

استخدام خوارزمية MCMC لإيجاد مقدر بايز لتوزيع باريتو بمعلمتين بالاعتماد على دالة خسارة متوازنة

لؤي فرح *، وفاء عيسى **، مصطفى مظهر رنة **

(الإيداع: 20 حزيران 2019 ، القبول: 3 آيلول 2019)

المخلص:

يعتبر توزيع باريتو من أكثر توزيعات الفشل انتشاراً لنماذج الإجهاد والمتانة ويستخدم بشكل واسع في تطبيقات نظرية الموثوقية ولهذا التوزيع تطبيقات مهمة في حقول متعددة منها بحوث العمليات وفي بناء نماذج نظرية الطوابير وكذلك في الاتصالات والهندسة الميكانيكية وفي علم الاقتصاد كنموذج لدخل الأفراد. تم في هذا البحث إيجاد مقدرات بايز لتوزيع باريتو بمعلمتين بالاعتماد على دوال خسارة متماثلة وغير متماثلة أوجدنا بإيجاد مقدرًا جديدًا لتوزيع باريتو بمعلمتين مجهولتين باستخدام سلسلة ماركوف مونت كارلو (MCMC) بالاعتماد على دالة خسارة جديدة (متوازنة) وتوزيعات قبلية مرافقة. وللتحقق من كفاءة أداء المقدر المقترح تم مقارنة هذا المقدر مع مقدرات بايز ومقدر الإمكانية العظمى باستخدام بيانات مولدة من توزيع باريتو بمعلمتين وبحجم عينة 100 وذلك بالاعتماد على مقياس متوسط مربعات الخطأ (MSE). دلت نتائج المقارنة أن المقدر المقترح كان الأفضل من بين المقدرات المدروسة فالاستدلالات التي حصلنا عليها لتوزيع باريتو ذي المعلمتين تعتبر نتائج جديدة تغطي كل النتائج السابقة التي تناولت هذا التوزيع أو أحد الحالات الخاصة المتعلقة به وتم حل مشكلة الحسابات المعقدة التي تنتج في استدلالات بايز باستخدام سلسلة ماركوف مونت كارلو (MCMC) .

الكلمات المفتاحية: توزيع باريتو، سلسلة ماركوف مونت كارلو (MCMC)، مقدر بايز، دالة الخسارة الأسية الخطية، دالة خسارة متوازنة

* طالب دراسات عليا (ماجستير)-قسم الإحصاء الرياضي-كلية العلوم-جامعة حلب

** قسم الإحصاء الرياضي-كلية العلوم-جامعة حلب

Using the MCMC algorithm to find the Bayes estimator of Pareto distribution with two Parameters based on a balanced loss function

Louay Farah *, Wafaa Issa **, Moustafa Mazhar Rene **

(Received: 20 June 2010 , Accepted: 3 September 2019)

Abstract:

Pareto's distribution is one of the most widespread failure distributions of stress and strength models and is widely used in reliability theory applications. This distribution is important in many fields, including operations research, queuing theory, communications, mechanical engineering, and economics as a model for individuals' income.

In this research, We found baysian estimator of Pareto distribution with two Parameters based on Symmetrical and asymmetrical loss functions. We have found a new estimator of Pareto distribution with two unknown Parameters using the Markov chain Monte Carlo (MCMC) based on a new balanced loss function and Conjugate Prior distributions, and to verify the efficiency of the performance of the proposed estimate, This estimate was compared with Bayes' estimations and maximum likelihood estimation using data generated from Pareto distribution with two parameters and sample size 100 based on the Mean Error Squared (MSE) scale.

The results of the comparison showed that the proposed estimate was the best of the studied estimations, since the inferences obtained for Pareto's distribution of the parameters are new results Covering all the previous results that dealt with this distribution or one of the special cases related to it and solved the problem of complex calculations that result in Bayes' inferences using Markov chain Monte Carlo (MCMC).

Keywords: Pareto's distribution, Markov chain Monte Carlo (MCMC), baysion estimator, Linear-exponential loss function, Balanced loss function

*Postgraduate Student (MSc.)–Dept. of Mathematical Statistics –Faculty of Science–University of Aleppo

**Dept. of Mathematical Statistics–Faculty of Science–University of Aleppo

1-مقدمة: Introduction

تستخدم التوزيعات الاحتمالية للتعبير عن المجتمعات الإحصائية التي تعتمد على معالم المجتمع تحت الدراسة وتعتبر عملية تقدير هذه المعالم من الأمور الأساسية في الاستدلال الإحصائي حيث يتم عادةً تقدير هذه المعالم باستخدام طرائق إحصائية تقليدية مثل طريقة الإمكانية العظمى والطرائق البايزية ، كما هو معلوم فإن التقدير البايزي يعتمد على مدى ملاءمة اختيارنا للتوزيع القبلي المقترح لهذه المعالم وأيضاً على دالة الخسارة المستخدمة التي تلعب دور أساسياً في المنهج البايزي حيث أن معظم التقديرات البايزية مبنية على دالة الخسارة التربيعية نظراً لسهولة الحصول على المقدرات المعتمدة عليها وهي دالة متناظرة أي ما فوق التقدير متساوي مع تحت التقدير حيث تعطي أهمية متساوية لحالتي التقدير الأعلى و الأدنى وهذا لا يتوافق مع عدة حالات يكون فيها التقدير الأعلى أهم من الأدنى كما في تقدير دالة الموثوقية مثلاً، (Lehmann; Casella, 2006).

2-مشكلة البحث:

سنحاول من خلال هذا البحث أخذ هذه الجزئية بعين الاعتبار وذلك باستخدام دالة الخسارة المتوازنة والتي تحتوي على دالة خسارة تربيعية أو أسية خطية كحالة خاصة لإيجاد مقدرات بايز لتوزيع باريتو بمعلمتين حيث لهذا التوزيع تطبيقات مهمة في حقول متعددة منها بحوث العمليات وفي بناء نماذج نظرية الطوابير وكذلك في الاتصالات و الهندسة الميكانيكية و في علم الاقتصاد كنموذج لدخل الأفراد، (Arnold,2015).

