مقارنة بين طرائق انحدار الحرف ونوع ليو في تقدير معلمات أنموذج انحدار ثنائي الحدين السالب في ظل وجود مشكلة التعدد الخطي باستخدام المحاكاة

أ.م. سهيل نجم عبود / جامعة بغداد / كلية الادارة والاقتصاد / قسم الاحصاء

الباحث/ ايناس صلاح خورشيد / جامعة بغداد / كلية الادارة والاقتصاد / قسم الاحصاء

تاريخ التقديم:2018/5/14

تاريخ القبول:2018/6/24

الخلاصة

ان مشكلة التعدد الخطي من المشاكل الشائعة والتي تتعامل الى حد كبير مع الارتباط الداخلي بين المتغيرات التوضيحية وتظهر هذه المشكلة خصوصا في الاقتصاد والبحوث التطبيقية، ويكون لمشكلة التعدد الخطي تاثير سلبي على أنموذج الانحدار مثل وجود درجة تباين متضخم وتقدير معلمات تكون غير مستقرة عندما نستخدم مقدرات المربعات الصغرى الاعتيادية (OLS) ، لهذا تم اللجوء الى استخدام طرائق اخرى لتقدير معلمات أنموذج ثنائي الحدين السالب منها طريقة مقدر انحدار الحرف ومقدر نوع ليو، ويعتبر أنموذج انحدار ثنائي الحدين السالب (Negative Binomial Regression Model) كأنموذج انحدار غير خطي او كجزء من العائلة الاسية المعممة و هذا ألانموذج الهيكل الاساسي لتحليل بيانات العد (Count Data) و الذي استخدم كبديل لنموذج بواسون عندما تكون هناك مشكلة فوق التشتت (Overdisperison) اي عندما تكون قيمة تباين متغير الاستجابة (Y) اكبر من وسطه الحسابي ، وتم تصميم دراسة محاكاة مونت كارلوا للمقارنة بين طريقتي تقدير انحدار الحرف (Ridge Regression Estimator) ومقدر نوع ليو (Liu Type Estimator) من خلال استخدام معيار مقارنة متوسط مربعات الخطأ (MSE)، حيث بينت نتيجة المحاكاة ان طريقة مقدر نوع ليو هي افضل من طريقة مقدر انحدار الحرف اذ جاءت متوسط مربعات الخطأ لها اقل في صيغته التقديرية الثالثة والرابعة .

المصطلحات الرئيسة للبحث / التعدد الخطي ، أنموذج انحدار ثنائي الحدين السالب ، مقدر انحدار الحرف ، مقدر نوع ليو .

المبحث الاول/ المقدمة العامة

1-1 المقدمة

تحليل الانحدار أو تحليل الارتباط هو طريقة إحصائية يتم فيها التنبؤ بمتوسط متغير عشوائي أو عدة متغيرات عشوائية اعتماداً على قيم وقياسات متغيرات عشوائية أخرى، وله عدة أنواع مثل الانحدار الخطي، الانحدار اللوجستي، انحدار بواسون، والانحدار ثنائي الحدين السالب وغيرها. ان تحليل الانحدار الخطي هو الشكل الأبسط لأنموذج الانحدار اذ يحوي على متغير تابع إضافة إلى متغيرات مستقلة، ويلاحظ من ذلك أن أنموذج الانحدار يعتمد دائماً على علاقة السببية بمعنى ان يكون التغير في المتغير المستقل مسبب رئيسي للتغير في المتغير التابع. وتعتبر طريقة المربعات الصغرى OLS واحدة من الطرق المهمة في تقدير معلمات أنموذج الانحدار الخطي المتعدد لما تتصف به هذه الطريقة من مواصفات تميزها عن طرق التقدير الاخرى المعروفة، اذ تمتاز بعدم تحيزها وأنها تمتلك اقل تباين ممكن، وتتصف بخاصية الـــ (Best Linear Unbiased Estimate (BLUE)). والمعروف ان اداء مقدرات المربعات الصغرى الاعتيادية (OLS) تكون غير مرضية في ظل وجود مشكلة التعدد الخطي، وبهذا تكون لها اخطاء معيارية كبيرة، لذا تم اللجوء الى استخدام طرائق اخرى للتقدير، في هذا البحث تطرقنا الى طرائق تقدير معلمات أنموذج انحدار ثنائي الحدين السالب، وهي طريقة مقدر انحدار الحرف (Ridge Regression Estimator) ومقدر نوع ليو (Liu Type Estimator).

2-1 مشكلة البحث

يعتبر الانحدار الخطي المتعدد من التقنيات الاحصائية الأكثر استخداما بين الباحثين في مختلف المجالات، وكثيرا ما يواجه الباحثين مشكلة التعدد الخطي (Multicollinearity) عند بناء أنموذج الانحدار الخطي المتعدد وذلك من خلال وجود علاقة ارتباط تامة بين متغيرين تفسيريين أو اكثر من المتغيرات المضمنة في الأنموذج، بحيث يصبح محدد مصفوفة المعلومات X'X مساوي للصفر اذ يستحيل ايجاد معكوس المصفوفة X'X، وبالتالي عدم امكانية استخدام طريقة المربعات الصغرى الاعتيادية OLS، أو أن يكون محدد المصفوفة X'X قريبا من الصفر في حالة الارتباط شبه التام، الذي معه تضمحل قدرة OLS على عكس الخصائص الحقيقية لمعلمات الأنموذج ويكون الأنموذج ذا قدرة تنبؤيه ضعيفة. ولتخطي مشكلة التعدد الخطي تم اقتراح كثير من الحلول منها طريقة انحدار الحرف (Ridge Regression Estimator) وطريقة نوع ليو(Liu Type Estimator). اذ تكمن مشكلة البحث في تقدير معلمات أنموذج انحدار ثنائي الحدين السالب في ظل وجود مشكلة التعدد الخطي.

3-1 هدف البحث

يهدف هذا البحث الى دراسة مشكلة التعدد الخطي وطرق تشخيصها وتأثيراتها على أنموذج انحدار ثنائي الحدين السالب (Negative Binomial Regression Model)، ومقارنة أداء طرق تقدير معلمات أنموذج ثنائي الحدين السالب وهي طرق انحدار الحرف (RR) ومقدر نوع ليو (LT).

المبحث الثاني/ الجانب النظري

1-2 المقدمة

في هذا المبحث سيتم عرض ودراسة أنموذج انحدار ثنائي الحدين السالب من حيث الصيغة العامة، والتطرق الى دراسة مشكلة التعدد الخطي من حيث نشأتها وتأثيراتها وكيفية الكشف عنها و دراسة طرائق تقدير معلمات أنموذج الأنحدار ثنائي الحدين السالب في ظل وجود التعدد الخطي فضلا عن طريقتي التقدير بوجود هذه مشكلة.

