# Exercise3

*Hukai Luo*

*21 September 2018*

**Abstract**

In this assignment, we try to use different methods to estimate the Maximum Likelihood Estimation of Cauchy distribution, including Newton-Raphson method, Fixedpoint method and Fisher scoring method. Every method has its advantage and disadvantage, we will comment them at the end of the assignment.

# 1 Question

Consider estimating the location parameter of a Cauchy distribution with a known scale parameter. The density function is

$$f(x;\theta) = \frac{1}{\pi[1 + (x - \theta)^2]}, \quad x \in R, \quad \theta \in R.$$

Let $X_1, \ldots, X_n$ be a random sample of size $n$ and $\ell(\theta)$ the log-likelihood function of $\theta$ based on the sample.

- Set the random seed as 20180909 and generate a random sample of size $n = 10$ with $\theta = 5$. Implement a loglikelihood function and plot against $\theta$.
- Find the MLE of $\theta$ using the Newton–Raphson method with initial values on a grid starting from $-10$ to 20 with increment 0.5. Summarize the results.
- Apply fixed-point iterations using $G(\theta) = \alpha\ell'(\theta) + \theta$, with scaling choices of $\alpha \in \{1, 0.64, 0.25\}$ and the same initial values as above.
- First use Fisher scoring to find the MLE for $\theta$, then refine the estimate by running Newton-Raphson method. Try the same starting points as above.
- Comment on the results from different methods (speed, stability, etc.).

# 2 Calculate Derivative and Fisher information

Show that

$$\ell(\theta) = -n \ln \pi - \sum_{i=1}^{n} \ln[1 + (\theta - X_i)^2],$$

$$\ell'(\theta) = -2 \sum_{i=1}^{n} \frac{\theta - X_i}{1 + (\theta - X_i)^2},$$

$$\ell''(\theta) = -2 \sum_{i=1}^{n} \frac{1 - (\theta - X_i)^2}{[1 + (\theta - X_i)^2]^2},$$

$$I_n(\theta) = \frac{4n}{\pi} \int_{-\infty}^{\infty} \frac{x^2 dx}{(1 + x^2)^3} = n/2,$$

where $I_n$ is the Fisher information of this sample.

In this question, we will calculate $\ell(\theta), \ell'(\theta), \ell''(\theta), I_n(\theta)$ as bellow:

$$\ell(\theta) = \ln \prod_{i=1}^{n} f(X_i; \theta)$$

$$= \sum_{i=1}^{n} \ln f(X_i; \theta)$$

$$= -n \ln \pi - \sum_{i=1}^{n} \ln[1 + (X_i - \theta)^2]$$

$$\ell'(\theta) = \frac{d}{d\theta} \ell(\theta)$$

$$= -2 \sum_{i=1}^{n} \frac{\theta - X_i}{1 + (\theta - X_i)^2}$$

$$\ell''(\theta) = \frac{d}{d\theta} \ell'(\theta)$$

$$= -2 \sum_{i=1}^{n} \frac{1 - (\theta - X_i)^2}{[1 + (\theta - X_i)^2]^2}$$

The Fisher information of this sample can be defined as:

$$I_n(\theta) = \int \ell''(\theta) f(X_i; \theta) dx$$

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

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

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

$$= \frac{2n}{\pi} \frac{x}{1 + x^2} \frac{1}{1 + x^2} \Big|_{-\infty}^{\infty} - \frac{2n}{\pi} \int_{-\infty}^{\infty} \frac{x}{1 + x^2} d(\frac{1}{1 + x^2})$$

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

Then we should convert $x$ to $\tan t$, the domain of $x$ is $-\infty$ to $\infty$, so the domain of t is $-\frac{\pi}{2}$ to $\frac{\pi}{2}$:

$$I_n(\theta) = \frac{4n}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{(\tan t)^2}{[(1 + \tan t)^2]^3} d\tan t$$

$$= \frac{4n}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\sin t)^2 (\cos t)^2 dt$$

