Probability, Computation and Simulation Homework 5

Brian Cervantes Alvarez

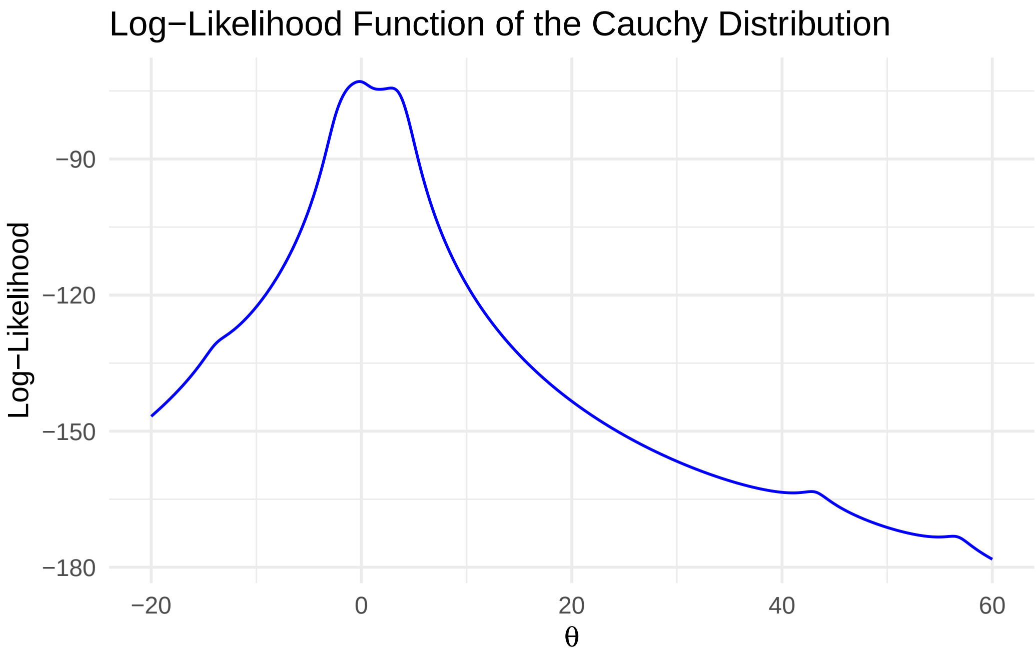
November 15, 2024

Problem 2.1

Part A

The log-likelihood function for the Cauchy distribution with scale parameter 1 is:

$$\ell(\theta) = -n\log(\pi) - \sum_{i=1}^{n} \log\left[1 + (x_i - \theta)^2\right]$$



The log-likelihood function is multimodal, reflecting the heavy-tailed nature of the Cauchy distribution. The presence of outliers causes the function to have multiple local maxima.

The Newton–Raphson update formula is:

$$\theta_{n+1} = \theta_n - \frac{\ell'(\theta_n)}{\ell''(\theta_n)}$$

First derivative:

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

Second derivative:

$$\ell''(\theta) = \sum_{i=1}^{n} \left[\frac{2}{1+(x_i-\theta)^2} - \frac{4(x_i-\theta)^2}{(1+(x_i-\theta)^2)^2} \right]$$

```
  Start Point  MLE of   Iterations
1       -11.0  3.776843      10000
2        -1.0  2.798120      10000
3         0.0 -2.648270      10000
4         1.5  4.492135      10000
5         4.0 -9.389038      10000
6         4.7  3.523097      10000
7         7.0  8.261298      10000
8         8.0 20.480057      10000
9        38.0 -2.004971      10000
```

```
$theta
[1] 1.39483

$iterations
[1] 10000
```

The mean is not ideal as a starting point for the Cauchy distribution because it is heavily influenced by outliers, causing suboptimal convergence.

Part B

```
$theta
[1] -0.1922865

$iterations
[1] 21
```

```
Error in bisectionMethod(scoreFunc, lower = 10, upper = 20, sample = sample) :
  The function must change sign over the interval.
```

The bisection method works only when the score function changes sign over the interval. Choosing the wrong interval can lead to failure, as shown above

Part C

```
  Alpha   MLE of   Iterations
1  1.00 -0.1922916         74
2  0.64 -0.1922951        114
3  0.25 -0.1923108        274
```

Smaller values of α lead to more iterations but can improve numerical stability. Larger (α) may fail to converge if too large!

Part D

```
$theta
[1] -0.1922866

$iterations
[1] 6

$theta
[1] -3

$iterations
[1] 1000
```

The secant method converges faster than bisection but requires two valid starting points. If the points are too close or identical, the method may fail.

