مجلة كلية التجارة للبحوث العلمية
كلية التجارة
جامعة الإسكندرية

العدد الثانى – المجلد الخامس والخمسون – يوليو ٢٠١٨

تأثير اختلاف قيمة نقطة الانكسار على الكفاءة التقاربية لمقدر S فى نموذج انحدار الشرائح المعاقبة*: دراسة محاكاة

أ/ صابرين محمد دسوقى ابراهيم
مدرس مساعد
قسم الإحصاء والرياضة والتأمين
كلية التجارة – جامعة دمنهور

The Effect of the Breakdown Point Value on the Asymptotic Efficiency of the S-Estimator for Penalized Regression Splines Model: A Simulation Study

الملخص

يعد نموذج انحدار الشرائح المعاقبة أحد اشهر النماذج التي يتم استخدمها لتمهيد البيانات التي يصعب فيها تحديد الشكل الدالي الملائم للتعبير عن العلاقة التي تربط المتغير التابع بالمتغير المستقل في نماذج الانحدار الخطي البسيط. إلا أنه في الواقع العملي يمكن أن تحتوي البيانات على مشاهدات شاذة لذلك يفضل استخدام مقدرات متينة "robust estimators" لذلك النموذج مثل مقدر S. ومن الخصائص العامة لمقدر S في حالة نماذج الانحدار الخطي بصفة عامة أن لهذا المقدر نقطة انكسار "breakdown point" يمكن أن تصل إلى ٥٠% إلا أن ذلك يصاحبه انخفاض في الكفاءه التقاربية للمقدر "lowasymptotic efficiency", وإذا ما تم استخدام مقدر S له نقطة انكسار أقل من ٥٠% فسوف يصاحب ذلك أرتفاع نسبي في الكفاءة التقاربية للمقدر. ولقد اهتم البحث الحالي بدراسة تأثير اختلاف نقطة الانكسار على أداء مقدر S لنموذج انحدار الشرائح المعاقبة وذلك من خلال عمل دراسة محاكاة.

Abstract

The penalized regression spline model is one of the most popular models that can be used for smoothing data in which it is difficult to determine the appropriate functional form that expresses the relationship between the dependent variable and the independent variable in the simple linear regression models. In practice, however, data containing outliers can be encountered, so there is a need for using robust estimators for that model such as S-estimator. A general feature of the S estimator in the linear regression models is that this estimator can have a breakpoint point 50% but this is accompanied by a low-asymptotic efficiency, and if a breakdown point of the estimator is less than 50% will be accompanied by a relative increase in the relative efficiency of the estimator. Therefore, the current research is concerned with the study of the effect of the difference of the breakdown point on the performance of the S estimator for penalized regression spline model by conducting a simulation study.

*البحث مشتق من رسالة دكتوراه بعنوان "التقدير المتين لنموذج انحدار الشرائح المعاقبة: دراسة محاكاة"، تحت اشراف الأستاذ الدكتور/ محمد على محمد أحمد، والدكتور/ أحمد صدقى محمد الديب.

١- مقدمة

يعـد نمـوذج انحـدار الشـرائح المعاقبـة "penalized regression spline (PRS)" من أكثـر الأسـاليب الإحصـائية المسـتخدمة فـي تمهيـد البيانـات المشوشـة "noisy data"، نظراً للمرونـة التي يتمتع بها هذا النموذج في التعبير عن العديد مـن العلاقـات الداليـة المعقـدة والتـي يصـعب فيهـا تحديـد الشـكل الـدالي الملائـم للعلاقـة التـي تـربط المتغير المفسر بالمتغير التابع في نموذج الانحدار البسيط. ولقد اتسع مجال استخدام ذلك النموذج في العديـد مـن النـواحي التطبيقيـة. فمـثلاً فـي النـواحي الطبية (Griggs,2013) ، وفي النواحي الاقتصادية (Greiner, 2009), وفي تطبيقات لظواهر طبيعية وغيرها (Ruppert et al., 2009) .

وفـي الواقـع العملـي, يمكـن أن يواجـه مسـتخدموا نمـوذج الانحـدار الخطـي البسـيط بصـفة عامـة مشـاهدات تختلـف مـن حيـث النسـق عـن بـاقي المشـاهدات, وهـي تلك التـي يطلـق عليهـا مشـاهدات شـاذة "outliers". وقـد يكـون لهـذه المشـاهدات الشـاذة تـأثير كبيـر علـى نمـوذج الانحـدار المقـدر, حيـث يمكـن أن تجعـل منحنـى المربعـات الصـغرى الموفق يتجه نحو المشاهدات الشاذة بدلاً من التعبير عن أغلبية المشاهدات والتي إن جاز التعبير يمكن أن يطلق عليها مشاهدات جيدة. ونفس الأمر يمكن أن يحـدث لمسـتخدمي نمـوذج انحـدار الشـرائح المعاقبـة, حيـث يمكـن أن يـؤثر وجـود بعـض المشاهدات الشاذة على منحنى PRS الموفق.

وكما ذكر Tharmaratnam et al. (2010) أن السـبب الأساسـي فـي تـأثر نمـوذج PRS بالمشاهدات الشاذة يرجع إلى تقديره باستخدام طريقة المربعـات الصـغرى والتـي مـن المعـروف شـدة حساسيتها لوجود مشاهدات شاذة في البيانات. لذلك بـات مـن الضـروري أن يـتم استخدام مقدرات متينـة لنموذج PRS وهي تلك المقدرات التي لا تتأثر بشدة في حالـة وجـود مشـاهدات شـاذة فـي البيانـات. ولقد تناولت الأدبيات الإحصائية السابقة مقدرين متينين لنموذج انحدار الشرائح المعاقبة وهما:

١- مقدر M والذي قدمه Lee and Oh (2007).

٢- مقدر S والذي قدمه Tharmaratnam et al. (2010).

وتعد نقطة الانكسار " breakdown point" أحـد أهـم المعـايير التـي يمكـن الاسـتناد عليهـا عنـد اختيار المقدر المتين الذي يمكن استخدامه للبيانات محل الاهتمام. حيث تشير نقطـة الانكسار إلى أقل نسبة من المشاهدات التي يؤدي تغييرها في عينة ما إلـى احـداث تغييـرات هائلـة علـى التقـديرات Rousseeuw and Leroy (1987). وتعد نقطة الانكسار ٥٠% هي أقصى قيمة يمكن أن يصل لها أي مقـدر متـين. وبصـفة عامـة, تتسـم مقـدرات M بانخفاض نقطة انكسارها في نماذج الانحدار بصفة عامـة إلا إنهـا تتميـز بـأن لهـا كفـاءة تقاربيـة عاليـة "high asymptotic efficiency" Rousseeuw and Leroy (1987). في حين أن نقطة انكسار مقـدرات S يمكـن أن تصـل إلـى ٥٠% فـي حـالات معينـة إلا أن هـذا سيكون مصـحوباً بانخفـاض فـي كفاءتها التقاربية, وكلما تم استخدام مقدر S بنقطـة انكسـار أقـل سـوف يصـاحب ذلـك زيـادة فـي قيمـة معيار الكفاءة التقاربية . وفيما يخص نموذج PRS, فإن مقدر S الذي قدمه Tharmaratnam et al. (2010) كان لـه أقصـي نقطـة انكسـار ممكنـة وهـي ٥٠% وهنـا يتبـادر سـؤال : هـل اسـتخدام مقـدر S