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

2

# 3 Generate a loglikelihood function.

Set the random seed as 20180909 and generate a random sample of size $n = 10$ with $\theta = 5$. Implement a loglikelihood function and plot against $\theta$.

```r
set.seed(20180909)
X <- rcauchy(10,5)                        #generate a sample of X using cauchy distribution
logfunction <- function(x) {                      #calculate the log likelihood function
  logfunction <- 0
  for (i in 1:length(X)){
    logfunction <- logfunction-log(pi)-log(1+(x- X[i])^2)
  }
  logfunction
}

library("ggplot2")                                    #plot the function against theta
ggplot(data.frame(x= c(-10,20)),aes(x = x))+
        stat_function(fun = function(x) logfunction(x))+
        labs(x=expression(~theta),y="log likelihood function")
```



3
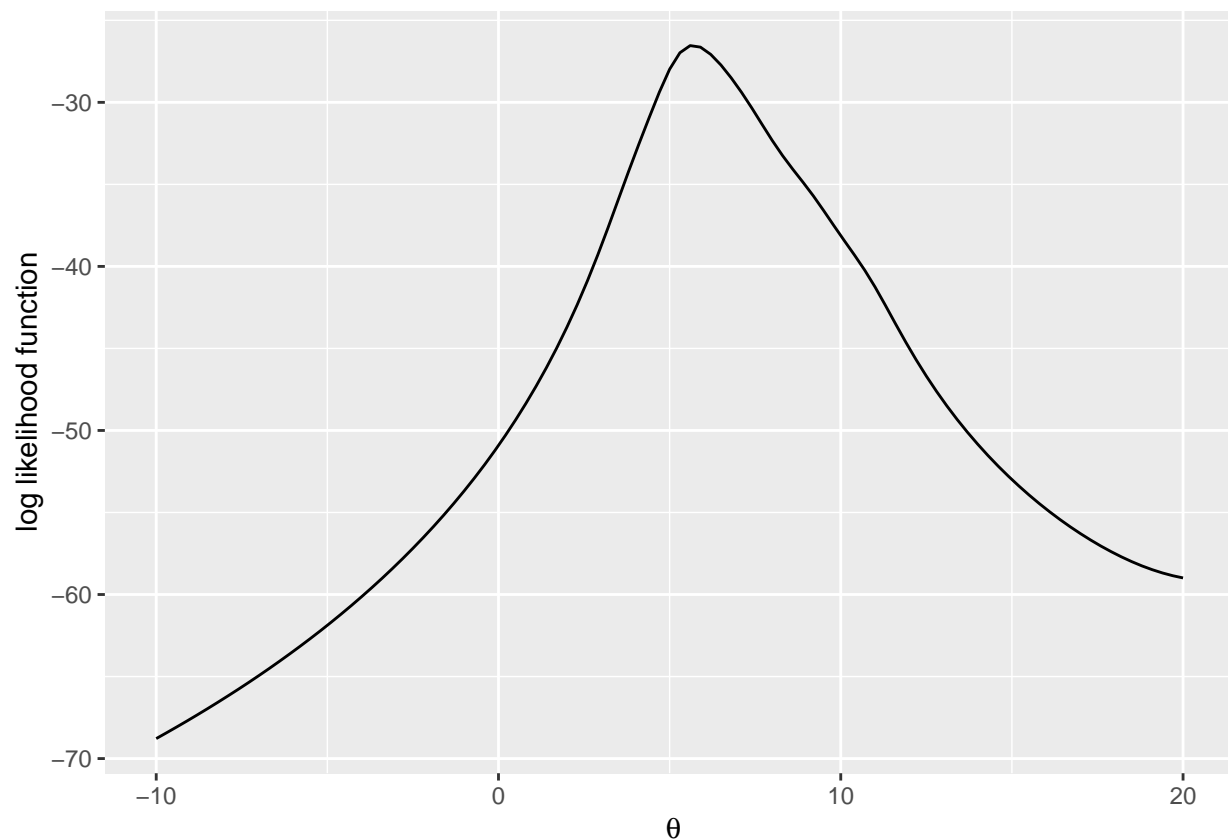
