نموذج Tobit Quantile Regression البيزي
باستعمال elastic net المكيفة المضاعفة وانحدار الحرف المكيفة

ا.د. محمود مهدي حسن البياتي/ كلية الادارة والاقتصاد/ جامعة بغداد
الباحث/ هيثم حسون ماجد

تاريخ التقديم:2017/10/29
تاريخ القبول:2017/12/20

المستخلص:

نموذج (Tobit Quantile Regression) انبثق حديثا كأداة احصائية مهمة في الكثير من التحليلات الاحصائية . وبغية تطوير عملية التقدير في هذا النموذج فقد تم في هذه الدراسة اقتراح النموذج البيزي الهرمي بتقنية elastic net المكيفة المضاعفة والنموذج الهرمي البيزي بتقنية انحدار الحرف المكيفة.

في تقنية elastic net المكيفة المضاعفة تم افتراض ان كل معلمة من معلمات الجزاء (penalty parameters λ_1, λ_2) تكون مختلفة لكل معلمة من معلمات النموذج ، كذلك في تقنية انحدار الحرف المكيفة فقد تم افتراض ان معلمة الجزاء((penalization parameter (λ) تكون ايضا مختلفة لكل معلمة من معلمات النموذج .

تم استخدام اسلوب المحاكاة في بيان كفاءة الطرق المقترحة واظهرت النتائج كفاءة هذه الطرق في التعامل مع عملية تقدير معلمات النموذج في حالة وجود ارتباطات كبيرة بين المتغيرات التوضيحية .

هذا هو العمل الاول (حسب علم الباحث) الذي يتم فيه مناقشة تقدير واختيار المتغيرات لنموذج Tobit Quantile Regression باقتراح النموذج الهرمي البيزي في تقنية elastic net المكيفة المضاعفة وتقنية ridge regression المكيفة.

المصطلحات الرئيسة للبحث/ نموذج الانحدار المجزئ توبت ، elastic net المكيفة ، انحدار الحرف المكيفة ، خوارزمية Gibbs Sampler ، خوارزمية Metropolis Hasting

مجلة العلوم
الاقتصادية والإدارية
العدد 107 المجلد 24
الصفحات 521ـ 537

*البحث مستل من اطروحة دكتوراة

المبحث الاول

1-1. المقدمة:

نموذج (Tobit Quantile Regression(TQR هو امتداد لنموذج Tobit الذي قدم من قبل (James Tobin 1958) وهو احد نماذج الانحدار المجزئ (Quantile Regression) المهمة المستخدمة في الدراسات والذي يكون فيه متغير الاستجابة (response variable) محدود (limited response variable) . هذا النموذج له اهمية كبير في كثير من الدراسات التطبيقية مثل الدراسات الطبية والدراسات الاقتصادية وغيرها .ولأهمية هذا النموذج نجد ان هناك الكثير من الدراسات الحديثة تولي اهتماما كبيرا في دراسة هذا النموذج .

وكما هو معروف ان احد اهم خطوات بناء نماذج الانحدار هو تحديد المتغيرات التوضيحية (explanatory variables) ذات العلاقة بمتغير الاستجابة (response variable) كما ان مشكلة التعدد الخطي (Maulticollinearity) هي احد المشاكل المهمة التي تواجه الكثير من الدراسات والتي تقود الباحث الى الاستنتاج الخاطئ حول حذف بعض المتغيرات ذات العلاقة في الدراسة وكذلك يمكن ان يتضمن النموذج بعض المتغيرات التوضيحية التي ليس لها اهمية في الدراسة.

الدراسات الحديثة اثبتت ان التنظيم (Regularization) في نماذج الانحدار له اهمية كبيرة في اختيار المتغيرات وايضا في التعامل مع مشكلة التعدد الخطي في النموذج فعلى سبيل المثال تقنية ridge regression التي قدمت من قبل (Hoeral & Bennard) في عام 1970 التي لها اهمية في معالجة مشكلة التعدد الخطي من خلال تقليل التباين للمقدرات مقابل التضحية بمقدار بسيط من التحيز و تقنية lasso التي تم اقتراحها من قبل Tibshirani في 1996 كأداة مهمة في اختيار المتغيرات التوضيحية للنموذج . وتقنية elastic net التي قدمت من قبل (Zou and Hastie 2005) التي لها اهمية كبيرة في معالجة مشكلة التعدد الخطي واختيار المتغيرات في نماذج الانحدار ، اعقب هذا التقديم التطورات التي حصلت على هذه التقنيات فمثلا (Kubokawa & Sivastava 2004) استخدم تقنية (Adaptive ridge regression)في نموذج الانحدار المتعدد وفي عام 2011 استخدم (Yu & Alhamzawi) (adaptive lasso) في نموذج Quantile Regression وتقنية (adaptive elastic net) التي قدمت من قبل (Zou and Zhang 2009) استخدموا فيها معلمة الجزاء λ_1 المختلفة لجزاء معلمات نموذج الانحدار حيث بينوا ان لهذا الاسلوب اهمية في التعامل مع مشكلة التعدد الخطي مقارنة بالطرق الاخرى هذه التقنية استخدمت ايضا من قبل (Alhamzawi 2014) في نموذج (TQR) وفي عام 2013 استخدم Deborah Gefang تقنية (Double Adaptive elastic net) في نموذج (vector autoregressive (VAR) models) استخدم في هذه التقنية الجزاء المختلف في كلا معلمتي التنظيم $\lambda_{1j}, \lambda_{2j}$ لجزاء كل معلمة من معلمات النموذج (βj) وبينو مرونة هذه الطريقة في التعامل مع مشكلة التعدد الخطي في هذا النموذج.

من ناحية اخرى ان الاسلوب البيزي له اهمية كبيرة في التعامل مع نماذج الانحدار بشكل عام ومع نموذج توبت بشكل خاص فالأسلوب البيزي له اهمية في الوصول الى الاستدلال الدقيق وحتى في حالة العينات الصغيرة وكذلك في كثير من الاحيان نجد ان الطرق الاعتيادية المستخدمة في تقدير نموذج توبت لا تزودنا بالخطأ المعياري standard error لمقدرات تقنية elastic net التي هي احد التقنيات المهمة المستخدمة في تقدير واختيار المتغيرات في نماذج الانحدار في حين ان الاسلوب البيزي يزودنا بذلك.

