نموذج Tobit Quantile Regression البيزي باستعمال المستوى الرابع من التوزيعات الأولية

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

تاريخ التقديم:2017/10/3
تاريخ القبول:2017/12/6

الملخص

تم في هذا البحث مناقشة تقدير معلمات واختيار المتغيرات في نموذج (Tobit Quantile Regression) في ظل وجود مشكلة التعدد الخطي. ولأهمية تقنية (elastic net) في التعامل مع مشكلة التعدد الخطي واختيار المتغيرات في نماذج الانحدار فقد تم اعتماد هذه التقنية في الدراسة . تم في هذه الدراسة اقتراح النموذج الهرمي البيزي (Bayesian Tobit hierarchical model) في المستوى الرابع من التوزيعات الاولية بالاعتماد على البيانات وبافتراض ان كلا معلمتي التنظيم (tuning parameter) في تقنية elastic net هي متغيرات عشوائية ويتم تقديرها مع بقية المعلمات في النموذج . تم استخدام اسلوب المحاكاة في بيان كفاءة الطريقة المقترحة كما تم مقارنة الطريقة المقترحة مع بعض الطرائق الحديثة المستخدمة في تقدير معلمات نموذج (Tobit Quantile regression) واظهرت النتائج الاداء الافضل للطريقة المقترحة .

هذا هو العمل الاول (بحسب علم الباحث) الذي يتم فيه مناقشة تقدير واختيار المتغيرات لنموذج TQR باقتراح النموذج الهرمي البيزي في المستوى الرابع من التوزيعات الاولية وبافتراض ان كلا معلمتي التنظيم في تقنية elastic net هي متغيرات عشوائية .

المصطلحات الرئيسة للبحث/النموذج الهرمي البيزي ، معلمتي التنظيم .

مجلة العلوم
الاقتصادية والإدارية
العدد 105 المجلد 24
الصفحات 487ـ 513

*البحث مستل من اطروحة دكتوراه

1- المقدمة

ان موضوع الانحدار هو احد المواضيع الاحصائية المهمة المستخدمة في كثير من الدراسات العلمية وله تطبيقات واسعة في كثير من المجالات . والانحدار الاعتيادي او ما يطلق عليه احيانا انحدار المتوسط (Mean Regression) هو احد الطرائق الاحصائية المهمة التي تستخدم في دراسة العلاقة بين المتغيرات التوضيحية X ومتغير (او متجه متغيرات) الاستجابة Y وذلك من خلال نقطة المتوسط في توزيع متغير الاستجابة، بكلام اخر انه في تحليل انحدار المتوسط الاعتيادي فان الاهتمام يكون منصب حول تقدير المتوسط الشرطي لتوزيع متغير الاستجابة {E(Y/X)}.

وكما هو معروف ان تحليل الانحدار الاعتيادي يستند الى فروض التحليل التي من ابرزها ان الاخطاء العشوائية تتوزع بصورة مستقلة توزيعا طبيعيا بمتوسط مساوي للصفر وتباين ثابت هو $\sigma^2 \mathrm{I}$ ، I: is identity matrix اي ان : $\varepsilon \sim N(0, \sigma^2 \mathrm{I})$.

حديثا ظهرت تقنية Quantile Regression التي اعطت دراسة اكثر شمولا للعلاقة بين متغير الاستجابة والمتغيرات التوضيحية وذلك من خلال تقدير الاقسام الشرطية $\{Q_p(Y/X), 0 < p < 1\}$ المختلفة في توزيع متغير الاستجابة بدلا من الاقتصار على تقدير التوقع الشرطي {E(Y/X)} كما هو الحال في انحدار المتوسط الاعتيادي . وتقنية Quantile Regression تستخدم عندما يراد تقدير العلاقة في اقسام مختلفة من التوزيع الشرطي لمتغير الاستجابة .

ولأهمية موضوع Quantile Regression في دراسة العلاقة بين متغير الاستجابة والمتغيرات التوضيحية في كثير من الدراسات وللمميزات المهمة التي ينفرد بها عن الانحدار الاعتيادي نجد ان الدراسات الحديثة تولي اهتماما كبيرا لهذا الموضوع وفي كثير من المجالات، كما ان تقنية Quantile Regression في المجالات التطبيقية تكون ملائمة لكثير من البيانات حيث ان هذه التقنية لا تشترط الحالة الطبيعية او التماثل في توزيع البيانات .

ومن المميزات المهمة لتقنية Quantile Regression هو عدم التأثر بمشكلة عدم تجانس التباين والقيم الشاذة والالتواء في التوزيع كما ان لهذه التقنية اهمية في الكشف عن مشكلة عدم تجانس التباين في الانحدار البسيط عن طريق الرسم .وبالرغم من هذه المميزات المهمة لهذه التقنية الا انه تبقى مشكلة التعدد الخطي من المشاكل المهمة التي لها تأثير واضح على كفاءة المقدرات في نموذج Quantile Regression وتزداد هذه المشكلة صعوبة عندما يكون متغير الاستجابة من النوع المحدود Limited response variable كما هو الحال في نموذج Tobit Quantile Regression .

من ناحية اخرى ان مسالة اختيار المتغيرات (variable selection) في نماذج الانحدار لها اهمية كبيرة في دقة التنبؤ (prediction accuracy) وفي ايجاد نماذج تفسيرية (interpretation models) للظاهرة قيد الدراسة.

ان الطرائق الشائعة المستخدمة في اختيار المتغيرات و تحسين دقة التنبؤ في النموذج (مثل طريقة ridge regression وطريقة subset selection وطريقة stochastic search variable selection (SSVS) هي طرائق يعتريها بعض العيوب فمثلا ان طريقة ridge regression لها اهمية في تحسين دقة التنبؤ من خلال تقليص (shrink) معاملات الانحدار لكن هذه العملية لا تجعل المعاملات الغير مهمة مساوية للصفر لذا فهي لا تعطي نموذج تفسيري (interpretation models) للظاهرة قيد الدراسة. وبالرغم من ان تقنية subset selection لها اهمية في الوصول الى نموذج تفسيري الا انها تعد طريقة اقل استقرارا من ridge regression فالتغييرات البسيطة في البيانات قيد الدراسة قد تؤدي الى اختيار نماذج مختلفة بشكل كبير وهذا الامر يؤثر في دقة التنبؤ لهذه الطريقة. كذلك ان طريقة (SSVS) تحتاج الى عمليات حسابية كبيرة مع ازدياد عدد المتغيرات التوضيحية فضلا عن ان زيادة عدد المتغيرات قد يؤدي الى عدم اختيار النموذج الامثل لتمثيل البيانات قيد الدراسة.

الدراسات الحديثة اظهرت ان التنظيم (Regularized) في نماذج الانحدار له اهمية كبيرة في تحسين دقة التنبؤ وتقنية elastic net التي وضعت من قبل (Zou & Hastie) في عام 2005 هي احد اشكال التنظيم التي اظهرت الدراسات ان لها اهمية كبيرة في التقدير واختيار المتغيرات عند وجود مشكلة التعدد الخطي في نماذج الانحدار .

ومن ناحية اخرى ان الاسلوب البيزي في تقنية elastic net له اهمية كبيرة في التعامل مع النماذج المحدودة (مثل Tobit Quantile Regression) ففي الطرق الاعتيادية ان ايجاد الخطأ المعياري (standard error) لمقدرات elastic net في هذه النماذج هو غير ممكن . في حين ان الطرق البيزية تمكننا من ذلك .

كذلك ان الاسلوب البيزي له اهمية في الوصول الى الاستدلال الدقيق حتى في حالة العينات الصغيرة وله اهمية ايضا في ايجاد (credible interval). حديثا تم اقتراح طرائق تقليدية وبيزية لتقدير معلمات نموذج Tobit Quantile regression بوجود مشكلة التعدد الخطي.

يهدف هذا البحث الى اقتراح النموذج الهرمي البيزي بتقنية elastic net وفي المستوى الرابع من التوزيعات الاولية لتقدير معلمات واختيار المتغيرات التوضيحية في نموذج Tobit Quantile regression وبوجود مشكلة التعدد الخطي ويتم مقارنة الطريقة المقترحة مع بعض الطرائق التي استخدمت في الدراسات الحديثة.

2- **نموذج** Tobit Quantile regression :

نموذج Tobit هو احد النماذج الاحصائية المهمة الذي اقترح من قبل James Tobin في عام 1958 يصف هذا النموذج العلاقة بين متغير الاستجابة y_i (الغير سالب) ومتغير او مجموعة من المتغيرات التوضيحية x_i ، يمكن النظر لهذا النموذج على انه حالة خاصة من نموذج الانحدار المبتور (censored regression) .

ان متغير الاستجابة في نموذج توبت احيانا قد يكون مقيد بشكل اخر (بدل من عدم سلبية المتغير) فقد يكون هذا المتغير اما مقيد بحدود دنيا كما في حالة وجود حد ادنى للأجور او قيود عليا كما في حالة وجود سقف للأسعار او كليهما[1].

