



## اختيار المتغيرات الاكثر تأثير على الخثرة الدموية باستخدام انحدار لاسو القسيمي البيزي

اميرة صفر مذخور<sup>1</sup> ، أ.م.د. فاضل حميد هادي<sup>2</sup>

### انتساب الباحثين

<sup>2,1</sup> كلية الادارة والاقتصاد، جامعة القادسية، العراق، الديوانية، 58006

<sup>1</sup>admin.stat21.15@qu.edu.iq

<sup>2</sup>fadel.alhusiny@qu.edu.iq

### المؤلف المراسل

### معلومات البحث

تاريخ النشر : آب 2022

### الخلاصة

تؤدي التوزيعات السابقة (Prior Distributions) دوراً مهماً في مقدرات النظرية البيزية، فهي التي تكون مسؤولة عن نوع التوزيعات اللاحقة (Posterior Distributions) للمعالم المراد تقديرها. ومن المعلوم عند تنفيذ عملية اختيار المتغيرات المهمة في نماذج الانحدار بتقنية الاسو البيزية لا بد من استخدام توزيع لابلاس (Laplace Distribution) كتوزيع مسبق، لكن استخدام توزيع لابلاس مباشرة يجعل خوارزمية (Markov Chain Monte Carlo) صعبة جداً وغير كفوءة وتحتاج الى وقت كثير جداً للتقارب بين المعلمة المقدره خلال دورات تقدير تلك المعالم. وللتغلب على هذه المشكلة في البحث الحالي تم توظيف تحويلي من تحويلات توزيع لابلاس التي تضمن لنا أن تكون خوارزمية MCMC كفوءة وسريعة التقارب بين المعالم المقدره خلال دورات أقل. ويمتلك انحدار القسيمي (Quantile Regression) ميزات جيدة جعلته من النماذج المميزة في تمثيل علاقة الاثر بين المتغير المعتمد ومجموعة من المتغيرات المستقلة، وفي البحث الحالي سيتم نمذجة العلاقة بين الخثرة الدموية كمتغير معتمد ومجموعة من المتغيرات المستقلة باستخدام نموذج الانحدار القسيمي، وتقدير معالم هذا النموذج بأسلوب تقنية الاسو Lasso (least absolute shrinkage and selection operator) البيزية.

**الكلمات المفتاحية:** الانحدار القسيمي، تقنية لاسو البيزية، التوزيعات السابقة، التوزيعات اللاحقة، توزيع لابلاس.

### Selection More Important Variables on D. Dimer by Using Bayesian Lasso Quantile Regression

Amira Safer Mathkhor<sup>1</sup> , Fadel Hamid Hadi<sup>2</sup>

### Abstract

The study aims to Prior distributions play an important role in the estimations of Bayesian theory, It is responsible for the type of Posterior Distributions for the parameters to be estimated. It is known that when the process of selecting important variables is carried out in the regression models using the Bayesian Lasso (least absolute shrinkage and selection operator) technique, the Laplace distribution must be used as a prior distribution. But using the Laplace distribution directly makes the Markov Chain Monte Carlo (MCMC) algorithm very difficult and inefficient, and it requires a very large time to converge between the estimated parameters during the iterations of estimating those parameters. In order to overcome this problem in the current paper, two transformation for Laplace distribution has been employed, which ensures that the MCMC algorithm is efficient and quick to converge between the estimated parameters within fewer iterations. Quantile Regression has good features that make it one of the distinguished models in representing the effect relationship between the dependent variable and a set of independent variables. In the current paper, the relationship between the **D. dimer** as a dependent variable and a set of independent variables will be modeled using the quantile regression model and estimating the parameters of this model using the Bayesian (Lasso) technique.

**Keywords:** Quantile regression , Bayesian Lasso technique, prior distribution , Posterior Distributions, Laplace distribution

### Affiliation of Authors

<sup>1,2</sup> College of Administration and Economics, University of Al-Qadisiya, Iraq, Al-Diwaniyah, 58002

<sup>1</sup>admin.stat21.15@qu.edu.iq

<sup>2</sup>fadel.alhusiny@qu.edu.iq

### <sup>1</sup> Corresponding Author

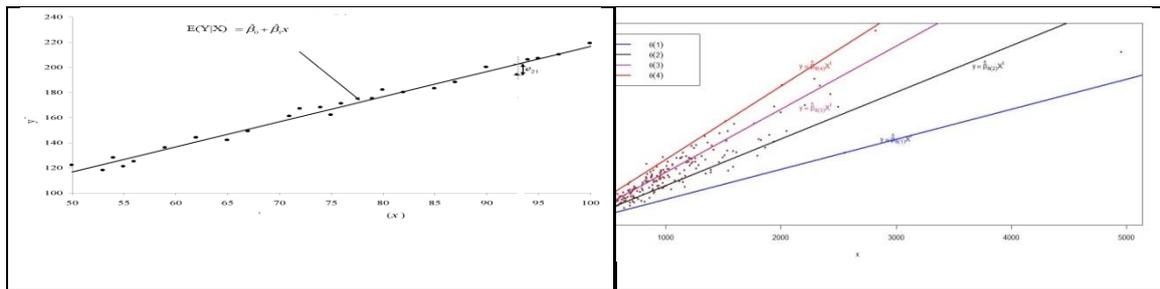
### Paper Info.