يمكن القول ان بداية استخدام الاسلوب البيزي في نماذج (Quantile Regression(QR كانت من قبل (Yu & Moyeed 2001) اتبع ذلك استخدام هذا الاسلوب في نموذج TQR من قبل (Yu &Stander) للمرة الاولى في عام 2007 . وفي الحقيقة ان استخدام الاسلوب البيزي في نماذج QR مبني على افتراض (بغض النظر عن التوزيع الحقيقي للبيانات) توزيع لابلاس الملتوي Asymmetric Laplace(AL)) في صياغة دالة الامكان للنموذج .

وبالرغم من الصعوبات التحليلية الناتجة في هذا الاسلوب الا انهم بينوا امكانية استخدام اسلوب Markov chain Monte Carlo (MCMC) في اجراء المعاينة من التوزيعات الشرطية اللاحقة (conditional posterior distribution). اعقب هذه الدراسة عدد من الدراسات التطويرية في هذا المجال مثل (Kozumi & Kobayashi 2011) ، (Alhamzawi 2014) ، (Alhusseini Fadel 2017).

2-1 الهدف من البحث :

يهدف البحث الى تقدير نموذج (TQR) في حالة وجود ارتباطات كبيرة بين المتغيرات التوضيحية وذلك من خلال اقتراح النموذج الهرمي البيزي وبأسلوب جديد في التعامل مع تقنية elastic net وتقنية ridge regression من خلال افتراض الدوال الاولية الملائمة والتي تسمح بان تكون معلمات الجزاء لكل تقنية مختلفة لكل معلمة من معلمات النموذج .

المبحث الثاني/ الجانب النضري

1-2 الاسلوب الكلاسيكي في تقدير نموذج TQR:

لتوضيح هيكلية نموذج Tobit Quantile Regression لنفرض ان y^* , y هي متغيرات عشوائية ترتبط بالعلاقة التالية[9]:

$$y = \max\{y^0, y^*\} \ldots\ldots(1)$$

حيث ان y^0 تمثل نقطة المراقبة المعلومة .

لنفرض ان هناك عينة من المشاهدات المستقلة $y=(y_1,y_2,y_3,\ldots y_n)$ المرتبطة بالمتغيرات التوضيحية $x=(x_1,x_2,x_3,\ldots x_n)$

حيث ان x_i هو متجه لـ k من المتغيرات التوضيحية .

لـ (pth quantile) المتغير اللاتيني (latent variable) y^* يمكن ان يعبر عنه بالمعادلة التالية[9] :

$$y_i^* = \acute{x}_i \beta_p + \varepsilon_i \ldots\ldots(2)$$

ε_i : تمثل البواقي والتي تكون مقيدة بحيث ان (pth quantile is zero)

β_p : هي متجه المعلمات المجهولة.

للسهولة ولبقية الدراسة فان متجه المعلمات β_p سيعبر عنه بالشكل β.

الاسلوب الكلاسيكي في تقدير معلمات نموذج TQR يمكن ان يعبر عنه بالشكل التالي[1]:

$$\min_\beta = \sum_{i=1}^{n} \rho_p(y_i - \max\{y^0, \acute{x}_i\beta\}) \ldots\ldots\ldots(3)$$

حيث ان ρ_p يمكن ان يعبر عنها بالشكل التالي [12]:

$$\rho_p(u) = \begin{cases} pu & if u \geq 0 \\ -(1-p)u & if u < 0 \end{cases} \ldots\ldots(4)$$

2-2 **الاسلوب البيزي وتقنية** Double Adaptive elastic net **في تقدير نموذج** TQR:

قبل التعرض الى الاسلوب البيزي في التقدير لابد من التعرض الى دالة الكثافة الاحتمالية (pdf) لتوزيع لابلاس الملتوي Asymmetric Laplace والتمثيل المختلط (mixture representation) لهذا التوزيع.

دالة الكثافة الاحتمالية pdf لتوزيع لابلاس الملتوي (AL) يمكن ان يعبر عنها بالشكل التالي [12] :

$$f_p(\varepsilon \backslash \tau) = p(1-p)\tau \exp[-\tau\, \rho_p(\varepsilon)] \ldots\ldots\ldots.(5)$$

ان (pth quantile) لهذا التوزيع هو صفر $(the\ pth\ quantile\ is\ zero)$ **.**

وبغية تطبيق الاسلوب البيزي في تقدير نموذج TQR افترض (Yu &stander) في عام 2007 وبغض النظر عن التوزيع الحقيقي للبيانات توزيع AL في صياغة دالة الامكان للنموذج.

ان هذا الافتراض سهل عملية ربط الاسلوب الكلاسيكي الى الاسلوب البيزي في تقدير النموذج . دالة الامكان لنموذج (TQR) يمكن ان يعبر عنها بالشكل التالي[9] :

$$f(y\backslash \boldsymbol{X},\beta,\tau) = p^n(1-p)^n\tau^n \exp\{-\tau \textstyle\sum_{i=1}^{n} \rho_p(y_i - \max(y^0, \acute{x}_i\beta)) \ldots (6)$$

وبرغم الاهمية النظرية من الافتراض اعلاه لكن هناك صعوبات تحليلية كبيرة في الجانب التطبيقي ناتجة عن الدوال الشرطية اللاحقة الناتجة من هذا الافتراض .

ومن اجل تسريع وزيادة كفاءة اداء الاسلوب البيزي في تحليل نموذج Tobit Quantile Regression فقد استخدم الباحثان (Kozumi & Kobayshi (2011)) التمثيل المختلط لتوزيع لابلاس الملتوي (AL) وذلك باستخدام التوزيع الطبيعي المختلط (mixed normal distribution) مع التوزيع الاسي (Exponential distribution) [6].

