تقدير معلمة الدليل الذيلي لتوزيع باريتو
بإستخدام مقدر الوسيط العام

أ. م. د. حمزة إسماعيل شاهين   أ. م. د. رعد فاضل حسن

الجامعة المستنصرية– كلية الادارة والاقتصاد

قسم الاحصاء

المستخلص

توسع اهتمام الباحثين في السنوات الاخيرة بموضوع تقدير معلمة الدليل الذيلي α لتوزيع باريتو نظرا لتوسع تطبيق هذا التوزيع في مجال علم القياس الاقتصادي ونظرية المعولية.

لقد تم في هذا البحث محاولة لإستخدام مقدر جديد يعرف بمقدر الوسيط العام وتمت مقارنته مع مقدرات طريقتي العزوم والامكان الاعظم باستخدام مقياس متوسط مربعات الخطأ (MSE) وبتوظيف اسلوب المحاكاة لتحقيق المقارنة. وتم التوصل الى ان مقدر الوسيط العام يكون الأكفأ من بين كل المقدرات المدروسة.

Abstract

Estimation of the tail index parameter of a one - parameter Pareto model has wide important by the researchers because it has awide application in the econometrics science and reliability theorem.

Here we introduce anew estimator of "generalized median" type and compare it with the methods of Moments and Maximum likelihood by using the criteria, mean square error.

The estimator of generalized median type performing best over all.

١- المقدمة

توسع في السنوات الاخيرة اهتمام الباحثين بموضوع تقدير معلمة الدليل الذيلي (α) لتوزيع باريتو نظرا لكثرة استخدام هذا التوزيع بشكل واسع في وصف السلوك المتطرف لقيمة الخسارة في المجال الاقتصادي، كما ان لهذا التوزيع تطبيقات مهمة في موضوع نظرية المعولية بوصفه احد توزيعات الفشل حيث غالبا ما يفترض توزيع باريتو كنموذج شبه معلمي[٤] (semi parametric model) مناسب بشكل تقريبي للمشاهدات العليا من العينة في الحالات التي يكون فيها الذيل العلوي غير مرتبط معلميا مع الاجزاء الدنيا والمتوسطة من العينة، إذ يتم التقدير بشكل منفصل باستخدام المشاهدات الأكثر علواً فقط من العينة.

ان هذا البحث يهدف بالدرجة الاساس الى استخدام مقدر جديد يعرف بمقدر الوسيط العام (Generalized Median Estimator) لتقدير معلمة الدليل الذيلي (α) لتوزيع باريتو واجراء المقارنة بين مقدر الوسيط العام وبعض مقدرات الطرائق التقليدية كطريقتي الامكان الاعظم والعزوم باستخدام مقياس متوسط مربعات الخطأ (MSE) وذلك بتوظيف اسلوب المحاكاة لتحقيق المقارنة تحت افتراض عدد من القيم للمعلمة α المراد تقديرها في هذا البحث ولحجوم عينات مختلفة.

١. توزيع باريتو Pareto Distribution:

من بين النماذج المعلمية المفيدة والتي تتصف بإحتمالية عالية نسبيا في الذيل العلوي هو توزيع باريتو[٤] ، والذي يعبر عنه بالرمز $P(\sigma,\alpha)$ والتي دالة كثافته الاحتمالية (p.d.f) هي[٢]:

$$\mathbf{f}\,(\chi,\sigma,\alpha)=\alpha\,\frac{\sigma^{\alpha}}{\chi^{\alpha+1}}\quad,\ \chi\geq\sigma \qquad \text{...... (1)}$$