لنموذج انحدار الشرائح المعاقبه بنقطة انكسار أقل من ٥٠% يمكن أن تكون نتائجه أفضل –من الناحية التطبيقية– من مقدر S ذا نقطة الانكسار ٥٠%؟

ولقد اهتم البحث الحالي بمحاولة الإجابة عن ذلك السؤال وذلك من خلال القيام بدراسة محاكاة والذي سيتم فيها المقارنة بين المقدرات التالية:

١- مقدر المربعات الصغرى: وذلك لمقارنة أداء المقدرات المتينة التي تتضمنها دراسة المحاكاة بمقدر المربعات الصغرى غير المتين.
٢- مقدر M الذي قدمه Lee and Oh (2007).
٣- مقدر S ذا نقطة انكسار ٥٠%.
٤- مقدر S ذا نقطة انكسار ٣٠%.

وسوف يتم تقييم أداء الأربع مقدرات السابق ذكرهم في تسع حالات مختلفة تعكس تسعة توزيعات مختلفة للخطأ .

ولقد تم تنظيم المتبقي من البحث كالتالي. يتناول القسم الثاني عرضاً مختصراً لنموذج انحدار الشرائح المعاقبة. كما يتناول القسم الثالث مقدر M لنموذج انحدار الشرائح المعاقبة. في حين يتناول القسم الرابع مقدر S لنموذج انحدار الشرائح المعاقبة. بينما يتناول القسم الخامس نتائج دراسة المحاكاة التي تم القيام بها للمقارنة بين نتائج استخدام أربعة مقدرات لنموذج انحدار الشرائح المعاقبة. وتم القسم السادس لتناول خلاصة البحث.

٢- نموذج انحدار الشرائح المعاقبة

تقوم الفكرة الأساسية لنموذج انحدار الشرائح – المعاقبة أو غير المعاقبة– من الدرجة p على تقسيم المدى الذى يأخذه المتغير المستقل X إلى فترات جزئية غير متداخلة. ويفصل بين تلك الفترات الجزئية نقاط يطلق عليها عقد "knots". ثم يتم توفيق كثيرة حدود من الدرجة p في الفترات الجزئية المختلفة مع فرض قيد يجعل قطع كثيرات الحدود الموفقة في الفترات الجزئية المختلفة متصلة ومشتقاتها متصلة إلى الدرجة (p−1) عند العقد.

وتعد الصيغة العامة التالية هي أكثر الصيغ شيوعاً للتعبير عن نموذج انحدار الشرائح Ruppert (2002).

$$\boldsymbol{y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p + \sum_{k=1}^{K} \beta_{p+k}(x_i - \xi_k)_+^p + \varepsilon_i} \quad (١)$$

حيث تشير p إلى درجة نموذج انحدار الشرائح, بينما تشير K إلى عدد العقد المستخدمة. وتشير ξ_k إلى العقدة رقم k .

وحيث يعبر المكون الأول من المعادلة (١) " $\boldsymbol{\beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p}$" عن كثيرة حدود من الدرجة p .

كما يعبر المكون الثاني من المعادلة (١) " $\boldsymbol{\sum_{k=1}^{K} \beta_{p+k}(x_i - \xi_k)_+^p}$" عن الحد الذي يضمن أن تكون قطع كثيرات الحدود ومشتقاتها متصلة إلى الدرجة *(p−1)* عند العقد.

كما يشير الرمز ε_i إلى حد الخطأ العشوائي.

وحيث

$$(x_i - \xi_k)_+^p = \begin{cases} (x_i - \xi_k)^p & \text{إذا كانت } x_i > \xi_k \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

ومن الواضح أن y_i تعد توليفة خطية في عدد $(p + K + 1)$ من الدوال التالية: $1, x, \dots\dots, x^p, (x - \xi_1)_+^p, (x - \xi_2)_+^p, \dots\dots, (x - \xi_K)_+^p$

ويطلق على تلك الدوال اسم دوال الأساس "Basis Function" حيث يمكن التعبير عن النموذج كتوليفة خطية في هذه الدوال. ويطلق على دوال الأساس الموضحة في معادلة (١) تحديداً دوال أساس القوى المبتورة "truncated-power basis function" . ويمكن التعامل مع النموذج الموضح في معادلة (١) كنموذج انحدار خطي متعدد ويمكن تقديره باستخدام المربعات الصغرى العادية Ruppert (2002).

وتعد عملية اختيار عدد ومواقع العقد الملائمة من اهم المشاكل التي تقيد استخدام نموذج انحدار الشرائح حيث باختلاف عدد ومواقع العقد يختلف النموذج الناتج. فبزيادة عقدة واحدة لنموذج معين أو بحذفها نحصل على نموذج آخر مختلف. علاوة على أنه عند نفس العدد من العقد يمكن بتحريك مكان أي عقدة الحصول على نموذج آخر.

وقد استخدمت كثير من الدراسات السابقة مثل دراسة Eilers and Marx (1996) ودراسة Ruppert (2002) مدخل الانكماش المعاقب "penalized shrinkage approach" للتغلب على مشكلة اختيار عدد ومواقع العقد الملائمة عن طريق استخدام عدد كبير –بدرجة كافية– من العقد، بحيث إذا تم تقدير معالم النموذج باستخدام المربعات الصغرى العادية يتم الحصول على نموذج زائد التوفيق "overfitted" . والذي يعني الحصول على توفيق يقترب فيه المنحنى الموفق من مشاهدات العينة والذي قد لا تمثل المجتمع تمثيلاً جيداً. وللحصول على توفيق أكثر تمهيداً للبيانات يتم استخدام ما يطلق عليه المربعات الصغرى المعاقبة "Penalized Least Squares" PLS حيث يتم الحصول على المقدرات $\widehat{\beta}_\lambda$ التي تدني معيار المربعات الصغرى المقيدة، وعادة يأخذ ذلك المعيار الشكل التالي Ruppert (2002) :

$$\sum_{i=1}^{n}(y_i-(\beta_0+\beta_1x_i+\beta_2x_i^2+\cdots+\beta_px_i^p+\sum_{k=1}^{K}\beta_{p+k}(x_i-\xi_k)_+^p))^2+\lambda\sum_{k=1}^{K}\beta_{k+p}^2 \qquad (٢)$$

حيث يكون متجه المعلمات المقدرة لطريقة المربعات الصغرى المعاقبة على الشكل التالي:

$$\widehat{\beta}_\lambda=(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{D})^{-1}\mathbf{X}^T\mathbf{y} \qquad (٣)$$

حيث $\mathbf{D}$ عبارة عن مصفوفة عقاب "Penalized Matrix" والتي عادةً تكون "مصفوفة قطرية" بحيث يكون عدد $p+1$ من عناصرها القطرية الأولى مساوية للصفر, وباقي العناصر القطرية والتي عددها K تساوى الواحد وبذلك يكون العقاب على المعاملات الخاصة بالعقد.