Published: Aug. 2022

**المقدمة:**

القسمي لا يحتاج الى اي افتراضات لتوزيع الاخطاء. على العكس من نموذج الانحدار الكلاسيكي الذي يشترط ان يتوزع الاخطاء بحسب التوزيع الطبيعي بمتوسط صفر وتباين  $\sigma^2$ . كذلك يمتلك نموذج الانحدار القسيمي مرونة عالية في وصف العلاقة بين متغيرات الظاهرة المدروسة لانه يقدر مجموعة غير محددة من خطوط انحدار القسيمي, لكن هذه الميزة غير متوفرة في نماذج الانحدار الكلاسيكية لانه يركز على تقدير خط انحدار واحد, وكما هو مبين في الشكل رقم (1) [4]

منذ ان تم اقتراح نموذج الانحدار القسيمي (quantile regression) بواسطة [1] اصبح من النماذج المهمة جداً في دراسة الظواهر الطبية [2] والبيئية [3] وغيرها من الدراسات. حيث يمتلك نموذج الانحدار القسيمي مميزات جذابة جعلته في مقدمة نماذج الانحدار, اذ يعدّ نموج الانحدار القسيمي حصيناً جداً ضد القيم الشاذة (outlier values) والقيم المتطرفة (extreme values), وبالعكس من نموذج الانحدار الكلاسيكي الذي يعدّ حساساً جداً للقيم الشاذة والمتطرفة. كذلك فإن نموذج الانحدار



شكل رقم (1) يبين خطوط الانحدار المقدره بواسطة نموذج الانحدار الكلاسيكي ونموذج الانحدار القسيمي

. ومن ثمّ فان نموذج الانحدار القسيمي يعرف وفق الصيغة الرياضية رقم (1)

$$y_i = x_i^T \beta_\theta + \varepsilon_i, \quad \theta \in (0,1), \dots \dots \dots (1)$$

الاعتماد على دالة الخسارة المبينة في ادناه في تقدير معالم نموذج الانحدار القسيمي [1], [5],

حيث أن  $(y_i)$  هو المتغير المعتمد (متغير الاستجابة),  $x_i^T$  هو متجه المتغيرات التوضيحي,  $\beta$  هو متجه المعالم غير معلومة القيم,  $\varepsilon_i$  حد الخطأ العشوائي و  $\theta$  هو مستوى القسيميات. يمكن

$$\min_{\beta_\theta} \sum_{i=1}^n \rho_\theta(y_i - x_i^T \beta_\theta) \quad (2)$$

مختصر الى المصطلح الانكليزي (Least Absolute Shrinkage and Selection Operator) الذي اقترح من قبل الباحث [7]. في عام 2008, وقام الباحث [8] بتوظيف تقنية (Lasso) مع دالة الخسارة المبينة في المعادلة رقم (2), التي تقوم بدورها بتقدير المعالم واختيار المتغيرات لنموذج الانحدار القسيمي كما مبين في معادلة رقم (3):

حيث إن دالة الخسارة في المعادلة (2) تكون غير قابلة للاشتقاق عند النقطة (0), ويمكن تصغير دالة الخسارة اعلاه بواسطة طريقة البرمجة الخطية (linear programming) التي اقترحت بواسطة [6]. في نهايات القرن العشرين اخذت تقنية اختيار المتغيرات وتقدير المعالم نموذج الانحدار الكلاسيكي اهتمام كبير في الادبيات الاحصائية, حيث اقترحت طريقة (Lasso) وهي

$$\min_{\beta_{\theta}} \sum_{i=1}^n \rho_{\theta}(y_i - x_i^T \beta_{\theta}) + \lambda \|\beta_{\theta}\|, \quad (3)$$

( Bayesian approach ), وان فلسفة أسلوب بيز هي البحث عن إيجاد التوزيعات اللاحقة ( Posterior distributions ) الذي يتم الحصول على هذه التوزيعات من حاصل ضرب الامكان الاعظم ( likelihood function ) في التوزيعات السابقة ( prior distributions ), كما هو مبين في المعادلة رقم (4)

$$f(\beta|y) = \frac{f(y, \beta)}{f(y)} = \frac{f(y|\beta) * g(\beta)}{f(y)} = \propto f(y|\beta) g(\beta). \quad (4)$$

حيث ان  $\lambda$  ( $\lambda \geq 0$ ) هي معلمة التضبيب وتكون مسؤولة عن تقليص المعامل (shrinkage parameter) نحو الصفر. لكن المعادلة (3) هي ايضاً غير قابلة للاشتقاق عند الصفر ويمكن تقدير معالم نموذج المعادلة (3) باستخدام البرمجة الخطية. ويمكن تقدير معالم النموذج المعادلة (2) باستخدام أسلوب بيز.