حيث ان $\sigma>0\,,\,\alpha>0$ وان α هي معلمة الشكل (shape parameter) وتمثل معلمة الدليل الذيلي (Tail Index) لتوزيع باريتو والمطلوب تقديرها في هذا البحث. في حين تمثل σ معلمة القياس (Scale parameter). والتي سوف تفترض معلومة إذ تستبدل بأصغر قيمة في العينة العشوائية المسحوبة من مجتمع مفرداته تتوزع باريتو بالمعلمتين ($\alpha\ \sigma$,) بعد ان يتم ترتيب قيم $X_1\,,X_2,\ldots,X_n$ العينة كالآتي:

$$\chi_{(1)}<\chi_{(2)}<\ldots\ldots<\chi_{(n)}$$

عندئذ نحصل على مقدر معلمة القياس بحسب الصيغة الاتية:

$$\hat{\sigma}=\chi_{(1)}=\min_i\,(X_i) \qquad \text{...... (2)}$$

لان الدالة ln L يمكن تعظيمها تحت شرط $\hat{\sigma}\leq\min_i\,(X_i)$ وواضح من المتباينة ان قيمة $\hat{\sigma}$ التي تجعل لوغارتم دالة الامكان في نهايتها العظمى $\hat{\sigma}=\min_i\,(X_i)$ علما بان $\frac{\partial\ln L}{\sigma}=\frac{n\alpha}{\sigma}$

كما ان دالة التوزيع التراكمي (c.d.f) لهذا التوزيع هي:

$$\mathrm{F}(\chi,\sigma,\alpha)=1-(\frac{\sigma}{\chi})^{\alpha} \qquad \text{...... (3)}$$

وعلى اعتبار ان العزوم حول نقطة الاصل يعبر عنها بـ $\mu'_r=E\left(X\right)^r$ فبذلك عندما r =1,2,.... فأن العزوم لتوزيع باريتو تكون:

$$\mu'_r=E\left(X\right)^r=\frac{\sigma^r\alpha}{\alpha-r}\ ,\ r<\alpha' \qquad \text{...... (4)}$$

وبالتالي فأن المتوسط والتباين لتوزيع باريتو تعطى كالآتي:

$$\mu_1' = E(X) = \frac{\sigma\alpha}{(\alpha - 1)} \;, \alpha > 1 \qquad \text{...... (5)}$$

$$\text{Var}(X) = \frac{\sigma^2 \alpha}{(\alpha - 1)^2 (\alpha - 2)} \;, \alpha > 2 \qquad \text{...... (6)}$$

٢. بعض الطرائق التقليدية لتقدير معلمة الدليل الذيلي (α):

١.٣ طريقة مقدر العزوم Method of Moment Estimator

تعد طريقة العزوم من الطرائق التقليدية الشائعة الاستخدام في تقدير المعلمات، إذ تتصف بالسهولة، وان اسلوب التقدير وفق هذه الطريقة يقوم على ايجاد المقدر عن طريق حل عدد من المعادلات باعتماد العزوم الدنيا لكل من العينة والمجتمع. وكما تقدم في الفقرة (٢) من هذا البحث فان العزم الأول للمتغير العشوائي X حول نقطة الاصل الذي يتبع توزيع باريتو $P(\sigma, \alpha)$ حسب الصيغة (1) هو:

$$\mu_1 = E(X) = \frac{\sigma\alpha}{\alpha - 1} \;, \alpha > 1$$

وبجعل متوسط العينة يساوي متوسط المجتمع اي ان:

$$\mu_1 = \hat{m}_1$$

يكون:

$$\bar{X} = \frac{\hat{\sigma}\hat{\alpha}}{\hat{\alpha} - 1}$$

حيث $\bar{X}$ يمثل الوسط الحسابي للعينة المسحوبة، ومن المعادلة اعلاه:

$$\hat{\alpha}_{mom} = \frac{\bar{X}}{\bar{X} - \chi_{(1)}} \qquad \text{...... (7)}$$

٢.٣ طريقة مقدر الامكان الاعظم

Method of Maximum Likelihood Estimator

تعد هذه الطريقة من طرائق التقدير المهمة لأنها تحتوي على خصائص جيدة، منها الثبات والاتساق وعلى فرض ان معلمة القياس (σ) تكون معلومة وتستبدل بأصغر مشاهدة في العينة $\chi_{(1)}$ فان دالة الامكان في حالة لدينا عينة عشوائية من المشاهدات $X_1, X_2, ..., X_n$ وتتبع توزيع باريتو $P(\sigma, \alpha)$ بالصيغة الاتية :

$$L(\chi_1, ..., \chi_n; \sigma, \alpha) = \frac{\alpha^n \sigma^{n\alpha}}{\prod_{i=1}^{n} \chi_i^{\alpha+1}} \qquad \text{...... (8)}$$

