# ON ESTIMATING THE LOCATION PARAMETER OF THE CAUCHY DISTRIBUTION



by

RICHARD EARL THOMAS

B. S., Creighton University, 1968

---

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1973

Approved by:

Ray Waller
Major Professor

# TABLE OF CONTENTS

PAGE

## 1. Introduction

The one parameter Cauchy distribution is defined by the density function

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + (x - \theta)^2}, \quad -\infty < x < \infty, \tag{1.1}$$

and distribution function

$$F(x) = \tfrac{1}{2} + \frac{1}{\pi} \tan^{-1} (x - \theta), \quad -\infty < x < \infty, \tag{1.2}$$

where $\theta$ is the population median, a location parameter.

The Cauchy distribution is of interest due to the similarities between it and the normal distribution. The Cauchy curve is symmetric around the value of $\theta$ and has a general bell shape. The principal difference is that the Cauchy curve has more weight in the tails of the distribution than does the normal. By comparison between the standard normal distribution and the standard Cauchy with $\theta = 0$, the values of the abscissa which bracket central portions of the two curves may be presented as below:

| Proportion | .80 | .90 | .95 | .99 |
|---|---|---|---|---|
| Normal | 1.282 | 1.645 | 1.960 | 2.375 |
| Cauchy | 3.078 | 6.34 | 12.706 | 63.657 |

As can be seen, once the distribution begins to fall away from the median value, it does so quite rapidly. This property is readily observed by noting that the standard Cauchy is identical to Student's t - distribution with 1 degree of freedom.

This paper deals with techniques for estimating the location parameter, $\theta = 0$, in the standard Cauchy given by density function

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + x^2}, \quad -\infty < x < \infty, \tag{1.3}$$

and distribution function

$$F(x) = \tfrac{1}{2} + \frac{1}{\pi} \tan^{-1} x, \quad -\infty < x < \infty . \tag{1.4}$$

The usual estimation techniques fail to provide useful results when used to estimate the median of a Cauchy distribution. The maximum likelihood estimate can only be obtained by an iterative technique in each separate case. The method of moments does not provide a solution at all. The sample mean as an estimate is inconsistent, and therefore not recommended.

<u>1.1 Maximum Likelihood Estimation</u>. The likelihood function for the one parameter Cauchy is given by

$$L(\underline{x}, \theta) = \prod_{i=1}^{n} f(x_i/\theta)$$
$$= \left(\frac{1}{\pi}\right)^n \cdot \prod_{i=1}^{n} \frac{1}{1 + (x_i - \theta)^2}$$

Thus, the log likelihood is

$$\ell n \quad L(\underline{x}, \theta) = -n \, \ell n \pi + \sum_{i=1}^{n} - \ell n \, [1 + (x_i - \theta)^2] .$$

Differentiating, we get

$$\frac{\partial \, \ell n \, (\underline{x}, \theta)}{\partial \, \theta} = \sum_{i=1}^{n} \frac{2 \, (x_i - \theta)}{1 + (x_i - \theta)^2} . \tag{1.5}$$

Therefore, the maximum likelihood estimate, $\hat{\theta}$, satisfies

$$\sum_{i=1}^{n} \frac{(x_i - \hat{\theta})}{1 + (x_i - \hat{\theta})^2} = 0 \tag{1.6}$$

There is no explicit form for the solution to this equation, so that an iterative technique must be used. Thus, maximum likelihood estimation is less than ideal.

1.2 Method of Moments. Estimation by the method of moments is not possible in the Cauchy distribution. The density function is so tenuous that the integrals which define the ordinary moments do not converge. The ususal calculations show that

$$\mu_1 = E(X^1) = \int_{-\infty}^{\infty} \frac{1}{\pi} \cdot \frac{x}{1 + (x - \theta)^2} dx$$

$$= \frac{1}{\pi} \left\{ (x - \theta) \tan^{-1} (x - \theta) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \tan^{-1} (x - \theta) \, dx \right\}, \tag{1.7}$$

which is not defined. So $\mu_1$ does not exist, and therefore higher order moments also do not exist [5]. The equating of sample and population moments, then, is not usable.