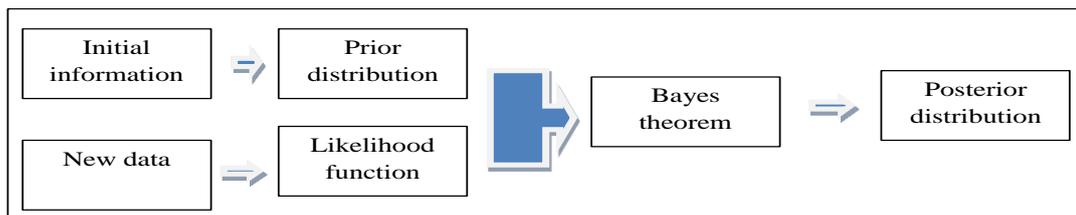
### المبحث الاول

#### 1- اسلوب بيز وتحولات توزيع لابلاس

يعد أسلوب بيز من افضل انواع اساليب التقدير ولا سيما في حجوم العينات الصغيرة [9] , لكن طرق التقدير الاخرى تتطلب حجوم عينات كبيرة ولا سيما في طريقة الامكان الاعظم , لذا فان نقص المعلومات الذي يحصل في دالة الامكان الاعظم في أسلوب بيز يمكن تعويضه من دالة التوزيعات المسبقة , كذلك تعمل على التحديث المستمر للمعالم المقدرة , وأخيراً فان الأسلوب يتعامل مع المعالم المراد تقديرها على انها متغيرات عشوائية تملك توزيعات احصائية , واسلوب بيز يمكن التعبير عنه من خلال المخطط في الشكل رقم (2) [10]

حيث ان  $f(\beta|y)$  التوزيعات اللاحقة , الامكان الاعظم  $g(\beta)$  هي التوزيعات السابقة في نموذج الانحدار القسيمي فان دالة الامكان الاعظم يتبع توزيع لابلاس غير المتماثل (asymmetric Laplace distribution) (ALD), والتوزيع السابق هو توزيع لابلاس ( Laplace distribution ).

ان هذا البحث سيتناول اسلوب بيز وتحولات توزيع لابلاس في المبحث الاول . وفي المبحث الثاني سيتم التركيز على التوزيعات اللاحقة للمعالم المقدرة , وفي المبحث الثالث تتم المقارنة ما بين طريقتي التقدير باستخدام بيانات حقيقية. اما المبحث الرابع سيركز على دراسة الاستنتاجات والتوصيات



شكل رقم (2) يمثل مخطط اسلوب بيز

#### 1.1 دالة الامكان الاعظم Likelihood function

ان توزيع الاخطاء العشوائية في نموذج الانحدار القسيمي يتبع توزيع لابلاس غير المتماثل (ALD) [11] الذي يملك دالة كثافة احتمالية

لتقدير واختيار معالم نموذج الانحدار القسيمي بواسطة الطريقة البيزية لا بد من تحديد دالة الامكان الاعظم ( Likelihood function ) وتحديد التوزيع المسبق ( prior distribution ) المناسب .

$$f(\epsilon_i|\sigma) = \theta(1 - \theta)\sigma \exp\{-\sigma\rho_\theta(\epsilon_i)\}, \quad (5)$$

فان دالة الامكان الاعظم للدالة (5) تصبح وفق الصيغة رقم (6)

$$f(\epsilon_i|\sigma) = \theta^n(1 - \theta)^n\sigma^n \exp\left\{-\sum_{i=1}^n \sigma\rho_\theta(\epsilon_i)\right\}, \quad (6)$$

المعتمد سيتبع التوزيع (ALD) ، وبحسب دالة الكثافة الاحتمالية المعادلة رقم (7):

وبحسب الاعتمادية الخطية بين الخطأ العشوائي والمتغير المعتمد في نموذج الانحدار القسيمي فان توزيع المتغير

$$f(y|\sigma) = \theta^n(1 - \theta)^n\sigma^n \exp\left\{-\sum_{i=1}^n \sigma\rho_\theta(y_i - x_i^T\beta_\theta)\right\}, \quad (7)$$

تمكنوا من برهان ان توزيع لابلاس غير المتماثل (ALD) هو توزيع مختلط من التوزيع الطبيعي ( Normal Distribution ) وتوزيع الاسي (exponential Distribution) , ومن ثم فان الخطأ يأخذ الشكل  $\epsilon_i \sim N((1 - 2\theta)Z_i, 2\sigma^{-1}Z_i)$  تفاصيل اكثر انظر [13] وحسب هذه المعطيات فان نموذج انحدار القسيمي يعرف بالصيغة رقم (8):

لكن استخدام الدالة في المعادلة رقم (7) مباشرة يكون صعب جداً للحصول على النموذج الهرمي البيزي ( Bayesian hierarchical model ) للتوزيعات اللاحقة ، فضلاً عن أن خوارزمية تكون غير كفوة , لذا فان كثير من الباحثين استخدموا تحويلات الى دالة (ALD) كتوزيعات مختلطة انظر [12] . واشهر تحويلات (ALD) هي التي قام بها الباحثون [13] , الذين

$$y_i = x_i^T\beta_\theta + (1 - 2\theta)Z_i + \sqrt{2\sigma^{-1}Z_i} \epsilon_i, \quad (8)$$