3- أهمية البحث:

تكمن أهمية البحث في تسليط الضوء على أهمية تطبيق سلسلة ماركوف مونت كارلو (MCMC) لمعالجة مشكلة الحسابات المعقدة التي تنتج في استدلالات بايز.

4-أهداف البحث:

يهدف هذا البحث بشكل أساسي إلى إيجاد مقدرات جديدة لمعلمتي توزيع باريتو المجهولتين وقيام بحسابات معقدة باستخدام سلسلة ماركوف مونت كارلو (MCMC) بالاعتماد على دالة خسارة جديدة (متوازنة) وتوزيعات قبيلية مرافقة.

5-المواد وطرائقُ البحث: Materials and Methods

5-1- توزيع باريتو بمعلمتين:

يستخدم توزيع باريتو بشكل واسع في تطبيقات نظرية الموثوقية لأنه أحد توزيعات الفشل لنماذج الإجهاد والمتانة وكذلك يستخدم في وصف توزيع السلوك المتطرف لقيمة الخسارة في مجال الاقتصاد ويستخدم أيضاً في مختلف مجالات العلوم منها الفيزياء و الجيولوجيا وأنظمة الاتصالات ، وينسب هذا التوزيع إلى العالم الاقتصادي الإيطالي الجنسية Fleverdu Pareto عندما حاول وضع قانون شمولي للتعامل مع توزيع الدخل لمجتمع معين وانطلق في نظريته واشتق صيغة توزيع باريتو. يوجد عدة أشكال لتوزيع باريتو منها توزيع بمعلمة واحدة (معلمة قياس) ومنها بمعلمتين (معلمة شكل ، معلمة قياس) وقد أنجزت عدة بحوث حول توزيع باريتو بمعلمة واحدة وسوف يتضمن بحثنا تقدير معلمتي توزيع باريتو. نقول عن X أنه يخضع لتوزيع باريتو بمعلمتين إذا كانت له دالة كثافة التالية:

$$f(x; \alpha, \beta) = \frac{\alpha\beta^\alpha}{(x + \beta)^{\alpha+1}} ; x > 0, \alpha > 0, \beta > 0 \quad (1)$$

حيث α معلمة الشكل و β معلمة القياس والتوقع الرياضي للتوزيع باريتو يعطى بالعلاقة:
 $E(X) = \frac{\beta}{\alpha-1}$ و التباين يعطى بالعلاقة $V(X) = \left[\frac{\alpha}{\alpha-2} - \frac{3\alpha-1}{\alpha-1} - \frac{1}{(\alpha-1)^2} \right] \beta^2$ حيث $\alpha > 2$ ، (Abdel-Ghaly, Attia, Aly, 1998)، (Parikh, 2011).

5-2- طرائق التقدير:

هناك منهجين للتقدير المنهج البايزي والمنهج الكلاسيكي حيث تعتبر معلمة التوزيع الاحتمالي للمجتمع مجهولة ثابتة ويتم تقديرها من خلال عينة عشوائية مسحوية من هذا المجتمع.

هناك عدة طرق للتقدير الكلاسيكي من أهم وأشهر الطرق المستخدمة في هذا الاتجاه طريقة الإمكانية العظمى التي سيتم التعرض لها بهدف المقارنة، (Lehmann, Casella, 2006).

5-2-1- تعريف مقدر الإمكانية العظمى:

يسمى المقدر $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r)$ بمقدر الإمكانية العظمى لمتجه المعالم $\theta = (\theta_1, \theta_2, \dots, \theta_r)$ إذا كانت دالة الإمكان $L(x, \hat{\theta}) \geq L(x, \theta)$ لجميع قيم θ ، أي أن مقدر الإمكانية العظمى $\hat{\theta}$ هو قيمة $\hat{\theta}_{MLE}$ التي تعظم دالة الإمكانية أو لوغاريتم دالة الإمكانية ونحصل عليها بحل المعادلات التالية، (MUKHOPADHYAY, 2000):

$$\frac{\partial}{\partial \theta_i} \ln L(\theta_1, \theta_2, \dots, \theta_r | x) = 0 ; i = 1, 2, \dots, r \quad (2)$$

5-2-2- التقدير البايزي :

أهم ما يميز التقدير البايزي تفسير θ على أنها متغير عشوائي له الكثافة الاحتمالية $\pi(\theta)$ التي تعبر عما يعتقد الفرد حول قيمة حدوث θ قبل أي مشاهدة تؤخذ ويدعى توزيع قبلي للمعلمة . إن طريقة بايز تعتمد على دمج المعلومات الأولية مع معطيات العينة، حيث يتم تحويل الاحتمالات الأولية إلى احتمالات لاحقة حيث نحصل على التوزيع الاحتمالي البعدي $\pi(\theta | x)$ وهو التوزيع الاحتمالي للمعلمة θ بشرط الحصول على العينة X وهو يصف درجة اعتقادنا في القيم المختلفة للمعلمة θ بعد الحصول على العينة أي أن قراءات العينة تغير درجة اعتقادنا في القيم المختلفة للمعلمة θ وذلك بتغيير التوزيع القبلي إلى توزيع بعدي، (Ghosh, Delampady, Samant, 2006).

5-3- مقدر بايز النقطي:

في تقدير بايز توجد هناك مشكلة في اتخاذ القرار حيث نواجه بموقف اختيار قرار decision أو إجراء action معين من بين مجموعة من الإجراءات في ظل عدم التأكد أو في ظل عدم وجود معلومات كاملة فلو اتخذنا قراراً بأن إجراء معيناً من فراغ الإجراءات هو الذي سيعبر عن القيمة الممكنة للمعلمة المراد تقديرها فلا بد من التفكير في عواقب مثل هذا القرار، وهذه تعتبر من أهم الصعوبات العملية في نظرية اتخاذ القرار theory decision، وهذا ما يقودنا إلى حتمية تقييم نتائج هذا الإجراء بناءً على مقياس كمي ولو بالتقريب حتى يكون هناك أساس نظري منطقي معقول لنظرية اتخاذ القرارات. لذلك سنفترض أن هناك دالة معطاة تبين الخسارة الناتجة من اتخاذ قرار أن $\hat{\theta}$ هي مقدر للمعلمة θ ، أي سنفترض أن الدالة $L(\hat{\theta}, \theta)$ تقيس الخسارة التي نتكبدها إذا اخترنا الإجراء $\hat{\theta}$ وهذه الدالة هي ما تسمى بدالة الخسارة، وفي نظرية التقدير بعد تحديد دالة الخسارة المناسبة يكون هدفنا هو اختيار المقدر الذي يجعل مخاطرة ببيز risk Bayes أقل ما يمكن عندما نقدر المعلمة θ بالمقدر $\hat{\theta}$ أي أن مقدر بايز $\hat{\theta}$ هو القيمة التي تجعل القيمة المتوقعة لدالة الخسارة بالنسبة للتوزيع البعدي أقل ما يمكن أي أن، (Lehmann; Casella, 2006):