ولغرض تقدير دالة الامكان (٨) يجب تحويلها الى الشكل الخطي وذلك من خلال أخذ اللوغارتم الطبيعي لطرفي الصيغة (٨) فنحصل على :

$$\operatorname{Log} L(\chi_1,....,\chi_n\ ;\sigma,\alpha) = n \operatorname{Log}\alpha + n\,\alpha \operatorname{Log}\sigma - (\alpha-1)\sum_{i=1}^{n}\operatorname{Log}\chi_i \quad ......\ (9)$$

وبأخذ المشتقة الجزئية الاولى للصيغة (٩) بالنسبة لمعلمة الدليل الذيلي (α) ومن خلال مساواة المشتقة الاولى للصفر، نحصل على الاتي (١) (٢):

$$\hat{\alpha}_{\mu L} = \frac{n}{\sum_{i=1}^{n} Log\ \chi_i - n\operatorname{Log}\hat{\sigma}} \quad ......\ (10)$$

وبالتعويض عن $\hat{\sigma}$ بـ $\chi_{(1)}$ في الصيغة (١٠) فان مقدر الامكان الاعظم للمعلمة α يكون بالصيغة الآتية :

$$\hat{\alpha}_{mL} = \frac{n}{\sum_{i=1}^{n}[Log\ \chi_i - \ Log\ \chi_{(1)}]} \quad ......\ (11)$$

والصيغة (١١) يمكن ان تكتب بشكل آخر كالآتي :

$$\hat{\alpha}_{mL} = \frac{n}{\sum_{i=1}^{n}\operatorname{Log}\ (\frac{\chi_i}{\chi_{(1)}})} \quad ......\ (12)$$

حيث ان الصيغة (١٢) حالة خاصة من الحالة العامة للمقدر والتي توصف بالصيغة التالية: (انظر المصدر (٢))

$$\hat{\alpha}_{mL} = \frac{n-s}{\sum_{i=1}^{n} Log\ (\frac{\chi_i}{\chi_{(1)}})} \quad ......\ (١٣)$$

حيث ان $s=0,1,2,3,...$

والتي تأخذ الحالات التالية:

عندما $s=0$ نحصل على مقدر الإمكان الأعظم.

عندما $s=1$ نحصل على المقدر غير المتحيز.

عندما $s=2$ نحصل على المقدر المنتظم غي المتحيز بأصغر تباين.

عندما $s=3$ نحصل على مقدر بيز القياسي.

عندما $s=4$ نحصل على مقدر بيز ولكن لا يجعل دالة المخاطرة في نهايتها الصغرى.

٣. مقدرات الوسيط العام A Generalized Median Estimators:

١.٤ الصيغة العامة لمقدر الوسيط العام

يعود استخدام مقدرات الوسيط العام (GM) الى الباحث Serfling (1984)[3] [5] ، ضمن دراسة قام بها لفئة واسعة من المقدرات اطلق عليها تسمية مقدرات – L العامة (Generalized L – estimators). وان فكرة مقدرات الوسيط العام (GM) يمكن توضيحها بشكل عام كالآتي:

لتكن $X_1 , X_2 ,....,X_n$ عينة عشوائية ذات حجم (n) من توزيع معين بالمعلمة θ ، فانه لأي عدد صحيح موجب k، الدالة اللبية (Kernel) بالصيغة $h(\chi_1 ,....,\chi_k)$ يتم اختيارها بحيث ان $h(X_1 ,....,X_k)$ تكون وسيط غير متحيز للمعلمة θ.

وبالتالي فان مقدر الوسيط العام (GM) للمعلمة θ سوف يكون عبارة عن الوسيط (Median) للتقييمات $h(X_{i_1} ,...., X_{i_k})$ للدالة اللبية h عند كل فترة، أي وفقا الى توافيق $\binom{n}{k}$ من المجموعات الجزئية المناظرة لـ $(i_1 ,...., i_k)$ للقيم الفاصلة من $\{1,....,n\}$. وبالنتيجة فان احصاءة الوسيط العام لتقدير المعلمة θ تكون وفق الصيغة النهائية الآتية [4][٥] :

$$\hat{\theta}_{GM} = Median\{ h(X_{i_1} ,...., X_{i_k})\} \qquad ...... (14)$$

حيث ان الاختيارات المختلفة لكل من الثابت k والدالة اللبية h يؤدي الى الحصول على قيم مختلفة لمقدرات الوسيط العام للمعلمة θ.