حيث إن  $Z_i$  يتوزع وفق التوزيع الاسي كما مبين في دالة الكثافة الاحتمالية  $\epsilon_i \sim \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\epsilon_i^2\right\}$  , ومن ثم فان توزيع المتغير المعتمد ( $y_i$ ) المبين في المعادلة (8) يتوزع وفق التوزيع الطبيعي  $y_i \sim N(x_i^T\beta_\theta + (1 - 2\theta)Z_i, 2\sigma^{-1}Z_i)$  , ويملك دالة الكثافة الاحتمالية في معادلة رقم (9)

فان دالة رقم (9) هي دالة الامكان الاعظم التي سوف تستخدم مع دوال التوزيعات السابقة للحصول على التوزيعات اللاحقة

$$f(y|Z_i, \beta, \theta, \sigma) = \frac{1}{\sqrt{4\pi\sigma^{-1}Z_i}} \text{Exp} \frac{-(y_i - x_i^T\beta - (1 - 2\theta)Z_i)^2}{4\sigma^{-1}Z_i} \quad (9)$$

### 1.2 التوزيعات السابقة (prior distributions)

في اسلوب النظرية البيزية فان التوزيعات السابقة (prior distributions) تؤدي دوراً مهماً جداً في مستقبل طريقة تقدير معالم النموذج المدروس. اشارة الباحث [6] ان انسب توزيع سابق (prior distributions) للمعالم في ظل تقنية (Lasso) هو توزيع لابلاس (Laplace distribution) , لانه

فان الدالة رقم (9) هي دالة الامكان الاعظم التي سوف تستخدم مع دوال التوزيعات السابقة للحصول على التوزيعات اللاحقة

التوزيع الوحيد الذي يتطابق مع الصيغة العامة لطريقة لابلاس. الذي يملك دالة كثافة احتمالية في معادلة رقم (10).

$$f(\beta_{\theta}|\lambda) = (\lambda/2)^p \exp\left(-\lambda \sum_{j=1}^p |\beta_{\theta}| \right) \quad (10)$$

في البحث الحالية سيتم توظيف صيغتين لتحويل توزيع لابلاس (Laplace distribution), التحويل الاول نفذ من قبل [15] الذي تمكن من برهنة ان توزيع لابلاس يمكن إعادة صياغته من خلط التوزيع المنتظم (uniform distribution) و توزيع كما عند تكون  $\Gamma 2$  (scale mixture uniform distribution) (SMU) والتي تعرف وفق دالة الكثافة الاحتمالية رقم (11)

$$\frac{\lambda}{2} e^{-\lambda|\beta_j|} = \int_{-u_j < \beta_j < u_j} \frac{1}{2u_j} \frac{\lambda^2}{\Gamma(2)} u_j^{2-1} \exp\{-\lambda u_j\} du_j, \quad (11)$$

يكون عبارة عن توزيع مختلط بين وزيع المنتظم (uniform distributions) والتوزيع الاسي القياسي (standard exponential distributions) (SMUSE) (Scale mixture uniform with standard exponential distribution), الذي يعرف وفق دالة المبينة في (12).

$$\frac{\sigma\lambda}{2} e^{-\sigma\lambda|\beta|} = \int_{w > \sigma\lambda|\beta|} \frac{\sigma\lambda}{2} e^{-w} dw \quad (12)$$

### المبحث الثاني

#### 2- التوزيعات اللاحقة Posterior distribution

بعد الحصول على دالة الامكان الاعظم والتوزيع السابق للمعالم النموذج المدروس يمكن بناء توزيعات النموذج الهرمي البيزي وبحسب الآتي:

##### 2.1 النموذج الهرمي البيزي في حالة (SMU) كتوزيع مسبق

من خلال توظيف دالة التوزيع المختلط (SMU) يمكن الحصول على توزيعات اللاحقة لمعالم نموذج الانحدار القسيمي

لكن استخدام توزيع لابلاس مباشرة يعد مشكلة كبيرة جداً تسبب تعقيد في عمل الخوارزمية (MCMC), مما تؤثر على كفاءتها في تقدير معالم النموذج, ولتغلب على هذه المشكلة يمكن استخدام تحويلات التوزيعات المختلطة توزيع لابلاس (Laplace distribution), واهم هذه التحويلات هي التحويل الذي قام به [14]

حيث ان  $\Gamma(2) = (2-1)! = 1$ . وحيث تم استخدام التحويل في المعادلة (11) من تقدير معالم نموذج الانحدار الاعتيادي وفق اسلوب الطريقة البيزية. تم توظيف تحويل المبين في المعادلة (11) لتقدير معالم نموذج الانحدار القسيمي من قبل [5] لتفاصيل اكثر انظر [5]. وفي هذا البحث سنستخدم التحويل اخر لتوزيع لتوزيع لابلاس (Laplace distribution) الذي

تفاصيل اكثر انظر [15], حيث ان هذه الصيغ المحولة تملك مرونة عالية في الحصول على نموذج هرمي بيزي عالي الجودة في تقدير التوزيعات اللاحقة, وتمتلك قابلية عالية في بناء خوارزمية (MCMC) جيدة وكفاءة.