كما يشير الرمز $\mathbf{X}$ إلى مصفوفة دوال الأساس المستخدمة. ويقيس الحد الأول من المعيار الموضح في معادلة (٢) مدى اقتراب المنحنى الموفق من البيانات, فكلما زاد عدد العقد كلما أقترب المنحنى الموفق من البيانات والعكس صحيح. بينما يعاقب الحد الثاني الانحناءات في المنحنى الموفق. ويطلق على المعلمة λ اسم معلمة التمهيد "smoothing parameter" وتعمل تلك المعلمة على الموازنة بين الحدين السابقين، حيث يكون لها تأثيراً على تقليص معاملات الحدود الخاصة بالعقد تجاه منحنى أكثر تمهيداً. فكلما زادت قيمة تلك المعلمة كلما زادت قيمة الحد الثاني من معادلة (٢) وبالتالي زاد العقاب على الانحناءات في المنحني الموفق ومن ثم الحصول على منحنى أكثر تمهيدا, والعكس صحيح.

ويطلق على انحدار الشرائح الموفق باستخدام مدخل االانكماش المعاقب اسم **الشرائح المعاقبة "penalized spline"** واختصارا "P-spline"

Eilers and Marx (1996) . ويمكن القول, أنه باستخدام نموذج انحدار الشرائح المعاقبة يكون قد تم التغلب على مشكلة اختيار عدد ومواقع العقد الملائمة التي كانت تحد من استخدام نموذج انحدار الشرائح الغير معاقبة. حيث يتطلب الأمر في النموذج المعاقب تحديد ملائم لمعلمة التمهيد λ بدلاً من التحديد الملائم لعدد ومواقع العقد في النموذج غير المعاقب. وعادةً ما يتم التحديد الملائم لتلك المعلمة عن طريق استخدام احد المعايير مثل معيار الملاءمة المقطعية المعمم "generalized cross validation (GCV)" لاختيار القيمة المثلى لمعلمة التمهيد من بين مجموعة من القيم المرشحة لتلك المعلمة.

٣- مقدر M لنموذج انحدار الشرائح المعاقب

بصفة عامة, تعتمد مقدرات M في نماذج الانحدار الخطية على تدنية دالة خسارة ما في البواقي, وتتحدد مدى متانة "robustness" المقدر الناتج على الوزن الذي تأخذه المشاهدات في كل حالة. وعلى ذلك يمكن النظر لمقدر المربعات الصغرى العادية OLS على إنه حالة خاصة من مقدرات M حيث يعتمد مقدر OLS على تدنية مجموع مربعات الخطأ كدالة خسارة, وتتساوي الأوزان التي تأخذها المشاهدات بالنسبة لهذا المقدر. ومن المعروف, أن مقدر OLS لا يعد من المقدرات المتينة حيث أن له نقطة انكسار "breakdown point" مساوية للصفر والذي يعني أنه يمكن أن يتأثر بشدة في حالة وجود نسبة صغيرة من المشاهدات الشاذة. وكخطوة لجعل المقدر الناتج أكثر متانة يمكن استخدام دالة خسارة أقل تأثراً بالمشاهدات الشاذة من دالة خسارة مربعات الخطأ.

ويعد ما قدمه Huber (1973) من ضمن أقدم ما تم اقتراحه في هذا الصدد. حيث يمكن على سبيل المثال استخدام دالة خسار هوبر "Huber" بدلاً من دالة خسارة مربعات الخطأ, وفي تلك الحالة يكون للمشاهدات التي لها قيم صغيرة للبواقي وزناً أكبر من المشاهدات الشاذة التي لها قيم كبيرة للبواقي أو بمعنى أخر تكون الأوزان متناقصة للمشاهدات التي تبعد عن خط الانحدار .

وفيما يتعلق بنموذج انحدار الشرائح المعاقبة, قدم Lee and Oh (2007) توفيقاً لإنحدار شرائح معاقب متين يعتمد على مقدر M عن طريق استخدام دالة متينة تتميز بانها أقل تأثراَ بالمشاهدات الشاذة والموضحة في معادلة (٤).

$$\hat{f}_{robust}(x) = argmin_{\beta}\{\sum_{i=1}^{n} \rho_c(y_i - \beta^T x_i) + \lambda\beta^T D\beta\} \qquad (٤)$$

حيث يشير الرمز $\rho_c(x)$ إلى دالة خسارة هوبر "Huber loss function" والتي تأخذ الشكل التالي:

$$\rho_c(x) = \begin{cases} x^2 & |x| \le c \\ 2c|x| - c^2 & |x| > c \end{cases}$$

وحيث $c = k\hat{\sigma}$, وتم اختيار قيمة $k = 1.345$ لتحقيق كفاءة تقاربية 95% بالنسبة للتوزيع المعتدل المعياري. وحيث $\hat{\sigma}$ هو مقدر متين للمعلمة σ "الانحراف المعياري للخطأ".

وإذا كانت ψ_c تشير إلى تفاضل "derivative" ρ_c , فيمكن الحصول على $\hat{f}_{robust}(x)$ عن طريق حل المعادلة التالية:

$$-\sum_{i=1}^{n} \psi_c\{y_i - \mathbf{\beta}^T \mathbf{x}_i\}\,\mathbf{x}_i + \frac{\partial\lambda\mathbf{\beta}^T D\mathbf{\beta}}{\partial\beta} = 0$$

ولقد أوضح Lee and Oh (2007) أن عملية تدنية معادلة (٤) للوصول إلى حل للمعادلة السابقة لا يعد أمراً سهلاً نظراً للطبيعة غير الخطية للدوال ρ_c و ψ_c , ولــذلك قامـا بانشـاء بيانـات زائفـة "pseudo data" $\tilde{y}_i$ كالتالي:

$$\tilde{y}_i = f(x_i) + \frac{\psi_c\{y_i - f(x_i)\}}{2}$$

ولقد عرفا المقدر الزائف $\tilde{f}_{pseudo}(x)$ كحل لمشكلة تدنية معيار المربعات الصغرى التالي والتي لها حل ذو صيغة صريحة "closed form solution" في حالة معرفة قيم المتغير التابع الزائفة $\tilde{y}_i$:

$$\tilde{f}_{pseudo}(x) = argmin_{\beta}\{\sum_{i=1}^{n}\{\tilde{y}_i - \beta^T x_i\}^2 + \lambda\beta^T D\beta\}$$

وعملياً لا يمكن حساب $\tilde{y}_i$ و بالتالي $\tilde{f}_{pseudo}(x)$ نظراً لإنهما يتطلبان معرفة الدالة المجهولة $f(x_i)$, إلا أن Lee and Oh (2007) اقترحا خوارزم للوصول إلى $\hat{f}_{robust}(x)$ نظراً لاثباتهم نظرياً أن هناك تقارب احتمالي "convergence in probability" بين $\hat{f}_{robust}(x)$ و $\tilde{f}_{pseudo}(x)$. وفيما يلي نعرض خطوات خوارزم حساب مقدر M لنموذج انحدار الشرائح المعاقب وفقاً لدراسة Lee and Oh (2007):

١- الحصول على مقدر مبدئي $\hat{f}^{(0)}(x)$ للدالة المراد تقديرها $f(x)$ وليكن مقدر OLS على سبيل المثال.