لذا فان المتغير العشوائي اللاتيني ((latent variable (y*) في (2) باستخدام التمثيل المختلط لتوزيع الخطأ يمكن ان يعبر عنه بالشكل التالي[6]:

$$y_i^* = \acute{x}_i\beta + \xi_1 e_i + \tau^{-1/2}\xi_2\sqrt{e_i}\; z_i \ldots\ldots\ldots\ldots(7)$$

Where $\xi_1 = (1-2p)/p(1-p)$ **,** ξ_2^2 **= 2/p(1-p)**

$z_i \sim N(0,1)\ ,\quad e_i \sim exp(1/\tau)\ ,\ \tau > 0$

بشكل عام الاسلوب الكلاسيكي بتقنية elastic net في تقدير نماذج Quantile Regression يمكن ان يعبر عنه بالشكل التالي[12]:

$$min_\beta \textstyle\sum_{i=1}^{n} \rho_p(y_i - \acute{x}_i\beta) + \delta_1 \sum_{j=1}^{k}|\beta_j| + \delta_2 \sum_{j=1}^{k}\beta_j^2 \ldots\ldots\ldots (8)$$

$$\delta_1 = \tau\lambda_1, \delta_2 = \tau\lambda_2$$

وبغية ربط الاسلوب البيزي في تقدير معاملات نموذج Linear Quantile regression بتقنية elastic net فقد افترض (Li & Lin 2010) الدالة الاولية الاتية للمعلمة β_k **[12] :**

$$f(\beta_k) = C(\delta_1,\delta_2)\,\frac{\delta_1}{2}\exp\{-\delta_1|\beta_k| - \delta_2\beta_k^2\} \ldots\ldots\ldots(9)$$

حيث ان :

$C(\delta_1,\delta_2)$ **: هو ثابت الدالة تعتمد قيمته على** δ_1 **,** δ_2

$\delta_1, \delta_2 \geq 0$**: هما معلمتي التنظيم (tuning parameters).**

$\sum_{j=1}^{k}|\beta_j|$ **: يمثل (**L_1**norm penalty).**

$\sum_{j=1}^{k}\beta_j^2$: يمثل (L2norm penalty).

ان الافتراض للدالة الاولية (9) اعطى اهمية ربط اسلوب التقدير الكلاسيكي مع الاسلوب البيزي في تقنية **elastic net** لتقدير معلمات نموذج **Linear Quantile regression** .

برهن (Andrews and Mallows 1974) ان لكل $d \geq 0$ فان المعادلة التالية تكون صحيحة[12]:

$$\frac{d}{2}\exp(-d|s|) = \int_0^\infty \frac{1}{\sqrt{2\pi a}}\exp\left(-\frac{s^2}{2a}\right)\frac{d^2}{2}\exp\left(\frac{-d^2}{2}a\right)da \ldots\ldots(10)$$

لذا فالدالة الاولية (9) يمكن ان يعبر عنها بالشكل التالي[12]:

$$f(\beta_k) = C(\delta_1,\delta_2)\int_0^\infty \frac{1}{\sqrt{2\pi\, a_k}}\exp\left\{-\frac{1+2\delta_2 a_k}{2a_k}\beta_k^2\right\}\frac{{\delta_1}^2}{2}\exp\left(-\frac{{\delta_1}^2}{2}a_k\right)da_k \ldots(11)$$

let $r_k=1+2\,\delta_2 a_k$**.....(12)**

$$f(\beta_k) = C(\delta_1,\delta_2)\int_1^\infty \frac{r_k^{-1\backslash 2}}{\sqrt{2\pi(r_k-1)\backslash(2\delta_2 r_k)}}\exp\left\{-\frac{1}{2}\left(\frac{r_k-1}{2\delta_2 r_k}\right)^{-1}\beta_k^2\right\}\times$$

$$\frac{{\delta_1}^2}{4\delta_2}\exp\left(-\frac{{\delta_1}^2}{4\delta_2}(r_k-1)dr_k\ldots\ldots(13)\right.$$

كما برهن (Li & Lin 2010) ان قيمة $C(\delta_1,\delta_2)$ يمكن ايجادها بالشكل التالي:[12]

$$C(\delta_1,\delta_2) = \Gamma^{-1}(1\backslash 2\,,\tilde{\delta_1})(\tilde{\delta_1})^{-1\backslash 2}exp(-\tilde{\delta_1})\ldots.(14)$$

$$\tilde{\delta_1} = {\delta_1}^2/4\delta_2$$

اما في نموذج **Tobit Quantile Regression** الاسلوب الكلاسيكي لتقنية **elastic net** في تقدير معاملات النموذج يمكن ان يعبر عنها بالشكل التالي[1] :

$$min_\beta \sum_{i=1}^{n}\rho_p(y_i-\max\{y^0,\acute{x}_i\beta\})+\delta_1\sum_{j=1}^{k}|\beta_j|+\delta_2\sum_{j=1}^{k}\beta_j^2 \ldots\ldots\ldots(15)$$

الحمزاوي رحيم (2014) استخدم الدالة الاولية (9) في الاسلوب البيزي لتقدير معاملات نموذج **Tobit Quantile Regression** بملاحظة ان تعظيم الدالة اللاحقة **posterior** الناتجة من حاصل ضرب الدالة الاولية في (9) ودالة الامكان في (6) τ(1=) يكافئ تصغير المعادلة (15) .افترض في هذه الدراسة ان معلمة التنظيم δ_1 هي متغير عشوائي اما δ_2 فهي مقدار ثابت كما استخدم قيم مختلفة لـ δ_1 في جزاء معلمات النموذج.

يقترح البحثان استخدام الاسلوب البيزي بتقنية **elastic net** في تقدير واختيار المتغيرات لنموذج **TQR** وذلك بافتراض ان كلا معلمتي التنظيم **(tuning parameters**δ_1,δ_2) هي متغير عشوائي وتكون مختلفة في كل معلمة من معلمات النموذج .

3-2 **النموذج الهرمي البيزي في تقنية** Double adaptive elastic net

(Bayesian hierarchical model in double adaptive elastic net technique)

النموذج الهرمي المقترح في الدراسة يستند الى النموذج المستخدم من قبل Li & Lin (2010) في تقدير معلمات نموذج Linear Quantile Regression [12]

تم اختيار الدوال الاولية الملائمة لتقدير معلمات نموذج Tobit Quantile Regression بوجود مشكلة التعدد الخطي وبافتراض ان كلا معلمتي التنظيم $(\delta_{1k}, \delta_{2k})$ هي متغيرات عشوائية ويتم تقديرها لكل معلمة من معلمات النموذج.