2-2 أنموذج ثنائي الحدين السالب

هناك عدة توزيعات احتمالية متقطعة منها توزيع برنولي (Bernoulli distribution)، توزيع ثنائي الحدين (Binomial distribution)، توزيع ثنائي الحدين السالب (Negative Binomial distribution)، توزيع كثير الحدود، توزيع الهيبروجيومتريك (Hypergeometric distribution)، وتوزيع بواسون (Poisson distribution) تستخدم في كثير من التطبيقات الاقتصادية والتجارية وكثيرا ما نكون معنيين بدراسة الاختبارات التي يكون لها نتيجتين ممكنتين فقط وتدعيان عادة بالنجاح أو الفشل، مثل هذه الاختبارات تعرف بالاختبارات ثنائية الحدين، ولكننا سندرس توزيع ثنائي الحدين السالب نظرا لتطبيقاتها النادرة الاستخدام في التجارة والاقتصاد والصحة والذي يسمى في بعض الاحيان بـ (توزيع باسكال) نسبة للعالم الرياضي الفرنسي (Blaise Pascal 1962)، ويعتبر هذا التوزيع واحدا من التوزيعات المتقطعة ذات الاهمية التطبيقية في الكثير من المجالات العملية في العلوم الصحية والزراعية وعلم البكتريا [(i)]، ان أنموذج ثنائي الحدين السالب (Negative Binomial Regression Model) نادر الاستخدام في البحوث التطبيقية عندما يكون المتغير التابع y_i في شكل الاعداد الصحيحة غير سالبة، وهذا الأنموذج مميز لانه يتمكن من التعامل مع البيانات التي تعاني من مشكلة فوق التشتت (Over dispression) عندما يتجاوز قيمة التباين لمتغير الاستجابة وسطه الحسابي ويمكن اعتباره تعميم لأنموذج انحدار بواسون، اذ يتم الحصول على نماذج انحدار ثنائي الحدين السالب بنفس طريقة نماذج انحدار بواسون عن طريق ربط المتوسط μ بالمتجه من المتغيرات التوضيحية X ويمكن كتابتها بالشكل الاتي[(xi)(viii)]:

$$\mu = \exp^{(x_i' \beta)} \qquad (2-1)$$

اذ ان:

x_i': مصفوفة المتغيرات التوضيحية ذات الدرجة (*(nx(p+1)*).

β : موجه معلمات ذو الدرجة (1×(p+1)).

والمتوسط المشروط والتباين المشروط للتوزيع تعطي كالاتي[(ii)]:

$$E(Y_i/X_i) = \mu_i \qquad (2-2)$$

$$Var - Cov(Y_i/X_i) = \mu_i(1 + \theta\mu_i) \qquad (2-3)$$

لنحصل على دالة الكثافة الاحتمالية لتوزيع ثنائي الحدين السالب بمعلمتين (μ_i, θ) بافتراض أن $(\theta = \frac{1}{\delta})$ والتي تعطى بالصيغة التالية [(vi)]:

$$f(y_i, \mu_i, \theta) = \frac{\Gamma(y_i + \theta^{-1})}{\Gamma(y_i + 1)\Gamma\theta^{-1}} \left(\frac{\theta^{-1}}{\mu_i + \theta^{-1}}\right)^{\theta^{-1}} \left(\frac{\mu_i}{\mu_i + \theta^{-1}}\right)^{y_i} \qquad (2-4)$$

Γ :- دالة كاما هي دالة معرفة عند جميع الأعداد المركبة باستثناء الأعداد الصحيحة السالبة. فللعدد z الذي يتكون من جزء حقيقي موجب تعرف دالة كاما .

3 الصيغة العامة لأنموذج انحدار ثنائي الحدين السالب

General Form *of* Negative Binomial Regression Model

يمكن التعبير عن أنموذج انحدار ثنائي الحدين السالب بالصيغة التالية [vii] :

$$\underline{Y} = e^{X\underline{\beta}+\underline{U}} \qquad (2-5)$$

إذ أن:

Y : موجه متغير الاستجابة ذي درجة **(nx1)**.

X: مصفوفة المتغيرات التوضيحية ذات الدرجة **(n x (p+1))**.

β: موجه المعلمات ذو الدرجة **((P+1) x 1)**.

U: موجه الاخطاء العشوائية **(nx1)**.

4-2 مشكلة التعدد الخطي

ان مشكلة التعدد الخطي تظهر فقط عندما تكون هناك علاقة خطية بين بعض أو جميع المتغيرات التوضيحية، وان الارتباطات بين هذه المتغيرات تعرف بالتعدد الخطي، واحد الشروط الواجب توفرها في أنموذج الانحدار هو شرط الرتبة **(Rank Condition)**، أي أن:

$$\text{rank}(x)=m \qquad (2\text{-}6)$$

اذ ان:

X : مصفوفة من مرتبة **(n×m)** لمشاهدات المتغيرات التوضيحية .

وعليه في حالة كون المتغيرات التوضيحية **(Explanatory Variables)** مستقلة خطيا **(linearly independent)** فانه يمكن إيجاد معكوس مصفوفة المعلومات $(X'X)$ ويمكن ايضاً إيجاد مقدرات طريقة المربعات الصغرى الاعتيادية **((Ordinary Least Square (OLS))**، أما إذا كانت هناك علاقة خطية تامة بين اثنين أو أكثر من المتغيرات التوضيحية فان ذلك سيؤدي إلى انتهاك شرط الرتبة، أي أن[ii] :

$$\text{rank}(x)<m \qquad (2\text{-}7)$$

وعليه فانه لا يمكن إيجاد معكوس مصفوفة المعلومات $(X'X)$ وبالتالي لا توجد إمكانية لتقدير معلمات الأنموذج وهذه الحالة تسمى بالتعدد الخطي التام **(Perfect Multicollinearity)**، ولغرض معالجة هذه المشكلة لابد من حذف المتغيرات التوضيحية المسببة للتعدد الخطي، ومن ثم تقدير معلمات الأنموذج، أما إذا كان محدد مصفوفة المعلومات لا يساوي الصفر وإنما قريب منه وتظهر هذه الحالة عندما تميل المتغيرات للتحرك سوية بالزيادة او النقصان او في حالة استخدام المتغيرات المرتدة زمنيا **(Lagged variables)** ، ففي هذه الحالة يمكن تقدير معلمات الأنموذج ولكن هذه التقديرات سوف تكون غير دقيقة وغير ممثلة لواقع المشكلة المدروسة ، وان تباين المعلمات المقدرة ستكون كبيرة جدا وبالتالي سيظهر اختبار **t** عدم معنوية معلمات الأنموذج وهذه الحالة تسمى بالتعدد الخطي شبه التام [iii].

5-2 اختبار وجود مشكلة التعدد الخطي:

ان من اهم الاختبارات للكشف عن مشكلة التعدد الخطي هو اختبار فراير وكلوبر **(Farrar- Glauber)** عام **(1967)** ويستند هذا الاختبار على احصاءة مربع-كاي (x^2) حيث يتم اختبار فرضية العدم التالي:

$$H_o: (X_j)\, Orthogonal$$

مقابل الفرضية البديلة:

$H_1:(X_j) Not\ Orthogonl$

اما صيغة الاختبار فتأخذ الشكل التالي:

$$x_0^2 = -\left[n - 1 - \frac{1}{6}(2p+5)\right]\ln|R| \qquad (2-8)$$

n: حجم العينة .
p: عدد المتغيرات المستقلة .
ln|R| : اللوغاريتم الطبيعي لمحدد مصفوفة معاملات الارتباط.

ثم نقارن قيمة مربع- كاي المحسوبة (x_o^2) مع القيمة النظرية (الجدولية) بدرجة حرية مساوية الى ($\frac{k\times(k-1)}{2}$) ومستوى معنوية معين، فاذا كانت القيمة المحسوبة اكبر من القيمة الجدولية ترفض فرضية العدم H_o وتقبل الفرضية البديلة H_1، اي ان هناك مشكلة التعدد الخطي بين المتغيرات المستقلة والعكس صحيح، وبموجب هذا الاختبار يمكن اثبات وجود مشكلة التعدد الخطي بين المتغيرات التوضيحية [ii].