1.3 Sample Mean. The following argument shows that the sample mean, $\overline{x} = \sum_{i=1}^{n} x_i / n$, is of no value as an estimator of $\theta$. First, the characteristic function of the Cauchy density defined in (1.1) is [6]

$$\phi_x (t) = \exp \{ it\theta - |t| \} . \tag{1.8}$$

Second, the characteristic function of $\bar{x}$ is

$$\phi_{\bar{x}}(t) = \phi_{\frac{\sum_{j=1}^{n} X_j}{n}}(t)$$

$$= \phi_{\sum_{j=1}^{n} X_j}\left(\frac{t}{n}\right)$$

$$= \prod_{j=1}^{n} \phi_x\left(\frac{t}{n}\right)$$

$$= \prod_{j=1}^{n} \{ \exp\{ i \frac{t}{n} \theta - \left|\frac{t}{n}\right| \}$$

$$= \exp\{ \sum_{j=1}^{n} \frac{it\theta}{n} - \sum_{j=1}^{n} \left|\frac{t}{n}\right| \}$$

$$= \exp\{\frac{nit\theta}{n} - n \cdot \frac{1}{n} |t|\}$$

$$= \exp\{ it\theta - |t|\} = \phi_x(t) .$$

Thus, the distribution of the sample mean is identical to the distribution of any one of the x's, and the sample mean is an inconsistent estimator. Further, the expected value of $\bar{x}$ does not exist.

<u>1.4 Variance Bound</u>. The Cramér-Rao lower bound for the variance of an unbiased estimator of $\theta$ is given by

$$- \frac{1}{E\left\{\frac{\partial^2 \ln L(\underline{x}, \theta)}{\partial \theta^2}\right\}} .$$

By differentiating (1.5) we obtain

$$\frac{\partial^2 \ell n L}{\partial \theta^2} = \sum_{i=1}^{n} \frac{-2 \quad 1 + (x - \theta)^2 \ (x_i - \theta) + 2\,(x_i - \theta)\ 2\,(x_i - \theta)}{\left\{ 1 + (x_i - \theta)^2 \right\}^2}$$

$$= \sum_{i=1}^{n} \frac{-2 + 2\,(x_i - \theta)^2}{\left\{ 1 + (x_i - \theta)^2 \right\}^2}$$

Thus, by the identical distributions for $x_i$

$$E \left\{ \frac{\partial^2 \ \ell n\ L}{\partial \ \theta^2} \right\} = n \int_{-\infty}^{\infty} \frac{1}{\pi} \cdot \frac{-2 + 2\,(x - \theta)^2}{\{1 + (x - \theta)^2\}^2} \cdot \frac{1}{1 + (x - \theta)^2} \quad dx$$

$$= \frac{2n}{\pi} \int_{-\infty}^{\infty} \frac{(x - \theta)^2 - 1}{\{1 + (x - \theta)^2\}^3} \ dx$$

$$= \frac{4n}{\pi} \int_{-\infty}^{\infty} \frac{y^2 - 1}{(1 + y^2)^3} \ dy$$

$$= \frac{4n}{\pi} \quad \left( \frac{-\pi}{8} \right)$$

$$= -\frac{n}{2} \qquad [6] \qquad (1.9)$$

Hence, the variance of any unbiased estimator of $\theta$, say $\tilde{\theta}$, must satisfy

$$\mathrm{Var}\ (\tilde{\theta}) \geqslant - \frac{1}{\frac{-n}{2}} = \frac{2}{n} \ . \qquad (1.10)$$

1.5 Purpose. The problem is to find estimators of $\theta$ that provide an efficient unbiased estimate. This study looks at estimators which have been developed and examines their statistical properties. A comparison is made in two cases of estimation in small samples, both among the estimators, and with their own asymptotic properties.

Since the usual techniques fail, we turn to "unusual" methods to provide solutions. The estimators should have a high asymptotic relative efficiency (ARE), and should be unbiased, at least asymptotically. The estimators also should be such that there is an explicit usable form. The types of estimators that are examined here are obtained by two methods. First, we find the simple average of a central portion of a sample. Secondly, we work with the order statistics of the sample and obtain a weighted average.

## 2. Estimators and Their Properties

We wish to find estimators of $\theta$ which provide a good estimate in various sample sizes. To do this, the estimator must be such that the extreme values have low (or even negative) weight. The estimators discussed in this report can be classified as belonging either to the class of trimmed mean estimators, or to the class of weighted order statistics estimators.

