Maximum Likelihood Estimates of the Parameters of the Cauchy Distribution for Samples of Size 3 and 4

Author(s): Thomas S. Ferguson

Source: *Journal of the American Statistical Association*, Mar., 1978, Vol. 73, No. 361 (Mar., 1978), pp. 211-213

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

Stable URL: https://www.jstor.org/stable/2286549

Maximum Likelihood Estimates of the Parameters of the Cauchy Distribution for Samples of Size 3 and 4

THOMAS S. FERGUSON*

Expressions are given for the joint maximum likelihood estimates of the location and scale parameters of a Cauchy distribution based on samples of size 3 and 4.

KEY WORDS: Maximum likelihood estimate; Cauchy distribution.

For the problem of estimating the location parameter μ of a Cauchy distribution when the scale parameter σ is given, it is known that the likelihood function is occasionally multimodal. In fact, for a sample of size n and for σ sufficiently small, it will have n local maxima, one close to each observation. In general, for a sample of size n the maximum likelihood estimate of location for fixed scale is a root of a polynomial of degree $2n - 1$. See Barnett (1966a) for a discussion of numerical methods for finding a global maximum.

One would expect it to be more difficult to find the joint maximum likelihood estimates for location and scale. However, Copas (1975) has recently shown that the joint likelihood function for location and scale is unimodal. This was conjectured by Haas, Bain, and Antle (1970) who developed numerical methods for solving the likelihood equations. See also Dumonceaux, Antle, and Haas (1973). Thus the two parameter situation is somewhat easier than the one parameter case in that even simple-minded approximation schemes, such as steepest ascent, serve to find the maximum likelihood estimates.

A similar situation exists in the asymptotic theory of estimation of the parameters of the Cauchy distribution by linear combinations of a fixed number of well-chosen order statistics (spacings). For estimation of μ with fixed σ, the optimal spacings are difficult to obtain (see Bloch 1966; Chan 1970; and Balmer, Boulton, and Sack 1974), while for joint estimation of μ and σ, the optimal spacings and weights can be written in closed form (Balmer et al. 1974; Cane 1974).

Curiously, closed-form expressions exist for the joint maximum likelihood estimates of the location and scale parameters of a Cauchy distribution for samples of size 3 and 4. These are presented below. For samples of size 5 or greater, it seems doubtful that such closed form expressions for the estimates exist.

Let $x_1, \ldots, x_n$ be a sample of size n from a Cauchy distribution with density $\sigma/[\pi\{\sigma^2 + (x - \mu)^2\}]$, where μ and $\sigma > 0$ are unknown parameters. The log likelihood function is

$$L = c + n \log \sigma - \sum_{i=1}^{n} \log \{\sigma^2 + (x_i - \mu)^2\} \ . \tag{1}$$

The likelihood equations are, therefore,

$$\begin{aligned} \frac{\partial L}{\partial \mu} &= \sum_{i=1}^{n} \frac{2(x_i - \mu)}{\sigma^2 + (x_i - \mu)^2} = 0 \ , \\ \frac{\partial L}{\partial \sigma} &= \frac{n}{\sigma} - \sum_{i=1}^{n} \frac{2\sigma}{\sigma^2 + (x_i - \mu)^2} = 0 \ . \end{aligned} \tag{2}$$

When $n = 2$, it was pointed out in Haas et al. (1970) that the maximum likelihood estimate is not unique and that all maximum likelihood estimates $(\hat{\mu}, \hat{\sigma})$ for observations x and y with $x < y$ are given by

$$x \le \hat{\mu} \le y \quad \text{and} \quad \hat{\sigma}^2 = (\hat{\mu} - x)(y - \hat{\mu}) \ . \tag{3}$$

For arbitrary sample sizes, Copas (1975) shows that the only time the maximum likelihood estimate is not unique is when half the observations are at a point x and the other half at a point y, when (3) again gives the estimates. When otherwise half or more of the observations are at a point x, the maximum of the likelihood function occurs at $\hat{\mu} = x$ and $\hat{\sigma} = 0$, a singular point for equations (2). In all other cases, however, the maximum likelihood estimate of (μ, σ) may be found as the unique root of the likelihood equations (2).