ومن اثار مشكلة التعدد الخطي[iii]:

1. القيم المقدرة لمعاملات الانحدار سوف تكون غير محددة وغير دقيقة.
2. الأخطاء المعيارية للقيم المقدرة لمعاملات الانحدار سوف تكون كبيرة جدا.

2-6 تقدير معلمات أنموذج انحدار ثنائي الحدين السالب بافتراض وجود مشكلة التعدد الخطي

Estimation of the Parameters for Negative Binomial Regression Model under Multicollinearity Problem

هناك عدة طرق لتقدير معلمات أنموذج انحدار ثنائي الحدين السالب ومنها:

2- 6 -1- طريقة مقدر انحدار الحرف Ridge Estimator Method Regression

لقد تم توضيح طريقة انحدار الحرف المتحيزة في الكثير من البحوث، وقد أظهرت هذه الطريقة فعالية في التخلص من مشكلة تعدد العلاقة الخطية، اذ إن مشكلة التعدد الخطي تؤدي إلى كبر حجم تباين المقدرات في اختلاف العلاقات النسبية بين المتغيرات التنبوية ومتغير الاستجابة باستخدام طريقة المربعات الصغرى (OLS). وتم التطرق الى طريقة انحدار الحرف لأول مرة من قبل (Hoerl & Kennard) عام (1970) اذ ثبت في الوقت الحاضر بانها الطريقة الاكثر كفاءة لبيان كيفية التعامل مع المشاكل العامة التي تسببها مشكلة تعدد الخطي. وتتلخص هذه طريقة بإضافة كمية صغيرة موجبة الى عناصر قطر مصفوفة المعلومات ($X'X$) حسب اقتراح الباحثان (Hoerl & Kennard)، ويعتبر اسلوب انحدار الحرف احد بدائل طرق التقدير عندما يكون هناك تعدد الخطي بين المتغيرات التوضيحية لأنموذج الانحدار ثنائي الحدين السالب والنادر الاستخدام في البحوث التطبيقية عند تحليل البيانات الكمية، اذ يستخدم عادة مقدر الامكان الاعظم (Maximum likelihood) كونه حساس جدا للمتغيرات التوضيحية المترابطة، لذلك تم اقتراح مقدر انحدار الحرف لأنموذج ثنائي الحدين السالب كخيار قوي لتقدير معلمات الأنموذج في ظل وجود مشكلة التعدد الخطي، لهذا تم اختيار مقدر ليكن $\underline{B}_{RR}$ فيمكن كتابة مجموع مربعات الاخطاء الموزون على النحو التالي [xii]:

$$u'u = (\underline{y}-\hat{\beta})'(\underline{y}-\hat{\beta}) = (\underline{y}-X\underline{\beta}_{ML})'(\underline{y}-X\underline{\beta}_{ML}) + (\underline{\beta}_{RR}-\underline{\beta}_{ML})'\left(X'\widehat{W}\left(\underline{\beta}_{ML}\right)X\right)\left(\underline{\beta}_{RR}-\underline{\beta}_{ML}\right)$$
$$= \emptyset_{min} + \emptyset\left(\underline{\beta}_{RR}\right) \qquad (2-9)$$

$\emptyset_{min}$ تمثل الزيادة في متوسط مربعات الأخطاء الموزون في حال استبدال المعلمات المقدرة بطريقة الإمكان الأعظم $(\underline{\beta}_{ML})$ بالمعلمات المزمع ايجادها $(\underline{\beta}_{RR})$ ، يلاحظ زيادة متوسط مربعات الخطأ أي تضخم التباين لمجموع مربعات الاخطاء المؤزونة عندما استخدام طريقة الامكان الاعظم لتقدير معلمات الأنموذج $\underline{\beta}_{ML}$ لهذا تم استبدله بمقدر انحدار الحرف .

ووجد ان مقدر انحدار الحرف لأنموذج ثنائي الحدين السالب عن طريق تقليل طول $\hat{\beta}'\hat{\beta}$ من خلال تطبيق القيد ادناه(ii):

$$\emptyset\left(\underline{\beta}_{RR}\right) = \emptyset_0 \qquad (2-10)$$

وعليه فان صيغة التقدير للمعالم في أنموذج يمكن اشتقاقها باستخدام مضاعف لاكرانج **(Langrange Multiplier)** لغرض تصغير مجموع مربعات الاخطاء المؤزون وفقا للقيد اعلاه في الصيغة **(2-10)**(ix):-

$$Minimize\ F = \underline{\beta}'_{RR}\underline{\beta}_{RR} + (\frac{1}{K})(\underline{\beta}_{RR} - \underline{\beta}_{ML})'(X'\widehat{W}(\underline{\beta}_{ML})X)\left(\underline{\beta}_{RR} - \underline{\beta}_{ML} - \emptyset_0\right) \qquad (2-11)$$

اذ ان (***1/ k***) هو مضاعف لاكرانج وبذلك سيكون مجموع مربعات الأخطاء المؤزون وباشتقاق المعادلة **(2-9)** وباستخدام مضاعف لاكرانج نحصل على مقدر انحدار الحرف لأنموذج ثنائي الحدين السالب لنحصل على:

$$\underline{\hat{\beta}}_{RR} = (X'\widehat{W}X + KI)^{-1}\left(X'\widehat{W}X\right)\underline{\hat{\beta}}_{ML} = Z\underline{\hat{\beta}}_{ML} \qquad (2-12)$$

اذ ان :

$$Z = (X'\widehat{W}X + KI)^{-1}\left(X'\widehat{W}X\right) \qquad (2-13)$$

I **:** هو مصفوفة احادية ذي الدرجة ***(p+1) x (p+1)*** .

K **:** هو ثابت غير سالب ، تسمى المتحيز او معلمة الحرف مع ملاحظة انه عندما ***K=0*** نحصل على تقديرات المربعات الصغرى الاعتيادية (***OLS***) ، اذا أن مقدرات الانحدار حرف متحيزة بتزايد قيمة ***K*** حيث انها تعطي تقديرات اكثر دقة من مقدرات ***OLS*** لموجه المعلمات . وبعد ان تم ايجاد مقدرات انحدارالحرف تبين ان تلك المقدرات متحيزة لكنها اكفأ من نظيرتها في طريقة الامكان الاعظم حيث ان تناقص معاملات القيم الموجبة الصغيرة من ***k*** تحسن تكييف للمشكلة وتقليل تباين لمقدرات على الرغم من التحيز.

كما أن مصفوفة التباين والتباين المشترك لمقدرات انحدار الحرف تكون كمايلي :

$$Var - Cov\left(\underline{\hat{\beta}}_{RR}\right) = Z\ Var - Cov\left(\underline{\hat{\beta}}_{ML}\right)\ Z'$$
$$= Z\ \sigma_u^2\ (X'\widehat{W}X)^{-1}\ Z' = \sigma_u^2\ Z\ (X'\widehat{W}X)^{-1}\ Z' \qquad \textbf{(2-14)}$$

اذ بين **(Mansson & Shukur 2011)** ان متوسط مربعات الخطأ **(MSE)** لمعلمات أنموذج الثنائي الحدين السالب المقدرة يساوي(ix):

$$MSE\left(\hat{\beta}_{RR}\right) = \sum_{j=1}^{J}\frac{\lambda_j}{(\lambda_j + K)^2} + K^2\sum_{j=1}^{J}\frac{\alpha_j^2}{(\lambda_j + K)^2}$$
$$= \gamma_1(K) + \gamma_2(K) \qquad (2-15)$$

حيث α_j تعرف كعنصر j_{th} من معلمات أنموذج انحدار ثنائي الحدين السالب عند اخذ اللوغاريتم للأنموذج المبين في الصيغة (2-5) تساوي مقدار $\gamma\underline{\hat{\beta}}_{ML}$ اي ان :-

$$\alpha_j = \gamma\underline{\hat{\beta}}_{ML} \qquad (2-16)$$

γ : تمثل المتجه المميز (*Eigen Vector*) للمصفوفة $X'\widehat{W}X$ المناظر للجذر المميز (□).