ليكن :

$\mathbf{e} = (e_1, e_2 \dots e_n)'$, $\boldsymbol{z} = (z_1, z_2 \dots . z_n)'$, $\boldsymbol{r} = (r_1, r_2 \dots r_k)'$

النموذج الهرمي يمكن ان يعبر عنه بالتالي:

$$y_i = \max\{y^0, y_i^*\}$$
$$y_i^* = \acute{x}_i \beta + \xi_1 e_i + \tau^{-1/2} \xi_2 \sqrt{e_i}\, z_i$$
$$\boldsymbol{e} \backslash \tau \sim \prod_{i=1}^{n} \tau \exp(-\tau\, e_i) \qquad (16)$$
$$\boldsymbol{z} \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right)$$
$$(\beta_k \backslash r_k, \delta_{2k}) \overset{ind}{\sim} \frac{1}{\sqrt{2\pi (r_k - 1) \backslash (2\delta_{2k} r_k)}} \exp\left\{-\frac{1}{2}\left(\frac{r_k - 1}{2\delta_{2k} r_k}\right)^{-1} \beta_k^2\right\}$$
$$(r_k \backslash \widetilde{\delta_{1k}}) \overset{ind}{\sim} \Gamma^{-1}(1\backslash 2, \widetilde{\delta_{1k}}) r_k^{-1\backslash 2} \widetilde{\delta_{1k}}^{1\backslash 2} \exp\{-\widetilde{\delta_{1k}}\, r_k\}, I(r_k > 1)$$
$$\tau \sim \tau^{-1}$$
$$(\widetilde{\delta_{1k}} \backslash \theta_1, \alpha_1) \overset{ind}{\sim} \frac{\theta_1^{\alpha_1}}{\Gamma(\alpha_1)} \widetilde{\delta_{1k}}^{\alpha_1 - 1} \exp(-\theta_1 \widetilde{\delta_{1k}})$$
$$\delta_{2k} \overset{ind}{\sim} \theta_3 \exp(-\delta_{2k} \theta_3)$$

4-2 **اشتقاق التوزيعات اللاحقة الشرطية** (conditional posterior distribution) **للمتغيرات العشوائية في النموذج الهرمي** (16) *

1- التوزيع الشرطي الكامل للمتغير العشوائي y^* :

بالاستناد الى التمثيل المختلط لتوزيع الخطأ المتضمن في المعادلة (7) فان التوزيع الشرطي للمتغير العشوائي y^* يمكن ان يعبر عنه بالشكل التالي:

$$f(y^* \backslash \boldsymbol{X}, \boldsymbol{e}, \boldsymbol{\beta}, \boldsymbol{r}, \tau, \delta_{2k}, \widetilde{\delta_{1k}}) = \prod_{i=1}^{n} \frac{\tau^{1\backslash 2}}{\sqrt{2\pi \xi_2^2 e_i}} \exp\{-\tau (y^* - \acute{x}_i \beta - \xi_1 e_i)^2 \backslash 2\xi_2^2 e_i\} \quad \dots (17)$$

2- التوزيع الشرطي الكامل للمتغير العشوائي e_i :

$$f(e_i \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}_{-i}, \boldsymbol{\beta}, \boldsymbol{r}, \tau, \delta_{2k}, \widetilde{\delta_{1k}}) \propto f(y^* \backslash \boldsymbol{X}, \boldsymbol{e}, \boldsymbol{\beta}, \boldsymbol{r}, \tau, \delta_{2k}, \widetilde{\delta_{1k}}) f(e_i) \dots (18)$$
$$\propto e_i^{-1\backslash 2} \exp\{-\tau (y^* - \acute{x}_i \beta - \xi_1 e_i)^2 \backslash 2\xi_2^2 e_i\} \exp(-\tau e_i) \dots (19)$$
$$f(e_i \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}_{-i}, \boldsymbol{\beta}, \boldsymbol{r}, \tau, \delta_{2k}, \widetilde{\delta_{1k}}) \propto e_i^{-1\backslash 2} \exp\{-\frac{1}{2}[(\tau \xi_1^2 \xi_2^{-2} + 2\tau) e_i +$$

$\xi_2^{-2}\tau(y_i^* - \acute{x}_i\beta)^2\ e_i^{-1}]\}\ldots(20)$

حيث ان : e_{-i} يمثل المتغير e باستبعاد e_i

يمكن ملاحظة ان التوزيع الشرطي للمتغير e_i هو **generalized inverse Gaussian**

3- التوزيع الشرطي الكامل للمتغير العشوائي β_k :

$$f(\beta_k \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \beta_{-k}, \boldsymbol{r}, \tau, \delta_{2k}, \widetilde{\delta_1}) \propto f(y^* \backslash \boldsymbol{X}, \boldsymbol{e}, \beta, \boldsymbol{r}, \tau, \delta_2, \widetilde{\delta_1}) f(\beta_k) \ldots (21)$$

حيث ان : β_{-k} يمثل المتجه β باستبعاد β_k

$$\propto\ \exp\{-\tau\sum_{i=1}^n (y_i^* - \acute{x}_i\beta - \xi_1 e_i)^2 \backslash 2\xi_2{}^2 e_i\ \ \exp\left\{-\frac{1}{2}\left(\frac{r_k-1}{2\delta_{2k} r_k}\right)^{-1}\beta_k^2\right\}\ldots(22)$$

$$\propto \exp\{-\frac{1}{2}\ [(\tau\xi_2^{-2}\sum_{i=1}^n x_{ik}^2 e_i{}^{-1} + 2\delta_{2k} r_k (r_k-1)^{-1})\beta_k^2 - 2\sum_{i=1}^n (\tau w_{ik} x_{ik}\xi_2^{-2} e_i^{-1})\beta_k \ldots (23)$$

$$w_{ik} = y_i^* - \xi_1 e_i - \sum_{j=1, j\neq k}^p x_{ij}\beta_j \ldots (24)$$

التوزيع الشرطي الكامل لـ β_k هو **Normal**

بمتوسط

$$\mathrm{Me_k} = \mathrm{V_k^2}\xi_2{}^{-2}\tau\sum_{i=1}^n \mathrm{w_{ik}}(\mathrm{x_{ik}}\backslash \mathrm{e_i})\ldots(25)$$