٢.٤ مقدر الوسيط العام لمعلمة الدليل الذيلي (α) لتوزيع باريتو $P(\sigma,\alpha)$

كما بينا سابقاً ان مقدر الوسيط العام (GM) يعتمد بشكل كبير على شكل الدالة اللبية h وعلى قيم الثابت k المختارة، ومن خلالهما يتم تحديد مقدر الوسيط العام المطلوب للتقدير.

فعندما يكون لدينا $X_1 ,...., X_n$ عينة عشوائية من توزيع باريتو $P(\sigma,\alpha)$ فإنه لتقدير المعلمة α باستخدام طريقة (GM) فانه يتم استخدام الدالة اللبية h بالاعتماد على طريقة الامكان الاعظم.

ونتيجة لذلك وحسب الصيغة (١٢) فان الدالة اللبية تكون كالآتي [٤] :

$$h_0(\chi_1,....,\chi_k) = \frac{k}{\sum_{i=1}^{k} Log\left(\frac{\chi_i}{\chi_{(1)}}\right)} \quad , \ k \geq 2 \qquad ...... (15)$$

وهذه الصيغة تعرف بـ (الامكان الاعظم اللبي)[٤]، وهي تستخدم في تقدير المعلمة α بالأعتماد على المشاهدات $X_{i_1}, X_{i_2} ,...., X_{i_k}$.

وبناءا على ذلك فإن المقدر الطبيعي للمعلمة α لتوزيع باريتو $P(\sigma,\alpha)$ يصبح عبارة عن الوسيط للتقييمات $h_0(X_{i_1},....,X_{i_k})$ للدالة **h** حسب الصيغة (١٤) من خلال توافيق $\binom{n}{k}$ من المجموعات الجزئية المناظرة لـ $\{i_k,....,i_1\}$ للقيم الفاصلة $\{n,....,1\}$ للعينة

٤. مقارنة الطرق المستخدمة:

للمقارنة بين مقدر الوسيط العام مع مقدري الامكان الاعظم والعزوم لمعلمة α لتوزيع باريتو $P(\sigma,\alpha)$ فقد اعتمد مقياس متوسط مربعات الخطأ **(MSE)** لكونه مقياس يجمع بين التباين **(Variance)** والتحيز **(Bias)** للمعلمة المقدرة وحسب الصيغة الآتية :

$$MSE(\hat{\alpha}) = \frac{1}{R}\sum_{i=1}^{r}(\hat{\alpha}_i - \alpha)^2 \qquad ......\ (16)$$

حيث **R** عدد مكررات التجربة **(Replication)**

$\hat{\alpha}$ هو مقدر α حسب الاسلوب المستخدم

والمقدر الجيد هو المقدر الذي يمتلك اصغر متوسط مربعات خطأ لحجوم العينات المختلفة.

٥. المحاكاة:

تم توظف اسلوب المحاكاة في هذا البحث لتوليد تجربة بياناتها تتبع توزيع باريتو $P(\sigma,\alpha)$ ، حيث

تم كتابة برنامج المحاكاة بالاعتماد على نظام **(Libteary Basic).**

وقد جرى توليد بيانات التجربة من خلال استخدام اسلوب معكوس الدالة التجميعية $X=F^{-1}(u)$ ، حيث **u** متغير عشوائي يتبع التوزيع المنتظم المستمر المعرف على الفترة المفتوحة **(0,1)** وبعد ذلك تم توليد بيانات تتبع توزيع باريتو $P(\sigma,\alpha)$ وفق الصيغة $X=\sigma(1-u)^{1/\alpha}$ وقد كررت التجربة ١٠٠٠ مرة وبأحجام عينات مختلفة **100,50,25,10** ولقيم افتراضية مختلفة لمعلمتي النموذج (σ,α) ولكل عينة من هذه العينات الاربعة تم تقدير معلمة الدليل الذيلي α لتوزيع باريتو باستخدام طرائق التقدير المبحوثة في الفقرتين **(4,3)** من هذا البحث وحسب الصيغ **(15,12,7)** كذلك تمت المقارنة بين افضلية هذه المقدرات باستخدام مقياس **(MSE)** وقد عرضت النتائج في الجدولين **(2,1)**