λ_j : هي القيمة المميزة للعنصر (j_{th}) للمصفوفة $(X'\widehat{W}X)$.

K : معلمة التحيز.

7-2 مقدرات معلمة الحرف Ridge Parameter Estimators

اقترح (Hoerl & Kennard) مقدر انحدار حرف كبديل لمقدر المربعات الصغرى الاعتيادية (*OLS*) في ظل وجود مشكلة التعدد الخطي، وان احد العقبات الرئيسية في استخدام انحدار الحرف هي اختيار قيمة مناسبة لمعلمة التحيز(*k*) وقد تم تطوير العديد من التقنيات لتقدير معلمة حرف *K* وتم استعراضها وتصنيفها الى اشكال مختلفة، منها :

1- تقدير (1970) Hoerl & Kennard [ix][xii] :

تعتبر هذه الصيغة اول واقدم المقترحات لتقدير معلمة الحرف *K* اذ اقترح الباحثان (*Hoerl & Kennard*) تقدير معلمة حرف مع اقل متوسط مربعات الخطأ (*MSE*)عن طريق تعظيم الثابت $\left(\alpha_i^2\right)$ اذ اقترح معلمة الحرف الامثل كالاتي :

$$\hat{k}_{HK}^{FM} = \frac{\hat{\sigma}^{\,2}}{Max(\hat{\alpha}_i^2)} \qquad i = 1,2,3,\ldots\ldots p \qquad (2-17)$$

اذ ان:

$\hat{\sigma}^2$: متوسط مربعات الخطأ (*MSE*) المقدر والذي يحسب في أنموذج انحدار ثنائي السالب وفق الصيغة ادناه:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-p-1} \qquad (2-18)$$

$\hat{\alpha}_{\max}^2$: تمثل اكبر عنصر في الصيغة (2-16).

$\sum_{i=1}^{n} e_i^2$: تمثل مجموع مربعات الاخطاء العشوائية.

n : تمثل حجم العينة.

p : تمثل عدد متغيرات مستقلة .

2- تقدير Hoerl واخرون عام (1975) .

اقترح تقدير مختلف لمعلمة التحيز K من خلال اخذ المتوسط التوافقي لمعلمة الحرف k_{HK} وصيغته كالاتي :

$$\hat{k}_{HK}^{HM} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{n} \alpha_i^2} \qquad (2-19)$$

3-تقدير مقدرات مستند على مقترح ((Lawless& Wang(1976):

في تحسين للمقترح الذي اقترحه وقدمه الباحثون (Herol واخرون عام (1975)) اقترح الباحثان (Lawless & Wang) مقدر مختلف لمعلمة (k) ناتجة من خلال اخذ الوسط التوافقي لمعلمة الحرف (k) معرفة كالاتي [x]:

$$\widehat{K}_{LW}^{HM} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p} \lambda_j \hat{\alpha}_j^2} \qquad (2-20)$$

اذ أن: λ_j : تمثل القيم المميزة (Eigen Values) للمصفوفة $X'\widehat{W}X$.

4- تقدير kibria عام (2003)

اقترح بعض المقدرات الجديدة لمعلمة التحيز (k) من خلال اخذ المتوسط الهندسي، المتوسط الحسابي، الوسيط لقيم $\hat{\alpha}_i^2$ $(P \geq 3)$ لمعلمة k_{HK} وهذه المقدرات على التوالي تعرف كالاتي:

$$\hat{k}_{HK}^{GM} = \frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p} \hat{\alpha}_i^2\right)^{\frac{1}{p}}} \qquad (2-21)$$

$$\hat{k}_{HK}^{AM} = \frac{1}{p}\sum_{i=1}^{p}\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\right) \qquad (2-22)$$

$$\hat{k}_{HK}^{M} = \text{Median}\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\right) \qquad (2-23)$$

5- تقدير (AL Khamisi واخرون) عام 2006:

ان ربط هذا المقدر بين القيم المميزة (Eigen Values) وتباين الاخطاء العشوائية فضلا عن الأخذ بالاعتبار أثر الموجهات المميزة (Eigen Vectors) عبر احتساب اكبر قيمة، وإن القيمة المثلى لمعلمة التحيز عند تحليل أنموذج انحدار ثنائي الحدين السالب هي :

$$\widehat{K}_{AKS_i}^{FM} = Max\{s_j\} \qquad (2-24)$$

إذ أن:

$$S_j = \frac{\lambda_j \hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + \lambda_j \hat{\alpha}_j^2} \qquad (2-25)$$

6-تقدير Muniz & Kibria في عام (2009).

اقترح بعض المقدرات لمعلمة k من الجذر التربيعي للمتوسط الهندسي لمعلمة k_{HK_i} ولمعكوسها، الجذر التربيعي لمقدر (Horel & Kennard) للوسيط ومعكوسها وهذه المقدرات تعرف على التوالي:

$$\hat{k}_{HK}^{GMSR} = \sqrt{\frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p} \hat{\alpha}_i^2\right)^{\frac{1}{p}}}} \qquad (2-26)$$

$$\hat{K}_{HK}^{GMRSR} = \frac{1}{\sqrt{\frac{\hat{\sigma}^2}{(\prod_{i=1}^{p} \hat{\alpha}_i^2)^{\frac{1}{p}}}}} \qquad (2-27)$$

$$\hat{K}_{HK}^{MSR} = Median\left(\sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}}\right) \qquad (2-28)$$

$$\hat{k}_{HK}^{MRSR} = Median\left(\frac{1}{\sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}}}\right) \qquad (2-29)$$

7-تقدير(Muniz & Kibria) عام 2009.

اقترح مقدر معلمة حرف K كمتوسط هندسي لمعلمة حرف K_{AKS_i} التي قدمها (AL Khamisi واخرون) مقدر معرف كالاتي :

$$\hat{K}_{AKS}^{GM} = (\prod_{j=1}^{p} S_j)^{\frac{1}{p}} \qquad (2-30)$$

8-2 **طريقة مقدر نوع ليو** (Liu Type Estimator Method)

يعرف أنموذج الانحدار الخطي معرف كالاتي $Y = XB + \varepsilon$ حيث X هو مصفوفة بيانات من الدرجة (nxp)، B هو متجه معاملات من الدرجة (px1)، ε متجه الخطأ العشوائي من الدرجة (nx1) يحقق $\varepsilon_i \sim N(0,\sigma^2)$، Y هو متغير معتمد من الدرجة (nx1)، وبوجود مشكلة التعدد الخطي تتأثر طريقة المربعات الصغرى اذ ان معاملات الانحدار تبقى غير متحيزة ولكن تبايناتها تكون كبيرة وبعيدة كل البعد عن قيمها الحقيقية، اذ أن مصفوفة $X'X$ تاتي (*ill-conditioned*) يعني هناك ارتباط بين المتغيرات التوضيحية وبعض القيم المميزة لمصفوفة $X'X$ تكون مغلقة للصفر، ومقدرات مربعات الصغرى الاعتيادية تكون غير مستقرة [vi]:

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y \qquad (2-31)$$

فيكون هناك طريق واحد الى قياس عامل تضخم بطريقة كمية هو عدد شرط $K=(\vartheta_{max} = \vartheta_{min})$ واذا كان عدد شرط كبيرة هذا يعني هناك الارتباط بين المتغيرات التوضيحية ، حيث ϑ_j هو القيم المميزة لمصفوفة $X'X$ وتناقص دالة K في مقدر انحدار الحرف :

$$\hat{\beta}_{Ridge} = (X'X + KI)^{-1}X'Y$$

لذلك يجب السيطرة على الشرط الى اقل مستوى ينبغي استخدام اكبر قيم لمعلمة K التي يفرض زيادة التحيز الى انحدار الحرف و يتم الحصول عليها من خلال دمج المعادلة $0 = K^{1/2}\beta + \varepsilon'$ مع المعادلة الاصلية للانحدار $Y = X\beta + \varepsilon$ ومن ثم استخدام طريقة مربعات الصغرى (*OLS*)، وباستعمال اكبر قيمة لمعلمة التحيز K يجعل مسافة كبيرة بين $K^{1/2}$ و *0* ولايزال هناك الارتباط بين المتغيرات ولذلك اقترح عالم ليو عام (2003) مقدر جديد من خلال الجمع بين مقدر انحدار حرف الاعتيادية مع اي مقدر اخر β واقترح استبدل الطرف الايسر من المعادلة $0 = K^{1/2}\beta + \varepsilon'$ بمقدار $(-d/k^{1/2})\hat{\beta}_{OLS}$ لنحصل على معادلة جديدة:

$$(-d/K^{1/2})\hat{\beta}_{OLS} = K^{1/2}\beta + \varepsilon'$$

سمي هذا مقدر $\hat{\beta}_{k,d}$ بمقدر نوع ليو (Liu Type Estimator) باعتباره تعميم لمقدر ليو التي قدمها العالم (Liu) عام (1993) اذ يتم اختيار اكبر قيمة لمعلمة انحدار الحرف K وذلك بتحديد معلمة مقدر ليو d لجعل معادلة تعطي تقدير جيد ودقيق وتعرف كالاتي[viii][vii]:

$$\hat{\beta}_{k,d} = (X'X + KI)^{-1}(X'X - dI)\hat{\beta} \qquad (2-32)$$

اذ *k>0* ، $-\infty < d < \infty$ و $\hat{\beta}$ اي مقدر.

ان مقدر نوع ليو $\hat{\beta}_{k,d}$ لها افضل اداء من طريقة مربعات الصغرى OLS ومقدر انحدار حرف من خلال معيار مقارنة متوسط المربعات الخطأ . ويعرف مقدر نوع ليو (LT) في أنموذج انحدار ثنائي الحدين السالب لحل مشكلة التعدد الخطي كالاتي:

$$\hat{\beta}_{LT} = (X'\widehat{W}X + KI)^{-1}\left(X'\widehat{W}X - dI\right)\hat{\beta} \qquad (2-33)$$

اذ ان مقدر نوع ليو LT مقدر عام يتضمن مقدر انحدار الحرف RR كالاتي:

$$\lim_{d\to 0}\hat{\beta}_{LT} = \hat{\beta}_k \qquad (2-34)$$

ومن اجل ان نرى تفوق مقدر نوع ليو (LT)، حيث ان متوسط مربعات الخطأ (MSE) تحتوي على معلومات ذات الصلة فيما يتعلق بالمقدرات، ويمكن استعمالها كمعيار مقارنة متوسط مربعات الخطأ MSE هي الاثر لأي مقدر $\tilde{\beta}$ وتعرف على التوالي كالاتي[iv]:

$$MSE(\tilde{\beta}) = tr\left(MMSE(\tilde{\beta})\right) = E[(\tilde{\beta} - \beta)'(\tilde{\beta} - \beta)]$$

$$MSE(LT) = \sum_{j=1}^{p+1}\frac{(\lambda_j - d)^2}{\lambda_j(\lambda_j + k)^2} + \frac{(d+k)^2\alpha_j^2}{(\lambda_j + k)^2} \qquad (2-35)$$
$$= f_1(k,d) + f_2(k,d)$$

حيث $f_1(k,d)$ تمثل دالة التباين و$f_2(\mathrm{k,d})$ تمثل مربع التحيز لمقدر $\hat{\beta}_{k,d}$ وبالتالي يتم اختيار قيم مناسبة لمعلمتين k وd من اجل الحصول على اقل قيمة لمتوسط مربعات الخطأ (MSE) مقارنة ببقية طرائق(v) .

المبحث الثالث/ الجانب التجريبي

3-1 المقدمة

يتضمن هذا المبحث استعمال اسلوب المحاكاة مونت– كارلو وصف مراحل المحاكاة ومقارنة طرائق تقدير معلمات أنموذج الانحدار ثنائي الحدين السالب (Negative Binomial Regression Model) والتوصل الى الطريقة الأفضل من خلال متوسط مربعات الخطأ (MSE) للأنموذج.

3-2 مفهوم المحاكاة

المحاكاة هو مصطلح أطلق على تلك البرمجيات أو نماذج التي يقوم الخبراء بتصميمها لتمثل أو لتقلد النظام الحقيقي سواء الموجود أو الذي ينوى إنشاؤه، وتهدف إلى معرفة والإلمام بجميع النتائج المتوقع الحصول عليها ويعتبر علما قائما بذاته، ولكنه يستند على مجموعة من العلوم الأخرى من أبرزها علم الرياضيات وعلم المعلومات وعلم الفيزياء وغيرها من العلوم . والتي تتم باستخدام برامج حاسوبية معينة، وتعتبر برمجية الماتلاب (Matlab Program) من أكثر البرمجيات شهرة، ومع تطور الحواسيب ازدادت المحاكاة الحاسوبية فعالية وإثارة في تدريس المفاهيم والمواضيع العلمية المختلفة وتنوعت لغات المحاكاة واستخداماتها في التدريس وهذا ما جعله أكثر مرونة وحيوية من ذي قبل، كما استخدمت المحاكاة في التقليل من الخسائر المادية والمعنوية وهذا ما جعلها من النشاطات الفاعلة والممتعة في إرساء أسس التعلم لبعض المهارات والمواضيع الصعبة التي يصعب التعامل معها دون مخاطر في الواقع، فهي تبسيط لبعض المواقف الحياتية أو لعملية ما يكون لكل فرد فيها دورا يتفاعل من خلالها مع الآخرين في ضوء عناصر الموقف المحاكي.

3-3 وصف مراحل تجربة المحاكاة : Descriptionad simulation stages experiment

بالاعتماد على انموذج الانحدار ثنائي الحدين السالب الذي تم ذكره سابقاً في المبحث الثاني من المعادلة (2-5) سوف يتم وصف مراحل تجربة المحاكاة من خلال كتابة برنامج بلغة البرمجة ماتلاب اذ يتم وصف مراحل تجربة المحاكاة من خلال الخطوات الاتية: تعيين القيم الافتراضية للمعالم وهذه المرحلة من اهم المراحل التي يعتمد عليها لاحقاً، اذ تم اختيار قيم المعلمات تم اختيار قيم معلمات الأنموذج حسب القيد ادناه :

$$\sum_{j=1}^{p} \beta_j^2 = 1 \qquad (3-1)$$

والقيم الافتراضية التي تم اخذها للمعلمات هي:

$(\beta_0 = 0.6, \beta_1 = 0.4, \beta_2 = 0.2, \beta_3 = 0.3, \beta_4 = 0.4, \beta_5 = 0.5)$

1-توليد قيم المتغيرات التوضيحية من خلال أسلوب مونت- كارلو في المحاكاة حيث يتم توليد ثلاث وخمسة متغيرات توضيحية وفق التوزيع الطبيعي القياسي بافتراض وجود مشكلة التعدد الخطي كما مبين بالمعادلة الاتية : -

$$X_{ij} = \rho Z_{ij} + (1-\rho^2)^{1/2} Z_{ip} \ .. \qquad (3-2)$$

اذ ان: ρ : معامل الارتباط البسيط بين المتغيرات التوضيحية .