# 4 Newton-Raphson method

Find the MLE of $\theta$ using the Newton–Raphson method with initial values on a grid starting from $-10$ to $20$ with increment $0.5$. Summarize the results.

```r
library(pracma)
derivative_1 <- function(x) {          #calculate the derivative of log likelihood function
  dev1 <- 0
  for (i in 1:length(X)){
    dev1 <- dev1-2*(x-X[i])/(1+(x-X[i])^2)
  }
  dev1
}
derivative_2 <- function(x) {  #calculate the second derivative of log likelihood function
  dev2 <- 0
  for (i in 1:length(X)){
    dev2 <- dev2-2*(1-(x-X[i])^2)/(1+(x-X[i])^2)^2
  }
  dev2
}
x_initial <- seq(-10,20,by=0.5)
result <- seq(-10,20,by=0.5)
for (i in 1:length(x_initial)){
  result[i] <- newtonRaphson(derivative_1,x_initial[i],dfun = derivative_2)
}
plot(x_initial, result)
```
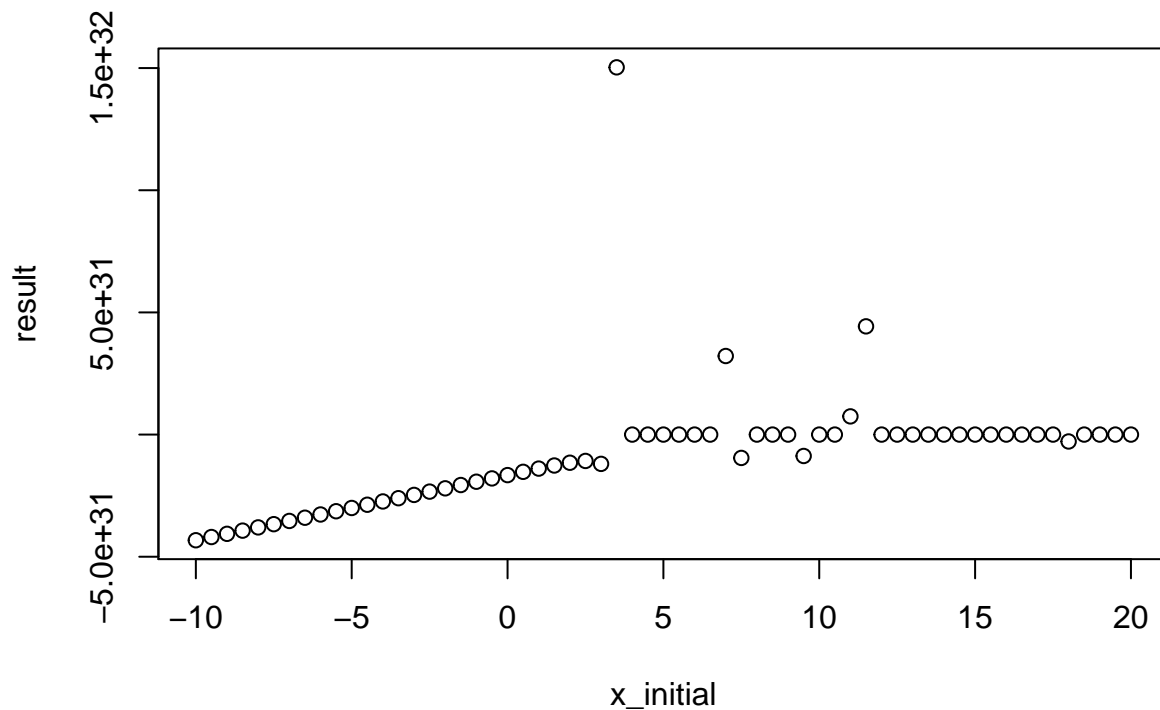
# 5 Fixed-point method