ونموذج Tobit Quantile regression الذي هو امتداد لنموذج Tobit (1958) يمكن ان ينظر اليه كنموذج Linear Quantile regression مع متغير استجابة مستمر لاتيني (latent continuous response variable (y^*)) الذي يكون مشاهد بصورة غير كاملة[8] لنفرض ان هناك عينة من المشاهدات المستقلة $y=(y_1,y_2,y_3,\dots y_n)$ المرتبطة بالمتغيرات التوضيحية $x = (x_1, x_2, x_3, \dots x_n)$ حيث ان x_i هو متجه لـ k من المتغيرات التوضيحية

لذا فنموذج Tobit Quantile regression يمكن ان يعبر عنه بالشكل الاتي [2]:

$$y_i^* = \acute{x}_i\beta_p + \varepsilon_i \quad \text{and} \quad y_i = \max\{y^0, y_i^*\} \ldots\ldots\ldots\ldots (2\text{-}1)$$

حيث ان:

y^0 : هي نقطة المراقبة المعلومة

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

β_p : هي متجه المعلمات المجهولة * التي يمكن ان تقدر بشكل متسق من المعادلة الاتية[2]:

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

$$\min_\beta = \sum_{i=1}^{n} \rho_p(y_i - \max\{y^0, \acute{x}_i\beta\}) \ldots\ldots\ldots (2-2)$$

حيث ان ρ_p يمكن ان يعبر عنها بالشكل التالي [7]:

$$\rho_p(u) = \begin{cases} pu & if u \geq 0 \\ -(1-p)u & if u < 0 \end{cases} \quad ......(2\text{-}3)$$

3- **دالة الامكان والاسلوب البيزي في تقدير معاملات نموذج** Tobit Quantile Regression :

مما لاشك فيه ان احد الدوال المهمة في الاسلوب البيزي لتقدير معاملات نماذج Quantile Regression هي دالة الامكان، حيث انها تمثل احد الموارد المهمة للمعلومات عن المعاملات المجهولة في النموذج ، لذا فتعيين دالة الامكان خطوه مهمة في الشروع بعملية التقدير. ولما كانت تقنية Quantile Regression لا تستند في تحليلها الى شكل محدد لتوزيع الخطأ لذا فان عملية تعيين دالة الامكان هو غير ممكن.

ولهذا السبب نجد ان استخدام الاسلوب البيزي في تقدير معاملات نماذج Quantile Regression بقي بعيدا عن الاستخدام في الدراسات والبحوث التطبيقية الى سنة 2001 حيث افترض Yu & Moyeed (وبغض النظر عن التوزيع الحقيقي للبيانات) توزيع لابلاس الملتوي (Asymmetric Laplace Distribution) لدالة الامكان في نموذج Linear Quantile regression[6]، وفي الحقيقة ان عملية اختيار هذه الدالة لم يكن الا افتراضيا من اجل ربط الطريقة البيزية في التقدير الى الطريقة الكلاسيكية وذلك بملاحظة ان تصغير دالة الخسارة التحقيقية (check loos function) المستخدمة في الطريقة الكلاسيكية لتقدير معاملات النموذج يقابل تعظيم دالة الامكان باعتماد توزيع لابلاس الملتوي (AL) للخطأ العشوائي في نموذج Linear Quantile Regression . وعلى الرغم من اهمية هذا الافتراض في تطبيق الاسلوب البيزي في التقدير الا انه ظهرت هناك صعوبات كثيرة في الوصول الى التوزيعات اللاحقة (posterior distribution) ولكن الباحثان {Yu & Moyeed (2001)} اشارو الى امكانية استخدام طريقة MCMC في الوصول الى النتائج فطريقة MCMC لها اهمية في الوصول الى التوزيعات اللاحقة حتى في الحالات المعقدة[6] .

دالة الكثافة الاحتمالية pdf لتوزيع لابلاس الملتوي (AL) يمكن ان يعبر عنها بالشكل الاتي [7] :

$$f(\varepsilon \backslash \tau) = p(1-p)\tau \exp[-\tau\, \rho_p(\varepsilon)] \ldots\ldots\ldots (3-1)$$

وبالاستناد الى المصدر(5) وتحت افتراض توزيع الأخطاء في (3-1) ان دالة الامكان للنموذج (2-1) يمكن ان يعبر عنها بالشكل الاتي :

$$f(y \backslash \boldsymbol{X}, \beta, \tau) = p^n(1-p)^n\tau^n \exp\{-\tau \textstyle\sum_{i=1}^{n} \rho_p(y_i - \max(y^0, \acute{x}_i\beta)) \ldots (3-2)$$

يمكن ملاحظة ان تصغير المعادلة (2-2) المستخدمة في الاسلوب الكلاسيكي لتقدير معاملات نموذج Tobit Quantile Regression يكافئ تعظيم دالة الامكان في (3-2) .

ومن اجل تسريع وزيادة كفاءة اداء الاسلوب البيزي في تحليل نموذج Tobit Quantile Regression فقد استند الباحثان على اقتراح Kozumi & Kobayshi (2009)[3] في استخدام التمثيل المختلط لتوزيع لابلاس الملتوي (AL) للخطأ العشوائي في نماذج Quantile Regression.

لذا فان المتغير ىالعشوائي اللاتيني (y^*) latent variable في نموذج (2-1) وباستخدام التمثيل المختلط لتوزيع الخطأ يمكن ان يعبر عنه بالشكل الاتي[3]:

$$y_i^* = \acute{x}_i\beta + \xi_1 e_i + \tau^{-1/2}\xi_2\sqrt{e_i}\, z_i \ldots\ldots\ldots\ldots(3\text{-}3)$$

Where $\xi_1 = (1-2p)/p(1-p)$, $\xi_2^2 = 2/p(1-p)$

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

4- تقنية elastic net والدالة الاولية لمتجه المعلمات β:

بشكل عام ان التنظيم في نماذج **Quantile regression** له اهمية كبيرة في تحسين دقة التنبؤ[7] وتقنية **elastic net** هي احد اشكال الانحدار المنظم **(Regularized Regression)** المهمة التي اقترحت من قبل **(Zou &Hastie 2005)** كتوسعة لتقنية **lasso** والتي لها اداء محسن عن تقنية **lasso** وذلك من خلال معالجة مشكلة الارتباط الذاتي وامكانية اختيار اكثر من **n** من المتغيرات في حالة كون عدد المتغيرات التوضيحية **p> n** [7] . وتظهر اهمية تقنية **elastic net** عندما يكون هناك مجموعة من المتغيرات التوضيحية المرتبطة مع بعضها بشكل كبير فان هذه التقنية تساهم في جعل هذه المتغيرات سوية في داخل او خارج النموذج[4].

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 [7] :

$$f(\beta_k) = C(\delta_1,\delta_2)\,\frac{\delta_1}{2}\exp\{-\delta_1|\beta_k| - \delta_2\beta_k^2\}\text{.........(4-1)}$$

اذ 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_1norm penalty)**.

$\sum_{j=1}^{k}\beta_j^2$: يمثل **(L_2norm penalty)**.

ان الافتراض للدالة الاولية (4-1) اعطى اهمية ربط اسلوب التقدير الكلاسيكي مع الاسلوب البيزي في تقنية **elastic net** لتقدير معلمات نموذج **Linear Quantile regression** .

(Andrews and Mallows 1974) برهن ان لكل **d ≥ 0** فان المعادلة التالية تكون صحيحة[7] **(d** هي: متغير عشوائي) :

$$\frac{d}{2}\exp(-d|s|) = \int_0^\infty \frac{1}{\sqrt{2\pi a}}\exp\left(-\frac{s^2}{2a}\right)\frac{d^2}{2}\exp\left(\frac{-d^2}{2}a\right)da \ldots\ldots(4-2)$$

لذا فالدالة الاولية (4-1) يمكن ان يعبر عنها بالشكل التالي[7]:

$$f(\beta_k) = C(\delta_1,\delta_2)\int_0^\infty \frac{1}{\sqrt{2\pi a_k}}\exp\left\{-\frac{1+2\delta_2 a_k}{2a_k}\beta_k^2\right\}\frac{{\delta_1}^2}{2}\exp\left(-\frac{{\delta_1}^2}{2}a_k\right)da_k\text{...(4-3)}$$

كما برهن **(Li & Lin 2010)** ان قيمة $C(\delta_1,\delta_2)$ يمكن ايجادها بالشكل التالي:[7]

let $r_k=1+2\,\delta_2 a_k$**.....(4-4)**

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

$$\frac{{\delta_1}^2}{4\delta_2}\exp\left(-\frac{{\delta_1}^2}{4\delta_2}(r_k-1)dr_k\ldots\ldots(4-5)\right.$$

$$\int_{-\infty}^{\infty} f(\beta_k)\, d\beta_k =$$
$$C(\delta_1,\delta_2)\int_1^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi(r_k-1)\backslash(2\delta_2 r_k)}}\exp\left\{-\frac{1}{2}\left(\frac{r_k-1}{2\delta_2 r_k}\right)^{-1}\beta_k^2\right\}d\beta_k \times$$
$$\frac{{\delta_1}^2}{4\delta_2} r_k^{-1\backslash 2}\exp\left(-\frac{{\delta_1}^2}{4\delta_2}(r_k-1)\right)dr_k \;\ldots.(4-6)$$
$$\int_{-\infty}^{\infty} f(\beta_k)\, d\beta_k = C(\delta_1,\delta_2)\int_1^{\infty}\frac{{\delta_1}^2}{4\delta_2} r_k^{-1\backslash 2}\exp\{-\frac{{\delta_1}^2}{4\delta_2}(r_k-1)\}dr_k \ldots(4-7)$$
$$\int_{-\infty}^{\infty} f(\beta_k)\, d\beta_k = 1 \quad \textit{by Fubini's theorem}$$
$$1 = C(\delta_1,\delta_2)\int_1^{\infty}\frac{{\delta_1}^2}{4\delta_2} r_k^{-1\backslash 2}\exp\{-\frac{{\delta_1}^2}{4\delta_2}(r_k-1)\}dr_k \ldots(4-8)$$