وتباين

$$V_k^2 = \frac{1}{\xi_2{}^{-2}\tau\sum_i^n (x_{ik}^2 e_i{}^{-1}) + 2\delta_{2k} r_k (r_k-1)^{-1}}\ldots(26)$$

4-التوزيع الشرطي الكامل للمتغير العشوائي τ :

$$f(\tau \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \delta_{2k}, \widetilde{\delta_{1k}}) \propto f(y^* \backslash \boldsymbol{X}, \boldsymbol{e}, \boldsymbol{\beta}, \boldsymbol{r}, \tau, \delta_{2k}, \widetilde{\delta_{1k}}) f(e_i \backslash \tau)\ f(\tau) \ldots (27)$$

$$\propto \tau^{n\backslash 2}\exp\{-\tau\sum_{i=1}^n (y_i^* - \acute{x}_i\beta - \xi_1 e_i)^2 \backslash 2\,\xi_2{}^2 e_i\}\tau^n \exp(-\tau\sum_{i=1}^n e_i)\,\tau^{-1}\ \ldots(28)$$

$$\propto \tau^{(3n\backslash 2)-1}\exp\{-\tau\sum_{i=1}^n ([(y_i^* - \acute{x}_i\beta - \xi_1 e_i)^2 \backslash 2\,\xi_2{}^2 e_i] + e_i)\}\ \ldots(29)$$

لذا فالتوزيع الشرطي الكامل لـ τ هو **Gamma**

5- التوزيع الشرطي الكامل للمتغير العشوائي $(r_k - 1)$:

$$f(r_k - 1 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}_{-k}, \beta_k, \tau, \delta_{2k}, \widetilde{\delta_{1k}}) \propto$$

$$(r_k-1)^{-1\backslash 2}\exp\left\{-\frac{1}{2}\left[\frac{2\delta_{2k} r_k}{r_k-1}\right]\beta_k^2\right\}\exp\{-\widetilde{\delta_{1k}} r_k\} I(r_k > 1)\ldots(30)$$

$$\propto (r_k-1)^{-1\backslash 2} \propto \exp\left\{-\frac{1}{2}\left[2\widetilde{\delta_{1k}}(r_k-1) + \frac{2\delta_{2k}\beta_k^2}{r_k-1}\right]\right\} I(r_k - 1 > 0)\ldots(31)$$

حيث ان : $\boldsymbol{r}_{-k}$ يمثل المتغير العشوائي $\boldsymbol{r}$ باستبعاد المتغير العشوائي r_k

لذا فان توزيع المتغير العشوائي هو generalized inverse Gaussian

6- التوزيع الشرطي الكامل للمتغير العشوائي $\widetilde{\delta_{1k}}$:

$$f(\widetilde{\delta_{1k}} \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \delta_{2k}) \propto f(\widetilde{\delta_{1k}}) \Gamma^{-1}(1\backslash 2, \widetilde{\delta_{1k}}) \widetilde{\delta_{1k}}^{1\backslash 2} \exp\{-\widetilde{\delta_{1k}} r_k\} \ldots (32)$$

$$\propto \tilde{\delta_1}^{\alpha_1 - 1} \exp(-\theta_1 \tilde{\delta_1}) \Gamma^{-1}(1\backslash 2, \widetilde{\delta_{1k}}) \widetilde{\delta_{1k}}^{1\backslash 2} \exp\{-\widetilde{\delta_{1k}} r_k\} \ldots (33)$$

$$\propto \Gamma^{-1}(1\backslash 2, \widetilde{\delta_{1k}}) \widetilde{\delta_{1k}}^{(1\backslash 2) + \alpha_1 - 1} \exp\{-\widetilde{\delta_{1k}} [\theta_1 + r_k]\} \ldots (34)$$

7- التوزيع الشرطي الكامل للمتغير δ_{2k}:

$$f(\delta_{2k} \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \widetilde{\delta_{1k}}) \propto f(\delta_{2k}) \delta_{2k}^{1\backslash 2} \exp\left\{-\frac{1}{2}\left[\frac{2\delta_{2k} r_k}{r_k - 1}\right] \beta_k^2\right\} \ldots (35)$$

$$\propto \theta_3 \exp(-\theta_3 \delta_{2k}) \delta_{2k}^{1\backslash 2} \exp\left\{-\frac{1}{2}\left[\frac{2\delta_{2k} r_k}{r_k - 1}\right] \beta_k^2\right\} \ldots (36)$$

$$f(\delta_{2k} \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \widetilde{\delta_{1k}}) \propto \delta_{2k}^{(1\backslash 2)} \exp\left\{-\delta_{2k} \left(\left[\frac{r_k}{r_k - 1}\right] \beta_k^2 + \theta_3\right)\right\} \ldots (37)$$

لذا فالتوزيع الشرطي للمتغير δ_{2k} هو Gamma .

***ان اشتقاق الدوال الشرطية في (2,3,5,6) يستند الى المصدر (12).**

5-2 **الاسلوب البيزي وتقنية** Ridge regression (RR) :

طريقة انحدار الحرف (ridge regression) التي قدمت من قبل (Hoeral & Bennard) في عام 1970 هذه الطريقة تعالج مشكلة التعدد الخطي من خلال تقليل التباين للمقدرات مع التضحية بمقدار بسيط من التحيز فهي تساهم في تحسين دقة التنبؤ [11] .

يطلق على هذه الطريقة احيانا بـ (L_2norm) للدلالة على القيد المستخدم حول معلمات النموذج في عملية التقدير.

ان تقدير معلمات نموذج الانحدار (المتوسط او الاعتيادي) في تقنية انحدار الحرف هو الحل للمعادلة التالية [14]:

$$min_\beta \sum_{i=1}^{n} (y_i - \acute{x}_i \beta)^2 \ subject\ to\ \sum_{k=1}^{p} \beta_k^2 \leq t \ldots..(38)$$

او بالشكل :

$$min_\beta \sum_{i=1}^{n} (y_i - \acute{x}_i \beta)^2 + \lambda \sum_{k=1}^{p} \beta_k^2 \ldots..(39)$$

$$t, \lambda \geq 0 \quad (\lambda : tuning\ parameter)$$

الطريقة البيزية في تقنية ridge regression يمكن اجراءها تحت افتراض التوزيع الطبيعي لمتجه المعلمات [3].