Z : تمثل الأرقام العشوائية المولدة وفق التوزيع الطبيعي القياسي.

2-توليد قيم متغير الخطأ العشوائي في انموذج الانحدار ثنائي الحدين السالب .
4- حساب المتغير التابع y ثنائي الاستجابة حيث يتم توليد متغير معتمد لنموذج انحدار ثنائي الحدين السالب باستخدام الارقام عشوائية لتوزيع ثنائي الحدين السالب $NB(\mu_i, \mu_i + \theta\mu_i^2)$ وحسب أنموذج ثنائي الحدين السالب يتم حساب قيم متغير الأستجابة (Y_i) المذكور في الصيغة (2-1):

$$Y_i = \mu_i = EXP(X_i\, \beta)$$

5- ومن اهم العوامل الأخرى التي يتم اختيارها والمؤثرة هي احجام العينات، اذ ستؤخذ اربعة احجام من العينات (25, 50, 100, 250) اما العامل الاخر الذي سيتم اخذة بنظر الاعتبار معامل الارتباط البسيط ، اذ تم اختيار ثلاث قيم لاختبار مدى قوة العلاقة بين المتغيرات المستقلة وهي (0.90, 0.95, 0.99).
6- تقدير معالم انموذج الانحدار أنموذج ثنائي الحدين السالب وفق طرائق التقدير التي تم عرضها في المبحث الثاني وبعد ذلك يتم المقارنة بين طرائق التقدير من خلال معيار مقارنة متوسط مربعات الخطأ (MSE) للأنموذج.

r= 1,2,…....,1000 (3-3)

$$MSE = \frac{\sum_{r=1}^{R}(\hat{\beta}-\beta)'_r(\hat{\beta}-\beta)_r}{R}$$

4-3 **نتائج تجربة المحاكاة**: The Results of the Simulation Experiment

سيتم عرض نتائج تجربة المحاكاة وحسب القيم الافتراضية لمعلمات هذه أنموذج من خلال طرائق تقدير معلمات أنموذج الانحدار ثنائي الحدين السالب ومتوسط مربعات الخطأ لكل من طرق السبعة في وجود مشكلة التعدد الخطي شبه التام (Quasi Complete Multicollinearity)، وعبر التعويض المباشر للصيغ (2-17)، (2-19)، (2-20)، (2-21)، (2-22)، (2-23)، (2-24)، (2-26)، (2-27)، (2-28)، (2-29)، (2-30)، في المعادلة (2-12) لطريقة انحدار الحرف (Ridge Regression) وطريقة مقدر نوع ليو (Liu Type Estimator) من خلال المعادلة (2-332) اذ تعكس النتائج قيم متوسط مربعات الخطأ (MSE) لكافة طرائق التقدير السابقة.

5-3 **تحليل النتائج طرائق التقدير عندما يكون عدد المتغيرات المستقلة** P=3

من خلال الجدول (1) نلاحظ ما يلي:
1-عند حجم العينة (25) ولقيم (ρ=0.90, 0.95, 0.99) نلاحظ ان مقدر نوع ليو (LT) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{Max}, D_{Min})}$ و $LT_{(K_{Max}, D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ مقارنة مع طريقة انحدار الحرف، k6 احد مقدرات انحدار حرف تمتلك ايضا اقل متوسط مربعات الخطأ (MSE) مقارنة مع باقي صيغ تقدير لمقدر انحدار الحرف.
2-عند حجم العينة (50) ولقيم (ρ=0.90, 0.95, 0.99) نلاحظ ان مقدر نوع ليو(LT) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{Max}, D_{Min})}$ و $LT_{(K_{Max}, D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنة مع طريقة مقدر انحدار الحرف، ولقيمة (ρ=0.90) تمتلك K6 اقل متوسط مربعات الخطا مقارنة مع صيغ الاخرى لتقدير مقدرات انحدار الحرف.
3-عند حجم العينة (100) ولقيم (ρ=0.90, 0.95, 0.99) نلاحظ ان مقدر نوع ليو (LT) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{Max}, D_{Min})}$ و $LT_{(K_{Max}, D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ مقارنة مع طريقة انحدار الحرف ، ومقدر انحدار حرف (K_4) في صيغتها ثانية و لقيمة معمل الارتباط (ρ=0.90,0.95) تمتلك اقل متوسط مربعات الخطأ مقارنة ببقية صيغ .

4-عند حجم العينة (250) ولجميع قيم ρ نلاحظ ان طريقة مقدر نـوع ليـو (LT) في صيغته التقديريـة الثالثـة والرابعـة $LT_{(K_{Max},D_{Min})}$ و $LT_{(K_{Max},D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنـة مـع طريقة مقدر انحدار الحرف ومقدر انحدار حرف (K_4) تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنـة مـع صيغ التقدير الاخرى.

جدول (1) قيم متوسط مربعات الخطأ لكافة طرائق التقدير عندما يكون عدد المتغيرات المستقلة P=3

n	ρ	Ridge Estimator							Liu Type Estimator			
		K_1	K_2	K_3	K_4	K_5	K_6	K_7	$LT_{(K_{Min},D_{Min})}$	$LT_{(K_{Min},D_{Max})}$	$LT_{(K_{Max},D_{Min})}$	$LT_{(K_{Max},D_{Max})}$
25	0.90	0.131080	0.129104	0.0981971 0.131635 0.132666	0.111137 0.125175 0.132666 0.1326666	0.132462	0.094232	0.121176	0.132666	0.132666	0.092273	0.092273
	0.95	0.140239	0.138667	0.123921 0.140523 0.142574	0.126371 0.130431 0.142574 0.142575	0.142307	0.121994	0.132731	0.1425749	0.1425749	0.1193911	0.1193911
	0.99	0.162898	0.1607674	0.130775 0.163355 0.165737	0.13982397 0.15375174 0.16573755 0.165737	0.165334	0.140701	0.153537	0.165737	0.165737	0.129645	0.129645
50	0.90	0.053638	0.0527306	0.05038125 0.05285750 0.052908	0.05073988 0.05131605 0.05290878 0.052908	0.052904	0.050011	0.051861	0.053652	0.053652	0.049839	0.049839
	0.95	0.057455	0.056469	0.05381299 0.05660685 0.056665	0.05425988 0.05496041 0.05666497 0.056665	0.056659	0.054925	0.055580	0.057472	0.057472	0.0533398	0.0533398
	0.99	0.066224	0.065137	0.062956 0.065294 0.065383	0.062928 0.062872 0.065383 0.0653834	0.065373	0.063041	0.064283	0.066249	0.066249	0.062164	0.062164