Let $\frac{{\delta_1}^2}{4\delta_2} = \tilde{\delta_1} \;\ldots.(4-9)$

وبشكل اكثر تفصيلا*:

$$1 = C(\delta_1,\delta_2)\int_1^{\infty}\tilde{\delta_1} r_k^{-1\backslash 2}\exp\{\tilde{\delta_1}\}\exp\{-\tilde{\delta_1}(r_k)\}dr_k \;\ldots.(4-10)$$

…(4-11) $let\, \tilde{\delta_1} r_k = r$

$$|\mathrm{J}| = 1\backslash\tilde{\delta_1} \ldots (4-12)$$
$$1 = C(\delta_1,\delta_2)(\tilde{\delta_1})^{1\backslash 2}\exp\{\tilde{\delta_1}\}\int_{\tilde{\delta_1}}^{\infty} r^{-1\backslash 2}\exp\{-r\}\, dr \;\ldots..(4-13)$$

حيث ان:

(upper incomplete gamma function) $\int_{\tilde{\delta_1}}^{\infty} r^{-1\backslash 2}\exp\{-r\}\, dr = \Gamma(1\backslash 2, \tilde{\delta_1})$ **…(4-14)**

$$\Rightarrow C(\delta_1,\delta_2) = \Gamma^{-1}(1\backslash 2, \tilde{\delta_1})(\tilde{\delta_1})^{-1\backslash 2} exp(-\tilde{\delta_1}) \;\ldots.(4-15)$$

في نموذج Tobit Quantile Regression (2-1) الاسلوب الكلاسيكي لتقنية elastic net في تقدير معاملات النموذج يمكن ان يعبر عنها بالشكل التالي[2] :

$$min_{\beta}\sum_{i=1}^{n}\rho_p(y_i - \max\{y^0, \acute{x}_i\beta\}) + \delta_1\sum_{j=1}^{k}|\beta_j| + \delta_2\sum_{j=1}^{k}\beta_j^2 \;\ldots\ldots\ldots(4-16)$$

الحمزاوي (2014) استخدم الدالة الاولية (1-4) لتقدير معاملات نموذج Tobit Quantile Regression بملاحظة ان تعظيم الدالة اللاحقة posterior الناتجة من حاصل ضرب الدالة الاولية في (1-4) ودالة الامكان في (2-3) τ(1=) يكافئ تصغير المعادلة (16-4). افترض في هذه الدراسة ان معلمة التنظيم δ_1 هي متغير عشوائي اما δ_2 فهي مقدار ثابت .

***: التوسع في الاشتقاق هو من قبل الباحثان.**
يقترح الباحثان استخدام الدالة الاولية (1-4) في النموذج الهرمي المستخدم في الدراسة لتقدير معلمات نموذج Tobit Quantile Regression بوجود مشكلة التعدد الخطي مع افتراض ان كلا معلمتي التنظيم هي متغير عشوائي و التوسع في اختيار الدوال الاولية الملائمة للمعلمات الفوقي بغية التوصل الى مقدرات كفوءة.

5- **النموذج الهرمي البيزي** (Bayesian hierarchical model) **المقترح في الدراسة**:

النموذج الهرمي المقترح في الدراسة يستند الى النموذج المستخدم من قبل **Li & Lin (2010)** في تقدير معلمات نموذج **Linear Quantile Regression** [7]

قام الباحثان بالتوسع في اختيار الدوال الاولية الملائمة لتقدير معلمات نموذج **Tobit Quantile Regression** بوجود مشكلة التعدد الخطي بافتراض ان المعلمات الفوقية هي متغيرات عشوائية ويتم تقديرها مع بقية المعلمات في النموذج وللمستوى الرابع من التوزيع الاولي .

ليكن :

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

النموذج الهرمي يمكن ان يعبر عنه بالتالي:

$$y_i = \max\{y^0, y_i^*\}$$

$$y_i^* = x'_i\beta + \xi_1 e_i + \tau^{-1/2}\xi_2\sqrt{e_i}\ z_i$$

$$\boldsymbol{e}\backslash\tau \sim \prod_{i=1}^{n} \tau \exp(-\tau\, e_i) \qquad \textbf{(5-1)}$$

$$\boldsymbol{Z} \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} z_i^2)$$

$$(\beta_k \backslash r_k, \delta_2) \overset{ind}{\sim} \frac{1}{\sqrt{2\pi(r_k-1)\backslash(2\delta_2 r_k)}} \exp\left\{-\frac{1}{2}\left(\frac{r_k-1}{2\delta_2 r_k}\right)^{-1} \beta_k^2\right\}$$

$$(r_k \backslash \tilde{\delta}_1) \overset{ind}{\sim} \Gamma^{-1}(1\backslash 2, \tilde{\delta}_1) r_k^{-1\backslash 2} \tilde{\delta}_1^{1\backslash 2} \exp\{-\tilde{\delta}_1\, r_k\} , I(r_k > 1)$$

$$\tau \sim \tau^{-1}$$

$$(\tilde{\delta}_1 \backslash \theta_1, \alpha_1) \sim \frac{\theta_1^{\alpha_1}}{\Gamma(\alpha_1)} \tilde{\delta}_1^{\alpha_1 - 1} \exp(-\theta_1 \tilde{\delta}_1)$$

$$\delta_2 \sim \theta_3 \exp(-\delta_2 \theta_3)$$

$$(\theta_1, \alpha_1) \sim \theta_2 \exp(-\theta_2 \theta_1)$$

$$\theta_2 \sim \exp(-\theta_2)$$

$$\theta_3 \sim \theta_3^{-1}$$

حيث ان : $(\theta_3, \theta_2, \theta_1, \alpha_1)$ تمثل المعلمات الفوقية **(hyper parameter)** في النموذج.

6- **اجراء المعاينة البيزية للمتغيرات العشوائية** Bayesian sampling of the random variables :

بغية تقدير المعلمات المجهولة واجراء اختيار المتغيرات في النموذج الهرمي البيزي المقترح في الدراسة فقد تم أجراء المعاينة من التوزيعات الشرطية الكاملة **(full conditional distributions)** في الملحق (1) وباستخدام الاسلوب البرمجي بطريقة **Markov Chain Monte Carlo (MCMC)** .

قام الباحثان و باستخدام برنامج R بكتابة الخوارزميات الخاصة بالدراسة حيث تم استخدام خوارزمية Gibbs Sampler في اجراء المعاينة من المتغيرات العشوائية التي لها توزيعات احتمالي معروفة اما بقية المتغيرات فقد تم استخدام خوارزمية Metropolis Hasting في اجراء المعاينة لها وكما في الخطوات التالية:

1- معاينة المتغير العشوائي: y^* , i=1,...n

باستخدام الدوال الخاصة في برنامج R تم توليد قيم عشوائية من y_i^* باعتماد الدالة الشرطية الكاملة (full conditional distributions) التالية[3]:

$$y_i^*\backslash \boldsymbol{y},\beta,\boldsymbol{e},\tau \sim y_i I(y_i>0)+\mp N_{(-\infty,0]}\left(\acute{x}_i\beta+\xi_1 e_i,{\xi_2}^2\tau^{-1}\, e_i\right) I(y_i=0) \ldots (6-1)$$

$\mp N_{(-\infty,0]}(\mu,\sigma^2)$ يرمز الى التوزيع الطبيعي المبتور **truncated normal distribution**

بمتوسط $\mu = \acute{x}_0\beta + \xi_1 e_0$ وتباين $\sigma = {\xi_2}^2 e_0/\tau$ حيث ان x_0 ، e_0 :تمثل قيم مصفوفة x و قيم متجه e المقابلة للأصفار في متجه y.