2.1 Trimmed Mean Estimators. The trimmed mean class of estimators contains estimators of the form

$$m(a) = \frac{1}{2a + 1} \sum_{i=-a}^{a} X_{(i)} \quad , \; a = 0, 1, \ldots, n \qquad (2.1)$$

where

$$X_{(-n)}, X_{(-n + 1)}, \ldots X_{(-1)}, X_{(0)}, X_{(1)}, \ldots, X_{(n)}$$

are the order statistics of a sample of size 2n + 1 and $X_{(0)}$ is the median. The estimator, then, is a simple average of the central portion of the sample. Included in the class are the median

$$m(0) = X_{(0)}$$

and the sample mean

$$m(n) = \frac{1}{2n + 1} \sum_{i=-n}^{n} X_{(i)} \quad .$$

The median is a consistent estimator, but shows an asymptotic relative efficiency (ARE) of only 81%. The sample mean is inconsistent, as we have shown in Section 1.2. Thus, we begin a search for a better (more efficient) estimator by considering the case $0 < a < n$.

Rothenberg, Fisher, and Tilanus [7] investigated this class of estimators to determine which member of the class gave the highest ARE. The asymptotic joint probability distribution of two order statistics, $\{X_{(r)}, X_{(s)}\}$ shows that the covariance of $\{X_{(r)}, X_{(s)}\}$ for any continuous density function is given by

$$\sigma_{rs} = \frac{(1 + \frac{r}{n})\;(1 - \frac{s}{n})}{4\; f(\hat{X}_{(r)})\; f(\hat{X}_{(s)})} \quad , \quad r < s. \qquad (2.2)$$

where

$$\hat{X}_{(j)} = E\;(X_{(j)}) \; .$$

For the Cauchy distribution, $\sigma_{rs}$ becomes

$$\sigma_{rs} = \frac{\pi^2}{4} \left(1 + \frac{r}{n}\right) \left(1 - \frac{s}{n}\right) \sec^2 \left(\frac{\pi r}{2n}\right) \sec^2 \left(\frac{\pi s}{2n}\right), \quad r < s . \tag{2.3}$$

Using $\sigma_{rs}$ in (2.1) shows that the variance of m(a) is

$$V(m(a)) = \frac{1}{(2a + 1)^2} \sum_{j=-a}^{a} \sum_{i=-a}^{a} \sigma_{ij} .$$

By using the central portion of the ordered sample defined by

$$k = a/n \qquad \text{or } a = k \cdot n$$

we get

$$\mathrm{Var}\,(m(k)) = \frac{\pi^2}{8k^2n^2} \sum_{j=-a}^{a} \sum_{i=-a}^{a} \left(1 + \frac{j}{n}\right)\left(1 - \frac{i}{n}\right) \sec^2 \left(\frac{\pi j}{2n}\right) \sec^2 \left(\frac{\pi i}{2n}\right)$$

which reduces to

$$\mathrm{Var}\,(m(k)) = \left\{\frac{1 - k}{k^2}\right\} \tan^2 \left(\frac{\pi k}{2}\right) + \left\{\frac{2}{\pi k^2}\right\} \tan \left(\frac{\pi k}{2}\right) - \frac{1}{k} .$$

Substituting values of k into this expression, the asymptotic sampling variances of the various estimators can be found. Rothenberg, et al., [7] provide a table which shows that as the value of k moves from 0 toward 1, Var (m(k)) first decreases, then increases. The minimum point is found to be at k = .24, where

$$\mathrm{Var}\,(m(.24)) = 2.278$$

Thus the ARE is given by

$$\left\{\frac{\text{Var }(m(a))}{\text{Var }(m(\hat{\theta}))}\right\}^{-1} = \left\{\frac{2.278/n}{2/n}\right\}^{-1}$$

$$= \frac{2}{2.278}$$

$$= .8779 \quad .$$

Based on the ARE, this estimator represents an improvement over the median. The optimum estimate of this class is thus the average or mean of the central 24% of the data,

$$m(.24n) = \frac{1}{.48n} \sum_{i=-.24n}^{.24n} X_{(i)} \tag{2.4}$$

The problem of extreme values is solved in the trimmed mean class of estimators by ignoring them, and only using those values which can be expected to be tightly clustered around the true population median. The optimum trimmed mean is thus easy to compute for various sample sizes. However, it is not always easy, in small samples, to determine what values actually comprise the central 24%.

<u>2.2 Weighted Order Statistics Estimators</u>. The class of weighted order statistics estimators deals with estimators of the form

$$\theta^{*} = \sum_{i=1}^{n} \alpha_i X_{(i)}$$

where

$\alpha_i$ : weight applied to the $i^{th}$ value in the ordered sample,

and

$$\sum_{i=1}^{n} \alpha_i = 1 \quad .$$

The following three estimators of this class are considered:

1) Best Linear Unbiased Estimator (BLUE) [1]

2) Quick Estimator (due to Bloch [2])

3) Optimal Estimator (due to Chernoff, et al. [3]).