Apply fixed-point iterations using $G(\theta) = \alpha \ell'(\theta) + \theta$, with scaling choices of $\alpha \in \{1, 0.64, 0.25\}$ and the same initial values as above.

```r
alpha <- c(1, 0.64, 0.25)
# define a fixedpoint function
fixed_point <- function(derivative_1, alpha, x_initial, maxiter = 100, tol = 1e-9){
  x <- x_initial
  for (i in 1:maxiter) {
   x1 <- alpha * derivative_1(x) + x
   if (abs(x1 - x) < tol) break
   x <- x1
  }
  if (i == maxiter)
    print("maximum iteration")
  x
}
result<- array(0, c(3,length(x_initial)))
for (j in 1:length(alpha)){
  for (i in 1:length(x_initial)){
    result[j,i] <- fixed_point(derivative_1,alpha=alpha[j],x_initial=x_initial[i])}}
#In the image, blue alpha=1, green alpha=0.64,red alpha=0.25
plot(x_initial, result[1,],col="blue",xlab = "x_initial", ylab = "result")
points(x_initial, result[2,],col="green")
points(x_initial, result[3,],col="red")
```
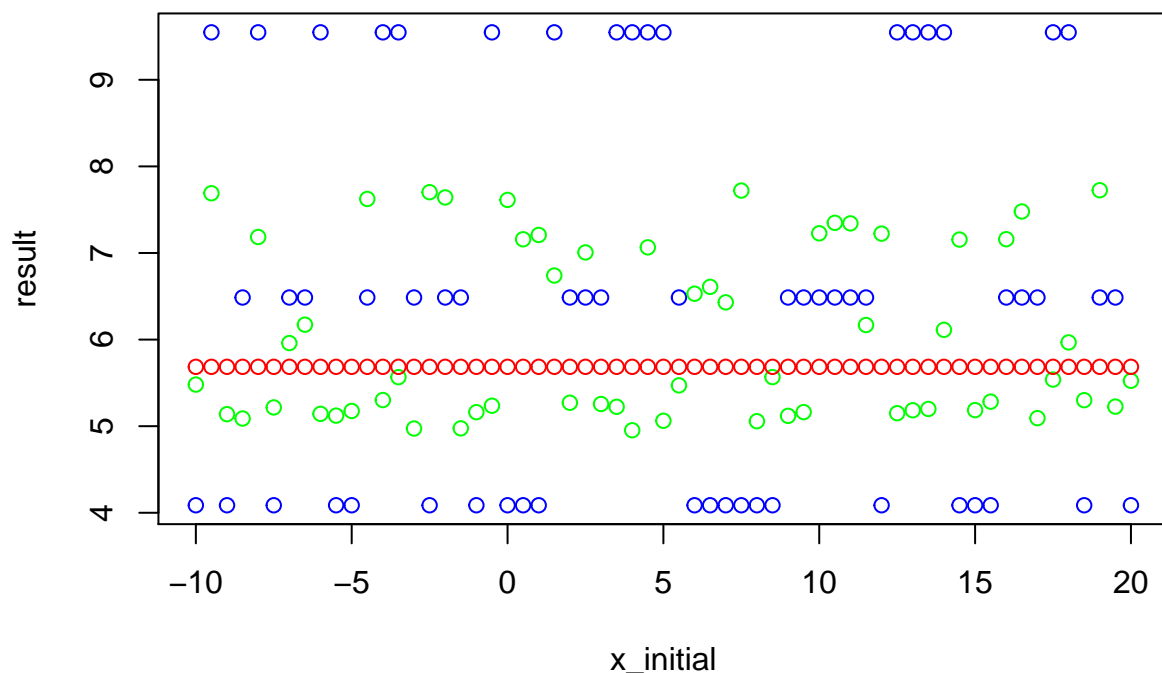
# 6 Fisher scoring method