2- معاينة المتغير العشوائي β_j , j=1,..k:

تم اجراء المعاينة باستخدام دالة التوزيع الشرطي الكاملة للمتغير العشوائي β_k حيث وجد ان التوزيع الشرطي للمتغير العشوائي β_j يتبع التوزيع الطبيعي بمتوسط Me_k وتباين V_k^2

حيث ان :

$$w_{ik}=y_i^*-\xi_1 e_i-\sum_{j=1,j\neq k}^{p} x_{ij}\beta_j \;\ldots.(6-2)$$

$$V_k^{-2}={\xi_2}^{-2}\tau\sum_i^n(x_{ik}^2\backslash e_i)+2\delta_2 r_k(r_k-1)^{-1} \quad \ldots.(6-3)$$

$$\mathrm{Me_k}=\mathrm{V_k^2}{\xi_2}^{-2}\tau\sum_{i=1}^{n}\mathrm{w_{ik}}(\mathrm{x_{ik}}\backslash \mathrm{e_i}) \;\ldots.(6-4)$$

3- معاينة المتغير العشوائي e_i , i= 1,...n :

تم اجراء المعاين للمتغير العشوائي e_i حيث وجد ان التوزيع الشرطي للمتغير العشوائي e_i هو generalized inverse Gaussian $GIG(\eta_1,\eta_2)$.

$$\eta_1=\sqrt{\frac{\eta_2}{\tau{\xi_2}^{-2}(y_i^*-\acute{x}_i\beta)^2}} \quad ,\eta_2=\tau{\xi_1}^2{\xi_2}^{-2}+2\tau \quad \ldots.(6\text{-}5)$$

الدالة الاحتمالية المستخدمة في توليد المتغير العشوائي e_i^{-1} هي [2] :

$$f(x\backslash\eta_1,\eta_2)=\sqrt{\frac{\eta_2}{2\pi}}\; x^{-3\backslash 2}\exp\{-\eta_2(x-\eta_1)\backslash 2{\eta_1}^2 x\}\;\; x>0 \ldots (6-6)$$

4- معاينة المتغير العشوائي (r_j-1) , j=1,...k:

تم استخدام المعادلة (6-6) في توليد المتغير العشوائي($r_j^{-1}-1$) وكما في الخطوة السابقة حيث ان التوزيع الشرطي للمتغير العشوائي هو GIG ايضا وبالمعلمات:

$$\eta_1=\sqrt{\frac{2\eta_2}{2\delta_2\beta_j^2}} \quad ,\eta_2 \;=\tilde{\delta_1}\ldots\ldots\ldots(6\text{-}7)$$

5- معاينة المتغير العشوائي τ:

تم استخدام دالة التوزيع الشرطية الكاملة للمتغير العشوائي τ في اجراء المعاينة حيث تم البرهنة ان التوزيع الشرطي للمتغير العشوائي هو **$Gama(\omega_1, \varphi_1)$.**

حيث ان :

$$\omega_1 = 3n\backslash 2 \ , \ \varphi_1 = \sum_{i=1}^{n} (y_i^* - \acute{x}_i\beta_p - \xi_1 e_i)^2 \backslash 2\xi_2{}^2 e_i + e_i \ \ldots (6-8)$$

6- معاينة المتغير العشوائي $\widetilde{\delta_1}$:

يمكن ملاحظة ان الدالة الشرطية للمتغير العشوائي $\widetilde{\delta_1}$ في الملحق (1)لا تشابه اي من التوزيعات الاحتمالية القياسية لذا فقد تم استخدام خوارزمية **Metropolis Hasting** في اجراء المعاينة للمتغير العشوائي $\widetilde{\delta_1}$.
الخطوات الخاصة باستخدام خوارزمية **Metropolis Hasting** تم باعتماد التوزيع المقترح الاتي [7] :

$$g(\widetilde{\delta_1} \backslash \boldsymbol{r}) \propto \widetilde{\delta_1}^{p+\alpha_1-1} \exp\{-\widetilde{\delta_1}[\theta_1 + \sum_{k=1}^{p}(r_k - 1)]\} \ldots . (6-9)$$

7- معاينة المتغير العشوائي δ_2:

تم اجراء المعاينة من التوزيع الشرطي للمتغير العشوائي الذي وجد ان التوزيع الشرطي الكامل له هو **$Gamma(\omega_2, \varphi_2)$.**

حيث ان :

$$\omega_2 = (p\backslash 2) + 1 \ , \varphi_2 = \sum_{k=1}^{p} r_k (\ r_k - 1)^{-1} \beta_k^2 + \theta_3 \ldots \ldots . (6-10)$$

8- معاينة المتغير العشوائي θ_1 :

بنفس اسلوب الفقرة السابقة حيث ان التوزيع الشرطي الكامل للمتغير هو

$Gamma(\omega_3, \varphi_3)$

$$\omega_3 = \alpha_1 + 1 \ , \ \varphi_3 = \theta_2 + \widetilde{\delta_1} \ldots \ldots (6-11)$$

9- معاينة المتغير العشوائي α_1:

من الملحق (1) دالة التوزيع الشرطي للمتغير α_1 هي:

$$\frac{\theta_1^{\alpha_1}}{\Gamma(\alpha_1)} \widetilde{\delta_1}^{\alpha_1 - 1}$$

ولصعوبة استخدام خوارزمية **Gibbs Sampler** في اجراء المعاينة من الدالة اعلاه فقد تم استخدام خوارزمية **(adaptive rejection sampling)** في المصدر **(hhh)** لاجراء المعاينة في المتغير العشوائي.

10- معاينة المتغير العشوائي θ_2 :

بنفس اسلوب الفقرة السابعة حيث ان التوزيع الشرطي الكامل للمتغير هو

Gamma (ω_4, φ_4)

$$\omega_4 = 2, \varphi_4 = \theta_1 + 1 \ldots \ldots . (6-12)$$

11- معاينة المتغير العشوائي θ_3 :

بنفس اسلوب الفقرة السابقة حيث ان التوزيع الشرطي الكامل للمتغير العشوائي هو

$$\text{Gamma}(\omega_5, \varphi_5)$$

$$\omega_5 = 1 , \varphi_5 = \delta_2 \dots \dots (6-13)$$

7- **الجانب التجريبي**:

تم في هذه الدراسة استخدام اسلوب المحاكاة لبيان كفاءة الطريقة المطبقة مع اجراء مقارنة مع طريقة (Bayesian elastic net Tobit Quantile Regression (Alhamzawi 2014) والطريقة القياسية (QR) .

تم توليد n من المشاهدات باستخدام النموذج التالي:

$$y_i = \max\{\boldsymbol{X}\beta + U, 0\} \dots . (7-1)$$

حيث ان :

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 $0.5^{|i-j|}$.

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

في هذا المثال تم افتراض ثلاث حالات من القيم لمتجه المعلمات الحقيقية وكالتالي :

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

$$\beta = (0.85,0.85,0.85,0.85,0.85,0.85,0.85,0.85)$$

$$\beta = (5,0,0,0,0,0,0,0)$$

وبغية تقدير كل متجه من 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 خوارزمية Gibbs Sampler مع خوارزمية Metropolis Hasting (13000 iteration, 3000 burn)، حيث تم الحصول على (10000) تقدير لمتجه المعلمات ومن ثمة تم حساب المتوسط لهذه التقديرات كما تم تكرار العملية هذه 50مرة من التكرارات وايضا تم حساب المتوسط لهذه التكرارات من اجل التوصل الى افضل تقدير لمتجه المعلمات.

الجدول رقم (2-1) في الملحق (2) يبين القيم التقديرية للمعلمات محسوبة بالطرق السابقة الذكر.

حيث ان :

BEH: تمثل الطريقة المقترحة في الدراسة .

BER:تمثل الطريقة المقترحة من قبل (Alhamzawi 2014).

QR: تمثل الطريقة القياسية

كما تم حساب MMAD لكل طريقة وكما في الجدول (2-2) من الملحق (2) وباستخدام حجم عينة (np=200)

n=nT+np

حيث ان [7]:

$$MMAD = median(1 \backslash n \sum_{i=1}^{n} |\acute{x_\imath}\hat{\beta} - \acute{x_\imath}\beta|) \dots . . (7-2)$$

تم الاشارة الى القيم الدنيا بالخط الغامق.

كما تم رسم خط الانحدار لنموذج (TQR) باستخدام القيم الحقيقية للمتغير المعتمد والقيم المقدرة له بطريقتي (BEH,BRE) وكما في الشكل (1,2,3) من الملحق (3).

اظهرت النتائج الاداء الافضل للطريقة المقترحة وفي جميع الاقسام (quantile) المطبقة و جميع حجوم العينات .

الاستنتاجات والتوصيات:

في هذه الدراسة تم اقتراح النموذج الهرمي البيزي بتقنية elastic net وفي المستوى الرابع من التوزيع الاولي لمناقشة مسالة تقدير واختيار المتغيرات في نموذج TQR بوجود مشكلة التعدد الخطي حيث تم بناء خوارزمية Gibbs sampler مع خوارزمية Metropolis Hasting لا جراء المعاينة من التوزيعات الشرطية اللاحقة.

تم استخدام اسلوب المحاكاة وتحت افتراض التوزيع الطبيعي الملتوي للخطأ العشوائي (الالتواء الموجب) من اجل بيان كفاءة الطريقة المقترحة ومقارنتها مع طريقة Bayesian elastic net Tobit Quntile Qegression (Alhamzawi 2014) والطريقة القياسية (QR).

اظهرت الدراسة وتحت الافتراض السابق لتوزيع الخطأ وبوجود مشكلة التعدد الخطي الاداء الافضل للطريقة المقترحة .

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

اولا: عندما يكون التوزيع الحقيقي للبيانات التي هي موضوع الدراسة هو توزيع ملتويا (التواء سالب او موجب) فهل ان لهذا اثرا على دقة الطريقة البيزية المستخدم في تقدير المعلمات لنموذج Tobit Quantile Regression بشكل خاص ولنماذج Quantile Regression بشكل عام التي شاع فيها افتراض ان توزيع الخطأ للنموذج هو التوزيع الملتوي للابلاس (Asymmetric Laplace).

ثانيا: هل ان دقة تقدير النموذج تتباين في الاقسام المختلفة لتوزيع متغير الاستجابة عندما يكون التوزيع الحقيقي للبيانات توزيعا ملتويا .