جدول رقم (١)

تقديرات معلمة الدليل الذيلي α

الحالات	حجم العينة	α	σ	MM	ML	GM: K =2	GM: K =٣	GM: K =٤
1	١٠	٢	٣	١.٩٣٢	١.١٧٥	١.٩٤١	٢.٣١٤	٢.٦١٦
	٢٥	٢	٣	٢.٧١١	٢.٢٦١	٢.٠٠١	١.٣٥٢	٢.٩١٧
	٥٠	٢	٣	٢.٨٤٨	١.٩٠٠	٢.٣٠٢	٢.١١١	٢.٠١٠
	١٠٠	٢	٣	٢.٤١٣	٢.١٣٦	٢.٥٠١	٢.٠٢٠	٢.٠٠٩
2	١٠	٢.٥	٣	٢.٦٤٣	٢.٩٧٤	٢.٧٧٧	٢.٦١٨	٢.٤٨١
	٢٥	٢.٥	٣	٢.٧٨١	٢.٣٣٣	٢.٩١٤	٢.٤٢٥	٢.٥٩٨
	٥٠	٢.٥	٣	٢.٢١٤	٢.٧١٢	٢.٣٩١	٢.٦١٣	٢.٥٤٩
	١٠٠	٢.٥	٣	٢.٤٤٤	٢.٤٤٨	٢.٦١٠	٢.٩٠١	٢.٤٥١
3	١٠	٢	٤	١.٢٧٦	٢.٧٨٩	٢.٨٨٨	٢.٦٩٣	٢.٤٢١
	٢٥	٢	٤	٢.٤١٠	١.٩٨٠	٢.٧٨١	٢.٦٠١	٢.٥٥١
	٥٠	٢	٤	٢.٩٣١	٢.٤٠١	٢.١١٧	٢.٠١٢	٢.١٠٩
	١٠٠	٢	٤	٢.١٠٣	٢.٢١٠	٢.٥٤٦	٢.٠٠٣	٢.٠٠٧
4	١٠	٢.٥	٤	٢.٧٨٣	٢.٩٤٢	٣.٠٠١	٣.٠١٩	٢.١٣٢
	٢٥	٢.٥	٤	٢.٦٢١	٢.١٦٢	٢.٩٨١	٢.٣٥٢	٢.٢٢٢
	٥٠	٢.٥	٤	٢.٨٠١	٢.٣٨٤	٢.٧٧٦	٢.٣٢٧	٢.١١٢
	١٠٠	٢.٥	٤	٢.٤٢٠	٢.٤٥١	٢.٥١٢	٢.٧٣٣	٢.٤٢٩

جدول رقم (٢)