• التوزيع اللاحق للمعلمة  $Z_i$

$$\text{InvGa}(Z_i^t, \vartheta^t), \text{ where } \vartheta^t = \frac{\sigma}{2}, Z_i^t = \frac{1}{\sqrt{(y_i - x_i^t \beta)^2}}$$

• التوزيع اللاحق للمعلمة  $\beta_j$  هو التوزيع الطبيعي المبتور بوسط  $\bar{\beta}_j$  وتباين  $\sigma_{\beta_j}^2$

$$\text{and } \sigma_{\beta_j}^2 = \left( \sum_{i=1}^n \frac{\sigma_{x_{ij}}^2}{2Z_i} \right)^{-1} \bar{\beta}_j = \left( \sigma_{\beta_j}^2 \sum_{i=1}^n \frac{\sigma_{x_{ij}} (y_i - (1-2\theta)Z_i - \sum_{j \neq i}^p x_{ij} \beta_j)}{2Z_i} \right) I\{|\beta_j| < u_j\}$$

• التوزيع اللاحق للمعلمة  $u_j$  التوزيع الاسي المبتور  $u_j | \beta, \lambda \sim \text{Exp}(\lambda) I\{u_j > |\beta_j|\}$

$$\sum_{i=1}^n \left( \frac{(y_i - x_i^t \beta + (1-2\theta)Z_i)^2}{4Z_i} + \theta(1-\theta)Z_i \right) + b \text{ معلمة قياس } (a + \frac{3n}{2})$$

• التوزيع اللاحق للمعلمة  $\lambda$  توزيع كما بمعالم

$$(c + 2p, d + \sum_{j=1}^p |\beta_j|).$$

من خلال التوزيعات اللاحقة للمعالم المبينة أعلاه يمكن بناء خوارزمية (MCMC) لتقدير هذه المعالم تفاصيل أكثر انظر [5].

## 2.2 النموذج الهرمي البيزي في حالة (SMUSE) كتوزيع مسبق.

من خلال توظيف دالة التوزيع المختلط (SMUSE) يمكن الحصول على توزيعات اللاحقة لمعالم نموذج الانحدار القسيمي

• التوزيع الاحق للمعلمة  $t$  هي  $\text{InvGa}(Z_i^t, \theta^t)$

$$Z_i^t = \frac{\sigma(y_i - x_i^T \beta)^2}{\xi_2^2} \text{ و } \theta^t = \frac{\sigma \xi_1^2}{\xi_2^2} + 2\sigma$$

• التوزيع اللاحق للمعلمة  $w$  هو توزع الاسي القياسي المبتور

$$e^{-w} I\{w_j > \sigma \lambda |\beta_j|\}$$

• التوزيع اللاحق للمعلمة  $(\beta)$  هو التوزيع الطبيعي المتعدد المبتور

$$N(x_i^T \beta + \xi_1 t_i, \sigma^{-1} \xi_2^2 t_i)$$

• التوزيع اللاحق للمعلمة  $(\sigma)$  هو توزع كما

$$\left( \sum_{i=1}^n \left( \frac{(y_i - x_i^T \beta - \xi_1 t_i)^2}{2 \xi_2^2 t_i} + t_i \right) + b \right) \text{ معلمة قياس } (a + \left( \frac{3n}{2} \right))$$

• التوزيع اللاحق للمعلمة  $(\lambda)$  هو توزع كما

$$\text{معلمة موقع } (k + c) \text{ معلمة قياس } \left\{ \prod_{j=1}^k I\left\{ \lambda_j < \frac{w_j}{\sigma_j |\beta_j|} \right\} \right\}$$

## المبحث الثالث:

### 3.1 البيانات الحقيقية

في هذا الفصل يتم التركيز علي ظاهرة طبية تحتوي

على متغير معتمد يمثل الخثرة الدموية

(D. dimer) ومجموعة من المتغيرات المستقلة التي

سيرد ذكرها لاحقاً، وتم استخدام طريقتين في تقدير معالم الانحدار

القسيمي هي (BSMUQReg) التي اقترحت من قبل الباحث [5]

وطريقة (B SMUSE QReg) التي اقترحت من قبل الباحثين

[16]، حيث تم تنفيذ تقدير المعالم من خلال استخدام خوارزمية

(MCMC) الخاصة بكل طريقة وكل خوارزمية تم تنفيذها

13000 الف دورة أول 3000 دورة تم حرقه لان المقدرات كانت

من خلال التوزيعات اللاحقة للمعالم المبينة أعلاه يمكن

تقدير واختيار المتغيرات في نموذج انحدار القسيمي من خلال

الخوارزمية المعدة لهذا الغرض. تفاصيل أكثر انظر [16]. في هذا

البحث سيتم تحديد أهم المتغيرات المؤثرة على الخثرة الدموية (D.

dimer) باستخدام طريقتي تقدير مختلفتين .

بعد توضيح التوزيعات اللاحقة للمعالم الموضحة

بالطريقتين أعلاه تم بناء خوارزمية MCMC للنموذج الهرمي

البيزي من خلال توظيف (Gibbs sampling)، حيث تميزت

خوارزمية الطريقتين بكفاءة عالية وسرعة في التنفيذ مقدرات بيز .