الملحق (1) : اشتقاق التوزيعات الشرطية الكاملة اللاحقة للمتغيرات العشوائية*

(The full conditional posterior distribution)

1- التوزيع الشرطي الكامل للمتغير العشوائي y^* :

بالاستناد الى التمثيل المختلط لتوزيع الخطأ المتضمن في المعادلة (3-3) فان التوزيع الشرطي للمتغير العشوائي y^* يمكن ان يعبر عنه بالشكل التالي:

$$f\left(y^*\backslash \boldsymbol{X},\boldsymbol{e},\boldsymbol{\beta},\boldsymbol{r},\tau,\delta_2,\widetilde{\delta_1},\alpha_1,\theta_1\right)=\prod_{i=1}^{n}\frac{\tau^{1\backslash 2}}{\sqrt{2\pi\xi_2^2 e_i}}\exp\{-\tau(y^*-\acute{x}_i\beta-\xi_1 e_i)^2\backslash 2\xi_2^2 e_i\}$$

2- التوزيع الشرطي الكامل للمتغير العشوائي e_i :

$$f(e_i\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e}_{-i},\boldsymbol{\beta},\boldsymbol{r},\tau,\delta_2,\widetilde{\delta_1},\alpha_1,\theta_1)\propto f\left(y^*\backslash \boldsymbol{X},\boldsymbol{e},\boldsymbol{\beta},\boldsymbol{r},\tau,\delta_2,\widetilde{\delta_1}\right)f(e_i)$$
$$\propto e_i^{-1\backslash 2}\exp\{-\tau(y^*-\acute{x}_i\beta-\xi_1 e_i)^2\backslash 2\xi_2^2 e_i\}\exp(-\tau e_i)$$

$$f\left(e_i\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e}_{-i},\boldsymbol{\beta},\boldsymbol{r},\tau,\delta_2,\widetilde{\delta_1},\alpha_1,\theta_1\right)\propto\ e_i^{-1\backslash 2}\exp\{-\frac{1}{2}[(\tau\xi_1^2\xi_2^{-2}+2\tau)e_i+\xi_2^{-2}\tau(y_i^*-\acute{x}_i\beta)^2\ e_i^{-1}]\}$$

حيث ان : $\boldsymbol{e}_{-i}$ يمثل المتغير $\boldsymbol{e}$ باستبعاد e_i

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

3- التوزيع الشرطي الكامل للمتغير العشوائي β_k :

$$f(\beta_k\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e},\beta_{-k},\boldsymbol{r},\tau,\delta_2,\alpha_1,\theta_1,\widetilde{\delta_1})\propto f\left(y^*\backslash \boldsymbol{X},\boldsymbol{e},\beta,\boldsymbol{r},\tau,\delta_2,\widetilde{\delta_1}\right)f(\beta_k)$$

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

$$\propto\ \exp\{-\tau\sum_{i=1}^{n}(y_i^*-\acute{x}_i\beta-\xi_1 e_i)^2\backslash 2\,{\xi_2}^2 e_i\ \ \exp\left\{-\frac{1}{2}\ (\frac{r_k-1}{2\delta_2 r_k})^{-1}\ \beta_k^2\right\}$$

$$\propto \exp\{-\frac{1}{2}\ [(\tau\xi_2^{-2}\sum_{i=1}^{n}x_{ik}^2 e_i^{-1} + 2\delta_2 r_k(r_k-1)^{-1})\beta_k^2 - 2\sum_{i=1}^{n}(\tau w_{ik}x_{ik}\xi_2^{-2}e_i^{-1})\beta_k$$

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

التوزيع الشرطي الكامل لـ β_k هو Normal بمتوسط

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

***ان اشتقاق الدوال الشرطية في (2,3,5,6) تم بالاستناد المصدر (7,pp552).**

وتباين

$$V_k^2 = \frac{1}{\xi_2^{-2}\tau\sum_i^n(x_{ik}^2 e_i^{-1}) + 2\delta_2 r_k(r_k-1)^{-1}}$$

4-التوزيع الشرطي الكامل للمتغير العشوائي τ :

$$f(\tau\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e},\boldsymbol{r},\beta_k,\alpha_1,\theta_1,\delta_2,\widetilde{\delta_1}) \propto f(y^*\backslash \boldsymbol{X},\boldsymbol{e},\boldsymbol{\beta},\boldsymbol{r},\tau,\delta_2,\widetilde{\delta_1})f(e_i\backslash\tau)\ f(\tau)$$

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

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

لذا فالتوزيع الشرطي الكامل لـ τ هو Gamma

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

$$f(r_k - 1\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e},\boldsymbol{r}_{-k},\beta_k,\tau,\alpha_1,\theta_1,\delta_2,\widetilde{\delta_1}) \propto$$

$$(r_k-1)^{-1\backslash 2}\exp\left\{-\frac{1}{2}\left[\frac{2\delta_2 r_k}{r_k-1}\right]\beta_k^2\right\}\exp\{-\widetilde{\delta_1}r_k\}I(r_k>1)$$

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

$$\propto (r_k-1)^{-1\backslash 2} \propto \exp\left\{-\frac{1}{2}\left[2\widetilde{\delta_1}(r_k-1) + \frac{2\delta_2\beta_k^2}{r_k-1}\right]\right\}I(r_k-1>0)$$

6- التوزيع الشرطي الكامل للمتغير العشوائي $\widetilde{\delta_1}$:

$$f(\widetilde{\delta_1}\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e},\boldsymbol{r},\beta_k,\tau,\delta_2,,\alpha_1,\theta_1) \propto f(\widetilde{\delta_1})\prod_{k=1}^{p}\Gamma^{-1}(1\backslash 2,\widetilde{\delta_1})\widetilde{\delta_1}^{1\backslash 2}\exp\{-\widetilde{\delta_1}r_k\}$$

$$\propto \frac{\theta_1^{\alpha_1}}{\Gamma(\alpha_1)}\ \widetilde{\delta_1}^{\alpha_1-1}\exp(-\theta_1\widetilde{\delta_1})\prod_{k=1}^{p}\Gamma^{-1}(1\backslash 2,\widetilde{\delta_1})\widetilde{\delta_1}^{1\backslash 2}\exp\{-\widetilde{\delta_1}r_k\}$$

$$\propto \Gamma^{-p}(1\backslash 2,\widetilde{\delta_1})\widetilde{\delta_1}^{(p\backslash 2)+\alpha_1-1}\exp\{-\widetilde{\delta_1}[\theta_1 + \sum_{k=1}^{p}r_k]\}$$

7- التوزيع الشرطي الكامل للمتغير δ_2:

$$f(\delta_2\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e},\boldsymbol{r},\beta_k,\tau,\alpha_1,\theta_1,\widetilde{\delta_1}) \propto f(\delta_2)\prod_{k=1}^{p}\delta_2^{\,1\backslash 2}\exp\{-\frac{1}{2}\left[\frac{2\delta_2 r_k}{r_k-1}\right]\beta_k^2\}$$

$$\propto \theta_3\exp(-\theta_3\delta_2)\delta_2^{\,p\backslash 2}\exp\{-\frac{1}{2}\sum_{k=1}^{p}\left[\frac{2\delta_2 r_k}{r_k-1}\right]\beta_k^2\}$$

$$f(\delta_2\backslash \boldsymbol{X},\boldsymbol{y}^*,\boldsymbol{e},\boldsymbol{r},\beta_k,\tau,\alpha_1,\theta_1,\widetilde{\delta_1}) \propto \delta_2^{\,(p\backslash 2)}\exp\{-\delta_2\sum_{k=1}^{p}\left[\frac{r_k}{r_k-1}\right]\beta_k^2 + \theta_3\}$$

لذا فالتوزيع الشرطي للمتغير δ_2 هو Gamma .

8-التوزيع الشرطي الكامل للمتغير العشوائي θ_1:

$$f(\theta_1 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_{pk}, \tau, \alpha_1, \delta_2, \tilde{\delta_1}) \propto f(\theta_1) f(\tilde{\delta_1})$$
$$\propto \theta_2 \exp(-\theta_1 \theta_2)\, \theta_1^{\alpha_1} \exp(-\theta_1 \tilde{\delta_1})$$
$$\propto \theta_1^{\alpha_1} \exp(-\theta_1(\theta_2 + \tilde{\delta_1}))$$

التوزيع الشرطي للمتغير العشوائي θ_1 هو Gamma

9- التوزيع الشرطي للمتغير العشوائي α_1:

$$f(\alpha_1 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \theta_1, \delta_2, \tilde{\delta_1}) \propto f(\alpha_1) f(\tilde{\delta_1})$$
$$\propto \frac{\theta_1^{\alpha_1}}{\Gamma(\alpha_1)} \tilde{\delta_1}^{\alpha_1 - 1}$$

10- التوزيع الشرطي للمتغير العشوائي θ_2 :

$$f(\theta_2 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \theta_1, \delta_2, \tilde{\delta_1}) \propto f(\theta_2) f(\theta_1, \alpha_1)$$
$$\propto \exp(-\theta_2)\ \theta_2 \exp(-\theta_2 \theta_1)$$
$$f(\theta_2 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \theta_1, \delta_2, \tilde{\delta_1}) \propto \theta_2 \exp(-\theta_2(\theta_1 + 1))$$

لذا فالتوزيع الشرطي للمتغير العشوائي θ_2 هو Gamma

11- التوزيع الشرطي للمتغير العشوائي θ_3 :

$$f(\theta_3 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \theta_1, \delta_2, \tilde{\delta_1}) \propto f(\theta_3) f(\delta_2)$$
$$\propto \theta_3^{-1} \theta_3 \exp(-\delta_2 \theta_3)$$
$$f(\theta_3 \backslash \boldsymbol{X}, \boldsymbol{y}^*, \boldsymbol{e}, \boldsymbol{r}, \beta_k, \tau, \theta_1, \delta_2, \tilde{\delta_1}) \propto exp(-\delta_2 \theta_3)$$

