Bilangan : Jurnal Ilmiah Matematika, Kebumian dan Angkasa
Volume. 2 Nomor. 6, Tahun 2024
e-ISSN :3032-7113; p-ISSN :3032-6389, Hal 29-33
DOI: https://doi.org/10.62383/bilangan.v2i6.300
Available online at: *https://journal.arimsi.or.id/index.php/Bilangan*

ROBUST M-ESTIMATORS' Parameter Scale is Highly Efficient on Base Q Grades

Zahraa Khaled Gaafar [1], Anwer Fawzi Ali [3]
Al-Qadisiyah Governorate Education Directorate, Ira
Email : zahraakhalid228@gmail.com anwerfawi49@gmail.com

Abstract** : Russiva and Crookes proposed the widely used, highly efficient, robust Q-score parameter scale in their 1993 work, and they approximated it using "fast" Huber's M-scores. Shown What proposed us* M-scores *are highly efficient and robust on arbitrary nom distributions, thanks to correct choice parameter approximations. We pay particular attention to the Cauchy divisions and cases of Gaussian distribution . Important terms: Cauchy's law, parameter scale, robustness, and Gaussian distribution With the Q-estimate as a foundation, the M-estimates of the scale parameter are very efficient and* ***robust*. Rousseeuw and Croux (1993) proposed the scale parameter of Q-estimate, which is often used with fast Huber M-estimates, and we were able to get rather close to it. By meticulously choosing the approximation parameters, we proved that the proposed M-estimates are efficient and resilient for any data distribution. In order to measure the robustness and efficiency of scale M estimates , we calculated their asymptotic variances; breakdown points , and influence functions. The Cauchy and Gaussian distributions were our main focus. Notably, the proposed robust estimate is consistent the Maximum likelihood estimation for the Cauchy distribution. Last but not least, these robust and efficient scale estimates need three to four times less computing time than their comparable Q-estimates .The research provides robust and efficient solutions for estimating distribution parameters more quickly and with less complexity, making them suitable for applications requiring fast and accurate computations.*

***Keywords:** Cauchy Distribution, Scale Parameter, M-Estimate*

1. INTRODUCTION

In statistical analysis, the mass headquarters parameter is a crucial metric for evaluating problems [1, 2, 6, 8, 10]. Currently, the best time is considered to be relatively highly efficient $\mathbf{Q_n}$ assessment-scale parameters (n-sample size) . This estimate defines the first quarter-style the distances between observations ;

$$\mathbf{Q_n = a\{|x_i - x_j|\}(k)}$$

A state of influence and effective activity evaluations are provided when A is constant

$$\mathbf{k = C_h^2(h = [n/2] + 1)}$$

We know that the $\mathbf{Q}_n$ following expression determines the influence function of the Q_n estimator : $\varepsilon^* = \mathbf{0,5}$.

the approach yields a high asymptotic efficiency of 82%. One of its drawbacks is that the computation procedure, which uses Q(N $\ln(n)$) time and same amount of memory, has a high asymptotic complexity

The computational cost of the scale parameter is minimal of robust M-estimates , on other hand . On top of that, there are chances to make them more efficient

This motivated the current effort, which aimed to build a computationally efficient and robust approximation of the ($\mathbf{Q_n}$) estimate and then tailor it to different distributions.

Received: Oktober 29, 2024; Revised: November 14, 2024; Accepted: November 25, 2024; Online Available: November 27, 2024;

2. THE PROBLEM

The issue Assuming we have solved the equation , we can look at the set of M-estimateions $S_{\wedge}$ for the scale parameter, The evaluation function X **(x)** is usually even and non-decreasing for all values of x greater than 0, as shown in (1), where

$$\sum X\left(x_i / S_{\wedge} \right) = 0 \qquad \ldots\ldots (1)$$

Estimator Q_n around which For statistical study of estimations

the impact function IF (X; S, F) is a valuable tool. Point X is "clogging" the evaluation function S = S(F), and this informs us how much the value of S has changed on the distribution F . There is a known impact function that describes the asymptotic variance of the estimate S.

The expression is used :

$$AV\left(S_{\wedge}, F\right) = \int IF\left(x; S, F\right)^2 dF(x).$$

An essential characteristic of the class of M estimates of the scale parameter (1) is that, up to a coefficient, the impact function IF (x; S, F) is identical to the evaluation function X (x).

$$IF\left(x; S, F\right) \infty \; X(x)$$

It means that any allowable impact function and efficiency may be used to generate an M-estimate. You may find the estimate's effect function in The following statement .