لمقارنة المفاضلة ما بين طريقتي التقدير في ظل البيانات الحقيقية عند ثلاثة مستويات للانحدار القسيمي، هي المستوى المنخفض (( $\theta_1 = 0.20$ )) والمستوى المتوسط ( $\theta_2 = 0.60$ ) والمستوى المرتفع  $\theta_3 = 0.95$  وكما موضح بالجدول رقم (1)

غير مستقرة، واستخدمت في الدراسة الحالية ثلاثة معايير هي متوسط الأخطاء المطلقة (MAE (Mean Absolute Error)) والانحراف المعياري (SD)(Stander Division) ومتوسط مربعات الخطأ (MSE)(Mean Square Error)

جدول رقم (1) يعرض قيم (MSE)(MAE)(SD) لكلا الطريقتين عند مستويات الانحدار القسيمي الثلاثة

المستوى القسيمي	طرق التقدير	MSE	SD	MAE
$\theta_1 = 0.20$	BSMUQReg	0.521	0.432	0.729
	B SMUSE QReg	0.462	0.370	0.679
$\theta_2 = 0.60$	BSMUQReg	0.445	0.332	0.667
	B SMUSE QReg	0.306	0.153	0.553
$\theta_3 = 0.95$	BSMUQReg	0.452	0.257	0.672
	B SMUSE QReg	0.278	0.299	0.527

المصدر: من عمل الباحثين .

(BSMUQReg) في تقدير المعالم واختيار المتغيرات المؤثرة على الخثرة الدموية (D. dimer) بعد تحديد الطريقة الأفضل في تقدير بياناتنا يمكن استخدامها في تقدير نموذج دراستنا عند المستويات الثلاثة المبينة أعلاه. لذا سنعتمد على الطريقة (B SMUSE QReg) التي تمتلك كفاءة عالية في تقدير معالم النموذج واختيار المتغيرات، كما هو مبين في الجدول رقم (2)

من خلال النتائج المعروضة في الجدول أعلاه نجد أن قيم (MSE)، (SD)، و (MAE) المحسوبة وفق الطريقة (B SMUSE QReg) اقل من قيم (MSE)، (SD)، و (MAE) المحسوبة وفق الطريقة (BSMUQReg) من خلال هذه النتيجة نجد ان طريقة (B SMUSE QReg) أفضل من طريقة (BSMUQReg) في تقدير معاملات المتغيرات المؤثرة على الخثرة الدموية خلال مستويات الكونتال القسيمي الثلاثة. حسب النتائج المعروضة في الجدول واحد سنعتمد على طريقة

جدول رقم (2) يبين تقدير معالم الانحدار القسيمي خلال المستويات الثلاثة باستخدام طريقة B SMUSE QReg

المتغيرات	اسم المتغير	رمز المتغير	تقدير المعالم $\theta_1 = 0.20$	تقدير المعالم $\theta_2 = 0.60$	تقدير المعالم $\theta_3 = 0.95$
X1	Haptoglobin	Hb	1.245	1.301	1.141
X2	Packed cell volume	PCV	0.854	1.451	1.008
X3	White blood cells	WBC	1.451	1.045	0.845
X4	Erythrocyte Sedimentation Rate	E . S .R	-0.045	0.000	0.000
X5	Blood group	Blood group	0.000	0.000	0.000
X6	Silurian uricacid	S. uric acid	1.745	1.542	0.048

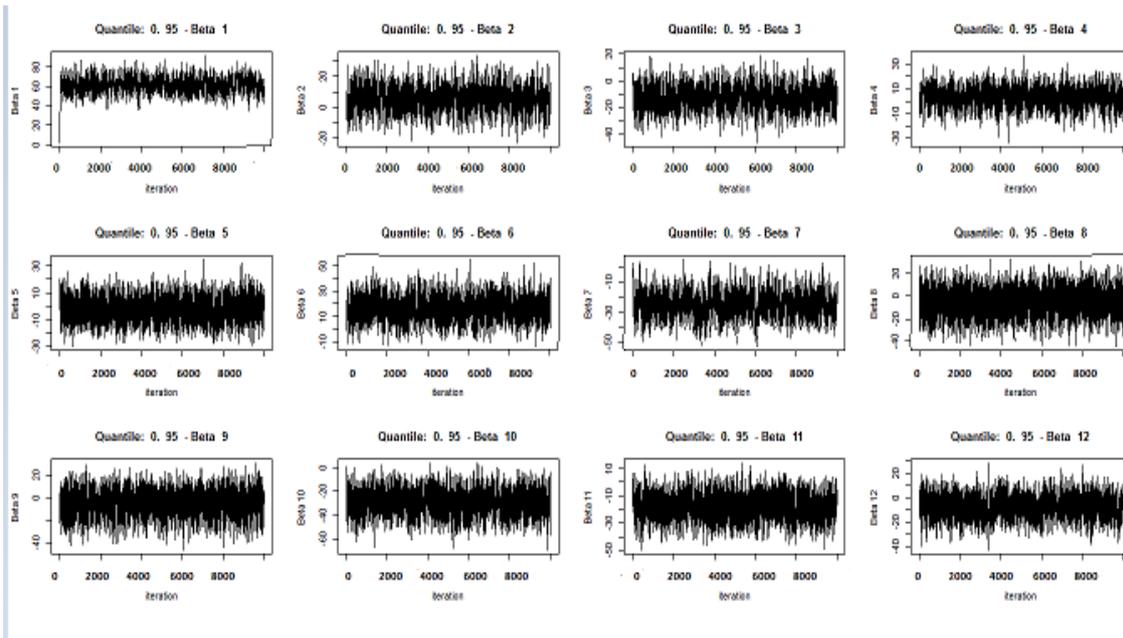
X7	Silurian cholesterol	S. cholesterol	1.418	1.473	1.542
X8	Silurian triglyceride	S. triglyceride	0.447	0.241	0.054
X9	Low-density lipoprotein	LDL	0.874	1.349	1.179
X10	high-density lipoprotein	HDL	1.457	1.761	1.245
X11	COVID-19	COVID-19	1.782	1.547	1.545
X12	Gender	Gender	0.000	0.000	0.000