لذا فاتوزيع الشرطي للمتغير العشوائي θ_3 هو exponential

جدول رقم (2-1)

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

p=0.5		n=25	
true β	(method) الطريقة		
	BER	QR	BEH
3	3.387764	3.308411	3.213477
1.5	1.66736	1.517248	1.463643
0	-0.17597	-0.07039	-0.02122
0	0.098458	-0.07362	0.103653
2	2.108437	2.210288	2.06533
0	0.173252	0.014256	0.11078
0	0.028122	-0.01807	-0.1062
0	0.079918	-0.05547	-0.01276
p=0.5		n=50	
true β	BER	QR	BEH
3	3.361983	3.194661	3.208525
1.5	1.794697	1.891027	1.790735
0	-0.30261	-0.36604	-0.2444
0	0.028179	-0.03111	-0.00129
2	2.276662	2.224323	2.19849

0	-0.01809	0.051262	0.058557
0	-0.0582	-0.19839	-0.12665
0	-0.02046	0.079263	0.050461
p=0.5		n=100	
true β	BER	QR	BEH
3	3.379269	3.341206	3.287548
1.5	1.61253	1.641467	1.598228
0	0.07775	-0.03983	0.051368
0	0.008559	0.01876	0.036023
2	2.217959	2.187426	2.160918
0	-0.07679	-0.07746	-0.05316
0	0.062582	0.085362	0.045267
0	0.03784	0.040964	0.054639
p=0.25		n=25	
true β	BER	QR	BEH
3	2.970307	3.019689	2.986783
1.5	1.73433	1.279086	1.497109
0	-0.04674	0.072513	0.034073
0	-0.05246	-0.14696	0.003535
2	1.604957	2.106706	1.813901
0	-0.06047	0.0934	0.007855
0	0.195533	-0.18771	-0.00956
0	0.318487	-0.0745	0.088114
p=0.25		n=50	
true β	BER	QR	BEH
3	3.148549	2.913533	3.038074
1.5	1.419433	1.629122	1.508869
0	-0.07931	-0.09505	-0.03764
0	0.108754	-0.01387	0.042553
2	2.015551	2.011117	2.000579
0	-0.19522	0.151931	-0.03316
0	0.238421	-0.11759	0.056126
0	-0.21538	0.092175	-0.05541
p=0.25		n=100	
true β	BER	QR	BEH
3	3.108696	2.979929	3.051067
1.5	1.488267	1.515357	1.500147
0	-0.06993	0.034473	-0.01188
0	0.055916	0.001686	0.038842
2	2.104731	1.984734	2.055053
0	-0.06194	-0.00193	-0.05864
0	0.005137	0.059776	0.050392
0	0.017496	-0.06161	-0.03426

p=0.05		n=25	
true β	BER	QR	BEH
3	3.64876	2.401722	3.171123
1.5	-0.21596	1.253341	1.044878
0	-0.32833	0.891054	-0.37088
0	1.519319	-0.54095	0.530397
2	2.499486	2.288318	2.167069
0	-0.25819	-0.1626	0.0108
0	0.344009	0.195266	-0.02785
0	0.57216	0.116018	0.093614
p=0.05		n=50	
true β	BER	QR	BEH
3	3.259705	1.573624	3.145363
1.5	1.262852	1.135983	1.240247
0	0.037889	-0.01978	0.02815
0	0.452215	-0.34849	0.220646
2	1.556058	1.421005	1.763037
0	-0.39509	0.284684	-0.18122
0	0.711495	-0.18476	0.353727
0	0.40181	-0.37373	0.167946
p=0.05		n=100	
true β	BER	QR	BEH
3	2.912854	1.995465	2.870589
1.5	1.287013	1.096642	1.31194
0	-0.03569	-0.06152	-0.05908
0	-0.09998	0.150791	-0.03643
2	2.075245	1.146106	1.956769
0	-0.07966	0.155346	-0.01063
0	-0.0355	0.159537	-0.03895
0	0.050531	-0.10325	0.043704
p=0.5		n=25	
true β	الطريقة (method)		
	BER	QR	BEH
0.85	1.050882	0.92412	0.926188
0.85	0.925163	0.940474	0.770782
0.85	1.350312	0.748424	1.146611
0.85	0.95939	0.921906	0.932654
0.85	0.774885	1.074401	0.897608
0.85	0.813554	0.784986	0.745881
0.85	1.006746	0.83484	0.840914
0.85	1.093107	1.094613	0.889427

p=0.5		n=50	
true β	BER	QR	BEH
0.85	1.0009	0.875951	0.912923
0.85	0.976035	1.095389	1.003739
0.85	0.88238	0.793985	0.824476
0.85	0.961503	0.913395	0.919707
0.85	0.977562	0.943485	0.943666
0.85	1.138401	1.11119	1.068557
0.85	0.911185	0.841164	0.854882
0.85	0.939934	0.968714	0.947751
p=0.5		n=100	
true β	BER	QR	BEH
0.85	0.988352	0.994944	0.95252
0.85	0.986414	0.956811	0.960874
0.85	0.899355	0.92725	0.894871
0.85	0.995112	0.90522	0.936675
0.85	0.938531	0.975972	0.944171
0.85	0.989693	0.933121	0.926962
0.85	0.8857	0.965891	0.935542
0.85	1.019225	0.970791	0.997904
p=0.25		n=25	
true β	BER	QR	BEH
0.85	1.196585	0.660173	0.934886
0.85	1.453851	0.678374	1.012781
0.85	-0.53327	1.228038	0.531654
0.85	1.420747	0.488377	0.918101
0.85	0.427759	0.823798	0.73925
0.85	1.661269	0.56583	1.011203
0.85	0.62507	0.742745	0.78389
0.85	1.307147	0.802494	0.999176
p=0.25		n=50	
true β	BER	QR	BEH
0.85	0.927033	0.796307	0.847807
0.85	0.759454	0.973689	0.861824
0.85	0.939226	0.748743	0.845919
0.85	0.879974	0.833676	0.867704
0.85	0.835472	0.843251	0.845526
0.85	0.868842	0.880605	0.859064
0.85	1.014203	0.885513	0.91795
0.85	0.639917	0.851492	0.760537

p=0.25		n=100	
true β	BER	QR	BEH
0.85	0.828381	0.847042	0.838328
0.85	0.864679	0.873494	0.871389
0.85	0.809602	0.830111	0.816706
0.85	0.839898	0.841148	0.849516
0.85	0.847483	0.856421	0.851421
0.85	0.812426	0.806826	0.81583
0.85	0.873819	0.878781	0.872673
0.85	0.838087	0.851288	0.851277
p=0.05		n=25	
true β	BER	QR	BEH
0.85	-0.02419	0.862731	0.670089
0.85	1.611815	0.238004	0.996138
0.85	1.183242	0.417153	1.045763
0.85	0.308356	0.404726	0.587261
0.85	-0.76199	0.912105	0.441674
0.85	-0.12431	0.575523	0.700328
0.85	1.665529	0.24649	1.002164
0.85	1.953561	0.204928	1.068588
p=0.05		n=50	
true β	BER	QR	BEH
0.85	0.912381	0.565951	0.803034
0.85	0.468942	0.882512	0.680538
0.85	1.082877	0.689557	0.928837
0.85	0.768877	0.582603	0.797345
0.85	0.693439	0.675552	0.7632
0.85	0.658357	0.815066	0.653002
0.85	1.157799	0.677173	1.022584
0.85	0.293029	0.757489	0.552167
p=0.05		n=100	
true β	BER	QR	BEH
0.85	0.848873	0.518454	0.819307
0.85	0.757428	0.646282	0.720186
0.85	0.762574	0.573003	0.750096
0.85	0.738788	0.657924	0.780718
0.85	1.004497	0.494657	0.884045
0.85	0.678297	0.641295	0.702233
0.85	0.806052	0.592002	0.767164
0.85	0.926268	0.502634	0.863336

p=0.5		n=25	
true β	(method) الطريقة		
	BER	QR	BEH
5	5.740133	5.494426	5.360564
0	-0.17974	-0.10009	-0.11418
0	0.056507	0.051916	0.068285
0	-0.02574	-0.08726	-0.02393
0	-0.45872	-0.08284	-0.12091
0	0.364228	0.167226	0.154138
0	-0.04207	0.00736	-0.08094
0	-0.15023	-0.08908	-0.15692
p=0.5		n=50	
true β	BER	QR	BEH
5	5.743039	5.512309	5.399089
0	0.04709	-0.03498	0.121392
0	-0.08506	0.103009	-0.04455
0	0.291174	-0.03806	0.287513
0	-0.14821	-0.07405	-0.15161
0	-0.05709	0.127191	-0.06291
0	0.078715	0.020388	0.00188
0	-0.04916	-0.0061	0.041056
p=0.5		n=100	
true β	BER	QR	BEH
5	5.391102	5.356014	5.299602
0	0.053075	-0.01744	0.066267
0	0.0657	0.014504	0.067181
0	-0.02203	0.030255	-0.00912
0	-0.06319	-0.06843	-0.08229
0	-0.00091	-0.00646	0.016472
0	0.026928	0.038888	0.036897
0	-0.01946	-0.0343	-0.02799
p=0.25		n=25	
true β	BER	QR	BEH
5	5.058689	4.974219	4.992144
0	0.090749	-0.01715	0.008884
0	-0.05189	-0.1128	-0.02497
0	-0.02398	-0.13495	0.022736
0	-0.63052	0.095685	-0.23438
0	0.157882	-0.02703	0.115487
0	0.07545	-0.13583	-0.10166
0	0.116188	-0.06325	0.020158