You may write it as :

$$IF(x;Q,F) = A[(1/4) - F(x + a^{-1}) + F(x - a^{-1})] \left[\int f(y + a^{-1}) f(y)\, dy\right]^{-1} \qquad (2)$$

To avoid include the normalising integral in expression (2), we may use expression (1) to find the evaluation function up to an arbitrarily large factor. The evaluation function then produces an M-estimate, which is the same as the estimate. The output is the product of the following integrals.

$$\chi\,(x) = \frac{a}{4} - A \cdot \left[F\left(x + a^{-1}\right) - F\left(x - a^{-1}\right)\right] \qquad (3)$$

Thus, the features obtained from it are consistent since, as the influence function IF (X; xQ, F) is congruent with IF (x; Q, F).

Let us perform (3): we will make the replacement $\alpha = a^{-1}$, while we will not fix a, considering this value as a parameter for adjusting the estimate , we will proceed to expand the distribution function F into a Taylor series, leaving only the initial three terms.

Let us perform (3): we will make the replacement $\alpha = a^{-1}$without modifying a. After that, we'll extract the first three terms of the Taylor series from the distribution function F and extend it. The following outcome is obtained by combining functions (3) and (4),

$$\mathbf{F}\ (\mathbf{x}\mp\propto) = \mathbf{F}(\mathbf{x}) \pm \boldsymbol{\alpha}\,\mathbf{f}(\mathbf{x}) + \frac{\mathbf{1}}{\mathbf{2}}\propto^{2} \acute{\mathbf{f}}(\mathbf{x}) \pm \frac{\mathbf{1}}{\mathbf{6}}\propto^{3} \mathbf{f}''(\mathbf{x}) + \mathbf{o}(\propto^{3}) \qquad (\mathbf{4})$$

- **Consider an analytic function on R that represents the probability density function f(x). A family of M-estimates is shown here, with evaluation functions of the type and a single parameter for each.**

$$\mathbf{X}_{\propto}(\mathbf{x}) = \mathbf{A}_{\propto} - \mathbf{2f}(\mathbf{x}) - \frac{\mathbf{1}}{\mathbf{2}}\propto^{2} \mathbf{f}''(\mathbf{x}) \qquad (\mathbf{5})$$

This set of $\mathbf{MQ_n}$ estimations is called a family. All estimations are guaranteed to be consistent by the constant ($\mathbf{A}_{\propto}$) in expression (5)

<u>The Gaussian distribution</u>

For a distribution density , we may apply the M-estimator as follows

$$\mathbf{f}\ (\mathbf{x}) = \boldsymbol{\varphi}(\mathbf{x}) = (\mathbf{2\pi})^{-1/2} \ \mathbf{exp}\ (-\mathbf{x}^2/\ \mathbf{2})$$

Afterwards , $\varphi''(x) = (x^2 - 1)\ \varphi(x)$

Here is the form that evaluation function takes:

$$\mathbf{X}_{\propto}(\mathbf{x}) = [\mathbf{A}_{\propto} - \tfrac{1}{3}(\mathbf{6} +\propto^{2} (\mathbf{x}^{2} - \mathbf{1})]\ \boldsymbol{\varphi}\ (\mathbf{x}) \quad , \quad \mathbf{A}_{\propto} = \frac{\mathbf{12}-\propto^{2}}{\mathbf{12}\ \sqrt{\boldsymbol{\pi}}} \qquad (\mathbf{6})$$

An important special case occurs when $\propto = \mathbf{0}$ expression (6) takes the form

When equation (6) is expressed as φ(x)... when $\propto = \mathbf{0}$, a significant exception arises :

$$\boldsymbol{X_0}(\boldsymbol{x}) = \frac{\mathbf{1}}{\boldsymbol{\pi}} - \mathbf{2}\square(\mathbf{x}) \quad \ldots\ldots..(7)$$

This result bears similarities to the generalized Welch estimator [3], as provided by

$$\mathbf{X}(\mathbf{x}) = \sqrt{\frac{\mathbf{d}}{\mathbf{d}+\mathbf{2}}} - \mathbf{exp}\left[-\frac{\mathbf{x}^{2}}{\mathbf{d}}\right], \mathbf{d} > \mathbf{0}$$

as shown by When d=2, this approximation lines up with the CM approximation of the scale parameter given by equation (7). According to equation (7), the asymptotic variance for the estimate can never be more than 95.9%.