المصدر: من عمل الباحثين.

من خلال النتائج المعروضة في الجدول (2) نلاحظ عند المستوى القسيمي المنخفض ( $\theta_1 = 0.20$ ) ان هناك عشرة متغيرات كانت ذات تأثير طردي وعكسي على متغير الخثرة الدموية، ومتغيري (gender) و (Blood Group) كان غير مؤثرين على متغير الخثرة الدموية. حيث نجد ان هناك 9 متغيرات ذات تأثير طردي على متغير الخثرة الدموية وهي (Hb, S., S. cholesterol, uric acid, WBC, PCV, triglyceride, LDL, HDL, COVID-19) اما بقية المتغيرات البالغة (3) متغيرات فهي (E . S .R, gender, blood group) هي متغيرات غير مؤثرة لان قيم المقدرة عند هذه المتغيرات الثلاثة مساوية الى الصفر.

ولمعرفة سلوك خوارزمية (MCMC) الخاصة في تحليل بياناتنا سنستخدم على رسم (Trace plot) و (Histogram) الخاص بمتغيرات بياناتنا الحقيقية وعند المستوى القسيمي المرتفع ( $\theta_3 = 0.95$ ) وحسب الشكل رقم (3)

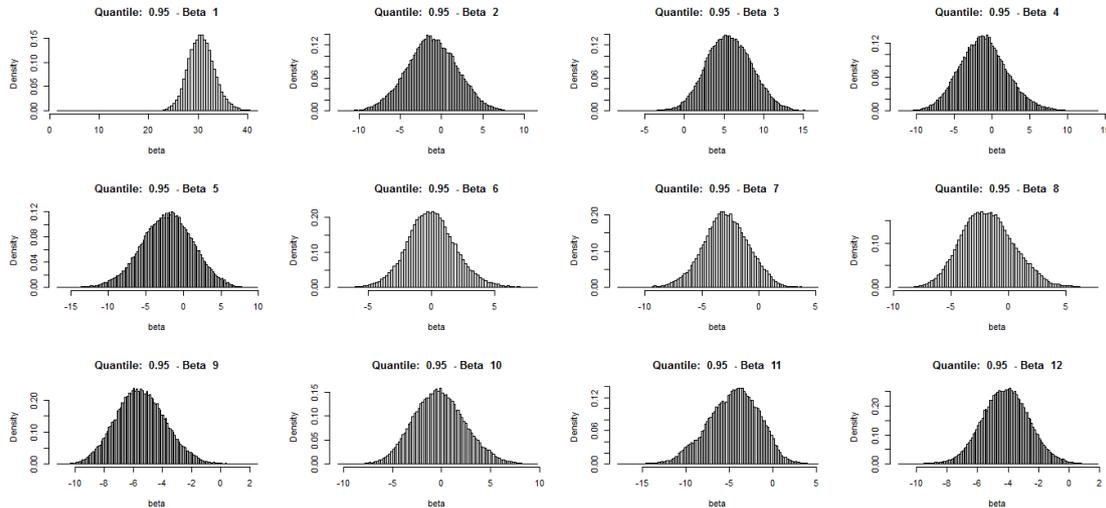
من خلال النتائج المعروضة في الجدول (2) نلاحظ عند المستوى القسيمي المنخفض ( $\theta_1 = 0.20$ ) ان هناك عشرة متغيرات كانت ذات تأثير طردي وعكسي على متغير الخثرة الدموية، ومتغيري (gender) و (Blood Group) كان غير مؤثرين على متغير الخثرة الدموية. حيث نجد ان هناك 9 متغيرات ذات تأثير طردي على متغير الخثرة الدموية وهي (Hb, S., S. cholesterol, uric acid, WBC, PCV, triglyceride, LDL, HDL, COVID-19) اما بقية المتغيرات البالغة (3) متغيرات فهي (E . S .R, gender, blood group) هي متغيرات غير مؤثرة لان قيم المقدرة عند هذه المتغيرات الثلاثة مساوية الى الصفر.