In the following, we assume that all observations are distinct. Then, it is easy to show, using the results of Copas, that any conjectured value of $(\hat{\mu}, \hat{\sigma})$ is a maximum likelihood estimate by merely showing that it satisfies (2).

Sample Size $n = 3$. Let the three observations be denoted by x, y, and z in the order $x < y < z$. The

* Thomas S. Ferguson is Professor, Department of Mathematics, University of California, Los Angeles, CA 90024. This research was supported in part by National Science Foundation Grant MCS 72-04591.

© Journal of the American Statistical Association
March 1978, Volume 73, Number 361
Theory and Methods Section

maximum likelihood estimates are given by

$$\hat{\mu} = \frac{x(z-y)^2 + y(z-x)^2 + z(y-x)^2}{(z-y)^2 + (z-x)^2 + (y-x)^2}, \qquad \hat{\sigma} = \frac{(3)^{\frac{1}{2}}(z-y)(z-x)(y-x)}{(z-y)^2 + (z-x)^2 + (y-x)^2}. \tag{4}$$

It may be observed that $\hat{\mu}$ is a weighted average of the three observations with weights proportional to the squares of the differences of the other two. To establish (4), it is sufficient to show these values satisfy equations (2), which reduce here to

$$\frac{(x-\mu)/\sigma}{1+[(x-\mu)/\sigma]^2} + \frac{(y-\mu)/\sigma}{1+[(y-\mu)/\sigma]^2} + \frac{(z-\mu)/\sigma}{1+[(z-\mu)/\sigma]^2} = 0,$$

$$\frac{1}{1+[(x-\mu)/\sigma]^2} + \frac{1}{1+[(y-\mu)/\sigma]^2} + \frac{1}{1+[(z-\mu)/\sigma]^2} = \frac{3}{2}.$$

Direct substitution and straightforward algebra show that these equations are satisfied by (4). As an aid in the computations, the following equations are offered.

$$1 + [(x-\hat{\mu})/\hat{\sigma}]^2 = 2((z-y)^2 + (z-x)^2 + (y-x)^2)/3(z-y)^2,$$

$$1 + [(y-\hat{\mu})/\hat{\sigma}]^2 = 2((z-y)^2 + (z-x)^2 + (y-x)^2)/3(z-x)^2,$$

$$1 + [(z-\hat{\mu})/\hat{\sigma}]^2 = 2((z-y)^2 + (z-x)^2 + (y-x)^2)/3(y-x)^2.$$

Sample Size $n = 4$. Let the four observations be denoted by w, x, y, and z in the order $w < x < y < z$. The maximum likelihood estimate is given by

$$\hat{\mu} = \frac{xz - yw}{z - y + x - w}, \qquad \hat{\sigma}^2 = \frac{(z-y)(y-x)(x-w)(z-w)}{(z-y+x-w)^2}. \tag{5}$$

Note that $\hat{\mu}$ is a weighted average of the two middle-order statistics with weights proportional to the distance of the opposite outlier from its neighbor, i.e., $\hat{\mu} = px + (1-p)y$, where $p = (z-y)/[(z-y)+(x-w)]$. It is sufficient to show that these values satisfy the equations (2), which reduce here to

$$\frac{w-\mu}{\sigma^2+(w-\mu)^2} + \frac{x-\mu}{\sigma^2+(x-\mu)^2} + \frac{y-\mu}{\sigma^2+(y-\mu)^2} + \frac{z-\mu}{\sigma^2+(z-\mu)^2} = 0,$$

$$\frac{1}{\sigma^2+(w-\mu)^2} + \frac{1}{\sigma^2+(x-\mu)^2} + \frac{1}{\sigma^2+(y-\mu)^2} + \frac{1}{\sigma^2+(z-\mu)^2} = \frac{2}{\sigma^2}.$$

Direct substitution and straightforward algebra show that these equations are satisfied by (5). As an aid in checking this, the following equations are offered.

