Estimation

Thomas Achia

5/15/2023

R Markdown

Cauchy Distribution A continuous random variable is said to follow standard Cauchy distribution if its probability density function(p.d.f) is given by

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

where x is the standard Cauchy variate. If $x = \frac{y-\theta}{\lambda}$

$$g(y) = \frac{\lambda}{\pi\left[1+\left(\frac{y-\theta}{\lambda}\right)^2\right]}$$

Letting $\lambda = 1$, the pdf of the Cauchy distribution is

$$g(y) = \frac{1}{\pi\left[1+(y-\theta)^2\right]}$$

Considering the random sample of

$$Y_1, \dots, Y_n$$

the likelihood function is

$$L(\theta; \boldsymbol{y}) = \prod_{i=1}^{n} g(y_i) = \prod_{i=1}^{n} \frac{1}{\pi\left[1+(y_i-\theta)^2\right]} = \frac{1}{\pi^n \prod_{i=1}^{n}\left[1+(y_i-\theta)^2\right]} = \pi^{-n}\left[\prod_{i=1}^{n}\left[1+(y_i-\theta)^2\right]\right]^{-1},$$

and the log-likelihood is

$$\ell(\theta; \boldsymbol{y}) = \ln L(\theta; \boldsymbol{y}) = \ln\left[\pi^{-n}\left[\prod_{i=1}^{n}\left[1+(y_i-\theta)^2\right]\right]^{-1}\right] = \ln\left[\pi^{-n}\right] + \ln\left[\left[\prod_{i=1}^{n}\left[1+(y_i-\theta)^2\right]\right]^{-1}\right] = -n\ln\pi - \sum_{i=1}^{n}\ln\left[1+(y_i -\right.$$

$$Y = \ln[g(x)]. \text{Let } g(x) = u \Rightarrow \frac{du}{dx} = g'(x) \text{ and } Y = \ln u \Rightarrow \frac{dY}{du} = \frac{1}{u}$$

$$\frac{dY}{dx} = \frac{dY}{du}\frac{du}{dx} = \frac{1}{u} \times g'(x) = \frac{g(x)}{g'(x)}$$

The score function is

$$S(\theta; \boldsymbol{y}) = \frac{\partial \ell(\theta; \boldsymbol{y})}{\partial \theta} = \sum_{i=1}^{n}\left[\frac{2(y_i-\theta)}{1+(y_i-\theta)^2}\right],$$

$$Y = \frac{U}{V} = \frac{2(y_i - \theta)}{1 + (y_i - \theta)^2} \Rightarrow \frac{dY}{dx} = \frac{V\frac{dU}{dx} - U\frac{dV}{dx}}{V^2}$$

Then the Hessian is

$$H(\theta; \boldsymbol{y}) = \frac{\partial S(\theta; \boldsymbol{y})}{\partial \theta} = \sum_{i=1}^{n} \left[\frac{-2\left[1 + (y_i - \theta)^2\right] + 4(y_i - \theta)^2}{\left[1 + (y_i - \theta)^2\right]^2} \right] = -2\sum_{i=1}^{n} \left[\frac{1 - (y_i - \theta)^2}{\left[1 + (y_i - \theta)^2\right]^2} \right]$$

The MLE should satisfy $\boxed{S(\theta; \boldsymbol{y}) = \sum_{i=1}^{n} \left[\frac{2(y_i - \theta)}{1 + (y_i - \theta)^2} \right] = 0.}$ This expression has no analytical solution

Using the Taylor expansion

$$\boxed{f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \ldots + \frac{(x-a)^r}{r!}f^{(r)}(a) + \ldots}$$

$$\boxed{f(x) \approx f(a) + (x-a)f'(a)}$$

Taylor's expansion for the score function is

$$S(\theta) = S(\theta^*) + (\theta - \theta^*)S'(\theta^*) + \frac{(\theta - \theta^*)^2}{2!}S''(\theta^*) + \ldots$$

$$\boxed{S(\theta) \approx S(\theta^*) + (\theta - \theta^*)S'(\theta^*)}$$

and substituting the MLE, $\hat{\theta}_{MLE}$, for which $S(\hat{\theta}_{MLE}; \boldsymbol{y}) = 0$, we have

$$S(\hat{\theta}_{MLE}) = S(\theta^*) + (\hat{\theta}_{MLE} - \theta^*)S'(\theta^*) + \ldots \Rightarrow 0 \approx S(\theta^*) + (\hat{\theta}_{MLE} - \theta^*)S'(\theta^*).$$

It follows that

$$\hat{\theta}_{MLE} = \theta^* - \frac{S(\theta^*; \boldsymbol{y})}{S'(\theta^*; \boldsymbol{y})}$$

That is,

$$\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} - \frac{S(\hat{\theta}^{(n)}; \boldsymbol{y})}{H(\hat{\theta}^{(n)}; \boldsymbol{y})} = \hat{\theta}^{(n)} - \frac{\sum_{i=1}^{n} \left[\frac{2(y_i - \theta^{(n)})}{1 + \left(y_i - \theta^{(n)}\right)^2} \right]}{-2\sum_{i=1}^{n} \left[\frac{1 - (y_i - \theta^{(n)})^2}{\left[1 + \left(y_i - \theta^{(n)}\right)^2\right]^2} \right]}$$