A note about the type of weights used by members of this class seems appropriate at this point in that the weight functions of the estimators exhibit similar characteristics. First, the weights are symmetric around the median value(s) of the sample. Second, the largest weight is applied to the median value(s) with successively smaller weights to the values in the tails, until the extreme values receive small negative weights. Thus, this class of estimators solves the problem of the extreme value by reducing the amount of its contribution to the estimate, and reversing the direction.

In the trimmed mean class, the weights were either $\frac{1}{2a + 1}$ or 0, depending on whether the sample point was to be included in the estimate or not. In this, a single extreme observation, either large or small, can throw the estimate off.

2.2.1 Best Linear Unbiased Estimate. The best linear unbiased estimate of $\theta$ is obtained by using the technique of ordered least squares estimation of location parameters. For a distribution function of the form

$$F(x - \theta)$$

the BLUE is

$$\theta^* = \frac{\underline{1}'\underline{\Omega}^{-1}\,\underline{x}}{\underline{1}'\underline{\Omega}^{-1}\,\underline{1}}, \quad [1] \tag{2.5}$$

where

$\underline{\Omega}$ : variance - covariance matrix of $\underline{x}$

$\underline{x}$ : column vector of order statistics

$\underline{1}$ : column vector of ones.

For the Cauchy distribution, the first two and last two order statistics have infinite variances [1], so we begin the search for a BLUE estimator by reducing the working sample size to n - 4. That is, zero weights are applied to $X_{(1)}$, $X_{(2)}$, $X_{(n-1)}$, and $X_{(n)}$. Then the joint distribution of the order statistics $\{X_{(r)}, X_{(s)}\}$, $r < s$

$$f_{rs}(x,y) = \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!}$$

$$\times F(x)\,[F(y) - F(x)]^{s-r-1}\,[1 - F(x)]^{n-s}\,f(x)f(y)$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!}\left[\tfrac{1}{2} + \frac{1}{\pi}\tan^{-1}(x)\right] \times$$

$$\left[\frac{1}{\pi}\tan^{-1}(y) - \tan^{-1}(x)\right]^{s-r-1} \times \left[1 - \tfrac{1}{2} - \frac{1}{\pi}\tan y\right]^{n-s} \times$$

$$\frac{1}{\pi^2} \cdot \frac{1}{1+x^2} \cdot \frac{1}{1+y^2}, \qquad r < s$$

is used to obtain

$$\sigma_{rs} = E(xy) - E(x)E(y)$$

$$= \frac{n!}{\pi^n (r-1)!\,(s-r-1)!\,(n-s)!} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\frac{\pi}{2}}^{y} \tan x \tan y \,\cdot$$

$$\cdot\, \left(\frac{\pi}{2} + x\right)^{r-1} (y - x)^{s-r-1} \left(\frac{\pi}{2} - y\right)^{n-s} dx\,dy \; -$$

$$\left[ \frac{n!}{\pi^n (r-1)!\,(n-r)!} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \tan x \left(\frac{\pi}{2} + x\right)^{r-1} \left(\frac{\pi}{2} - x\right)^{n-r} dx \right] \cdot$$

$$\left[ \frac{n!}{\pi^n (s-1)!\,(n-s)!} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \tan y \left(\frac{\pi}{2} + y\right)^{s-1} \left(\frac{\pi}{2} - y\right)^{n-s} dy \right]$$

These values are extremely difficult to compute, but Barnett [1] has tabled them for n = 5(1)16(2)20.

The BLUE is obtained by substituting these values in (2.5) and determining the weights, $\alpha_i$, for the various order statistics. Using these weights, the BLUE can be computed easily from sample data. The disadvantage to the BLUE is, of course, the costliness of computing the covariances for sample sizes other than those listed in [1].

2.2.2 Quick Estimator. The quick estimator was proposed by Bloch [2], who sought to find an estimator based on only a selected number of order statistics from the sample. The estimator should have the property of providing a good estimate regardless of the sample size.

The quick estimator is obtained by selecting k order statistics from the sample.