The following outcome may be inferred from this

Theory 1 : For a B-robust Gaussian distribution with $\alpha \in [0; 2]$, the formula gives a bounded impact function

$$\mathbf{IF}(\mathbf{x}; \mathbf{MQ}, \phi) = \frac{\mathbf{2}(\mathbf{12} -\propto^{2}) - \mathbf{8}\sqrt{\boldsymbol{\pi}}\left(\mathbf{6} +\propto^{2} (\boldsymbol{X}^{2} - \mathbf{1})\right)\phi(\boldsymbol{x})}{\mathbf{3}(\mathbf{4} -\propto^{2})}$$

The proof of the theorem may be found in [9]. It relies on including the $\mathbf{MQ_n}$ estimator's evaluation function.

With an evaluation function of type (7), the asymptotic efficiency of such estimators is 81%, which is 1% lower than the asymptotic efficiency of Qn-estimators on a Gaussian distribution. The result is a dramatic acceleration in processing speed.

The Cauchy distribution

let's think about evaluation functions of the kind (5) and estimates for the Cauchy distribution with "heavy tails"... As a result, the density distribution looks like this

$$\mathbf{f}(\mathbf{x}) = \frac{\mathbf{1}}{\boldsymbol{\pi}(\mathbf{1} + \mathbf{x}^{\mathbf{2}})}.$$

In the scenario when α = 0, we take into account the $\mathbf{MQ_n}$ -estimate in order to get a minimal algorithm complexity. The other parameter values provide inferior outcomes, therefore we aren't interested in them

Theory 2: $\mathbf{MQ_n}$ Evaluation of Distribution Equation (5) uses an evaluation function of the type to get the error as the greatest likelihood estimate for a given distribution.

$$\mathbf{X_0}(\mathbf{x}) = \frac{\mathbf{1}}{\boldsymbol{\pi}} \frac{\mathbf{x^2 - 1}}{\mathbf{x^2 + 1}}$$

In [9], we also present the proof of the theorem.

On a Cauchy distribution, the asymptotic effectiveness of such a scale parameter estimate may reach 100%; on a Gaussian distribution, it stays below 50%. Last thoughts and suggestions We provide accurate scale parameter estimations for a large class of distributions that are characterised by fast calculation and minimal method complexity .

On a Cauchy distribution, the asymptotic effectiveness of such a scale parameter estimate may reach 100%; on a Gaussian distribution, it stays below 50%.

3. CONCLUSIONS AND RECOMMENDATIONS

We provide accurate scale parameter estimations $\mathbf{MQ_n}$ for a large class of distributions that are characterised by fast calculation and minimal method complexity. The efficiency of the suggested estimates is comparable to that of the robust $\mathbf{Q_n}$ estimator, which is very efficient. We have theoretically studied important cases of Gaussian and Cauchy distributions in an asymptotic manner. We plan to apply the proposed approach to parametric families of t- and exponential power distributions , we will use the suggested method.

4. REFERENCES

Balakrishnan, N., & Cohen, A. C. (2014). *Order statistics and inference: Estimation methods*. Elsevier.

Cramér, H. (1946). *Mathematical methods of statistics* (Vol. 9). Princeton University Press.

Genton, G. (1998). Asymptotic variance of M-estimators for dependent Gaussian random variables. *Statistics and Probability Letters*, *38*(3), 255-261. https://doi.org/10.1016/S0167-7152(98)00048-4

Hampel, F., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986). *Robust statistics: The approach based on influence functions*. John Wiley & Sons.

Huber, J. (1981). *Robust statistics*. John Wiley & Sons.

Launer, L., & Wilkinson, G. N. (Eds.). (2014). *Robustness in statistics*. Academic Press.

Rousseeuw, J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, *88*(424), 1273-1283. https://doi.org/10.2307/2290275

Shevlyakov, G. L., & Oja, H. (Eds.). (2013). *Robust correlation: Theory and applications* (Vol. 3). John Wiley & Sons.

Smirnov, O., & Shevlyakov, G. L. (2014). Fast, highly efficient, and robust one-step M-estimators of scale based on Qn. *Computational Statistics and Data Analysis*, *78*, 153-158. https://doi.org/10.1016/j.csda.2014.04.009