقدير ، ووفقا لما تقدم فان المعادلة التقديرية للقيمة السوقية لشركة بغداد للمشروبات الغازية ستعتمد على المتغيّرات والمقدّرات لطريقة Elastic- net المقترحة الاستعمال؛ كونها أعطت أقلّ متوسّط مربّعات خطأً لأنموذج HDKR .

4- الاستنتاجات والتوصيات :

4-1 الاستنتاجات :

1- بالاعتماد على نتائج الجانب التطبيقي نستنتج أنّ لوجود نقاط القطع Kink أثرها الواضح في عملية تحديد المتغيّرات المؤثّرة فعلاً (البيانات المحاسبية)، في أنموذج HDKR على متغيّر الاستجابة، الذي يمثّل القيمة السوقية لشركة بغداد للمشروبات الغازية ، فقد تَمّ اختيار متغيّرات بقيم قبل القطع kink أو بعده، أو قد يكون للمتغيّر نفسه قبل القطع وبعده، ولم يكن بالضرورة اختيار المتغيّر نفسه قبل القطع وبعده.

2- أفضلية مقدّرات PLS المستعملة دالّة الجزاء Elastic- net المقترحة الاستعمال من قبل الباحثة، وهذا يعني أنّ هناك متغيّرات توضيحية في انموذج HDKR قد شكّلتْ مجاميع لمتغيّرات شديدة الارتباط، أثّرت على مقدّرات الطريقة بدوال الجزاء الأخرى في عملية الاختيار والتقدير؛ لأنّها امتلكت أقلّ متوسّط مربّعات الخطأ.

3- تمتاز طرائق التقدير المربّعات الصغرى الجزائية (PLS) بالتبعثر Sparsity ، وأنّها لا تحافظ على مقدّرات ثابتة؛ بسبب استعمال دوال جزاء مختلفة، تؤدّي إلى اختلاف في عملية الاختيار والتقدير لمَعْلَمَات أنموذج HDKR.

4-2 التوصيات :

1- توصي الباحثة باعتماد مقدّرات المربّعات الصغرى الجزائية، باستعمال دالّة الجزاء Elastic- net (المقترحة من قبل الباحثة) في التحليل الإحصائي عند تقدير مَعْلَمَات أنموذج HDKR المتضمّن عدداً كبيرا من المتغيّرات التوضيحية، كما هو الحال في بيانات القيمة السوقية لشركة بغداد للمشروبات الغازية، اعتماداً على تحديدها للمتغيّرات المؤثّرة، والتي تَمّ توضيحها في الجدولين (1) و (2) .

2- استعمال دوال جزاء جديدة، قد تكون خليطا من الدوال السابقة بعد الاستفادة من خواصهما المفيدة؛ لتحسين عملية اختيار المتغيّرات التوضيحية المؤثّرة، وتقدير مَعْلَمَاتها المرتبطة بها، ومقارنة نتائجها مع الطريقة PLS مع دالّة الجزاء Elastic- net سواء أكان لدراسة القيمة السوقية لشركة بغداد للمشروبات الغازية أم عند دراسة أيّ ظاهرة أخرى.

المصادر:

أولاً :- المصادر العربية:

1- السراي، علي حميد يوسف (2017) " استعمال المقدّرات الحصينة الجزائية لنموذج الانحدار الخطّي في ظلّ وجود مشكلتي الأبعاد والقيم الشاذة، مع تطبيق عملي " ، اطروحة دكتوراه، كُلّية الإدارة والاقتصاد – جامعة بغداد.

2- المسعودي، سعد جميل (2021) " مقارنة بين عدد من تقديرات الانتروبي العظمى العامّة في انموذج انحدار Kink، مع تطبيق عملي" رسالة ماجستير في الإحصاء، كُلّية الإدارة والاقتصاد، جامعة بغداد.

3- صالح، طارق عزيز (2016) " بعض الطرائق شبه المَعْلَمية في تقدير واختيار المتغيّر لأنموذج المؤشّر الواحد " اطروحة دكتوراه (إحصاء)، كُلّية الإدارة والاقتصاد، جامعة بغداد.

4- طالب، حيدر رائد (2022) " اختار المتغيّرات للأنموذج التجميعي شبه المَعْلَمي لبيانات عالية الأبعاد مع التطبيق "، اطروحة دكتوراه، كُلّية الإدارة والاقتصاد – جامعة بغداد.

ثانياً :- المصار الاجنبية

5- Ando, M. (2017) " How much should we trust regression-kink-design estimates? "Department of Economics, Uppsala University.

6- Brimacombe, M. (2014) "High-dimensional data and linear models: a review" Open Access Medical Statistics, 4 ,17–27.