See https://utstat.toronto.edu/keith/papers/cauchymle.pdf

```r
cauchy.mle <- function(x,start,eps=1.e-8,max.iter=50){
if (missing(start)) start <- median(x)
theta <- start
n <- length(x)
score <- sum(2*(x-theta)/(1+(x-theta)^2))
iter <- 1
```

```
conv <- T
while (abs(score)>eps && iter<=max.iter){
hessian <- -sum((2-2*(x-theta)^2)/(1+(x-theta)^2)^2)
theta <- theta - score/hessian
iter <- iter + 1
score <- sum(2*(x-theta)/(1+(x-theta)^2))
}
if (abs(score)>eps) {
print("No Convergence")
conv <- F
}
loglik <- -sum(log(1+(x-theta)^2))
hessian <- -sum((2-2*(x-theta)^2)/(1+(x-theta)^2)^2)
info<--hessian
r <- list(theta=theta,loglik=loglik,hessian=hessian,info=info,convergence=conv)
r
}
```

The function can now be used as follows:

```
x <- rcauchy(100) + 5 # 100 observations with theta = 5
r <- cauchy.mle(x,start=median(x))
r
```

```
## $theta
## [1] 5.138844
##
## $loglik
## [1] -139.9602
##
## $hessian
## [1] -47.17785
##
## $info
## [1] 47.17785
##
## $convergence
## [1] TRUE
```

Iterative approach to MLE

$$Y_1, \ldots, Y_n: \text{ A random sample of size n from } f(y; \boldsymbol{\theta})$$

$$\text{Likelihood: } L(\boldsymbol{\theta}; \boldsymbol{y}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta})$$

$$\text{log-Likelihood: } \ell(\boldsymbol{\theta}; \boldsymbol{y}) = \ln L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} \ln f(y_i; \boldsymbol{\theta})$$

$$\text{The Score Vector: } S(\boldsymbol{\theta}; \boldsymbol{y}) = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial \boldsymbol{\theta}}$$

$$\text{For the MLE of } \boldsymbol{\theta}, \; S(\boldsymbol{\theta}; \boldsymbol{y}) = \mathbf{0}$$

$$S(\hat{\boldsymbol{\theta}}_{MLE}; \boldsymbol{y}) = \mathbf{0}$$

The Hessian matrix: $H(\boldsymbol{\theta};\boldsymbol{y}) = \dfrac{\partial S(\boldsymbol{\theta};\boldsymbol{y})}{\partial \boldsymbol{\theta}} = \dfrac{\partial^2 \ell(\boldsymbol{\theta};\boldsymbol{y})}{\partial \boldsymbol{\theta}' \partial \boldsymbol{\theta}}$

In the case where $S(\boldsymbol{\theta};\boldsymbol{y}) = \mathbf{0}$ lacks an analytic solution, we use the Taylors expansion.

$$S(\boldsymbol{\theta}) = S(\boldsymbol{\theta}^*) + (\boldsymbol{\theta} - \boldsymbol{\theta}^*)S'(\boldsymbol{\theta}^*) + \ldots$$

$$S(\boldsymbol{\theta}) \approx S(\boldsymbol{\theta}^*) + (\boldsymbol{\theta} - \boldsymbol{\theta}^*)S'(\boldsymbol{\theta}^*)$$

Substituting the MLE

$$\underbrace{S(\hat{\boldsymbol{\theta}}_{MLE})}_{\mathbf{0}} \approx S(\boldsymbol{\theta}^*) + (\hat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}^*)S'(\boldsymbol{\theta}^*)$$

$$\begin{aligned}
S(\boldsymbol{\theta}^*) + (\hat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}^*)S'(\boldsymbol{\theta}^*) &\approx \mathbf{0} \\
S(\boldsymbol{\theta}^*) + (\hat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}^*)H(\boldsymbol{\theta}^*) &\approx \mathbf{0} \\
(\hat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}^*) \qquad \underbrace{H(\boldsymbol{\theta}^*)}_{\text{The Hessina matrix}} \qquad &\approx \boldsymbol{S}(\boldsymbol{\theta}^*) \\
(\hat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}^*) &\approx \boldsymbol{S}(\boldsymbol{\theta}^*)H^{-1}(\boldsymbol{\theta}^*) \\
\hat{\boldsymbol{\theta}}_{MLE} &\approx \boldsymbol{\theta}^* + \boldsymbol{S}(\boldsymbol{\theta}^*)H^{-1}(\boldsymbol{\theta}^*)
\end{aligned}$$

$$\hat{\boldsymbol{\theta}}^{(n+1)} \approx \boldsymbol{\theta}^{(n)} + \boldsymbol{S}(\boldsymbol{\theta}^{(n)})H^{-1}(\boldsymbol{\theta}^{(n)}) \text{ An iterative formula for estimating the MLE}$$