$$\min_{\hat{\theta}} E[l(\hat{\theta}, \theta)] = \min_{\hat{\theta}} \int l(\hat{\theta}, \theta) f(\theta | x) d\theta \quad (3)$$

4-5- تقدير بايز في حالة عدة معالم مجهولة:

بفرض لدينا مجموعة المعالم $\theta = (\vartheta_1, \vartheta_2, \dots, \vartheta_j)$ ونريد تقديرها ولها التوزيع البعدي المشترك $\pi(\theta|x)$ حيث X عينة المشاهدات فإن التوزيع الهامشي لأي معلمة مثلاً ϑ_j يمكن الحصول عليه من ايجاد تكامل $\pi(\theta|x)$ على المعالم بعد تثبيت ϑ_j ويعطى التوزيع الهامشي البعدي للمعلمة ϑ_j

$$\pi_j(\vartheta_j|x) = \int \dots \int \pi(\vartheta_1, \dots, \vartheta_j | X) d\vartheta_1 \dots d\vartheta_{j-1} d\vartheta_{j+1} \dots d\vartheta_j \quad (4)$$

بعد إيجاد التوزيع الهامشي لكل معلمة يكون المقدر بايز للمعلمة ϑ_j وفق العلاقة:

$$\hat{\vartheta}_j = \min_{\vartheta_j} E[l(\hat{\theta}, \vartheta_j)] = \min_{\vartheta_j} \int l(\hat{\theta}, \vartheta_j) \pi_j(\vartheta_j|x) d\vartheta_j \quad (5)$$

(Rowe, 2003).

5-5- دالة الخسارة التربيعية Squared error loss function:

عند تقدير المعلمة θ بالمقدر $\hat{\theta}$ فإن دالة الخسارة التربيعية وهي دالة متماثلة يكون فيها فوق التقدير (Overestimation) متساوي مع تحت التقدير (Underestimation) تعطى بالعلاقة :

$$l_1(\hat{\theta}, \theta) \propto (\hat{\theta} - \theta)^2 \quad (6)$$

ومخاطرة بايز أو التوقع الرياضي لدالة الخسارة بالنسبة لتوزيع البعدي هو:

$$R(\hat{\theta}, \theta) = E(l_1(\hat{\theta}, \theta)|x) = E((\hat{\theta} - \theta)^2|x) \quad (7)$$

مقدر بايز للمعلمة θ اعتماداً على دالة خسارة التربيعية هو قيمة $\hat{\theta}_{BS}$ التي تجعل المخاطرة أقل مايمكن هي:

$$\hat{\theta}_{BS} = E(\theta|x) \quad (8)$$

أي أن مقدر بايز اعتماداً على دالة الخسارة التربيعية هو القيمة المتوقعة لهذه المعلمة بالنسبة للتوزيع البعدي، (Lehmann; Casella, 2006).

5-6- دالة الخسارة الأسية الخطية Linear-exponential loss function:

عند تقدير المعلمة θ بالمقدر $\hat{\theta}$ فإن دالة الخسارة الأسية الخطية LINEX تعطى بالعلاقة:

$$l_2(\Delta) \propto e^{a\Delta} - a\Delta - 1 \quad (9)$$

حيث $\Delta = (\hat{\theta} - \theta)$ و $a \neq 0$ ثابت يمثل معلمة الشكل للدالة $l_2(\Delta)$ ومن أهم خواص هذه الدالة أنها غير متماثلة حيث تتحكم القيمة العددية لمعلمة الشكل a في درجة واتجاه عدم التماثل للدالة، أما إشارة المعلمة a فتعكس اتجاه عدم التماثل. بمعنى أنه لقيم $a > 0$ فإن التقدير الأعلى يكون أكثر خطورة من التقدير الأدنى، والعكس أيضاً صحيح، بمعنى أنه لقيم $a < 0$ يكون التقدير الأدنى أكثر خطورة من التقدير الأعلى. ولقيم $|a|$ الصغيرة جداً تؤول دالة الخسارة الخطية الأسية إلى دالة خسارة مربع الخطأ وتكون متماثلة.

تعطى مخاطرة بايز بالعلاقة:

$$l_2 R(\hat{\theta}, \theta) = E(l_2(\Delta)|x) \propto e^{a\hat{\theta}} E(e^{-a\theta}|x) - a[\hat{\theta} - E(\theta|x)] - 1 \quad (10)$$

ومقدر بايز للمعلمة θ الذي يجعل الطرف الأيمن أقل ما يمكن يعطى بالعلاقة:

$$\hat{\theta}_{BL} = -\frac{1}{a} \ln[E(e^{-a\theta} | \underline{x})] \quad (11)$$

بشرط $E(e^{-a\theta})$ موجود ومحدود، (Ahmadi, Doostparast, Parsian, 2005).