متوسط مربعات الخطأ (MSE) لتقديرات معلمة الدليل الذيلي α

الحالات	حجم العينة	MM	ML	GM: K = 2	GM: K = ٣	GM: K = ٤
		MSE	MSE	MSE	MSE	MSE
١	١٠	٣.٢١٩	٢.٩٠١	١.٩٨٤	١.٦٦٦	١.٥٦١
	٢٥	٣.١٢١	٢.٤٦١	١.٧٢٩	١.٤٩٤	١.٣٢١
	٥٠	٢.١٩٠	٢.٠٠١	١.٥٠٠	١.٤١٤	١.٢٢٩
	١٠٠	٠.٩١٩	٢.٠١٣	١.٠٠١	١.٠٠٠	١.٠٠٧
٢	١٠	٣.٠٠١	٣.٠٧٠	١.٧٨١	١.٨٢٩	١.٧٣١
	٢٥	٣.٢١٤	٢.٠٩١	١.٩٥٩	١.٦٠٠	١.٤٢١
	٥٠	٢.١٤٣	٢.٧٢٩	١.٣٣٢	١.٤٠٠	١.١١١
	١٠٠	١.٢٥١	١.٩٩٩	١.٢٠٣	١.١٠١	١.٠٣١
٣	١٠	٢.٩١٨	٢.٩٢١	٢.٠١٠	٢.٠٠٠	١.٩٧٨
	٢٥	٢.٨٨٨	٢.٨٠٦	١.٩٨٩	٢.١٠١	١.٩٩٢
	٥٠	٢.١٠٩	٢.١٢٣	١.٥٤١	١.٣٠١	١.٥٤٥
	١٠٠	١.٩٤١	١.٨٢١	١.١٢٥	١.٩٩٢	٠.٨٨١
٤	١٠	٢.٧٧١	٢.٤٩١	١.٧٠١	١.٨٩٣	١.٨٩٠
	٢٥	٢.٨٠٥	٢.٦٢٦	١.٤٠٤	١.٣٣٣	١.٧٨٢
	٥٠	١.٣٢٧	٢.١٥٢	١.٢١١	١.٣٢٥	٠.٩٩٩
	١٠٠	٠.٦٢٩	١.٩٨١	١.٠٠٩	١.٠٧١	٠.٩٧٩

ويظهر من هذه النتائج الاتي :

١- من خلال الجدول رقم (١) يلاحظ ان تقديرات المعلمة α للحالات الاربعة ولجميع احجام العينات المقترحة اظهرت تقديرات المعلمة α بطريقة الوسيط العام (GM) في حالات $(k = 2,3,4)$ أفضلية على طرائق التقدير الاخرى وهي طريقة الامكان الاعظم وطريقة العزوم من خلال قربها من القيم الافتراضية للمعلمة α .

٢- ومن خلال الجدول رقم (٢) يلاحظ ان طريقة الوسيط العام في حالات $(k=2,3,4)$ هي الأكفأ من طرائق التقدير الاخرى حيث تمتلك مقدرات معلمة الدليل الذيلي وفق هذه الطريقة أقل متوسط لمربعات الخطأ من بين المقدرات الاخرى.

٣- ومن خلال الجدول رقم (٢) يلاحظ انه بزيادة قيم k وجد الباحث تحسنا في افضلية المقدر بالنسبة لاستخدام طريقة الوسيط العام في التقدير.

٤- يلاحظ ان زيادة حجم العينة ادى الى اقتراب نتائج المحاكاة من القيم الحقيقية للمعالم وذلك تحقيقا لنظرية الغاية المركزية.

٦. الاستنتاجات والتوصيات

ان مقدر الوسيط العام (GM) يكون الأكفأ مقارنة بمقدر الامكان الاعظم ومقدر العزوم لمعلمة الدليل الذيلي α لتوزيع لامتلاكه أقل متوسط لمربعات الخطأ مقارنة بالمقدرين $\hat{\alpha}_{mL}$ و $\hat{\alpha}_{mom}$ ونوصي بإمكانية التوسع في استخدام طريقة مقدرات الوسيط العام في النماذج المعلمية ذات المعلمتين فأكثر.

٧. المصادر

١- صالح، ستار محمد (٢٠٠٦) "مقارنة اسلوب بيز مع طرائق اخرى (Ls,ML,M) لتقدير المعلمات ودالة المعولية لتوزيع باريتو من النوع الاول" رسالة ماجستير في الاحصاء كلية الادارة والاقتصاد جامعة بغداد.

٢- الاعظمي، علاء حسين (٢٠٠٥) "مقارنة طرائق تقدير معلمة الشكل لتوزيع باريتو باستخدام المحاكاة" رسالة ماجستير في الاحصاء كلية الادارة والاقتصاد جامعة بغداد.

3- Serfling, R (1984) "Generalized L,M and R statistics", Ann. Statist 12, 76-86.

4- Serfling, R (1999) " Robust and Efficient Estimation of the Tail Index of a one parameter pareto Distribution, Email: Serfling @ utdallas. edu.

5- Serfling, R (2002) "Efficient and Robust Fitting of Lognormal Distribution", Email: Serfling @ ut dallas. edu, website; www.utdallas. edu / Serfling.