٢- وضع $z_i^{(0)} = y_i \qquad i = 1, \dots, n$

حيث تشير z_i إلى البيانات الزائفة التجريبية "empirical pseudo data" والتي يمكن حسابها كالتالي:

$$z_i = \hat{f}(x_i) + \frac{\psi_c\left(y_i - \hat{f}(x_i)\right)}{2}$$

٣- تكرار الخطوات التالية لقيم .... $j = 0,1, \dots$ إلى أن يحدث التقارب :

(أ)- الحصول على تقدير متين للانحراف المعياري لحد الخطأ العشوائي $\hat{\sigma}^{(j+1)}$ باستخدام البواقي $\hat{\varepsilon}_i = y_i - \hat{f}^{(j)}(x_i) \quad , i = 1, \dots, n$ كالتالي: $\hat{\sigma}^{(j+1)} = 1.4826 \times MAD(\hat{\varepsilon}_i)$

حيث يشير الرمز MAD إلى وسيط الانحراف المطلق عن الوسيط "median of absolute deviation from median".

(ب)- حساب قيمة $z_i^{(j+1)}, i = 1, \dots, n$ باستخدام قيمة قطع $c = 1.345\hat{\sigma}^{(j+1)}$ عند حساب ψ_c.

(ج)- الحصول على تقدير للدالة $\hat{f}^{(j+1)}(x)$ كالتالي:

$$\hat{f}^{(j+1)}(x) = \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{D})^{-1}\mathbf{X}^T z^{(j+1)}$$

حيث يتم استخدام معيار GCV التالي للحصول على القيمة المثلى لمعلمة التمهيد λ وذلك من بين مجموعة مرشحة من القيم لتلك المعلمة.

$$GCV_{\lambda} = n\left\|y - \hat{f}(X)\right\|^2 / (n - trace(H(\lambda)))^2$$

وعادةً ما يطلق على المصفوفة $\mathbf{H}(\lambda)$ اسم مصفوفة التمهيد "smoothing matrix" وهي مصفوفة تناظر مصفوفة التقدير "hat matrix" في حالة استخدام المربعات الصغرى العادية ويمكن الرجوع إلى Hastie et al. 2001 لمعرفة بعض من أوجه الشبه والاختلاف بين هاتين المصفوفتين.

٤- يعد $\hat{f}^{(j+1)}(x)$ الذي يتم الحصول عليه بعد حدوث التقارب هو التقدير المتين النهائي $\hat{f}_{robust}(x)$ للدالة $f(x)$.

وبعد التعرف على مقدر M لنموذج انحدار الشرائح المعاقب وفقاً لما قدمه Lee and Oh

(2007), سيتناول القسم التالي مقدر S لذلك النموذج.

٤- مقدر S لنموذج انحدار الشرائح المعاقب

نظراً لانخفاض قيمة نقطة الانكسار لمقدرات M في نماذج الانحدار الخطية, قدم Rousseeuw and Yohai (1984) ما يعرف بمقدرات S . والتي عرفت بهذا الأسم نظراً لإنها تعتمد على مقدرات معلمة قياس "Scale parameter". وتتبع الفكرة الأساسية لمقدرات S من أن مقدر المربعات الصغرى العادية OLS لمتجه معاملات الانحدار β يعتمد على تدنية مجموع مربعات الخطأ ومن ثم الانحراف المعياري للبواقي, وكبديل متين لذلك يتم الاعتماد على تدنية مقياس متين لتشتت البواقي بدلاً من الانحراف المعياري.وبالتالي يمكن تعريف مقدر S لمتجه معاملات الانحدار كالتالي:

$$\hat{\beta}_n = \operatorname{argmin}_\beta\left(\hat{\sigma}_n(\beta)\right) \qquad (٥)$$

حيث يشير $\hat{\sigma}_n(\beta)$ إلى مقدر M المتين لمقياس التشتت "robust M-scale estimator" كما قدمه Huber(1964) والذي يتم الحصول عليه بحل المعادلة التالية:

$$\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{r_i}{\hat{\sigma}_n}\right) = b$$

وحتى يكون المقدر متسقاً في حالة ما إذا كان توزيع الأخطاء طبيعياً يتم تحديد قيمة b كالتالي:

$$b = E_\Phi[\rho(Z)]$$

حيث تشير Φ إلى التوزيع الطبيعي المعياري. وقد أوضح Tharmaratnam et al. (2010) أن حل المعادلة (٥) يعد مشكلة صعبة حيث يتضمن إيجاد القيمة الصغرى لدالة معرفة بشكل ضمني وغير محدبة "non-convex" في عدة متغيرات. ويعد الخوارزم الذي قدمه كلٍ من Salibian and Yohai (2006) أحد الاقتراحات لحل تلك المشكلة في حالة الانحدار غير المعاقب . وفيما يتعلق بنموذج انحدار الشرائح المعاقب اقترح Tharmaratnam et al. (2010) أن يكون مقدر S لمتجه معاملات الانحدار $\hat{\beta}_S$ كالتالي:

$$\widehat{\beta}_s = argmin_\beta(n\hat{\sigma}_n^2(\beta) + \lambda\beta^T D\beta) \quad (٦)$$

حيث – لكل متجه β – $\hat{\sigma}_n^2(\beta)$ تحقق المعادلة التالية:

$$\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{y_i - \mathrm{x}_i^T\beta}{\hat{\sigma}_n(\beta)}\right) = b$$

وأحد الأمثلة الشائعة الاستخدام للدالة ρ هي:

$$\rho_c(u) = \begin{cases} (u^2/2) - (u^4/2c^2) + (u^6/6c^4) & if\ |u| \le c \\ c^2/6 & if\ |u| > c \end{cases}$$

وتفاضل "derivative" تلك الدالة هو دالة "Tukey's biweight"" التي قدمها Beaton and Tukey (1974):

$$\psi(u) = \begin{cases} u(1-(u/c)^2)^2 & |u| \le c \\ 0 & |u| \ge c \end{cases}$$

وعند اختيار قيمة c=1.547 يكون لمقدر S للانحدار غير المعاقب أكبر نقطة انكسار تقاربية "maximal asymptotic breakdown point 50%" (Rousseeuw and Yohai (1984)) إلا أن ذلك سيصاحبه انخفاضاً في قيمة الكفاءة التقاربية "e" (انظر جدول ١). ويتضح من جدول (١) أنه بزيادة قيمة الثابت c ستزيد معه قيمة الكفاءة التقاربية وستنخفض قيمة نقطة الانكسار المناظرة, الأمر الذي يتطلب تحديد قيمة ملائمة للثابت c للحصول على نتائج مرضية للمقدر S. ولقد أوصى Rousseeuw and Yohai (1984) بعدم استخدام مقدرات S ذوات نقاط الانكسار الأقل من ٢٥% بصفة عامة .