بمعنى اخر وبالاستناد الى المصدر السابق يمكن التعبير عن الدالة الاولية للمعلمة بالشكل التالي :

for all k ……..(40)$\beta_k \sim N(0\ , \sigma^2/\lambda)$

وبشكل اخرعند افتراض $(\sigma^2 = 1)$

for all k ……..($\beta_k \sim N(0\ , 1/\lambda)$ **41)**

بالرغم من اهمية تقنية **ridge regression** في التعامل مع مشكلة التعدد الخطي وتحسين دقة التنبؤ **(predictive accuracy)** في نماذج الانحدار الا ان هناك انتقادات كثيرة حول هذه التقنية كونها لا تساهم في ايجاد نماذج تفسيرية **(interpretation models)** فهي تساهم بتقليص **(shrink)** معاملات النموذج الا انها لا تجعل المعاملات الغير مهمة في النموذج مساوية تماما للصفر.

في هذه الدراسة سيتم اقتراح النموذج الهرمي البيزي بتقنية **RR** في تقدير نموذج **(TQR)** وتحت افتراض ان $(\sigma^2 = 1)$ وقيمة المعلمة λ تكون مختلفة في كل معلمة من معلمات النموذج اي ان :

$$\beta_k \overset{ind}{\sim} N\left(0\ ,{}^{1}/_{\lambda_k}\right) \ldots\ldots (42)$$

6-2 **النموذج الهرمي البيزي في تقنية انحدار الحرف المكيفة** (Bayesian Hierarchical Adaptive ridge regression model):

Let $\mathbf{e} = (e_1, e_2 \ldots e_n)'$, $\mathbf{z} = (z_1, z_2 \ldots . z_n)'$, $\boldsymbol{r} = (r_1, \ldots r_k)'$

النموذج الهرمي يمكن ان يعبر عنه بالتالي:

$$\begin{aligned}
& y_i = \max\{y^0, y_i^*\} \\
& y_i^* = x_i'\beta + \xi_1 e_i + \tau^{-1/2}\xi_2\sqrt{e_i}\ z_i \\
& \mathbf{e}\backslash\tau \sim \textstyle\prod_{i=1}^{n} \tau \exp(-\tau\, e_i) \qquad\qquad \textbf{(43)} \\
& \mathbf{z} \sim \textstyle\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} z_i^2) \\
& (\beta_k\backslash\delta_k) \overset{ind}{\sim} \frac{1}{\sqrt{2\pi\backslash\delta_k}} \exp\left\{-\frac{1}{2}\ \delta\ \beta_k^2\right\} \\
& \delta_k \overset{ind}{\sim} \delta_k{}^{a-1} \exp(-d\delta_k) \\
& \tau \sim \tau^{-1}
\end{aligned}$$

7-2 **اشتقاق التوزيعات اللاحقة الشرطية** (Conditional posterior distribution) **للمتغيرات العشوائية في النموذج الهرمي** (43)

ان التوزيعات اللاحقة الشرطية للمتغيرات العشوائية $(y_i^*, \tau$ **, e ,z) هي كما في (4-2)**

<u>1- التوزيع اللاحق الشرطي للمتغير العشوائي β_k:</u>

$$f(\beta_k\backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \beta_{-k}, \tau, \delta_k) \propto f(y^*\backslash \boldsymbol{X}, \boldsymbol{e}, \overline{\beta}, \tau, \delta_k) f(\beta_k) \ldots (44)$$

$$\propto \exp\{-\tau \textstyle\sum_{i=1}^{n}(y_i^* - x_i'\beta - \xi_1 e_i)^2 \backslash 2\xi_2{}^2 e_i \quad \exp\left\{-\frac{1}{2}\ \delta_k\ \beta_k^2\right\} \ldots (45)$$

$$\propto \exp\{-\frac{1}{2}\ [(\tau\xi_2^{-2} \textstyle\sum_{i=1}^{n} x_{ik}^2 e_i{}^{-1} + \delta_k)\beta_k^2 - 2\sum_{i=1}^{n}(\tau w_{ik} x_{ik} \xi_2^{-2} e_i^{-1})\beta_k \ldots (46)$$

التوزيع اللاحق الشرطي لـ β_k هو **Normal** بمتوسط

$$\mathrm{Me_k} = \mathrm{V_k^2}\xi_2{}^{-2}\tau\sum_{i=1}^{n} w_{ik}(x_{ik}\backslash e_i) \dots (47)$$

وتباين

$$V_k^2 = \frac{1}{\xi_2{}^{-2}\tau\sum_i^n(x_{ik}^2 e_i{}^{-1}) + \delta_k} \dots .(48)$$

<u>2- التوزيع الشرطي اللاحق للمتغير العشوائي δ_k :</u>

$$f(\delta_k\backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \beta_k, \tau) \propto f(\delta_k)\delta_k{}^{1\backslash 2}\exp\left\{-\frac{1}{2}\ \delta_k\ \beta_k^2\right\} \dots .(49)$$

$$f(\delta_k\backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \beta_k, \tau) \propto \delta_k{}^{a-1}\exp(-b\delta_k)\,\delta_k{}^{1\backslash 2}\exp\left\{-\frac{1}{2}\ \delta_k\ \beta_k^2\right\} \dots (50)$$

$$f(\delta_k\backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \beta_k, \tau) \propto \delta_k{}^{a-\left(\frac{1}{2}\right)}\exp\left(-\delta_k\left(b + \frac{\beta_k^2}{2}\right)\right) \dots (51)$$

ثالثا: الجانب التجريبي:

تم في هذه الدراسة استخدام اسلوب المحاكاة لبيان كفاءة الطريقين المطبقتين .
تم توليد n من المشاهدات (n=nT+np) باستخدام النموذج التالي:

$$y_i = \max\{X\beta + U, 0\} \dots .(52)$$

حيث ان :
nT: العينة المستخدمة في تقدير المعلمات (training data).
np: العينة المستخدمة في التنبؤ (predictive data) .
X: هي مصفوفة المتغيرات التوضيحية بحيث يتم توليد هذه المصفوفة من التوزيع الطبيعي متعدد المتغيرات بمتوسط $\underline{0}$ ومصفوفة تباين $\sum$ عناصرها في الصف i والعمود j هي $0.5^{|i-j|}$.

$$U \sim skew\ normal(mean = 0, variance = 1, slant\ parameter = 0.95)$$

تم افتراض القيم التالية لمتجه المعلمات الحقيقية (true Beta) وكالتالي :

$$\beta = (3,1.5,0,0,2,0,0,0)$$

وبغية تقدير متجه المعلمات الحقيقية وباستخدام الطرق الانفة الذكر فقد تم استخدام ثلاثة قيم مختلفة لـ (quantile) هي (0.5,0.25,0.05) وثلاث حجوم للعينات (training data) (nT= 25,50,100) لكل من التوافيق السابقة .
تم اجراء المعاينة باستخدام برنامج R حيث تم كتابة خوارزمية Gibbs Sampler مع خوارزمية Metropolis Hasting من قبل الباحثان.
وبغية التوصل الى تقديرات دقيقة الى متجه معلمات النموذج فقد تم تكرار عملية التقدير 6000 من المرات ومن ثم تم حساب المتوسط لهذه التقديرات.
الجدول رقم (1) يبين القيم التقديرية للمعلمات المحسوبة بطريقة Bayesian Double Adaptive elastic net (BDAE) وطريقة Bayesian Adaptive Ridge Regression (BARR)
كما تم حساب قيمة الوسيط لمتوسط انحرافات القيم (median of mean absolute deviations (MMAD)) والانحرافات الخاصة به ولكل طريقة وباستخدام حجم عينة (np=200) وكما في الجدول (2)

حيث ان [12]:

$$MMAD = median(1\backslash n\sum_{i=1}^{n} |x_i'\hat{\beta} - x_i'\beta|) \ldots..(53)$$

اظهرت النتائج الاداء الافضل للطرق المقترحة وفي جميع الاقسام (quantile) المطبقة كما بينت النتائج تفوق الاداء لطريقة (BARR) عند حجم العينة (n=25) .

جدول رقم (1)

يمثل تقدير معلمات النموذج حسب الطريقة المستخدمة وحجم العينة (nT) والمجزئ (p) التقديرات تمثل متوسط نتائج 50 محاكاة لكل طريقة

p= 0.5 , n=25		
true β	الطريقة(method)	
	BDAE	BARR
3	3.002513	3.255978
1.5	1.754567	1.481159
0	0.050304	-0.06476
0	0.01314	-0.00107
2	2.056504	2.178079
0	0.081236	-0.12452
0	-0.03879	-0.13469
0	-0.09458	0.245194
p=0.5 , n=50		
true β	BDAE	BARR
3	3.272112	3.326375
1.5	1.728292	1.580214
0	-0.26315	-0.0054
0	0.342441	-0.00595
2	2.054548	2.235399
0	0.059034	-0.00709
0	-0.07489	-0.02066
0	-0.00666	0.075151
p=0.5 , n=100		
true β	BDAE	BARR
3	3.321842	3.271004
1.5	1.565429	1.742753
0	0.053755	-0.09188
0	-0.09621	0.132202
2	2.242391	2.057292
0	0.107652	0.033527
0	-0.09585	0.003438
0	0.009927	0.035986

p=0.25 , n=25		
BARR	**BDAE**	**true β**
2.859276	2.496226	3
1.603921	1.852084	1.5
0.089983	0.165746	0
-0.0433	-0.13798	0
1.989846	1.89814	2
-0.04624	-0.21566	0
0.145285	0.277647	0
-0.05007	-0.0719	0
p=0.25 , n=50		
BARR	**BDAE**	**true β**
3.004314	2.987093	3
1.57725	1.536435	1.5
-0.10252	-0.0274	0
0.121012	0.079822	0
2.057161	2.059255	2
-0.16777	0.019036	0
-0.00971	-0.1811	0
0.009584	-0.01393	0
p=0.25, n=100		
BARR	**BDAE**	**true β**
3.066014	3.019492	3
1.483461	1.600811	1.5
0.067002	-0.03799	0
-0.03939	-0.0167	0
2.019989	1.967229	2
0.037452	0.110229	0
0.010252	-0.01309	0
-0.02231	0.046191	0

n=25	p=0.05	
BARR	BDAE	true β
2.881747	2.876165	3
1.492243	1.414894	1.5
0.016536	-0.042	0
0.156494	0.412876	0
1.77764	1.513609	2
-0.07533	0.097745	0
0.581143	0.590592	0
-0.10106	-0.33621	0
n=50	p=0.05	
BARR	BDAE	true β
2.856937	2.699399	3
1.325482	1.482276	1.5
-0.13531	0.355142	0
0.230385	-0.10532	0
1.968103	1.701481	2
-0.31263	-0.00216	0
-0.01572	0.060392	0
0.098547	-0.1923	0
n=100	p=0.05	
BARR	BDAE	true β
2.819245	2.786355	3
1.403708	1.614718	1.5
-0.00444	-0.02971	0
0.002649	0.117965	0
1.76173	1.645806	2
0.0744	0.026191	0
0.10811	-0.12225	0
-0.0281	0.240398	0

جدول (2)

قيم الوسيط لمتوسط مطلق الانحرافات (MMAD) المحسوب لـ 50 محاكاة ولكل طريقة مستخدمة مع قيم الانحرافات المرافقة لها sd(MAD) وحسب احجام العينات (np) وقيمة القسيم (p)

true β =(3,1.5,0,0,2,0,0,0)			
method	MMAD	sd	np
BDAE	0.610536	0.165495	n=25
BARR	0.882772	0.292619	p=0.5
BDAE	0.608003	0.153025	n=50
BARR	0.480908	0.115002	p=0.5
BDAE	0.468621	0.102119	n=100
BARR	0.488226	0.095981	p=0.5
BDAE	0.720622	0.218525	n=25
BARR	0.57124	0.211808	p=0.25
BDAE	0.449213	0.130204	n=50
BARR	0.444538	0.131189	p=0.25
BDAE	0.331767	0.082318	n=100
BARR	0.265173	0.106452	p=0.25
BDAE	0.866854	0.174554	n=25
BARR	0.73879	0.238725	p=0.05
BDAE	0.690744	0.143268	n=50
BARR	0.625119	0.115564	p=0.05
BDAE	0.445332	0.086986	n=100
BARR	0.424043	0.105315	p=0.05