First use Fisher scoring to find the MLE for $\theta$ ,then refine the estimate by running Newton-Raphson method.

```r
I=length(X)/2
fisher_score <- function(derivative_1, x_initial, maxiter = 100, tol = 1e-9){
  x <- x_initial
  for (i in 1:maxiter) {
   x1 <- x + derivative_1(x)/I
   if (abs(x1 - x) < tol) break
   x <- x1
  }
  if (i == maxiter)
    print("maximum iteration")
  x
}
result <-array(0, c(length(x_initial)))
for (i in 1:length(x_initial)){
  result[i] <- fisher_score(derivative_1=derivative_1, x_initial = x_initial)}
#Then use Newton method to refine the estimate and plot the result.
finalresult <- array(0, c(length(result)))
for (i in 1:length(result)){
  finalresult[i] <- newtonRaphson(derivative_1,result[i],dfun = derivative_2)
}
plot(x_initial, finalresult)
```
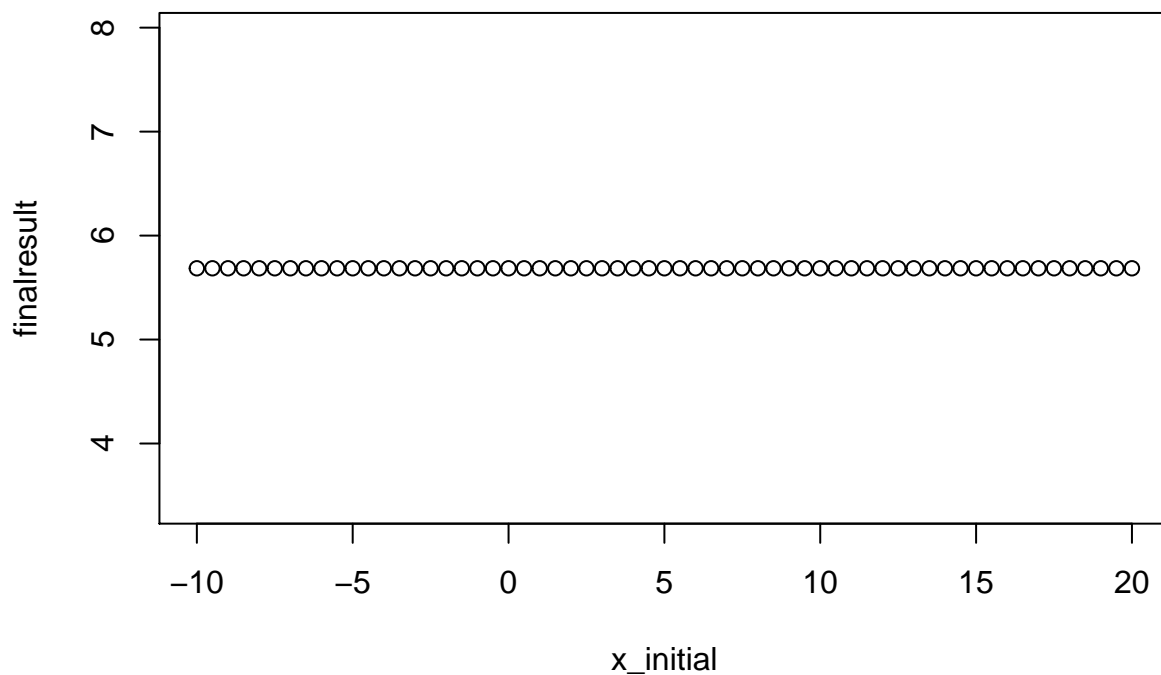
# 7 Comments

As we cab see from above images , Newton-Raphson method is the most unstable one, because the results change a lot when we choose different initial value, and the Fisher scoring method with results refined bt Newton method is the most stable one. However, Newton's method is the fastest due to its algorithms.