7-5- دالة الخسارة المتوازنة **Balanced loss function**:

تعطى دالة الخسارة المتوازنة بالعلاقة التالية:

$$l_{\rho, \omega, \xi}^q(\gamma(\theta), \delta) = \omega q(\theta) \rho(\xi(\underline{x}), \delta) + (1 - \omega) q(\theta) \rho(\gamma(\theta), \delta) \quad (12)$$

حيث δ هو تقدير للمعلمة $\gamma(\theta)$ ، $\xi(\underline{x})$ تقدير سابق للمعلمة $\gamma(\theta)$ ، $0 \leq \omega \leq 1$ ، دالة خسارة اختيارية لتقدير $\gamma(\theta)$ بالمقدر δ ، دالة وزن موجبة

وفي هذه البحث سيتم استخدام دالة الخسارة الخطية المتوازنة بمعلمة الشكل $a \neq 0$ ونحصل عليها بتعويض

$$\xi(\underline{x}) = \hat{\theta}_{MLE}, q(\theta) = 1, \rho(\gamma(\theta), \delta) = e^{a(\gamma(\theta) - \delta)} - a(\gamma(\theta) - \delta) - 1$$

العلاقة (10) كما يلي:

$$l_3(\gamma(\theta), \delta) = \omega [e^{a(\delta - \hat{\theta}_{MLE})} - a(\delta - \hat{\theta}_{MLE}) - 1] + (1 - \omega) [e^{a(\gamma(\theta) - \delta)} - a(\gamma(\theta) - \delta) - 1] \quad (13)$$

ويكون مقدر بايز للمعلمة $\gamma(\theta)$ اعتماداً على دالة الخسارة $l_3(\gamma(\theta), \delta)$ على الشكل:

$$\hat{\delta}_{new} = -\frac{1}{a} \ln[\omega e^{-a\hat{\theta}_{MLE}} + (1 - \omega) E(e^{-a\gamma(\theta)} | \underline{x})] \quad (14)$$

نلاحظ من أجل $\omega = 0$ نحصل على مقدر بايز اعتماداً على دالة الخسارة الخطية الأسية، ومن أجل $\omega = 1$ نصل على مقدر الإمكانية العظمى، (Ahmadi, Jozani, Marchand, Parsian, 2009).

مما سبق نجد أن مقدرات بايز تعتمد في حسابها على القيم المتوقعة لدوال في المعالم اعتماداً على التوزيعات البعدية والتي غالباً ما تحتوي على تكاملات معقدة يصعب حسابها حتى باستخدام طرق التكامل العددية المعروفة وقد ظهرت طرائق عدة لحساب التكاملات مباشرة من أهم تلك الطرق هي سلسلة ماركوف مونت كارلو والتي تكتب بشكل مختصر MCMC.

5-8- سلسلة ماركوف مونت كارلو وتطبيقها في التقدير البايزي:

إن طرائق سلسلة ماركوف مونت كارلو هي طريقة لمحاكاة التوزيع الاحتمالي البعدي $\pi(\theta | \underline{x})$ عن طريق توليد سلسلة ماركوف اركوديك $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(T)}$ من التوزيع البعدي $\pi(\theta | \underline{x})$ من ثم بتطبيق نظرية اركوديك فإن المقدرات تعطى بالعلاقة، (Casella, 2008)، (Ntzoufras, 2011):

$$\widehat{g(\theta)} = \frac{1}{T} \sum_{i=1}^T g(\theta^{(i)}) \quad (15)$$

حيث $\widehat{g(\theta)} \rightarrow E(g(\theta))$ عندما $T \rightarrow \infty$

من أشهر خوارزميات MCMC خوارزمية:

1- ميتروبوليس- هاستنغ (Metropolis-Hastings)

2- معاينة جيبس (Gibbs sampler algorithm)

3- متروبوليس داخل جيبس (Metropolis within Gibbs) وفي مايلي نعرض الخوارزميات:

أ-ميتروبوليس- هاستنغ (Metropolis-Hastings):

صمم ميتروبوليس (Metropolis) عام (1953) خوارزمية ميتروبوليس من خلال إدخال طرق المحاكاة المعتمدة على سلاسل ماركوف. وفي مابعد قدم Hastings عام 1970 تعميم لخوارزمية ميتروبوليس في ما يعرف باسم خوارزمية ميتروبوليس-

هاستنغ (Metropolis- Hastings) وتعتبر هذه الخوارزمية هي الصيغة العامة لجميع خوارزميات MCMC بفرض لدينا التوزيع البعدي $\pi(\theta|x)$ تدعى دالة هدف حيث نريد توليد عينة عشوائية منه بحجم T و $q(x^*|x^{(t)})$ دالة كثافة شرطية من الممكن محاكاتها بسهولة وتدعى التوزيع المقترح proposal distribution عندئذ يمكن تلخيص خوارزمية Metropolis-Hastings بالخطوات التالية، $\theta^{(t)}$ القيمة المولدة في تكرار t :

1- نفرض القيمة الابتدائية $\theta^{(0)}$ من أجل $t = 0$

2- نولد θ^* من توزيع $q(\theta^*|\theta^{(t-1)})$ و نولد $U \sim Uniform(0,1)$

3- نحسب احتمال القبول $\alpha(\theta^*, \theta^{(t-1)}) = \min(r, 1)$ حيث

$$r = \frac{p(\theta^*|X)q(\theta^*|\theta^{(t-1)})}{P(\theta^{(t-1)}|X)q(\theta^{(t-1)}|\theta^*)} = \frac{p(X|\theta^*)p(\theta^*)q(\theta^*|\theta^{(t-1)})}{P(X|\theta^{(t-1)})p(\theta^{(t-1)})q(\theta^{(t-1)}|\theta^*)}$$

4- إذا كان $U \leq \alpha(\theta^*, \theta^{(t-1)})$ فإن $\theta^{(t)} = \theta^*$ وإلا $\theta^{(t)} = \theta^{(t-1)}$

5- نكرر الخطوات السابقة من أجل $t = 1, 2, \dots, T$

(Ntzoufras, 2011)، (Gilks, Richardson, Spiegelhalter, 1995).