تابع جدول (1)

n	ρ	Ridge Estimator							Liu Type Estimator			
		K_1	K_2	K_3	K_4	K_5	K_6	K_7	$LT_{(K_{Min},D_{Min})}$	$LT_{(K_{Min},D_{Max})}$	$LT_{(K_{Max},D_{Min})}$	$LT_{(K_{Max},D_{Max})}$
100	0.90	0.026978	0.026963	0.026650 0.026982 0.026989	0.026576 0.026376 0.026989 0.026987	0.026989	0.026741	0.026773	0.026989	0.026989	0.026154	0.026154
	0.95	0.029064	0.029049	0.0286835 0.0290684 0.029078	0.028610 0.028422 0.029078 0.029076	0.029077	0.028719	0.028846	0.029078	0.029078	0.028044	0.028044
	0.99	0.0338905	0.033872	0.0333855 0.0338953 0.033908	0.033307 0.033113 0.0339087 0.033909	0.033907	0.032913	0.033659	0.0339087	0.0339087	0.032554	0.032554
250	0.90	0.0105681	0.010568	0.0105115 0.0105687 0.0105695	0.010503 0.0104825 0.010569 0.010568	0.010569	0.0105658	0.0105531	0.0105694	0.0105694	0.0104826	0.0104826
	0.95	0.011337	0.011336	0.0113276 0.0113373 0.011339	0.0113087 0.0111167 0.0113384 0.011338	0.0113383	0.011299	0.0113285	0.0113384	0.0113384	0.0111255	0.0111255
	0.99	0.0131214	0.0131206	0.0131124 0.0131214 0.0131231	0.0130903 0.0128206 0.0131231 0.0131123	0.0131231	0.0131077	0.0131132	0.0131233	0.0131233	0.0128366	0.0128366

6-3 تحليل النتائج طرائق التقدير عندما يكون عدد المتغيرات المستقلة P=5

من خلال الجدول (2) نلاحظ ما يلي:

1-عند حجم العينة (25) نلاحظ ان مقدر نوع ليو (LT) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{Max},D_{Min})}$ و $LT_{(K_{Max},D_{Max})}$ تمتلك متوسط مربعات الخطأ (MSE) مقارنة بطريقة مقدر انحدار الحرف، واحد مقدرات انحدار حرف (K_6) تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنة ببقية الصيغ ولقيمتي معامل الارتباط (0.90, 0.95).

2-عند حجم العينة (50) نلاحظ ان مقدر نوع ليو (LT) في صيغته تقديرية الثالثة والرابعة $LT_{(K_{Max},D_{Min})}$ و $LT_{(K_{Max},D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنة بطريقة مقدر انحدار الحرف (RR).

3- عند حجم العينة (100) نلاحظ ان مقدر نوع ليو(LT) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{Max},D_{Min})}$ و $LT_{(K_{Max},D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنة بطريقة مقدر انحدار الحرف (RR)، اما في القيمتين $(\rho = 0.90, 0.95)$ تمتلك احد مقدرات انحدار حرف K_4 اقل متوسط مربعات الخطأ (MSE) مقارنة بباقي صيغ تقدير.

4-عند حجم العينة (250) نلاحظ ان مقدر نوع ليو (LT) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{Max},D_{Min})}$ و $LT_{(K_{Max},D_{Max})}$ تمتلك اقل متوسط مربعات الخطأ (MSE) مقارنة بطريقة مقدر انحدار الحرف (RR).

7-3 تفسير النتائج تبعا لتغير حجم العينة (n)

إن الملاحظة بدقة للجدولين (1) و (2) يتبين التراجع الحاصل في قيم متوسط مربعات الخطأ (MSE) كلما ازداد حجم العينة، اذ يظهر ذلك بوضوح في كافة طرائق التقدير وهذا يعكس إحدى الخصائص الجيدة عندما تقترب القيمة المقدرة من القيمة الحقيقة للمعلمة بزيادة حجم العينة، وبذلك نستنتج أن ألافضلية لطريقة مقدر نوع ليو (LT) عند كافة احجام العينة حيث سجلت تاثير ايجابي لزيادة حجم العينة كلما ازداد حجم العينة تقل متوسط مربعات الخطأ (MSE) لطريقتي التقدير مقدر انحدار الحرف (RR) ومقدر نوع ليو(LT).

8-3 تفسير النتائج تبعا لتغيير قيمة معامل الارتباط (ρ)

من نتائج المحاكاة يمكن الملاحظة بسهولة تفوق طريقة مقدر نوع ليو (LT) عند ثبوت كافة العوامل المختلفة واختلاف قيمة معامل الارتباط ، بعض صيغ مقدرات انحدار حرف K_1 و K_2 جاء متوسط مربعات الخطأ (MSE) لها كبيرة، حيث كلما ازدادت قيمة معامل الارتباط تبدأ الفوارق بين قيم متوسط مربعات الخطأ (MSE) لكافة الطرائق بالازدياد، نستنتج من ذلك ان الافضلية دائما تكون لطريقة مقدر نوع ليو (LT) لكافة قيم معامل الارتباط.

9-3 تفسير النتائج تبعاً لتغير عدد المتغيرات التوضيحية P

سجلت قيم متوسط مربعات الخطأ (MSE) لمعلمات أنموذج انحدار ثنائي الحدين السالب المقدرة وفق كافة الطرائق عند احتسابها للأنموذج الذي يتضمن خمس متغيرات توضيحية كما في الجدول (3) زيادة ملحوظة عن نظيراتها المحتسبة لأنموذج انحدار ثنائي الحدين السالب المتضمن متغيرين توضيحيين، في حين أظهرت نتائج المقارنة ورغم اختلاف عدد المتغيرات المستقلة في الأنموذج تفوق طريقة تقدير معلمات أنموذج انحدار ثنائي الحدين السالب وهي طريقة مقدر نوع ليو LT اذ تبين هناك تأثير سلبي لزيادة عدد المتغيرات التوضيحية حيث كلما زادت عدد المتغيرات التوضيحية المدروسة يزداد متوسط مربعات الخطأ (MSE) لطرائق التقدير.

جدول (2) قيم متوسط مربعات الخطأ لكافة طرائق التقدير عندما يكون عدد المتغيرات المستقلة P=5

n	ρ	Ridge Estimator							Liu Type Estimator			
		K_1	K_2	K_3	K_4	K_5	K_6	K_7	$LT_{(K_{Min},D_{Min})}$	$LT_{(K_{Min},D_{Max})}$	$LT_{(K_{Max},D_{Min})}$	$LT_{(K_{Max},D_{Min})}$
25	0.90	0.160264	0.160222	0.137361 0.161101 0.161425	0.143415 0.151632 0.161425 0.161427	0.1612993	0.1227722	0.1510586	0.16142474	0.16142474	0.09593886	0.09593886
	0.95	0.173935	0.173905	0.150899 0.174880 0.175252	0.156129 0.163693 0.175252 0.175252	0.1750963	0.1553452	0.1645855	0.17525211	0.17525211	0.10852345	0.10852345
	0.99	0.205122	0.205395	0.186378 0.206412 0.207059	0.1874797 0.1895252 0.2070599 0.2070599	0.2068606	0.1761900	0.1970058	0.2070598	0.2070598	0.16527913	0.16527913
50	0.90	0.068129	0.068235	0.067319 0.068298 0.068327	0.0666774 0.0640110 0.0683267 0.0683267	0.0683234	0.0671345	0.0675124	0.068326711	0.068326711	0.06351977	0.06351977
	0.95	0.074461	0.074574	0.073645 0.074641 0.074675	0.0729275 0.0697699 0.0746748 0.0746748	0.0746704	0.0701529	0.0738488	0.074674830	0.074674830	0.06989277	0.06989277
	0.99	0.089327	0.089308	0.083577 0.089413 0.089442	0.0847579 0.0864786 0.0894418 0.089441	0.0894328	0.0840142	0.0877192	0.089441756	0.089441756	0.080813004	0.080813004