p=0.25		n=50	
true β	BER	QR	BEH
5	4.978365	4.940362	5.006786
0	-0.00121	-0.11402	0.015238
0	0.080095	0.001296	0.031366
0	-0.303	0.150215	-0.13246
0	0.122516	-0.05154	0.069914
0	-0.1307	0.01929	-0.06469
0	0.004705	0.184777	0.010912
0	-0.12439	-0.00738	-0.06757
p=0.25		n=100	
true β	BER	QR	BEH
5	5.298649	1.278666	5.184171
0	-0.17845	-0.00049	-0.08665
0	0.017513	0.027376	0.008555
0	0.005939	0.01494	0.028484
0	0.029821	-0.0314	0.003541
0	-0.01694	0.005258	-0.01196
0	0.093739	-0.02972	0.057906
0	-0.17223	0.007482	-0.10795
p=0.05		n=25	
true β	BER	QR	BEH
5	3.789497	3.994122	4.751848
0	1.134753	-0.27816	0.24439
0	-0.4316	0.093006	-0.08685
0	-0.44271	-0.13064	-0.08286
0	-1.69109	0.836411	-0.38666
0	-0.79256	-0.35683	-0.28469
0	1.194343	-0.44735	0.316731
0	0.853264	-0.27657	0.127645
p=0.05		n=50	
true β	BER	QR	BEH
5	4.489342	4.013772	4.657208
0	-0.26854	0.245895	-0.17289
0	0.649964	-0.31456	0.405676
0	-0.50483	0.703721	-0.21371
0	0.30351	-0.54996	0.061587
0	-0.29655	0.300864	-0.06731
0	-0.13242	0.020141	-0.04024
0	-0.05469	-0.35854	-0.02818

p=0.05		n=100	
true β	BER	QR	BEH
5	4.722371	4.134119	4.696854
0	-0.01891	0.011145	0.010836
0	-0.0136	-0.09913	-0.01407
0	0.225191	0.185121	0.176378
0	-0.32132	0.032741	-0.22946
0	0.055911	-0.05344	0.050521
0	0.125164	-0.17778	0.049809
0	0.02792	0.072491	0.00461

جدول (2-2)

قيم الوسيط لمتوسط مطلق الانحرافات (MMAD) المحسوب لـ 50 محاكاة ولكل طريقة مستخدمة مع قيم الانحرافات المرافقة لها sd(MAD) وحسب احجام العينات (n) وقيمة القسيم (p) حيث تم كتابة القيم الاصغر لـ (MMAD) في طرق التقدير باللون الغامق.

true β =3,1.5,0,0,2,0,0,0			
	sd	MMAD	method
n=25	0.196761	0.614411	BEH
p=0.5	0.258021	0.854431	QR
	0.266459	0.798115	BER
n=50	0.124069	0.527198	BEH
p=0.5	0.182402	0.666802	QR
	0.119687	0.627745	BER
n=100	0.106021	0.433815	BEH
p=0.5	0.128038	0.488605	QR
	0.109053	0.492805	BER
n=25	0.167345	0.556796	BEH
p=0.25	0.248932	0.805097	QR
	0.218327	0.799955	BER
n=50	0.114497	0.381474	BEH
p=0.25	0.173968	0.551782	QR
	0.122189	0.425657	BER
n=100	0.080848	0.294986	BEH
p=0.25	0.120985	0.368262	QR
	0.084661	0.276588	BER
n=25	0.34055	0.822232	BEH
p=0.05	0.38455	1.275926	QR
	0.410094	1.865025	BER
n=50	0.188234	0.648301	BEH
p=0.05	0.590519	1.736781	QR
	0.173125	0.916432	BER
n=100	0.120128	0.39251	BEH
p=0.05	0.908277	1.050885	QR
	0.096795	0.360084	BER

true β =0.85,0.85,0.85,0.85,0.85,0.85,0.85,0.85			
n=25	0.175787	0.648184	BEH
p=0.5	0.501178	1.218432	QR
	0.236709	0.95611	BER
n=50	0.121291	0.459063	BEH
p=0.5	0.159334	0.627217	QR
	0.151924	0.568543	BER
n=100	0.112824	0.437949	BEH
p=0.5	0.135786	0.522584	QR
	0.103102	0.489801	BER
n=25	0.170878	0.487204	BEH
p=0.25	0.536033	0.789239	QR
	0.291129	1.196559	BER
n=50	0.083968	0.373132	BEH
p=0.25	0.163865	0.491969	QR
	0.121845	0.41496	BER
n=100	0.056253	0.200158	BEH
p=0.25	0.067993	0.210939	QR
	0.055696	0.206818	BER
n=25	0.299455	0.833507	BEH
p=0.05	0.55568	1.720467	QR
	0.463033	2.155516	BER
n=50	0.177116	0.603041	BEH
p=0.05	0.288615	0.711985	QR
	0.165159	0.625066	BER
n=100	0.158059	0.367934	BEH
p=0.05	0.602066	1.037916	QR
	0.111558	0.310341	BER

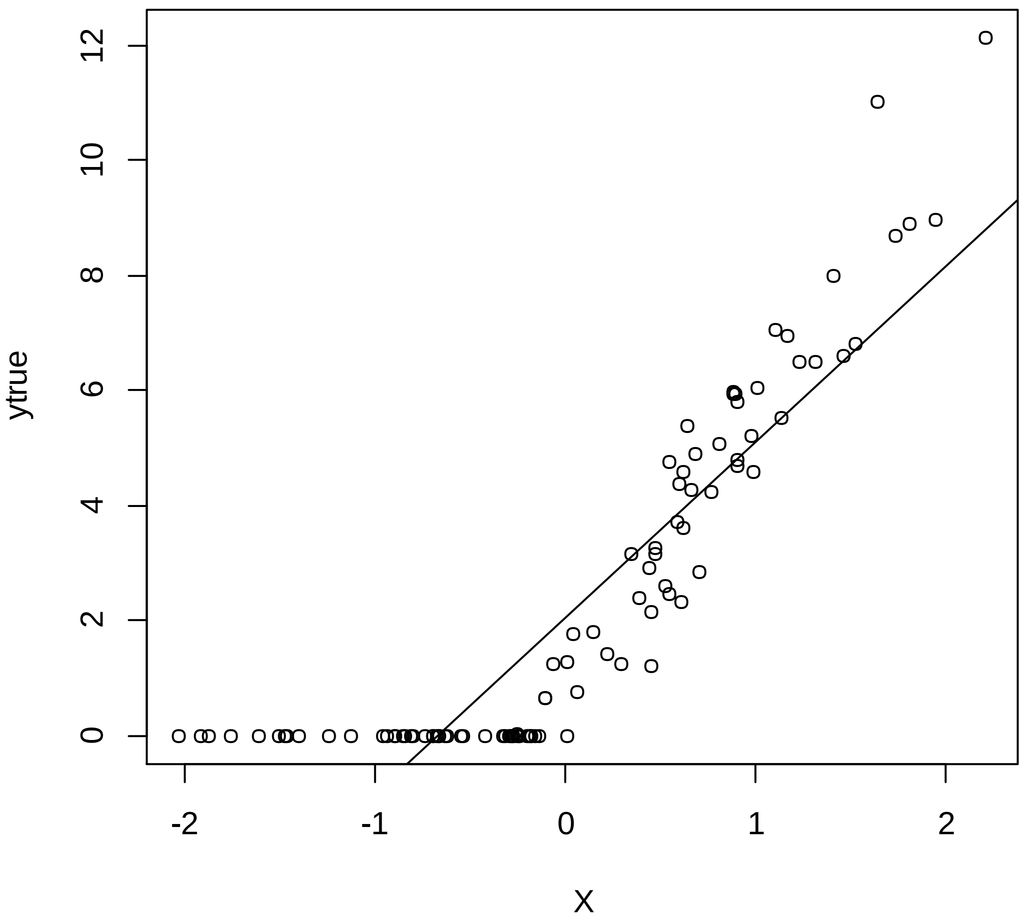
true β =5,0,0,0,0,0,0,0			
n=25	0.151557	0.621363	BEH
p=0.5	0.376907	0.91483	QR
	0.179026	0.863125	BER
n=50	0.236564	0.587153	BEH
p=0.5	0.531384	0.904119	QR
	0.310113	0.759793	BER
n=100	0.082251	0.356983	BEH
p=0.5	0.124456	0.415846	QR
	0.091301	0.426027	BER
n=25	0.120523	0.559904	BEH
p=0.25	0.386565	0.875458	QR
	0.16127	0.658333	BER
n=50	0.129227	0.368383	BEH
p=0.25	0.274753	0.746531	QR
	0.119854	0.422506	BER
n=100	0.105192	0.351254	BEH
p=0.25	1.755487	4.284827	QR
	0.092024	0.389019	BER
n=25	0.330663	0.788451	BEH
p=0.05	0.631808	1.468828	QR
	0.491463	2.266091	BER
n=50	0.201733	0.615953	BEH
p=0.05	0.269212	1.123975	QR
	0.141604	0.768693	BER
n=100	0.088019	0.413798	BEH
p=0.05	0.306674	0.774709	QR
	0.071995	0.420598	BER