رابعا :الاستنتاجات والتوصيات:

في هذه الدراسة تم اقتراح النموذج الهرمي البيزي بتقنية elastic net المكيفة المضاعفة (Double Adaptive elastic net) وطريقة انحدار الحرف المكيفة (Adaptive Ridge Regression) لمناقشة مسالة تقدير واختيار المتغيرات في نموذج TQR بوجود مشكلة التعدد الخطي حيث تم بناء خوارزمية Gibbs sampler مع خوارزمية Metropolis Hasting من قبل الباحثان لا جراء المعاينة من التوزيعات الشرطية اللاحقة.

تم استخدام اسلوب المحاكاة وتحت افتراض وجود مشكلة التعدد الخطي والالتواء في توزيع الخطأ العشوائي (الالتواء الموجب) للنموذج المفترض في توليد البيانات من اجل بيان كفاءة الطريقتين المقترحتين .

اظهرت الدراسة وتحت الافتراضات السابقة الاداء الافضل للطرق المقترحة .

هناك بعض التساؤلات التي تطرحها هذه الدراسة والتي تحتاج الى دراسات مكثفة للإجابة الدقيقة عليها:

اولا: كما مر علينا في الجانب النظري ان الطريقة البيزية المستخدمة في تحليل نماذج Quantile Regression تستند الى افتراض ان توزيع الخطأ للنموذج هو التوزيع الملتوي للابلاس (Asymmetric Laplace) وبغض النظر عن التوزيع الحقيقي للبيانات فهل ان افتراض حالة الالتواء في التوزيع الحقيقي للبيانات (التواء سالب او موجب) مع وجود مشكلة التعدد الخطي له اثرا على دقة الطريقة البيزية المستخدم في تقدير المعلمات لنموذج Tobit Quantile Regression بشكل خاص ولنماذج Quantile Regression بشكل عام .

ثانيا: هل ان دقة تقدير النموذج تتباين في الاقسام او المجزئات (quantile) المختلفة لتوزيع متغير الاستجابة عندما يكون التوزيع الحقيق للبيانات توزيعا ملتويا .

المصادر(Reference)

1- Alhamzawi Rahim (2014) "Bayesian Elastic Net Tobit Quantile Regression " Communication in statistics-Simulation and Computation 45, Issue 7, 2409-2427.

2- Alhamzawi Rahim & Kemming Yu (2012)" Bayesian adaptive Lasso quantile regression" SAGE , 12, Issue 3.

3- Arthur E Hoerl and Robert W Kennard (1970) ." Ridge regression: Biased estimation for nonorthogonal problems". Technometrics, 12, No. 1, 69-82.

4- DeborahGefang (2014)" Bayesian doubly adaptive elastic-net Lasso for VAR shrinkage"International Journal of Forecasting. 30, Issue 1, pp.1-11

5- Fadel Hamid Hadi Alhusseini (2017) " New Bayesian Lasso in Tobit Quantile Regression" Revista Română de Statistică - Supplement nr. 6 / 2017 .pp.213-229.

6- Hidio Kozumi & Genya kobayshi(2011) "Gibbs Sampling Methods for Bayesian Quantile Regression" Journal of Statistical Computation and Simulation. 81, No.11, 1565–1578.

7- Hui Zou & Hao Helen Zhang (2009) "On The Adaptive elastic net with A diverging Number of Parameters " The Annals of Statistics, 37,No.4,1733-1751

8- Hui Zou & Trevor Hastie (2005) "Regularization and variable selection via the elastic net"J. R. Statist. Soc. 67, Part 2, 301–320.

9- Keming Yu & Julian Stander (2007)" Bayesian analysis of a Tobit quantile regression model" Econometrics. 137, issue 1, 260-276
10- Kemming Yu & Rana A.Moyeed (2001)" Bayesian Quantile Regression" Statistical & Probability Letters 54,437-447.
11 -Pereira, Basto & Silva (2016) "The logistic lasso and ridge regression in predicting corporate failure" Procedia Economics and Finance 39, 634 – 641.
12 - Qing Li, Rubin Xi &Nan Lin (2010) "Bayesian Regularized Quantile regression" Bayesian Analysis 5, No. 3, 533-556.
13-Tatsuya Kubokawa & Muni S. Srivastava (2004) " Improved Empirical Bayes Ridge Regression Estimators Under Multicollinearity" Communications in Statistics - Theory and Methods. 33, Issue 8, 1943-1973
14-Tibshirani(1996)"Regression Shrinkage and selection via lasso" J. R.Statis.Soc.58,No.1, 267-288

Bayesian Tobit Quantile Regression Model
Using Double Adaptive elastic net and Adaptive Ridge Regression

Abstract:

Recently Tobit Quantile Regression(TQR) has emerged as an important tool in statistical analysis . in order to improve the parameter estimation in (TQR) we proposed Bayesian hierarchical model with double adaptive elastic net technique and Bayesian hierarchical model with adaptive ridge regression technique .

in double adaptive elastic net technique we assume different penalization parameters for penalization different regression coefficients in both parameters $\lambda_1$and λ_2 , also in adaptive ridge regression technique we assume different penalization parameters for penalization different regression coefficients in parameter λ .

Simulation study was used for explain the efficiency of the proposed methods .The result illustrated the efficiency of the proposed methods for dealing with the estimation of parameters model in present of high correlation in explanatory variables .

This is the first work that is discussing the parameter estimation in TQR model with double adaptive elastic net and adaptive ridge regression.

Keyword: Tobit Quantile Regression , Adaptive elastic net , Adaptive Ridge Regression , Gibbs Sampler algorithm, Metropolis Hasting algorithm.