جدول (١): الكفاءة التقاربية e لمقدرات S المناظرة لنقاط انكسار مختلفة ε^* وقيم الثوابت c و b المختلفة باستخدام دالة "Tukey's biweight".

ε^*	e	c	b
50%	28.7%	1.547	.1995
45%	37.0%	1.756	.2312
40%	46.2%	1.988	.2634
35%	56.0%	2.251	.2957
30%	66.1%	2.560	.3278
25%	75.9%	2.937	.3593
20%	84.7%	3.420	.3899
15%	91.7%	4.096	.4194
10%	96.6%	5.182	.4475

المصدر: Rousseeuw and Yohai (1984)

ولقد أثبت Tharmaratnam et al. (2010) نتيجة توضح إمكانية كتابة النقاط الحرجة "the critical points" لدالة الهدف الموضحة في معادلة (٦) على انها حل لمشكلة انحدار شرائح معاقب مرجح "weighted Penalized splines problem"، ومن ثم يمكن استخدام أسلوب معاودات "iterative procedure" لايجاد مقدر S المعاقب.

وبالتالي يمكن التعبير عن مقدر S كالتالي:

$$\hat{f}_S(x) = X\hat{\beta}_S$$

حيث

$$\hat{\beta}_S = \left\{X^T W(\hat{\beta}_S)X + \frac{\lambda}{\tau(\hat{\beta}_S)} D\right\}^{-1} X^T W(\hat{\beta}_S)y \quad (٧)$$

حيث

$$W(\beta) = diag\big(W_i(\beta)\big) \in \mathbb{R}^{n\times n}$$

$$W_i(\beta) = \rho'\big(\tilde{r}_i(\beta)\big)/\tilde{r}_i(\beta)$$

$$\tilde{r}_i(\beta) = \big(y_i - x_i^T\beta\big)/\hat{\sigma}_n(\beta)$$

$$\tau(\beta) = n\hat{\sigma}_n^2(\beta)/[(y - X\beta)^T W(\beta)(y - X\beta)]$$

وبالرغم من أن معادلة (٧) تقترح استخدام المعاودات لحساب النقطة الحرجة لمعادلة (٦) إلا أنه ينبغي توخي الحذر لإن الدالة $\hat{\sigma}_n$ تعد بصفة عامة غير محدبة كما ذكر. والذي قد يعني أن يكون للمعادلة (٦) أكثر من نقطة حرجة والتي تناظر قيم صغرى محلية "local minima" مختلفة , والذي يمكن أن يسفر عن وجود نقاط حرجة غير مثلى. ولذلك تم اقتراح أن يتم بدء المعاودات باستخدام العديد من القيم المبدئية واختيار افضلهم (وفقا لقيمة دالة الهدف ٨) . وفيما يلي نعرض خطوات خوارزم الحصول على تقدير S لنموذج PRS وفقاً لما قدمه Tharmaratnam et al.(2010):

١- الحصول على عدد R من التقديرات المبدئية المرشحة لمتجه المعالم $\boldsymbol{\beta}$ والتي سوف يشار لها بالرموز $\hat{\boldsymbol{\beta}}_1^{(0)}, \hat{\boldsymbol{\beta}}_2^{(0)}, \ldots\ldots, \hat{\boldsymbol{\beta}}_R^{(0)}$. ثم يتم تنفيذ الخطوات التالية لكل متجه $\hat{\boldsymbol{\beta}}_r^{(0)}$:

(أ)- حساب كلٍ من $\tau\left(\widehat{\boldsymbol{\beta}}_r^{(0)}\right)$ و $\hat{\sigma}_n\left(\widehat{\boldsymbol{\beta}}_r^{(0)}\right)$ و $\mathbf{W}\left(\widehat{\boldsymbol{\beta}}_r^{(0)}\right)$.

(ب)- يتم وضع j=0 ثم تكرار الخطوات التالية:

(I)- يتم حساب $\hat{\beta}_r^{(j+1)} = \left\{X^T W\left(\hat{\beta}_r^{(j)}\right) X + \lambda D \tau^{-1}\left(\hat{\beta}_r^{(j)}\right)\right\}^{-1} X^T W(\hat{\beta}_r^{(j)}) y$

(II)- إذا وصلت j للحد الأقصى من التكرارات "المحددة مسبقاً" ,

أو إذا تحقق أن $\left\|\widehat{\boldsymbol{\beta}}_r^{(j)} - \widehat{\boldsymbol{\beta}}_r^{(j+1)}\right\| < \epsilon \left\|\widehat{\boldsymbol{\beta}}_r^{(j)}\right\|$, حيث $\epsilon > 0$ هو قيمة ثابتة صغيرة تعبر عن مستوى السماح *"tolerance level"* , يتم التوقف عن التكرار ووضع $\widehat{\boldsymbol{\beta}}_r^B = \widehat{\boldsymbol{\beta}}_r^{(j)}$.

(III)- إذا لم يتحقق أي من الشرطين في الخطوة (II) يتم حساب كلٍ من $\hat{\sigma}_n\left(\hat{\beta}_r^{(j+1)}\right)$ و $\tau\left(\widehat{\boldsymbol{\beta}}_r^{(j+1)}\right)$ و $\mathbf{W}\left(\widehat{\boldsymbol{\beta}}_r^{(j+1)}\right)$, ثم يتم وضع $j \leftarrow j+1$.

٢- حساب قيمة دالة الهدف "objectivefunction" لكل المتجهات $\widehat{\boldsymbol{\beta}}_r^B$, حيث $r = 1,2,\ldots\ldots,R$ واختيار المتجه صاحب أقل قيمة لدالة الهدف ليكون هو تقدير S لنموذج PRS كالتالي:

$$\hat{\beta}_S = argmin_{1\le r\le R}\left[n\hat{\sigma}_n^2\left(\hat{\beta}_r^B\right) + \lambda\left(\hat{\beta}_r^B\right)^T D\hat{\beta}_r^B\right]$$

ولقد تم توضيح الخوارزم السابق بناءً على استخدام قيمة واحدة لمعلمة التمهيد λ , إلا أن الأمر يتطلب إلى استخدام معيار مثل معيار الملاءمة المقطعية *GCV* لاختيار قيمة مثلى لتلك المعلمة من بين عدد من القيم المرشحة لها. ويمكن ملاحظة أن الصيغة العادية لمعيار *GCV* تعامل جميع المشاهدات $y_i, i = 1,\ldots,n$ بنفس الأهمية, إلا أنه بالطبع في حالة وجود بعض المشاهدات الشاذة $y_j, 1 \le j \le n$ لا يحبذ أن يتم إعطاء تلك المشاهدات نفس أهمية المشاهدات الجيدة. وبمعنى أخر- كما ذكر *(2010)*Tharmaratnam et al. - أنه بغض النظر عن متانة المقدر $\hat{f}(x)$ فإن معيار *GCV* بصيغته العادية يمكن أن يختار قيمة للمعلمة λ ينتج عنها اقتراب تقدير $f(x_j)$ من قيمة y_j بشكل غير مرغوبٍ فيه. وفي ذلك الصدد اقترح كلٍ من Cantoni and Ronchetti (2001) أن يتم استخدام أوزان للمشاهدات وفقاً لقيمة البواقي *"residuals"* المناظرة, ومن ثم إعطاء المشاهدات الشاذة وزناً أقل من وزن المشاهدات الجيدة. وفيما يخص مقدر *S* لنموذج انحدار الشرائح المعاقب اقترح Tharmaratnam et al. (2010) صيغة متينة لمعيار *GCV* وأشار لها بالرمز *RGCV* وهي الموضحة في المعادلة التالية:

$$\mathrm{RGCV}_\lambda = \mathrm{n_w}\left\|\mathrm{W}(\hat{\beta})^{1/2}(\mathrm{y} - \mathrm{X}\hat{\beta})\right\|^2 / \left(\mathrm{n_w} - \mathrm{trace}(\mathrm{H_S}(\lambda))\right)^2 \quad (٨)$$