ب- معاينة جيبس (Gibbs sampler algorithm):

ظهرت معاينة جيبس في ورقة نشرها Geman&Geman عام 1984 في عمليات معالجة الصور وهي واحدة من أشهر الخوارزميات المستخدمة في الاستدلال البايزي في حالة المعالم المتعددة سنقوم بعرضها بشكل التالي: بفرض لدينا $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ وأن لها التوزيعات شرطية $\pi_j(\theta_j|\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_k, \underline{x})$ حيث $j = 1, 2, \dots, k$ وهذه التوزيعات تدعى توزيعات شريطة تامة عندئذ يمكن وصف خطوات الخوارزمية، $\theta^{(t)}$ القيمة المولدة في تكرار t :

1- نفرض القيمة الابتدائية $\underline{\theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$

2- نولد قيمة المعلمة θ_1 ولتكن $\theta_1^{(t)}$ من التوزيع الشرطي $\pi_1(\theta_1|\theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_k^{(t-1)}, \underline{x})$

3- نولد قيمة المعلمة θ_2 ولتكن $\theta_2^{(t)}$ من التوزيع الشرطي $\pi_2(\theta_2|\theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_k^{(t-1)}, \underline{x})$

4- وهكذا.... نستمر حتى نولد قيمة المعلمة θ_k ولتكن $\theta_k^{(t)}$ من التوزيع الشرطي

$$\pi_k(\theta_k|\theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{k-1}^{(t)}, \underline{x})$$

5- نكرر الخطوات السابقة من أجل $t = 1, 2, \dots, T$

بالتالي سنحصل على سلسلة ماركوف الإركوديك لكل معلمة $\theta_1, \theta_2, \dots, \theta_k$ من ثم نقوم بالتقديرات المطلوبة بتطبيق نظرية إركوديك، (Ntzoufras, 2011)، (Gilks, Richardson, Spiegelhalter, 1995).

د- متروبوليس داخل جيبس (Metropolis within Gibbs):

في بعض الأحيان من الصعب إيجاد التوزيعات الشرطية التامة عندئذٍ نستخدم خوارزمية متروبوليس-هاستغ لتوليد الكثافات الشرطية التامة في خوارزمية جيبس وهذه تسمى خوارزمية متروبوليس داخل جيبس والتي يمكن تلخيصها بالخطوات التالية. بفرض $\theta^{(t)}$ القيمة المولدة في تكرار t :

$$1- \text{نفرض القيمة الابتدائية } (\vartheta_2^{(0)}, \dots, \vartheta_k^{(0)}) = \underline{\theta}^{(0)}$$

2- نولد قيمة المعلمة ϑ_1 ولتكن $\vartheta_1^{(t)}$ من $\pi_1(\vartheta_1 | \vartheta_2^{(t-1)}, \vartheta_3^{(t-1)}, \dots, \vartheta_k^{(t-1)}, \underline{x})$ باستخدام خوارزمية ميتروبوليس-هاستغ.

3- نولد قيمة المعلمة ϑ_2 ولتكن $\vartheta_2^{(t)}$ من $\pi_2(\vartheta_2 | \vartheta_1^{(t)}, \vartheta_3^{(t-1)}, \dots, \vartheta_k^{(t-1)}, \underline{x})$ باستخدام خوارزمية ميتروبوليس-هاستغ.

4- وهكذا.... نستمر حتى نولد قيمة المعلمة ϑ_k ولتكن $\vartheta_k^{(t)}$ من $\pi_k(\vartheta_k | \vartheta_1^{(t)}, \vartheta_2^{(t)}, \dots, \vartheta_{k-1}^{(t)}, \underline{x})$ باستخدام خوارزمية ميتروبوليس-هاستغ.

5- نكرر الخطوات السابقة من أجل $t = 1, 2, \dots, T$ (Ntzoufras, 2011).

5-9- إيجاد مقدر الإمكانية العظمى ومقدر بايز لتوزيع باريتو بالمعلمتين:

إذا كان لدينا عينة عشوائية $X = (X_1, X_2, \dots, X_n)$ من توزيع باريتو بمعلمتين فإن لوغاريتم دالة الإمكانية يعطى بالعلاقة:

$$\ln L(\alpha, \beta) = n \ln(\alpha) + n\alpha \ln(\beta) - (\alpha + 1) \sum_{i=1}^n \ln(x_i + \beta) \quad (16)$$

نشق بالنسبة α, β ونساوي المشتقات بالصفر لنحصل على مقدر الإمكانية العظمى كما يلي:

$$\hat{\alpha}_{MLE} = \frac{n}{\sum_{i=1}^n \ln\left(\frac{x_i + \hat{\beta}_{MLE}}{\hat{\beta}_{MLE}}\right)} \quad (17)$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n\hat{\alpha}_{MLE}}{\hat{\beta}_{MLE}} - (\hat{\alpha}_{MLE} + 1) \sum_{i=1}^n \frac{1}{x_i + \hat{\beta}_{MLE}} = 0 \quad (18)$$

بنعويض (17) في المعادلة (18) نحصل على:

$$\frac{n^2}{\hat{\beta}_{MLE} \sum_{i=1}^n \ln\left(\frac{x_i + \hat{\beta}_{MLE}}{\hat{\beta}_{MLE}}\right)} - \left(\frac{n}{\sum_{i=1}^n \ln\left(\frac{x_i + \hat{\beta}_{MLE}}{\hat{\beta}_{MLE}}\right)} + 1 \right) \sum_{i=1}^n \frac{1}{x_i + \hat{\beta}_{MLE}} = 0 \quad (19)$$

لإيجاد قيمة $\hat{\beta}_{MLE}, \hat{\alpha}_{MLE}$ نوجد حل المعادلة اللاخطية (19) بإحدى طرق التحليل العددي واستخدمنا هنا في البحث طريقة نيوتن-رافسون، (Abdel-Ghaly, Attia, Aly, 1998)، (Parikh, 2011).