Let

$X_{(1)}', X_{(2)}', \ldots, X_{(k)}'$ : selected order statistics

$Y_{(1)}', Y_{(2)}', \ldots, Y_{(k)}'$ : corresponding population quantiles

$\lambda_{(1)}', \lambda_{(2)}', \ldots, \lambda_{(k)}'$ : $F(Y_{(i)}') = \lambda_i$

The ARE of the estimator is given [2] by

$$\varepsilon = \frac{K_1}{E\left\{\left\{\frac{f'(x)}{f(x)}\right\}^2\right\}}$$

$$= 2K_1$$

where

$$K_1 = \sum_{i=1}^{k+1} \frac{f(Y'_{(i)}) - f(Y'_{(i-1)})}{\lambda'_i - \lambda'_{i-1}} \qquad [2]$$

and

$$f(Y'_{(o)}) = F(Y'_{(k+1)}) = 0$$

$$\lambda'_o = 0, \qquad \lambda'_{k+1} = 1$$

If we select $k = 5$, then we need only maximize $K_1$ to get the most efficient estimate. Bloch [2] found that for $k = 5$, the value of $K_1$ was essentially maximized when

$$\underline{\lambda} = (\lambda_1, \quad \lambda_2, \quad \lambda_3, \quad \lambda_4, \quad \lambda_5)$$

$$= (.13, .40, .50, .60, .87) \qquad (2.6)$$

and

$$\varepsilon = .95161.$$

Thus, if we use the order statistics that approximate those population quantiles, we obtain the optimum estimator for 5 order statistics. The weights for those order statistics are obtained by using the general formula for the BLUE.

Using the values of $\underline{\lambda}$ above and the corresponding values of f(x), the weights are given as

$$\underline{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5,)$$
$$= (-.052, .3485, .407, .3485, -.052). \tag{2.7}$$

The quick estimator, then, for 5 selected order statistics is

$$\theta^* = \sum_{i=1}^{5} \alpha_i \; X'_{(i)} \tag{2.8}$$

with variance

$$V(\theta^*) = \frac{2.1017}{n} \tag{2.9}$$

so that

$$ARE = \frac{2}{2.1017} = .95161.$$

The advantage to the quick estimator over others is that it provides an easy estimate which can be rapidly obtained from any size of sample. In addition, it has eliminated over half of the efficiency loss associated with the other easy estimates, the optimum trimmed mean.

<u>2.2.3 Optimal Estimator</u>. The final estimator in this class to be investigated is due to Chernoff, Gastwirth and Johns [3]. This estimator provides an asymptotically efficient and optimum result based on the distribution of the order statistics. The general form of the Chernoff, <u>et al</u>. estimator is given by

$$T_n = \frac{1}{n} \sum_{i=1}^{n} J(u) \; X_{(i)} \tag{2.10}$$

where

$$u = \frac{i}{n + 1}$$

and

J is a weight function dependent on the distribution.

For the parameter $\theta$ in the Cauchy distribution,

$$\alpha_i = J\left(\frac{i}{n + 1}\right)$$

$$= \frac{\sin 4\pi \left(\frac{i}{n + 1}\right) - \frac{1}{2})}{\tan \pi \left(\frac{i}{n + 1} - \frac{1}{2}\right)} \quad . \qquad (2.11)$$

The weights for this estimator are easy to calculate on a computer, since they depend only on the sample size n. Thus, for large sample sizes, this estimator is usable when a computer is available for calculations.

## 3. Generation of Data and Discussion

Since most of the properties of these estimators (e.g., ARE) for comparison purposes are asymptotic, it is of interest to compare the characteristics of these estimators in relatively small sample sizes. Two sample sizes were chosen, $n = 8$ and $n = 16$. These choices were made to facilitate the use of the optimum trimmed mean, as they made the central 24% of the sample more readily discernable.

The sample data points were generated using the distribution function of the standard Cauchy distribution. Equating (1.4) to an observation r from the uniform [0, 1] distribution, we get

$$r = \tfrac{1}{2} + \frac{1}{\pi} \tan^{-1} x. \tag{3.1}$$

Solving (3.1) for x yields

$$x = \tan \{\pi (r-\tfrac{1}{2})\}$$

where x denotes a random observation from the standard Cauchy.

Utilizing a uniform random number generator, 100 samples of each size were obtained using and IBM 360 computer. Each sample thus generated was then used to obtain an estimate of $\theta = 0$ by each estimator discussed. The results were as presented below.

## 3.1 Samples of Size 8

3.1.1 Optimum Trimmed Mean (OTM). For $n = 8$, the optimum trimmed mean (2.4) is actually the median, i.e., the average of the middle two values, $X_{(4)}$ and $X_{(5)}$. The asymptotic variance for the trimmed mean is given by

$$\frac{2.278}{n} = \frac{2.278}{8} = .28475$$

The 100 estimates were then compiled and their properties determined. The sample mean of these estimators was 0.0501 with sample variance 0.58843. The estimator thus showed relatively unbiased for a sample as small as 8. The sample variance, though, was over twice its asymptotic value.