حيث

$$\mathrm{H}_S(\lambda) = \tilde{X}\left(\tilde{X}^T\tilde{X} + \left(\lambda/\tau(\hat{\beta}_S)\right)D\right)^{-1}\tilde{X}^T$$
$$= W(\hat{\beta}_S)^{1/2}X\left(X^T W(\hat{\beta}_S)X + \left(\lambda/\tau(\hat{\beta}_S)\right)D\right)^{-1}X^T W(\hat{\beta}_S)^{1/2}$$

وحيث تشير n_W إلى عدد الأوزان غير الصفرية.

٥- دراسة المحاكاة

تهدف دراسة المحاكاة في هذا البحث إلى تحديد مدى تأثير اختلاف نقطة انكسار مقدر S لنموذج انحدار الشرائح المعاقبة على أداء ذلك المقدر من الناحية العملية. ويعتمد تصميم دراسة المحاكاة في هذا البحث على التصميم الذي تم استخدامه في الدراسات التالية : Cantoni and Ronchetti (2001) و Lee and Oh (2007) و Wang et al. (2014) حيث كانت دالة الاختبار الحقيقية المراد تقديرها كالتالي:

$$y_i = sin\{2\pi(1-x_i)^2\} + \varepsilon_i, \quad i = 1, \dots, n$$

وحيث تم توليد قيم المتغير المفسر X من توزيع منتظم في الفترة (صفر, ١). وتم توليد الأخطاء $\varepsilon_i's$ من ثمانية توزيعات متماثلة وتوزيع واحد غير متماثل. والثمانية توزيعات المتماثلة مرتبة حسب درجة كثافة الأطراف والتوزيع غير المتماثل كانت كالتالي:

١- التوزيع المنتظم uniform(0,1)

٢- التوزيع الطبيعي N(0,1)

٣- التوزيع اللوجيستي logistic(0,1)

٤- التوزيع الأسي المضاعف double exponential (0,1)

٥- توزيع "contaminated normal I" $0.95N(0,1) + 0.05N(0,10)$

٦- توزيع "contaminated normal II" $0.90N(0,1) + 0.1N(0,10)$

٧- توزيع سلاش معرف كالتالي $N(0,1)/uniform(0,1)$

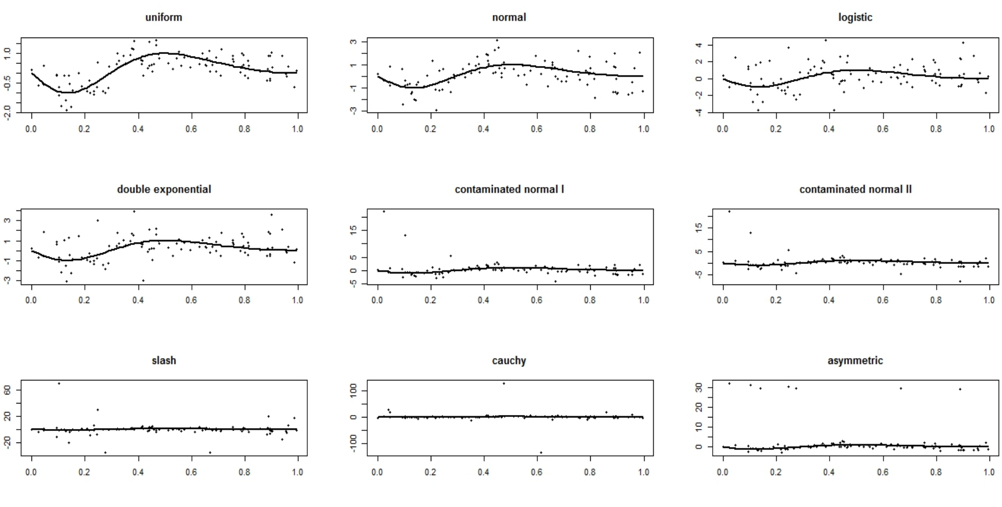
٨- توزيع كوشى Cauchy(0,1)

٩- التوزيع غير المتماثل $0.90N(0,1) + 0.1N(30,1)$

ونقلاً عن Lee and Oh (2007) , يمكن استخدام مقياس لدرجة كثافة أطراف أي توزيع كما عرفه Hoaglin et al. (1983) كالتالي:

$$\tau(F) = \frac{F^{-1}(0.99) - F^{-1}(0.5)}{F^{-1}(0.75) - F^{-1}(0.5)} \Big/ \frac{\Phi^{-1}(0.99) - \Phi^{-1}(0.5)}{\Phi^{-1}(0.75) - \Phi^{-1}(0.5)}$$

حيث يشير الرمز Φ إلى دالة التوزيع التجميعي الخاصة بالتوزيع الطبيعي المعياري "cumulative distributionfunction of standard normal" , ويشير الرمز F إلى دالة التوزيع التجميعي للتوزيع المراد قياس درجة كثافة أطرافه . ويمكن النظر إلى شكل (١) التالي الذي يعرض أشكال انتشار البيانات المولدة في دائرة معاودة واحدة من دراسة المحاكاة التي تم تنفيذها في البحث الحالي مرفق بها منحنى الدالة الحقيقية المراد تقديرها.



شكل (١): اشكال انتشار البيانات المولدة في أحد دوائر المعاودة في دراسة المحاكاة مرفق بها منحنى الدالة الحقيقية (الخط المتصل).

ولقد تم استخدام أربعة مقدرات لنموذج انحدار الشرائح المعاقبة في دراسة المحاكاة الحالية وهي: مقدر المربعات الصغرى LS, ومقدر M , ومقدر S ذا نقطة الانكسار ٥٠% وسيشار له بالرمز S_50%, ومقدر S ذا نقطة الانكسار ٣٠% والذي سيشار له بالرمز S_30%.