5-10- الطريقة المقترحة في تقدير توزيع بارينو من أجل α, β مجهولتين:

إذا كان لدينا عينة عشوائية $X = (X_1, X_2, \dots, X_n)$ من توزيع بارينو بمعلمتين سنفترض أن التوزيع القبلي المشترك للمعلمتين وتحت فرض وجود التوقع الرياضي و التباين موجودين التالي:

$$\pi(\alpha, \beta) = \theta e^{-\theta(\alpha-2)} \frac{1}{\Gamma(k)\eta^k} \beta^{k-1} e^{-\frac{\beta}{k}} ; \alpha \geq 2, \eta > 0, \beta > 0 \quad (20)$$

وبناءً عليه فإن التوزيع البعدي المشترك له الشكل:

$$\pi^*(\alpha, \beta | \underline{x}) \propto \alpha^n \beta^{n\alpha+k-1} e^{-(\alpha\theta + \frac{\beta}{\eta})} \prod_{i=1}^n \frac{1}{(x_i + \beta)^{\alpha+1}} \quad (21)$$

وبالتالي التوزيعات الشرطية التامة لكل من α, β لها الشكل التالي:

$$\pi_1(\alpha | \beta, \underline{x}) \propto \pi^*(\alpha, \beta | \underline{x}) \propto \alpha^n \beta^{n\alpha} e^{-\alpha\theta} \prod_{i=1}^n \frac{1}{(x_i + \beta)^{\alpha+1}} \quad (22)$$

$$\pi_2(\beta | \alpha, \underline{x}) \propto \pi^*(\alpha, \beta | \underline{x}) \propto \beta^{n\alpha+k-1} e^{-\frac{\beta}{\eta}} \prod_{i=1}^n \frac{1}{(x_i + \beta)^{\alpha+1}} \quad (23)$$

من الواضح أن حساب التكاملات الناتجة من أجل إيجاد التوزيعات البعدية الهامشية و التقديرات المتعلقة بها يصعب إيجادها تحليلياً لذلك سنقوم بتطبيق خوارزمية ميتروبوليس داخل جيبس لتوليد سلسلة ماركوف الإركوديك من قيم α, β من ثم نقوم بحساب التقديرات المطلوبة باستخدام العلاقة (5) كما يلي:

1- نفرض القيمة الابتدائية $\beta = \beta_0$

2- نولد قيمة المعلمة α ولتكن α_1 من $\pi_1(\alpha | \beta, \underline{x})$ باستخدام خوارزمية ميتروبوليس -هاستنغ باستخدام التوزيع $g_1(\alpha) = \theta e^{-\theta(\alpha-2)} ; \alpha \geq 2$ كتوزيع المقترح.

3- نولد قيمة المعلمة β ولتكن β_1 من $\pi_2(\beta | \alpha_1, \underline{x})$ باستخدام خوارزمية ميتروبوليس -هاستنغ باستخدام التوزيع القبلي $g_2(\beta) = \frac{1}{\Gamma(k)\eta^k} \beta^{k-1} e^{-\frac{\beta}{k}} ; \eta > 0, \beta > 0$ كتوزيع مقترح.

4- نكرر الخطوات السابقة T مرة.

من ثم نحصل على سلسلة ماركوف الإركوديك بطول T من أجل إيجاد المقدر α, β باعتماد على دالة الخسارة التربيعية نطبق العلاقة (8) و من ثم نظرية إركوديك نحصل على:

$$\hat{\alpha}_{BS} = E(\alpha | \underline{x}) \cong \frac{1}{T} \sum_{i=1}^T (\alpha_i) \quad (24)$$

$$\hat{\beta}_{BS} = E(\beta | \underline{x}) \cong \frac{1}{T} \sum_{i=1}^T (\beta_i)$$

من أجل دالة الخسارة الخطية الأسية نطبق العلاقة (11) ومن ثم نظرية إركوديك نحصل على:

$$\hat{\alpha}_{BL} = -\frac{1}{a} \ln[E(e^{-a\alpha} | \underline{x})] \cong -\frac{1}{a} \ln \left[\frac{1}{T} \sum_{i=1}^T (e^{-a\alpha_i}) \right] \quad (25)$$

$$\hat{\beta}_{BL} = -\frac{1}{a} \ln[E(e^{-a\beta} | \underline{x})] \cong -\frac{1}{a} \ln \left[\frac{1}{T} \sum_{i=1}^T (e^{-a\beta_i}) \right]$$

حيث $a \neq 0$ ثابت يمثل معلمة الشكل للدالة الخسارة الأسية الخطية.

من أجل دالة الخسارة الأسية المتوازنة (المقترحة) نطبق العلاقة (14) ومن ثم نظرية إركوديك نحصل على:

$$\hat{\alpha}_{new} = -\frac{1}{a} \ln[\omega e^{-a\hat{\alpha}_{MLE}} + (1 - \omega)E(e^{-a\alpha} | \underline{x})] \cong -\frac{1}{a} \ln \left[\omega e^{-a\hat{\alpha}_{MLE}} + (1 - \omega) \frac{1}{T} \sum_{i=1}^T (e^{-a\alpha_i}) \right] \quad (26)$$

$$\hat{\beta}_{new} = -\frac{1}{a} \ln[\omega e^{-a\hat{\beta}_{MLE}} + (1 - \omega)E(e^{-a\beta} | \underline{x})] \cong -\frac{1}{a} \ln \left[\omega e^{-a\hat{\beta}_{MLE}} + (1 - \omega) \frac{1}{T} \sum_{i=1}^T (e^{-a\beta_i}) \right]$$

حيث $a \neq 0$ ثابت يمثل معلمة الشكل للدالة الخسارة الأسية الخطية، $\hat{\alpha}_{MLE}, \hat{\beta}_{MLE}$ مقدر الإمكانية العظمى، $0 \leq \omega \leq 1$

5-11- الجانب التطبيقي:

باستخدام بيانات مولدة من توزيع باريتو بمعلمتين بحجم عينة $n = 100$ ومن أجل قيم حقيقية مختلفة لمعلمتين تم الحصول على تقدير الإمكان الأعظم ومقدرات بايز بالاعتماد على دوال الخسارة (دالة الخسارة التربيعية l_1 و دالة الخسارة الأسية الخطية l_2 و دالة الخسارة الأسية المتوازنة (المقترحة) l_3 وذلك عند أكثر من قيمة لمعلمة الشكل وقيمة $\omega = 0.5$) وباستخدام عدة قيم لمعالم لتوزيعات القبالية تم الحصول على تقديرات الإمكان الأعظم وتقديرات بايز باستخدام دوال خسارة متماثلة وغير متماثلة وتطبيق خوارزمية MCMC واستخدام العلاقات (17،19،24،25،26) حصلنا على النتائج الموضحة في الجدول التالي:

الجدول رقم (1) القيم الحقيقية والتقديرية المستخدمة في تقدير توزيع باريتو بمعلمتين

القيم الحقيقية	MLE		Bays(MCMC)			
		$\hat{\theta}_{BS}$	$\hat{\theta}_{BL}$		$\hat{\delta}_{new}$	
			a=1	a=-1	a=1	a=-1
$\alpha = 3$ $\beta = 1$	3.540404	2.3103806	2.2757874	2.3581975	2.720242	3.114686
	1.543563	0.8790305	0.8622196	0.8979515	1.145952	1.271978
$\alpha = 4.5$ $\beta = 2$	6.513176	2.739468	2.597193	2.957935	3.270615	5.848202
	3.778250	1.324478	1.259666	1.406798	1.875322	3.174345
$\alpha = 3$ $\beta = 3$	3.540404	2.250964	2.230239	2.276363	2.684546	3.096076
	4.630689	2.429816	2.341636	2.539879	2.938236	4.054068
$\alpha = 2.9$ $\beta = 3.5$	3.379609	2.254387	2.231757	2.283108	2.649306	2.974672
	5.335547	3.035485	2.906448	3.204720	3.515147	4.754602
$\alpha = 3$ $\beta = 6$	4.381269	2.258504	2.236250	2.285851	2.818692	3.804142
	8.670042	3.597526	3.432922	3.805945	4.120768	7.984584

المصدر: من إعداد الباحث اعتماداً على مخرجات اللغة البرمجية R

يظهر الجدول رقم (1) كيفية إيجاد المقدرات في التطبيقات الفعلية، ولكن من الواضح أنه لا يمكننا عمل أي مقارنات بين المقدرات المختلفة لمعلمتي التوزيع ورغبة منا في الوقوف على مدى كفاءة وسلوك طرق التقدير التي تم الحصول عليها يلزمنا عمل مقارنات عديدة باستخدام أسلوب المحاكاة، ومن المعلوم أيضاً أن عينة واحدة لا تعطي المطلوب لذلك تم توليد بيانات عشوائية وبأحجام مختلفة ($n = 100, n = 50, n = 25$) وبتكرار قدره ($L = 500$) وذلك من خلال كتابة برنامج نصي باللغة البرمجية R (لغة برمجية إحصائية)، حيث تستطيع هذه اللغة القيام بالعديد من عمليات تحليل البيانات بحيث يتم تنظيم هذه التحليلات ضمن ما يسمى بالحزم Packages مما يعني قدرة الباحثين على تطوير البرامج المختلفة الأمر الذي ساهم بانتشار استخدامها في المجالات الأكاديمية (Cotton، 2013، Matloff، 2011).

من أجل الحكم على جودة أداء كل طريقة بغرض المقارنة بين طرق التقدير والوقوف على مدى تأثرها بافتراضاتنا المختلفة للتوزيعات القبلية وبدوال الخسارة المختلفة (دالة الخسارة التربيعية l_1 و دالة الخسارة الأسية الخطية l_2 و دالة الخسارة الموزونة المقترحة) l_3 وذلك عند أكثر من قيمة لمعلمة الشكل) تم استخدام متوسط الخطأ التربيعي MSE للمقارنة بين المقدرات والذي يعطى بالعلاقة:

$$MSE = \frac{1}{L} \sum_{i=1}^L (\hat{\theta} - \theta)^T (\hat{\theta} - \theta) \quad (25)$$

حيث $\hat{\theta} = (\hat{\alpha} \hat{\beta})$ ، $\theta = (\alpha \beta)$ ، تم الحصول على النتائج الموضحة في الجدول التالي:

الجدول رقم (2): نتائج تطبيق الطرائق المستخدمة في تقدير معلمتي توزيع باريتو

القيم الحقيقية	حجم العينة	MSE_{MLE}	Bays(MCMC)				
			$MSE_{\hat{\theta}_{BS}}$	$MSE_{\hat{\theta}_{BL}}$		$MSE_{\hat{\delta}_{new}}$	
				a=1	a=-1	a=1	a=-1
$\alpha = 3$ $\beta = 1$	100	1.883708	0.523113	0.5857977	0.438468 5	0.3112482	1.130762
	50	3.169445	0.553233	0.6155512	0.471268 9	0.3300663	1.937213
	25	3.395239	0.578803	0.6432515	0.496679	0.3685052	2.165567
$\alpha = 4.5$ $\beta = 2$	100	3.118761	3.774339	4.413288	2.813718	2.725822	2.058623
	50	3.441711	4.195016	4.812168	3.186335	3.330861	2.42498
	25	4.183418	4.356258	4.98374	3.300464	3.880923	3.160943
$\alpha = 3$ $\beta = 3$	100	1.480534	1.601068	1.775621	1.396148	1.103666	1.008383
	50	2.06809	1.768935	1.997954	1.502855	1.395531	1.305096
	25	2.51745 7	1.979607	2.292155	1.609722	1.862795	1.559345
$\alpha = 2.9$ $\beta = 3.5$	100	1.618829	1.49283	1.732972	1.214881	1.182496	1.026054
	50	1.83932 3	1.608335	1.928532	1.248888	1.367932	1.101271
	25	2.13297 2	1.757057	2.19665	1.267048	1.657137	1.238537
$\alpha = 3$ $\beta = 6$	100	2.23273 4	1.735703	2.222693	1.199211	0.9333728	1.298935
	50	2.09883 5	2.181547	2.876564	1.417918	1.313951	1.210821
	25	2.46110 5	2.608401	3.57515	1.51868	1.735056	1.426324

المصدر: من إعداد الباحث اعتماداً على مخرجات اللغة البرمجية R

نلاحظ من الجدول رقم (2) عند المقارنة بين الأخطاء تبين أن قيم الأخطاء في تقديرات بايز باعتماد على دوال الخسارة المختلفة أصغر من مقابلاتها في تقدير الإمكان الأعظم ونلاحظ أيضاً أن قيم الأخطاء المقدرّة بالاعتماد على دالة الخسارة الأسية المتوازنة (المقترحة) l_3 أي أن المقدر المقترح له أصغر أخطاء من بين أخطاء مقدرات بايز بالاعتماد على الدوال الأخرى ونلاحظ أيضاً أن كفاءة المقدرات تزداد بازدياد حجم العينة.