The values of the estimates are represented by the frequency distribution below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 0 |
| -1.75 | -1.25 | 5 |
| -1.25 | - .75 | 4 |
| - .75 | - .25 | 23 |
| - .25 | .25 | 43 |
| .25 | .75 | 15 |
| .75 | 1.25 | 6 |
| 1.25 | 1.75 | 2 |
| above | 1.75 | 2 |

The estimates are strongly clustered around the true value, 0, but exhibit a slight negative tendency. This is offset in the mean by the two extreme positive values, 1.91 and 4.66.

3.1.2 BLUE. The BLUE estimates (2.5) were obtained using the weights tabled by Barnett [1], applied to the observations $X_{(3)}, \ldots, X_{(6)}$. The BLUE is an asymptotically efficient estimator, so its asymptotic variance is given by

$$\frac{2}{n} = \frac{2}{8} = .250 \quad .$$

The 100 estimates had a sample mean of -.02105, with sample variance 0.57063. Again, the estimates are essentially unbiased in the small sample case, but the variance exceeds twice the asymptotic value.

A frequency distribution of the obtained values is presented below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 0 |
| -1.75 | -1.25 | 4 |
| -1.25 | - .75 | 6 |
| - .75 | - .25 | 24 |
| - .25 | .25 | 42 |
| .25 | .75 | 15 |
| .75 | 1.25 | 5 |
| 1.25 | 1.75 | 2 |
| above | 1.75 | 2 |

These values differ only slightly from those obtained by the optimum trimmed mean, again exhibiting a slight negative tendency.

3.1.3 Quick Estimator. The quick estimator values were obtained using (2.8) above. The asymptotic variance for these estimates is given by

$$\frac{2.1017}{n} = \frac{2.1017}{8} = .26271$$

The 100 estimates had a sample mean of -.00044 and a sample variance of 2.19657. Thus, their distribution was centered around the true value of $\theta$, but has a variance of more than 8 times its asymptotic value. In comparison with the previous two estimators, we have almost twice as large a standard deviation of the estimators.

The frequency distribution presented below indicates how this extimator gives more extreme values.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 3 |
| -1.75 | -1.25 | 3 |
| -1.25 | - .75 | 14 |
| - .75 | - .25 | 25 |
| - .25 | .25 | 31 |
| .25 | .75 | 17 |
| .75 | 1.25 | 2 |
| 1.25 | 1.75 | 1 |
| above | 1.75 | 4 |

These values do not present the tight clustering that was observed in the previous two samples, but are also negatively inclined as were the previous two. The presence of the extreme values apparently stems from the fact that in using this estimator, we use the extreme values that are not used by the previous estimators.

3.1.4 *Optimal Estimator*. The optimal estimator values as described by Chernoff, et al. [3] were obtained first by using (2.10). On inspection of the weights generated according to (2.11), however, it was noted that $\sum_{i=1}^{n} \alpha_i = n + 1$, at least for n = 8, 16. In an effort to improve the estimate, the value of

$$\sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{(i)}$$

was normalized by dividing by n + 1 instead of n.

The asymptotic variance of this estimator is

$$\frac{2}{n} = \frac{2}{8} = .250$$

as in the BLUE, since this is an efficient estimator. The 100 estimates had a sample mean of -.10566, with a sample variance of 4.05003. The sample moments obtained differ from the asymptotic moments by amounts greater than for any of the other estimators. There were more extreme values, both positive and negative, than for any of the previous estimators, as can be seen in the frequency distribution below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 7 |
| -1.75 | -1.25 | 2 |
| -1.25 | - .75 | 8 |
| - .75 | - .25 | 19 |
| - .25 | .25 | 38 |
| .25 | .75 | 15 |
| .75 | 1.25 | 4 |
| 1.25 | 1.75 | 2 |
| above | 1.75 | 5 |

The negative trend of the value is somewhat present here, but not as sharply as in the previous estimators. The values which fall away from the central class do so more repidly than in the previous estimators.

3.1.5 Summary. As a means of comparison, we can determine the frequency in each estimator of values occurring no more than one class interval from the median class, i.e., between - .75 and .75. The table below summarizes these results.

| Estimator | Frequency |
|---|---|
| OTM | 81 |
| BLUE | 81 |
| Quick Estimator | 73 |
| Optimal | 72 |

From this, we can see that the strongest clustering around the true value occurred in the Optimum Trimmed Mean and the BLUE, while the Quick and Optimal estimators were almost identical.