7- Card, D., Lee, D. S., & Pei, Z. (2009). Quasi-Experimental Identification and Estimation in the Regression Kink Design (No. 1206).

8- Card, D., Lee, D. S., Pei, Z., & Weber, A. (2015). Inference on causal effects in a generalized regression kink design. Econometric, 83(6), 2453-2483.

9- Card, D., Lee, D. S., Pei, Z., & Weber, A. (2016). Regression kink design: Theory and practice. Emerald Publishing Limited.

10- Frank l.E., Friedman J.H. (1993) "A Statistical View of Some Chemometrics Regression Tools" TECHNMETRICS, vol.35, No.2,109-135.

11- Hansen, B., E. (2017) "Regression Kink with an Unknown Threshold" Journal of Business & Economic Statistics, 35:2, 228-240.

12- Huang, J., Ma, S., Li, H., Hui, C. (2011)" THE SPARSE LAPLACIAN SHRINKAGE ESTIMATOR FOR HIGH-DIMENSIONAL REGRESSION" The Annals of Statistics, vol. 39, No. 4, 2021-2046.

13- Huang, J., Jiao, Y., Liu, Y., Lu, X., (2018) "A Constructive Approach to L0 Penalized Regression" Journal of Machine Learning Research ,19, 1-37.

14- Jameel, S., O., Msallam, B., S., (2021) "Comparison between Renyi GME and Tsallis GME for Estimation Kink Regression" Int. J. Nonlinear Anal. Appl. (12) No. 1, 1235-1242.

15- Kim, J., Lee, S. (2017) "A convenient approach for penalty parameter selection in robust lasso regression" Communications for Statistical Applications and Methods 2017, Vol. 24, No. 6, 651–662.

16- Lee, S., Seo, M.H., Shin, Y. (2016) "The lasso for high dimensional regression with a possible change point "J. R. Statist. Soc. B , 78, Part 1, pp. 193–210.

17- Li,N., Yang,H. (2019) "Nonnegative estimation and variable selection under minimax concave penalty for sparse high-dimensional linear regression models "Springer-Verlag GmbH Germany.

18- Nielsen, H. S., Sørensen, T., & Taber, C. (2010). Estimating the effect of student aid on college enrollment: Evidence from a government grant policy reform. American Economic Journal: Economic Policy, 2(2), 185-215.

19- SHI, Y.Y., JIAO, Y. L., CAO, Y. X., LIU, Y. Y. (2018) " An Alternating Direction Method of Multipliers for MCP-penalized Regression with High-dimensional Data " Acta Mathematica Sinica, English Series, © Springer-Verlag GmbH Germany & The Editorial Office of AMS.

20- Tarkhamtham, P., Yamaka, W., & Sriboonchitta, S. (2018). The generalize maximum Tsallis entropy estimator in kink regression model, In Journal of Physics: Conference Series (Vol. 1053, No. 1, p. 012103). IOP Publishing.

21- Tarkhamtham, P., Yamaka, W., (2019). High-Order Generalized Maximum Entropy Estimator in Kink Regression Model, Thai Journal of Mathematics, 185-200.

22- Tibshirani, R. (1996)" Regression Shrinkage and Selection via the lasso" J. R. Statist. Soc., 58, No. 1, pp. 267-288.

23- Tibshirani, R. (2014) "Lasso and Sparsity in Statistics" chapter five in Statistics in Action: A Canadian outlook (book) edited by Jerald F., Tayler & Francis Group. LIC.

24- Wang, Y., Fan, Q., Zhu, L. (2018)" Variable selection and estimation using a continuous approximation to the L0 penalty" Ann Inst Stat Math 70:191–214.

25- Wang, X., Dunson, D., Leng, C., (2016) "No penalty no tears: Least squares in high-dimensional linear models" Proceedings of the 33 rd International Conference on Machine Learning, New York, NY, USA, 2016. JMLR: W&CP volume 48.

26- Xu, C., Fang, J., Shen, H., Wang, Y., Deng, H. (2018) "EPS-LASSO: test for high-dimensional regression under extreme phenotype sampling of continuous traits" Bioinformatics, 34(12), 1996–2003.

27- Yamaka.W.(2021) "Sparse estimation in kink regression model " Soft Computing, Springer.

28- Zhang, S., Wu, J., Jia, Y., Wang, Y., Zhang, Y., Duan, Q. (2021) " A temporal LASSO regression model for the emergency forecasting of the suspended sediment concentrations in coastal oceans: Accuracy and interpretability" Engineering Applications of Artificial Intelligence, vol. 100 (36):104206.