$$\hat{\sigma}^2 + (w-\hat{\mu})^2 = (x-w)(y-w)(z-w)/(z-y+x-w),$$

$$\hat{\sigma}^2 + (x-\hat{\mu})^2 = (x-w)(y-x)(z-x)/(z-y+x-w),$$

$$\hat{\sigma}^2 + (y-\hat{\mu})^2 = (y-w)(y-x)(z-y)/(z-y+x-w),$$

$$\hat{\sigma}^2 + (z-\hat{\mu})^2 = (z-w)(z-x)(z-y)/(z-y+x-w).$$

An interesting property of the best linear unbiased estimate of location based on order statistics noticed by Barnett (1966b), and others (Bloch 1966; Chan 1970; Balmer et al. 1974) for the problem of optimal spacings, is the negative effect of outlying observations. More precisely, changes in the values of observations in the lower and upper quantiles have the effect of pushing the estimate in the opposite direction. This phenomenon is a well-known property of Huber M estimators with redescending ψ functions at least for observations sufficiently far out in the tails. This property is visible also for the maximum likelihood estimate $\hat{\mu}$ of μ in (4) and (5).

Even for a sample of size 3 and equally spaced observations, this occurs. Let $x = -1$, $y = 0$, and $z > 0$ in the $\hat{\mu}$ of (4), so that

$$\hat{\mu}(z) = \frac{z(1-z)}{2(z^2+z+1)}.$$

At $z = 1$, we have $\hat{\mu}(z) = 0$ as it should be by symmetry; but for $z > 1$, we have $\hat{\mu}(z) < 0$. In fact, $(d/dz)\hat{\mu}(z) < 0$ for all $z > (3^{\frac{1}{2}} - 1)/2 \cong .366\ldots$.

For a sample of size 4, the phenomenon is even more striking. Put $w = -1$, $x = 0$, and $0 < y < z$ in (5) to obtain

$$\hat{\mu}(y, z) = y/(z - y + 1).$$

Here the estimate $\hat{\mu}(y, z)$ is decreasing in z for all $z > y$, with limiting value y as $z \to y$, and limiting value $0 (= x)$ as $z \to \infty$.

[*Received February 1977. Revised September 1977.*]

REFERENCES

Balmer, D.W., Boulton, M., and Sack, R.A. (1974), "Optimal Solutions in Parameter Estimation Problems for the Cauchy Distribution," *Journal of the American Statistical Association*, 69, 238–242.

Barnett, V.D. (1966a), "Evaluation of the Maximum Likelihood Estimator Where the Likelihood Equation Has Multiple Roots," *Biometrika*, 53, 151–165.

——— (1966b), "Order Statistics Estimators of the Location of the Cauchy Distribution," *Journal of the American Statistical Association*, 61, 1205–1218.

Bloch, Daniel A. (1966), "A Note on the Estimation of the Location of the Cauchy Distribution," *Journal of the American Statistical Association*, 61, 852–855.

Cane, Gwenda J. (1974), "Linear Estimation of Parameters of the Cauchy Distribution Based on Sample Quantiles," *Journal of the American Statistical Association*, 69, 243–245.

Chan, Lai K. (1970), "Linear Estimation of Location and Scale Parameters of Cauchy Distribution Based on Sample Quantiles," *Journal of the American Statistical Association*, 65, 851–859.

Copas, J.B. (1975), "On the Unimodality of the Likelihood for the Cauchy Distribution," *Biometrika*, 62, 701–704.

Dumonceaux, Robert, Antle, Charles E., and Haas, Gerald (1973), "Likelihood Ratio Tests for Discrimination Between Two Models with Unknown Location and Scale Parameters," *Technometrics*, 15, 19–27.

Haas, Gerald, Bain, Lee, and Antle, Charles (1970), "Inferences for the Cauchy Distribution Based on Maximum Likelihood Estimators," *Biométrika*, 57, 403–408.

Rothenberg, Thomas J., Fisher, Franklin M., and Tilanus, C.B. (1964), "A Note on Estimation from a Cauchy Sample," *Journal of the American Statistical Association*, 59, 460–463.