We can summarize our results, finally, by comparing the values of the sample moments to both each other, and to their own asymptotic values.

| Estimator | Sample Mean | Sample Variance | Asymptotic Variance |
|---|---|---|---|
| OTM | 0.05010 | 0.58843 | 0.28475 |
| BLUE | -0.02105 | 0.57063 | 0.25000 |
| Quick | -0.00044 | 2.19657 | 0.26271 |
| Optimal | -0.10566 | 4.05003 | 0.25000 |

The highest efficiency, then, is in the OTM and BLUE estimators, both relative to the other estimators, and to their own asymptotic variances.

## 3.2 Samples of Size 16

3.2.1 Optimum Trimmed Mean. The OTM estimator (2.4) for n = 16 is the average of the central 4 values, $X_{(7)}, \ldots, X_{(10)}$. The asymptotic variance is given by

$$\frac{2.278}{n} = \frac{2.278}{16} = 0.1438$$

For the 100 estimates obtained from samples of size 16, the sample mean was 0.01458 with a sample variance of 0.12705. Thus, the sample variance was actually below the asymptotic value, so that as n increased, the OTM estimates began to cluster more tightly about the true value of θ. The obtained estimates are represented by the frequency distribution below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 0 |
| -1.75 | -1.25 | 0 |
| -1.25 | - .75 | 1 |
| - .75 | - .25 | 21 |
| - .25 | .25 | 48 |
| .25 | .75 | 30 |
| .75 | 1.25 | 0 |
| 1.25 | 1.75 | 0 |
| above | 1.75 | 0 |

Thus only one value falls more than one class interval away from the central class. In addition, the values now show a positive tendency as opposed to the negative trend for n = 8.

3.2.2 BLUE. Again using Barnett's [1] weights on the values $X_{(3)}, \ldots, X_{(14)}$, in (2.5) we obtain an estimate of $\theta$ from each of the 100 samples of size 16. The asymptotic variance of the BLUE is given by

$$\frac{2}{n} = 0.125 .$$

The 100 estimates had a sample mean of 0.00582 with sample variance .11602. As was the case for the OTM, the mean is closer to the true value of 0 and the variance is slightly below its asymptotic value. The obtained estimates are represented by the frequency distribution below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 0 |
| -1.75 | -1.25 | 0 |
| -1.25 | - .75 | 1 |
| - .75 | - .25 | 19 |
| - .25 | .25 | 54 |
| .25 | .75 | 26 |
| .75 | 1.25 | 0 |
| 1.25 | 1.75 | 0 |
| above | 1.75 | 0 |

Again, only one value falls more than one class interval from the central class, and the curve shows a slight positive trend. There is a higher concentration in the central class than was the case for the OTM's.

3.2.3 Quick Estimator. For n = 16, Bloch's [2] quick estimator (2.8) has asymptotic variance given by

$$\frac{2.1017}{n} = \frac{2.1017}{16} = .13136 \; .$$

The 100 estimates obtained for n = 16 had a sample mean of -.05859 with sample variance of .15847. Thus, by increasing the sample size we have greatly improved the variance of this estimator. It is only slightly above its asymptotic value, while the sample mean remains close to the true value of zero. The obtained values of the estimate are represented by the frequency distribution below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 0 |
| -1.75 | -1.25 | 1 |
| -1.25 | - .75 | 3 |
| - .75 | - .25 | 26 |
| - .25 | .25 | 48 |
| .25 | .75 | 20 |
| .75 | 1.25 | 2 |
| 1.25 | 1.75 | 0 |
| above | 1.75 | 0 |

Although this does not approach the previous two estimators for clustering, it is a marked improvement over the case for n = 8. The values still display a negative tendency, but it is much less pronounced.

3.2.4 Optimal Estimator. The value of

$$\sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{(i)}$$

was again divided by n + 1 = 17 instead of n = 16 as in (2.10) in an attempt to improve the estimate. The asymptotic variance of this estimator is given by

$$\frac{2}{n} = \frac{2}{16} = .125 .$$

The 100 estimates obtained had a sample mean of -.10064 with sample variance 0.58844. The variance of this estimator is still above its asymptotic value, but has improved considerably over the case for n = 8. The obtained values of the estimator are represented by the frequency distribution below.

| Class | | Frequency |
|---|---|---|
| Below | -1.75 | 2 |
| -1.75 | -1.25 | 1 |
| -1.25 | - .75 | 3 |
| - .75 | - .25 | 23 |
| - .25 | .25 | 44 |
| .25 | .75 | 23 |
| .75 | 1.25 | 3 |
| 1.25 | 1.75 | 1 |
| above | 1.75 | 0 |

These values are virtually symmetric about the central class, with only the 2 extreme negative values out of balance.