4-الاستنتاجات والتوصيات:

4-1-الاستنتاجات: Conclusions

اعتماداً على دراسة المحاكاة تم الحصول على مقدرات معلمتي توزيع باريتو باستخدام طريقة الإمكان الأعظم وطرائق بايز اعتماداً على دوال خسارة متماثلة وغير متماثلة (دالة الخسارة التربيعية l_1 و دالة الخسارة الأسية الخطية l_2 و دالة الخسارة الموزونة(المقترحة) l_3 وذلك عند أكثر من قيمة لمعلمة الشكل) وتم اختبار دقة النتائج باستخدام دراسة المحاكاة وتوصلنا من خلال ذلك إلى النتائج التالية:

(1) إن دالة الخسارة الأسية المتوازنة(المقترحة) l_3 دالة معقدة يمكن الحصول من خلالها على دوال خسارة متماثلة وغير متماثلة حيث يمكننا تخصيصهما بسهولة إلى دالة خسارة التربيعية ودالة الخسارة الخطية الأسية، وعلى ذلك فإن النتائج التي تم التوصل إليها بالاعتماد على هذا النوع من دوال الخسارة تعتبر نتائج معقدة لجميع الدراسات السابقة التي اعتمدت على توزيع باريتو بمعلمتين أو أحد الحالات الخاصة منه واستخدمت فيها دوال الخسارة العادية.

(2) إن الاستدلالات التي حصلنا عليها لتوزيع باريتو ذي المعلمتين تعتبر نتائج جديدة تغطي كل النتائج السابقة التي تناولت هذا التوزيع أو أحد الحالات الخاصة منه.

(3) تم حل مشكلة الحسابات المعقدة التي تنتج في استدلالات بايز باستخدام إحدى طرق سلسلة ماركوف وحصلنا على نتائج جيدة.

(4) يتضح من النتائج أن قيم الأخطاء في تقديرات بايز بالاعتماد على الدالة المقترحة أصغر في جميع الحالات من تلك التي تقابلها في تقديرات بايز بالاعتماد على الدالة الأسية الخطية والتربيعية وتقديرات الإمكان الأعظم وذلك من أجل قيم α, β المختلفة، وأيضاً تقل قيم الأخطاء للتقديرات المختلفة بزيادة حجم العينات وبالتالي فإن أداء طرائق بايز يعطي نتائج أفضل من طريقة الإمكان الأكبر وتزداد كفاءة المقدرات بزيادة حجم العينات.

4-2-التوصيات: Recommendations

1- مما سبق نوصي باستخدام الطريقة المقترحة كونها تحل مشكلة التقدير الأعلى والأدنى علاوة على أن التقديرات الناتجة عن الطريقة المقترحة أعطت نتائج أفضل من جميع الطرائق السابقة.

2- نوصي بتطبيق مقدر بايز تحت دالة الخسارة المقترحة في تقدير معالم توزيعات احتمالية أخرى مثل توزيع الطبيعي وتوزيع غامبل ودراسة تأثير حجم العينة على المقدرات.

3- نوصي جميع الباحثين في كافة الاختصاصات ولاسيما الاقتصادية الطبية والبيئية بالعمل وفق تطبيق سلسلة ماركوف مونت كارلو (MCMC) في حل المشاكل التي تظهر في تقدير بعض التوزيعات الاحتمالية وخاصة إذا كان لدينا أكثر من معلمة مجهولة.

4- تطبيق الطريقة المقترحة على بعض التوزيعات الاحتمالية متعددة المتغيرات.

5- إيجاد فترة ثقة بالاعتماد على المقدر المقترح لتوزيع باريتو.

5-References:

- 1) **Abdel-Ghaly, A. A., Attia, A. F., & Aly, H. M.** (1998). Estimation of the parameters of pareto distribution and the reliability function usin accelerated life testing with censoring. *Communications in Statistics–Simulation and Computation*, 27(2), 469–484.
- 2) **Ahmadi, J., Doostparast, M., & Parsian, A.** (2005). Estimation and prediction in a two-parameter exponential distribution based on k-record values under LINEX loss function. *Communications in Statistics–Theory and Methods*, 34(4), 795–805.
- 3) **Ahmadi, J., Jozani, M. J., Marchand, É., & Parsian, A.** (2009). Bayes estimation based on k-record data from a general class of distributions under balanced type loss functions. *Journal of Statistical Planning and Inference*, 139(3), 1180–1189.
- 4) **Arnold, B. C.** (2015). *Pareto distributions*. Chapman and Hall/CRC.
- 5) **Atloff, N.**, (2011). *The Art of R Programming*, Malloy Incorporated, United States of America, 373.
- 6) **Casella, G.** (2008). *Monte Carlo Statistical Methods*. University of Florida.
- 7) **Cotton, R.**, (2013). *Learning R*, O'Reilly Media, Inc., United States of America, 377.
- 8) **Ghosh, J.; Delampady, M; Samant, T.** (2006). *An Introduction to Bayesian Analysis Theory and Methods*. springer science+business media, LLC, New york, USA.
- 9) **Gilks, W. R., Richardson, S., & Spiegelhalter, D.** (1995). *Markov chain Monte Carlo in practice*. Chapman and Hall/CRC. UK.
- 10) **Lehmann, E; Casella .G.** (2006)–*Theory of point estimation*. UAS. 2nd edition.
- 11) **MUKHOPADHYAY N.**, (2000)–*Probability and Statistical Inference* .Marcel Dekker, Inc. New York. Basel.
- 12) **Ntzoufras, I.** (2011). *Bayesian modeling using WinBUGS (Vol. 698)*. John Wiley & Sons, Inc. Hoboken, New Jersey.
- 13) **Parikh, R. V.** (2011). **Study on advanced Bayesian computing techniques in life testing experiments**. Thesis. Department of statistic, Bhavnagar University, Bhavnagar.
- 14) **Rowe, D. B.** (2003). *Multivariate Bayesian statistics: models for source separation and signal unmixing*. Chapman and Hall/CRC.