ولقد أعتمد أساس المقارنة بين طرق التقدير الأربع التي تم المقارنة بينها في دراسة المحاكاة على تنفيذ الخطوات التالية:

١- ايجاد متوسط مربعات الخطأ ASE المناظر لطرق التقدير الأربع في كل حالات دراسة المحاكاة ولكل دائرة معاودة والذي يتم حسابه كالتالي:

$$ASE_j = \frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - \hat{f}_j(x_i)\right)^2 \quad j = 1,2,\ldots\ldots,J$$

حيث يشير الرمز $f(x_i)$ إلى قيمة الدالة الحقيقية عند x_i , كما يشير الرمز n إلى حجم العينة المستخدم وهو ١٠٠ في دراسة المحاكاة الحالية, بينما يشير الرمز J إلى عدد مرات المعاودة "iterations" المستخدمة في دراسة المحاكاة وهي ٥٠٠ في دراسة المحاكاة الحالية.

٢- إيجاد الوسيط "median" والانحراف المطلق عن الوسيط "medianabsolute deviation" لمتوسط مربعات الخطأ ASE المناظر لطرق التقدير المختلفة في كل حالات دراسة المحاكاة. حيث تعد الطريقة الأفضل هي المناظرة لأقل وسيط وأقل انحراف مطلق عن الوسيط.

٣- رسم الصناديق البيانية للوغاريتم متوسط مربعات الخطأ $\log(ASE)$ الناتجة من كل مرات المعاودة وذلك لكل طرق التقدير التي يتم المقارنة بينها . حيث تعد الطريقة الأفضل هي المناظرة لأقل وسيط لوغاريتم متوسط مربعات الخطأ "$\text{median}(\log(ASE))$" .

٤- استخدام اختبار ويلكوكسون "Wilcoxon test" للمقارنة المزدوجة بين وسيطي $\log(ASE)$ المناظر لكل طريقتين من طرق التقدير لمعرفة ما إذا كان هناك فرق معنوي بينهما ام لا وهو ما استخدمه Wand (2000) في دراسة المحاكاة الخاصة به. فإذا لم يوجد فرق معنوي بين طريقتين من الطرق فيتم إعطاء كلاً منهما نفس الرتبة واذا كان هناك فرق معنوي فيتم اعطائهما رتباً مختلفة. ويعد الأسلوب الأفضل هو المناظر لأقل رتبة. ولقد تم اضافة نتائج استخدام اختبار ويلكوكسون اسفل الصناديق البيانية .

٥- رسم الخطوط البيانية لوسيط لوغاريتم متوسط مربعات الخطأ "$\text{median}(\log(ASE))$" المناظر لطرق التقدير الأربع ولكل التوزيعات التي اشتملت عليها دراسة المحاكاة. وذلك لاعطاء ملخص عام لاداء المقدرات في تلك الدراسة, حيث تعد الطريقة الأفضل هي المناظرة لأقل $\text{median}(\log(ASE))$.

ويلخص جدول (٢) والشكلين (٢, ٣) نتائج دراسة المحاكاة التي تم الحصول عليها والذي يتضح منها ما يلي:

١- أن أداء مقدر S_30 كان دائما أفضل من أداء مقدر S_50 في جميع حالات الدراسة (انظر جدول (٢)), كما كان ترتيب مقدر *S_30* بالنسبة للتوزيعات (المنتظم, الطبيعي, اللوجيستي, الأسي المضاف, contaminated normal I, contaminated normal II, سلاش, كوشي, غير المتماثل) هي على الترتيب

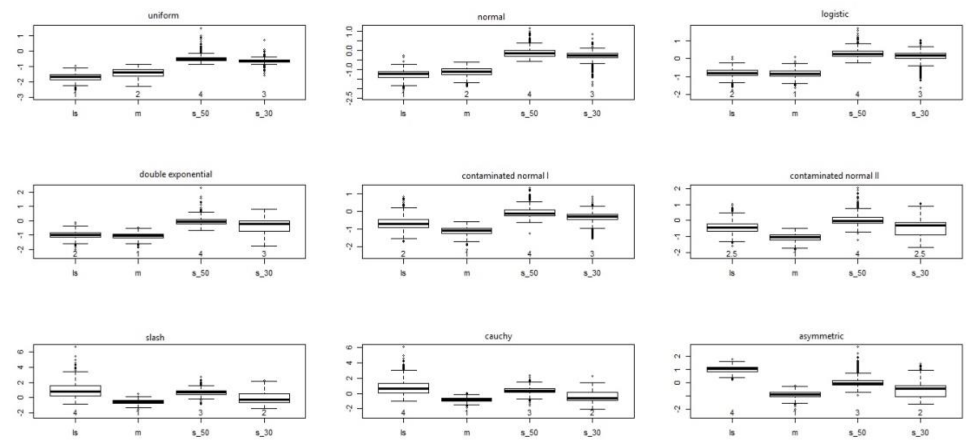
(٣, ٣, ٣, ٣, ٣, ٢,٥, ٢, ٢ , ٢) بينما كان ترتيب مقدر S_50 هي على الترتيب (٤, ٤, ٤, ٤, ٤, ٣, ٣, ٣, ٣) (انظرشكل (٢)) , كما كان الخط البياني المناظر للمقدر الأول أسفل الخط البياني للمقدر الثاني عند جميع التوزيعات (انظر شكل (٣)).

٢- أن مقدر LS كان هو الأفضل بالمقارنة بالمقدرات الثلاثة الأخرى في حالتين فقط هما حالة التوزيع المنتظم وحالة التوزيع الطبيعي حيث كان ترتيبه في الحالتين هو الأول في شكل (٢) وكان الخط البياني المناظر لمقدر *LS* اسفل الخطوط البيانية المناظرة للمقدرات الأخرى في الحالتين وهو ما كان متوقعاً نظرا لان في تلك الحالتين تحديداً لا يتطلب الأمر استخدام مقدر متين.

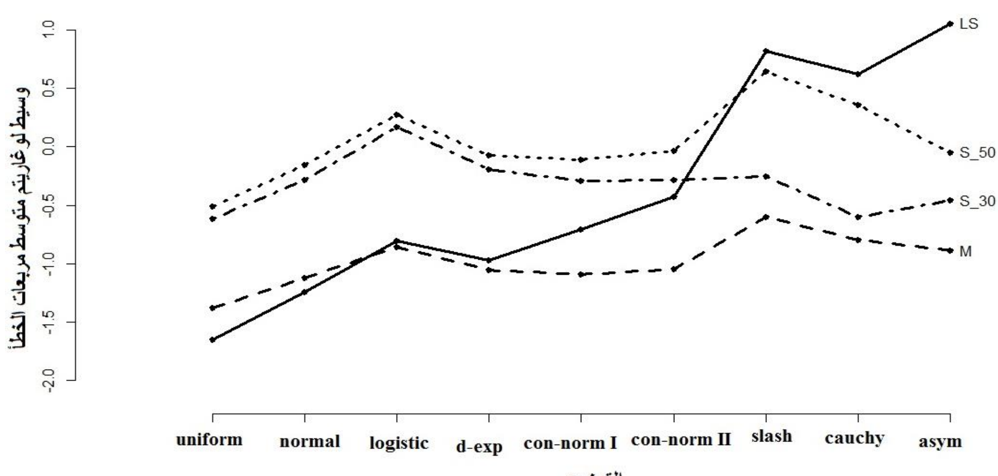
٣- أن مقدر M كان هو الأفضل بالمقارنة بالمقدرات الثلاثة الأخرى في سبع حالات وهي المناظرة للتوزيعات (اللوجيستي, الأسي المضاف, contaminated normal I, contaminated normal II, سلاش, كوشي, غير المتماثل) وهي توزيعات لها درجة كثافة أطراف أكبر من الواحد Lee and Oh (2007)حيث كان ترتيبها هو الأول في شكل (٢) بالنسبة لتك التوزيعات وكان الخط البياني المناظر لمقدر M اسفل جميع الخطوط المناظرة للمقدرات الأخرى بالنسبة لتلك التوزيعات والذي يعني أن مقدر M هو أفضل مقدر متين في حدود دراسة المحاكاة.