3.2.4 Summary. We can again use the frequency of the central three class intervals as a means of comparison among the estimators.

| Estimator | Frequency |
|---|---|
| OTM | 99 |
| BLUE | 99 |
| Quick | 94 |
| Optimal | 90 |

The highest concentration, then, occurs in the OTM and BLUE estimators, as was the case for $n = 8$. However, the Quick estimator, now, is somewhat better than the Optimal, indicating that it improves more rapidly.

A comparison of the sample moments and the asymptotic variances is presented below.

| Estimator | Mean | Sample Variance | Asymptotic Variance |
|---|---|---|---|
| OTM | 0.01458 | 0.12705 | 0.1438 |
| BLUE | 0.00582 | 0.11602 | 0.1250 |
| Quick | -0.05859 | 0.15847 | 0.13136 |
| Optimal | -0.10064 | 0.58844 | 0.1250 |

The BLUE and OTM perform best again, as for $n = 8$, but the quick estimator has become a viable alternative for $n = 16$.

## 4. Conclusions

In each case, we obtained estimates that give a good approximation to the known value of $\theta$, $\theta = 0$, from each of the four estimators. It would be expected that this property would hold as well for any value of $\theta$ as well. The comparisons of the estimators, then, will hinge on their efficiency and ease of computation.

For samples up to size 20, where the weights are available, the BLUE is obviously best. Little computation is required, and the BLUE has lowest sample variance. The OTM is an easy estimator to use, and performs well relative to the BLUE. However, as pointed out previously, it is not always easy to determine what comprises the central 24% in small samples, and may be somewhat subjective.

The quick estimator is easy to use since it only uses a number of selected order statistics, in this case, 5. Although it does not approach its asymptotic variance as rapidly as the previous two, it is fairly well-behaved. It becomes a valuable estimator for use in various sample sizes, since the weights do not change from one sample size to the next. It would be a more useful estimator, yet one which provided viable estimates.

The optimal estimator does not perform well in the small sample case, even with the improvement of the 1/n + 1 factor. This proved to be the most variable, and the only one to produce extreme values in the n = 16 case. For larger samples, though, the weights (2.11) are easily built on a computer, and this could prove to be a useful estimator.

## ACKNOWLEDGMENT

The author wishes to express his sincere thanks and appreciation to his major professor, Dr. Ray A. Waller for his assistance in the preparation of this report.

## REFERENCES

[1] Barnett, N. D. (1966). "Order statistics estimators of the location parameter of the Cauchy distribution", Journal of the American Statistical Association, 61, 1205-1217.

[2] Bloch, D. (1966). "A note on the estimation of the location parameter of the Cauchy distribution", Journal of the American Statistical Association, 61, 852-855.

[3] Chernoff, H., Gastwirth, J. L. and Johns, M. V. (1967). "Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation", Annals of Mathematical Statistics, 38, 52-72.

[4] Johnson, N.L. and Kotz, S. (1970). Continuous Univariate Distributions-1, Boston: Houghton Mifflin.

[5] Kendall, M. G. and Stuart, A. (1969). The Advanced Theory of Statistics, Vol. I, New York: Hafner.

[6] Kendall, M. G. and Stuart, A. (1967). The Advanced Theory of Statistics, Vol. II, London: Griffin.

[7] Rothenberg, R. J., Fisher, F. M. and Tilanus, C. B. (1964). "A note on estimation from a Cauchy sample", Journal of the American Statistical Association, 59, 460-463.

# ON ESTIMATING THE LOCATION PARAMETER OF THE CAUCHY DISTRIBUTION

by

RICHARD EARL THOMAS

B.S., Creighton University, 1968

---

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1973

ABSTRACT

Estimating the location parameter of the Cauchy distribution presents certain difficulties, since the usual techniques of estimation fail to give tractable results. As a result, unusual techniques have been proposed for estimating this parameter.

The two types of estimators investigated are the trimmed mean class and the weighted order statistics class. Estimators in these classes have been previously studied in relation to their asymptotic properties. The behavior of these estimators in sample sizes of 8 and 16 is investigated by compiling estimates from 100 samples of each size from the standard Cauchy distribution. Their resulting sample variances are compared, both to each other and to their known asymptotic values.