تابع لجدول (2)

n	ρ	Ridge Estimaor							Liu Type estimator			
		K_1	K_2	K_3	K_4	K_5	K_6	$\bar{K}$	$LT(K_{Min}, D_{Min})$	$LT(K_{Min}, D_{Max})$	$LT(K_{Max}, D_{Min})$	$LT(K_{Max}, D_{Min})$
100	0.90	0.0307548	0.030762	0.030598 0.030770 0.030774	0.0304757 0.0300294 0.0307736 0.0307736	0.030773	0.030153	0.030629	0.0307732	0.0307732	0.029967785	0.029967785
	0.95	0.0338646	0.033873	0.033692 0.033882 0.033886	0.03355577 0.03343491 0.03388594 0.03388594	0.033886	0.033591	0.033731	0.033885	0.033885	0.03297676	0.03297676
	0.99	0.0410688	0.041125	0.040483 0.041138 0.041142	0.04043766 0.04033878 0.04114239 0.04114238	0.041142	0.039939	0.040867	0.04114238	0.04114238	0.03911404	0.03911404
250	0.90	0.0132909	0.013292	0.013287 0.013291 0.013292	0.01327214 0.01312654 0.01329176 0.0132917	0.013291	0.013263	0.013287	0.013291	0.013291	0.013040698	0.013040698
	0.95	0.0146305	0.014631	0.014626 0.014631 0.014631	0.01460809 0.01425363 0.01463140 0.01463140	0.014631	0.014587	0.014625	0.01463140	0.01463140	0.013964835	0.013964835
	0.99	0.0177550	0.017756	0.017736 0.017756 0.017757	0.01770761 0.01767147 0.01775625 0.01775625	0.017756	0.017597	0.017744	0.01775625	0.01775625	0.0160240056	0.016024005

المبحث الرابع/ الاستنتاجات والتوصيات

يتضمن هذا المبحث أهم الاستنتاجات التي تم التواصل اليها من خلال المباحث الثلاثة السابقة في هذا البحث، وكذلك التوصيات التي يوصى بها الباحث قيام به بعد هذه البحث .

1-4 الاستنتاجات

1-اذ تبين من نتائج المحصلة من المحاكاة ان طريقة مقدر نوع ليو (Liu Type Estimator) في صيغته التقديرية الثالثة والرابعة $LT_{(K_{MAX},D_{Min})}$، $LT_{(K_{Max},D_{Max})}$ على التوالي انها افضل الطرائق المدروسة في تقدير معلمات أنموذج ثنائي الحدين السالب والتي اعطت اقل متوسط مربعات للخطأ عند كل حجوم العينة .

2-زيادة عدد المتغيرات التوضيحية ومعامل الارتباط تؤثر بشكل سلبي على تقديرات متوسط مربعات الخطأ (Mean Square Error) ومتمثلة بمعيار مقارنة بين طرائق التقدير.

3- لا تشكل زيادة حجم العينة أية عوائق تجاه كفاءة طريقة مقدر نوع ليو في تقدير معلمات أنموذج انحدار ثنائي الحدين السالب بينما تسبب العوامل الاخرى وهي معامل الارتباط وعدد المتغيرات التوضيحية تأثيرا على كفاءة طرائق التقدير السابقة.

2-4 التوصيات

1-ضرورة استخدام مقاييس النزعة المركزية كالوسيط في بناء مقدرات معلمة التحيز خصوصا عند وجود قيم شاذة في البيانات.

2-استعمال أنظمة البرمجيات الحديثة في تبويب وارشفة البيانات في المؤسسات الحكومية ولاسيما المؤسسات الطبية.

3-اعتماد طريقة مقدرات نوع ليو (LT) في تقدير معلمات أنموذج ثنائي الحدين السالب في حالة وجود مشكلة التعدد الخطي .

المصادر

References

المصادر العربية

i. هرمز، امير حنا (1990) م. "الاحصاء الرياضي"، مديرية دار الكتب للطباعة والنشر، العراق، نينوى.

ii. كاظم، اموري هادي، ومسلم، باسم شليبة (2002) م. "القياس الاقتصادي المتقدم النظرية والتطبيق"، مطبعة دنيا الأمل، العراق، بغداد.

iii. يحيى، م.م. مزاحم محمد (2005) م. "استخدام المكونات الرئيسة وانحدار الحرف في تقدير معادلة السعر العالمي للقمح للفترة من (1961-2002)، كلية الادارة والاقتصاد، جامعة تكريت.

المصادر الاجنبية

iv- Asar ,Y., (2016). " Liu–type Negative Binomial Regression: A comparison of Recent Estimators And Applications", Department Of Mathematics Necmettin Erbakan University Konya, Turkey.

v- Asar, Y., (2016). "Liu Type Logistic Estimators With Optimal Shrinkage Parameter", Necmettin Erbakan University, Journal Of Modern Applied Statistical Methods ,Konya, Turkey.

vi- Alheety, M. I. & Kibria, B. M. G. (2013). "Modified Liu–Type Estimator Based On (r-k) Class Estimator", [1] Department Of Mathematics, University Of Al-Anbar, Ramadi ,Iraq,[2] Department Of Mathematics & Statistics, Florida International University Miami, U.S.A.

vii- Eskelson, B. N. J. & Temesqen, H. & Barrett, Tara M. (2009). "Estimating Cavity Tree And Snag Abundance Using Negative Binomial Regression Models And Nearest Neighbor Imputation Methods", Oregon State University, U.S.A.

viii- Lord, D. & Park, B. J. (2012). "Negative Binomial Regression Models And Estimation Models", Texas A&M University, Korea Transport Institute.

ix- Levine, N. & Lord, D. & Park, B. J. (2010). ''Regression Modeling", Zachry Department of Civil Engineering Texas A&M University.

x- Lukman, A. F. & Ayinde, K. (2015). "Review And Classifications of the Ridge Parameter Estimation Techniques", Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.

xi-Lawless, J. F. (1987). "Negative Binomial and Mixed Poisson Regression", University of Waterloo.

xii- Månsson, K. (2012). "On Ridge Estimators for the Negative Binomial Regression Model", Department Of Economics And Statistics, Jồ nkỏ ping University, Sweden.

Comparison between the Methods of Ridge Regression and Liu Type to Estimate the Parameters of the Negative Binomial Regression Model Under Multicollinearity Problem by Using Simulation

Abstract

The problem of Multicollinearity is one of the most common problems, which deal to a large extent with the internal correlation between explanatory variables. This problem is especially Appear in economics and applied research, The problem of Multicollinearity has a negative effect on the regression model, such as oversized variance degree and estimation of parameters that are unstable when we use the Least Square Method (OLS), Therefore, other methods were used to estimate the parameters of the negative binomial model, including the estimated Ridge Regression Method and the Liu type estimator, The negative binomial regression model is a nonlinear regression model or part of the general exponential family. This is the basic structure of the Count Data Analysis, which was used as an alternative to the Poisson model when there is a problem with overdisperison Where the variation value of the response variable (Y) is greater than its arithmetic mean ,The Monte Carlo study was designed to compare the Ridge Regression Estimator and the Liu Type Estimator By using the standard Compare Mean Square Error (MSE), A simulation result showed that the method of the Liu Type estimator is better than the Ridge Regression Method, The Mean Square Error in Liu Type Estimator are lower in the third and fourth estimation formulas.

Research Key: linear multiplicity, negative binomial regression model, Ridge Rigression Estimator , Liu Type Estimator.