جدول (٢): الوسيط والانحراف المطلق عن الوسيط (بين قوسين) لمتوسط مربعات الخطأ ASE الناتجة من تطبيق طرق التقدير (LS, M, S_50%, S_30%)
على تسعة توزيعات مختلفة للخطأ

distribution	LS	M	S_50%	S_30%
uniform	0.02	0.04	0.31	0.24
	(0.01)	(0.03)	(0.10)	(0.05)
normal	0.06	0.08	0.70	0.52
	(0.03)	(0.04)	(0.31)	(0.18)
logistic	0.16	0.14	1.88	1.49
	(0.07)	(0.06)	(0.87)	(0.73)
d-exponantial	0.11	0.09	0.85	0.64
	(0.06)	(0.04)	(0.47)	(0.49)
Contaminated Normal I	0.19	0.08	0.77	0.51
	(0.14)	(0.04)	(0.39)	(0.26)
Contaminated Normal II	0.37	0.09	0.92	0.52
	(0.28)	(0.05)	(0.48)	(0.48)
Slash	6.57	0.25	4.43	0.56
	(8.53)	(0.16)	(2.90)	(0.67)
Cauchy	4.18	0.16	2.30	0.25
	(5.40)	(0.09)	(1.82)	(0.26)
Asymmetric	11.42	0.13	0.90	0.35
	(6.41)	(0.07)	(0.49)	(0.37)



شكل (٢): الصناديق البيانية للوغاريتم متوسط مربعات الخطأ المناظرة لاستخدام تسعة توزيعات للخطأ عند تطبيق طرق التقدير الأربع (LS و M و S_50% و S_30%). وموضح أسفل الصناديق الرتب التي أخذتها طرق التقدير الأربع وفقاً لاختبار ويلكوكسون.



شكل (٣): الخطوط البيانية لوسيط متوسط مربعات الخطأ الناتجة من تطبيق طرق التقدير (LS, M, S_30, S_50) على تسعة توزيعات مختلفة للخطأ.

٦- خلاصة البحث

أظهرت نتائج دراسة المحاكاة التي تم القيام بها في البحث الحالي أن استخدام مقدر *S* لنموذج انحدار الشرائح المعاقبة ذا نقطة الانكسار ٣٠% كان أفضل من مقدر *S* ذا نقطة انكسار ٥٠% في جميع حالات الدراسة. وحيث أن ثلاثة مقدرات من بين الأربعة مقدرات التي تم المقارنة بينهم تتطلب معاودات للوصول لنتائجها وهم (مقدر *M* ومقدر *S_30* ومقدر *S_50*) والذي من شأنه زيادة الوقت اللازم لتنفيذ دراسة المحاكاة حيث استغرق تنفيذ دائرة معاودة واحدة فقط حوالي ١٩ دقيقة على جهاز " Intel(R) Core (TM) i7-4700MQ CPU @2.40 GHz", والذي بدوره أدى إلى أن يكون الوقت اللازم للحصول على نتائج ٥٠٠ دائرة معاودة هو ٦ أيام و١٢ ساعة تقريباً في تلك الحالة. والذي قيد الباحثة في استخدام مقدرين فقط من مقدرات *S* ذوي نقاط انكسار مختلفة. ولذلك تنصح الباحثة مستخدموا مقدر *S* لنموذج انحدار الشرائح المعاقبة بأن يقوموا بتجريب عدد من مقدرات *S* ذوات نقاط انكسار مختلفة ثم النظر إلى شكل انتشار البيانات الذي يحتوي على المنحنى الموفق باستخدام المقدر في كل حالة لتحديد المقدر الأفضل الذي يستطيع أن يعبر عن أغلبية البيانات.

References

Beaton, A. E., and Tukey, J. W. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. Technometrics, vol. 16, no. 2, pp.147-185.

Cantoni,E., and Ronchetti, E. (2001). Resistant selection of the smoothing parameter for smoothing splines. Statistics and Computing, vol. 11, no. 2, pp. 141-146.

Eilers, P. H. C., and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties (with discussion). Statistical Science, vol. 11, pp. 89-121.

Fox, J. (2002). An R and S-Plus companion to applied regression. Sage.

Greiner, A. (2009). Estimating penalized spline regressions: Theory and application to economics. Applied Economics Letters,vol.16,no.18, pp. 1831-1835.

Griggs, W. (2013). Penalized spline regression and its applications. Available at http://www.whitman.edu/Documents/Academics/Mathematics/Griggs.pdf.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2001). The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer.

Hoaglin, D. C., Mosteller, F. and Tukey, J. W. (1983). Understanding robust and exploratory data anlysis. New York :Wiley.

Huber, P. J. (1973). Robust regression: asymptotics, conjectures and Monte Carlo. The Annals of Statistics, pp. 799-821.

Lee, T. C. M., and Oh, H. (2007). Robust penalized regression spline fitting with application to additive mixed modeling. Computational Statistics, vol. 22, no. 1, pp. 159-171.

Ruppert, D. (2002). Selecting the Number of Knots for Penalized Splines. Journal of Computational and Graphical Statistics, vol. **11**, no. 4, pp.735-757.

Ruppert, D., Wand, M., and Carroll, R. (2009). Semiparametric regression. UK :Cambridge.

Rousseeuw, P. J., and Leroy, A. M. (1987). Robust regression and outlier detection. New York: Wiley.

Rousseeuw, P., and Yohai, V. (1984). Robust regression by means of S-estimators. In Robust and nonlinear time series analysis . New York :Springer. pp. 256-272.

Salibian-Barrera, M., and Yohai, V. J., (2006). A fast algorithm for S-regression estimates. Journal of Computational and Graphical Statistics, vol. 15, no. 2, pp.414-427.

Tharmaratnam, K., Claeskens, G., Croux, C., and Salibian-Barrera, M. (2010). S-estimation for penalized regression splines. Journal of Computational and Graphical Statistics, vol. 19, no. 3, pp. 609-625.

Wand, M.P., (2000). A comparison of regression spline smoothing procedures. Computational Statistics, vol. 15, no. 4, pp.443-462.

Wang, B., Shi, W., and Miao, Z. (2014). Comparative analysis for robust penalized spline smoothing methods. Mathematical Problems in Engineering. Available at http://dx.doi.org/10.1155/2014/642475.