شكل رقم (3) يمثل رسم (Trace plot) لمتغيرات المؤثرة على الخثرة الدموية

المؤثرة على الخثرة الدموية خلال 10000 دورة. وكما موضح بالشكل رقم (4)

من خلال الشكل رقم 3 نجد ان خوارزمية (MCMC) التي تعود لطريقة (BSMUSE QReg) والخاصة بتقدير بياناتنا الحقيقية عند المستوى القسيمي المرتفع ( $\theta_3 = 0.95$ ) مستقرة خلال تقدير معالم المتغيرات



شكل رقم (4) يمثل رسم (Histogram) لمتغيرات المؤثرة على الخثرة الدموية

D. نجد ان هناك تطابقاً كبيراً جداً بين مقدرات معالم نموذجنا المدروس مع الجانب النظري الخاص بمعالم النموذج المقدر بطريقة التقدير (BSMUSE QReg), حيث نجد ان توزيع معالم النموذج المقدر يتبع التوزيع الطبيعي في الجانب النظري وفي تقدير معالم المتغيرات .

E. من خلال القيم المقدره عند مستويات القسيمية الثلاثة نجد ان علاقة المتغيرات التوضيحية هي علاقة طردية مع متغير الخثرة الدموية، يعني بزيادة اي من هذه المتغيرات ستسبب زيادة في الخثرة الدموية.

#### 4.2 التوصيات

A. توسعت الدراسة الحالية لتشمل متغيرات توضيحية أكثر وحجم عينة أكبر وتوظيف تقنية اختيار المتغيرات في تحديد المتغيرات ذات التأثير القوي على الخثرة الدموية.  
B. من خلال الدراسة الحالية يجب قياس تحليل الخثرة الدموية (D. dimer) عند اصابة الشخص بارتفاع نسبة الدهون الثلاثية (HDL, LDL, S. cholesterol), او اصابة الشخص (COVID-19) لان النتائج تظهر ان هناك علاقة اثر قوية جداً بين هذه المتغيرات التوضيحية والخثرة الدموية

من خلال الشكل رقم (4) نجد ان المعالم المقدره لكل المتغيرات التوضيحية الخاصة ببياناتنا قيد الدراسة تتوزع وفق التوزيع الطبيعي وهذا يعني ان هناك تطابقاً مع توزيعات معالم النموذج الانحدار القسيمي .

#### المبحث الرابع

#### الاستنتاجات والتوصيات

#### 4.1 الاستنتاجات

A. من خلال النتائج المعروضة نجد ان طريقة (BSMUSE QReg) هي افضل من طريقة (BSMUQReg) في تحليل اثر العلاقة بين متغيرات الدراسة الحالية .

B. نجد ان المتغيرين التوضيحيين (Gender) (Blood Group) ذات تأثير صفري على الخثرة الدموية عند المستويات القسيمية الثلاثة المستخدمة في الدراسة الحالية لكن المتغير التوضيحي (ESR) كان متغيراً ذا تأثير صفري عند المستوى القسيمي المتوسط والمرتفع .

C. نجد ان مقدرات المتغيرات التوضيحية الخاصة بنموذجنا المدروس بحسب خوارزمية (MCMC) التي تعود لطريقة التقدير (BSMUSE QReg) كانت مستقرة خلال دورات التقدير، وذلك واضح من خلال شكل رقم 2

## المصادر

- [9] Alhamzawi, R. (2014). Bayesian elastic net tobit quantile regression. *Communications in Statistics - Simulation and Computation* 45 (7), 2409–2427.
- [10] Draper, N., & Smith, H. (1981). *Applied regression analysis*. Series in probability and mathematical statistics. Wiley.
- [11] Yu, K. and R. A. Moyeed (2001). Bayesian quantile regression. *Statistics & Probability Letters* 54, 437–447.
- [12] Kotz, S., Kozubowski, T. J., and Podg`orski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkh` auser, Boston.
- [13] Kozumi, H. and Kobayashi, G. (2011). "Gibbs sampling methods for Bayesian quantile regression." *Journal of statistical computation and simulation*, 81(11): 1565–1578. 3.
- [14] Andrews, D. F. and C. L. Mallows (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, Series B* 36, 99–102. 7.
- [15] Mallick, H. and Yi, N. (2014). "A new Bayesian lasso." *Statistics and Its In-Terface*, 7(4): 571–582. 2, 3, 4.
- [16] Hassan.R.O and Alhusseini .F.H.H (2021)" Bayesian lasso in quantile regression with a new prior" *journal of Al-Qadsia*.
- [1] Koenker, R. and G. J. Bassett (1978). Regression quantiles. *Econometrica* 46, 33–50.
- [2] Wu, Y., & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica*, 801-817.
- [3] Cade, B. S. and Noon, B. R. (2003). "A gentle introduction to quantile regression for ecologists." *Frontiers in Ecology and the Environment*, 1(8): 412–420. 1.
- [4] Alhusseini F. H.H.(2018):. "Bayesian variables selection and coefficients estimation in Tobit Quantile Regression." *Romanian university of Craiova, Romanian*.
- [5] Alhusseini F. H. H (2017) " Bayesian Quantile Regression With Scalemixture Of Uniform Prior Distributions " *International Journal of Pure and Applied Mathematics* pp 77-91.
- [6] Koenker, R. and V. D'Orey (1987). Algorithm AS 229: Computing regression quantiles. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 36, 383–393.
- [7] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., & Knight, K. (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1), 91-108.
- [8] Li, Y., & Zhu, J. (2008). L 1-norm quantile regression. *Journal of Computational and Graphical Statistics*, 17(1), 163-185.