ملحق (3)

شكل رقم (1)

يمثل خط الانحدار لنموذج (TQR) باستخدام القيم الحقيقية للمتغير المعتمد (y)

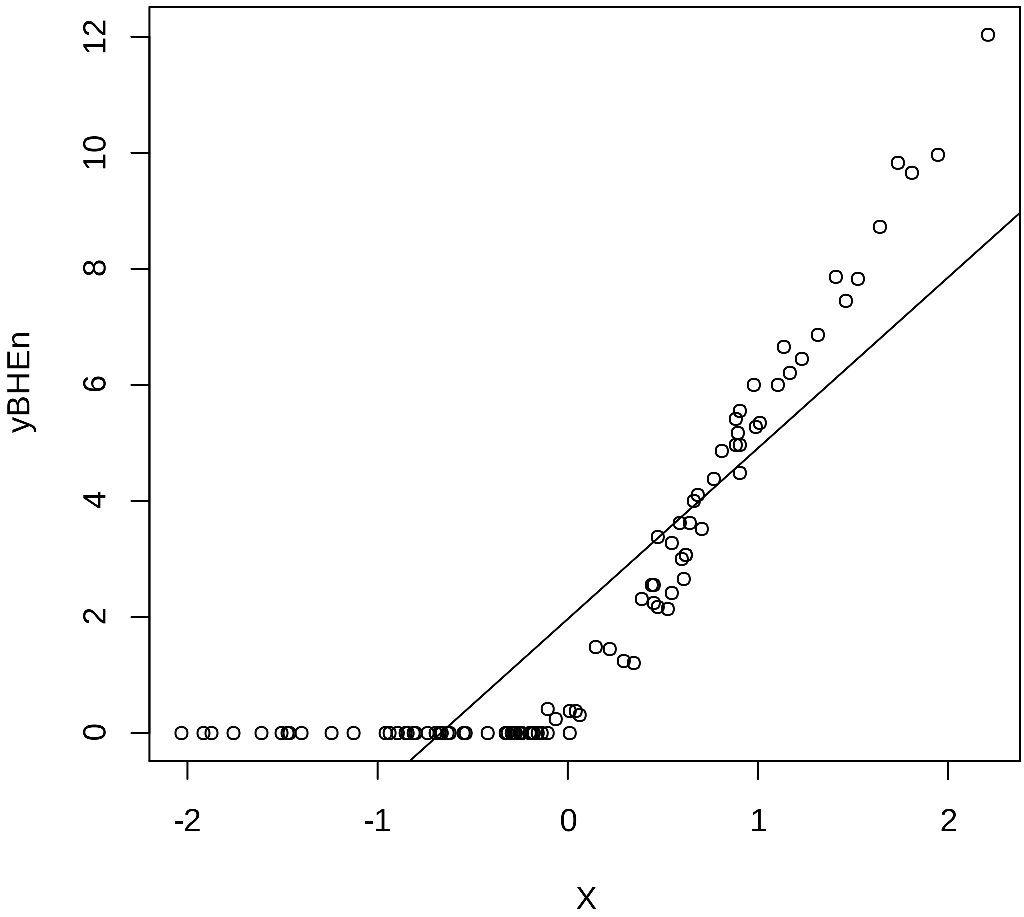
β=(5,0,0,0,0,0,0,0) , p=0.5 , n=10



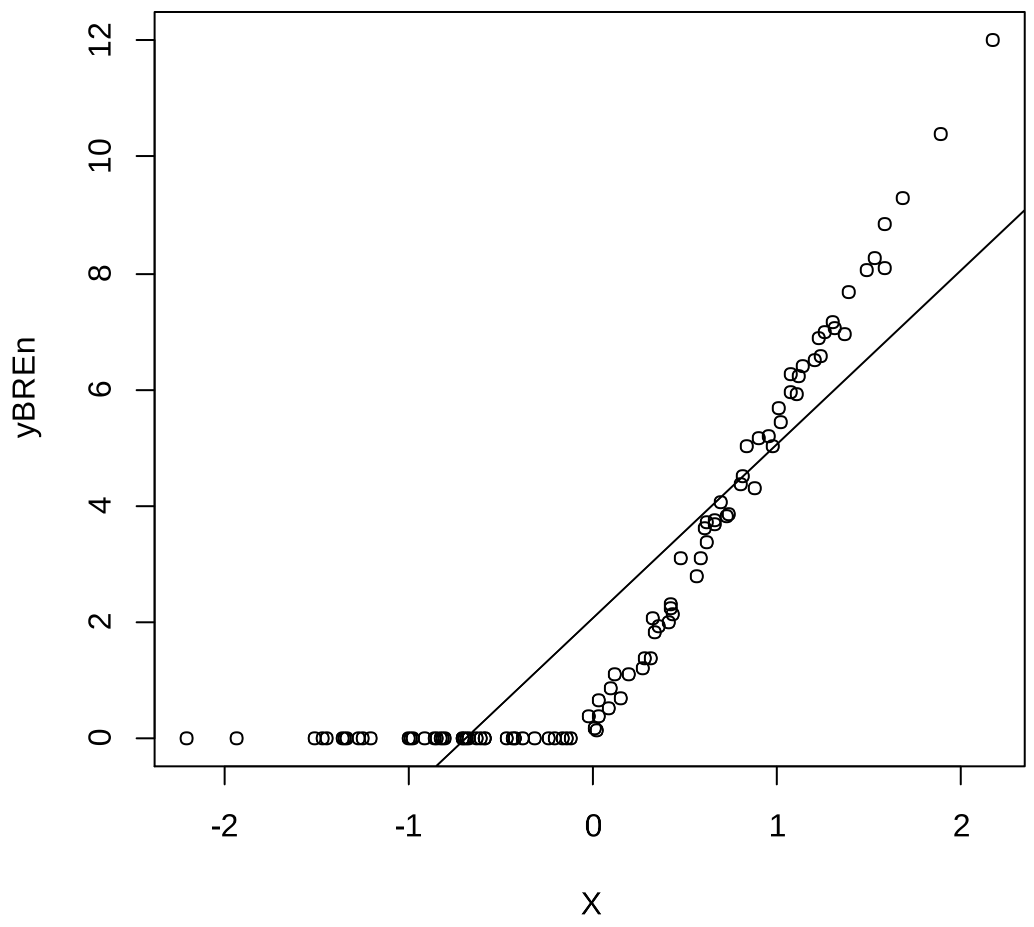
شكل رقم (2)

يمثل خط الانحدار لنموذج (TQR) باستخدام القيم التقديرية للمتغير المعتمد (y) المحسوبة بطريقة BEH

β=(5,0,0,0,0,0,0,0) , p=0.5 , n=10



شكل رقم (3)
يمثل خط الانحدار لنموذج (TQR) باستخدام القيم التقديرية للمتغير المعتمد (y) المحسوبة بطريقة BER
$\beta=(5,0,0,0,0,0,0,0)$, p=0.5 , n=10



المصادر(Reference)

المصادر العربية:

1- د. سمر حسن الباجوري (2016) مشكلة الإحصاءات في إفريقيا جنوب الصحراء وتحليل أثرها على فاعلية الأداء الحكومي باستخدام نموذج توبت. مجلة البحوث الاقتصادية والمالية 9822 – 2352.

المصادر الاجنبية:

2- Alhamzawi Rahim (2014) Bayesian Elastic Net Tobit Quantile Regression " Communication in statistics-Simulation and Computation Volume 45, 2016 - Issue 7..

3- Hidio Kozumi & Genya kobayshi(2009) "Gibbs Sampling Methods for Bayesian Quantile Regression" Graduate School of Business Administration ,Kobe University 2-1Rokko, Kobe 657-8501,Jaban.

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

5-Keming Yu & Julian Stander (2005) Bayesian analysis of a Tobit quantile regression model.Econometrics.vol.137,issue 1,p 260-276

6- Kemming Yu& Rana A.Moyeed (2001)" Bayesian Quantile Regression" Statistical & Probability Letters 54:437-447.

7- Qing Li, Rubin Xi &Nan Lin "Bayesian Regularized Quantile regression" Bayesian Analysis (2010) ,5.Number 3,pp. 533-556.

8- Yonggang Ji, Nan Lin, Baoxue Zhang (2012) " Model selection in binary and tobit quantile regression using the Gibbs Sampler" Computational Statistics and Data Analysis 56,827-839

Bayesian Tobit Quantile Regression Model Using Four Level Prior Distributions

Abstract:

In this research we discussed the parameter estimation and variable selection in Tobit quantile regression model in present of multicollinearity problem. We used elastic net technique as an important technique for dealing with both multicollinearity and variable selection. Depending on the data we proposed Bayesian Tobit hierarchical model with four level prior distributions . We assumed both tuning parameter are random variable and estimated them with the other unknown parameter in the model .Simulation study was used for explain the efficiency of the proposed method and then we compared our approach with (Alhamzwi 2014 & standard QR) .The result illustrated that our approach was outperformed.

This is the first work that suggested Bayesian hierarchical model with four level prior distribution in estimating and variable selection for TQR model.

Keywords: Bayesian Tobit hierarchical model , tuning parameter.