Part E

Cauchy Distribution: When applying different estimation methods to the Cauchy distribution, we observed that some techniques were unreliable and took varying amounts of time to find the correct value. This inconsistency was mainly because the Cauchy data included extreme values, which made it challenging for the methods to consistently arrive at the right answer. Only certain methods, when started with good initial guesses, were able to work effectively. Other approaches were more dependable but tended to be slower. Overall, the unusual nature of the Cauchy data made it harder for these methods to perform consistently without careful setup.

Normal Distribution: In contrast, when we used the same estimation methods on data from a Normal distribution, all techniques performed smoothly and quickly. The methods reliably found the correct value with minimal effort because the Normal distribution is more balanced and doesn't have extreme outliers. This made the estimation process straightforward and efficient, demonstrating that these numerical methods work best with well-behaved data like that from the Normal distribution.

Comparison of Numerical Methods on Cauchy$(\theta, 1)$

Method	Converged	Theta Estimate	Iterations	Time Taken (secs)
Newton-Raphson (start = -11)	FALSE	0.5449998	1000	0.010
Newton-Raphson (start = -1)	FALSE	-0.5655976	1000	10.002
Newton-Raphson (start = 0)	FALSE	-3.2913955	1000	20.002
Newton-Raphson (start = 1.5)	FALSE	2.8366942	1000	30.003
Newton-Raphson (start = 4)	FALSE	0.4951102	1000	40.002
Newton-Raphson (start = 4.7)	FALSE	1.9993329	1000	50.002
Newton-Raphson (start = 7)	FALSE	3.6492607	1000	60.002
Newton-Raphson (start = 8)	FALSE	2.6277720	1000	0.002
Newton-Raphson (start = 38)	FALSE	-3.6444995	1000	0.000
Bisection (-20, 60)	TRUE	-0.1922864	25	0.001
Fixed-Point (alpha = 1)	TRUE	-0.1922917	74	0.002
Fixed-Point (alpha = 0.64)	TRUE	-0.1922951	114	0.000

Method	Converged	Theta Estimate	Iterations	Time Taken (secs)
Fixed-Point (alpha = 0.25)	TRUE	-0.1923108	274	0.001
Secant (start = -2, -1)	TRUE	-0.1922867	2	0.004
Secant (start = -3, -3)	FALSE	NA	NA	0.000

Comparison of Numerical Methods on Normal$(\theta, 1)$

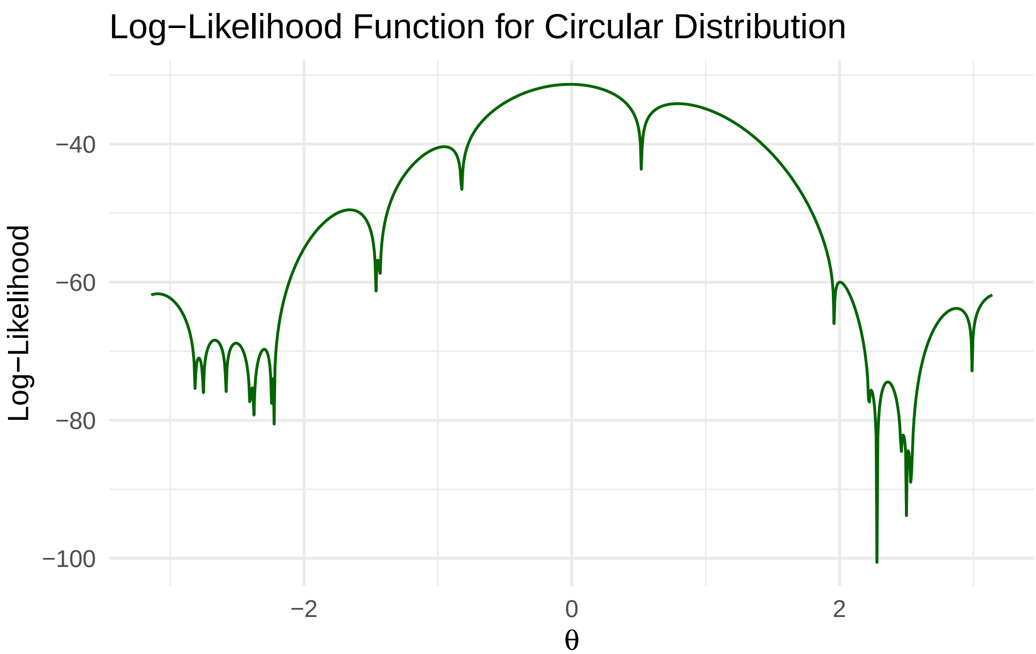
Method	Converged	Theta Estimate	Iterations	Time Taken (secs)
Newton-Raphson (Normal)	TRUE	0.6416238	1	0.004
Bisection (Normal)	TRUE	0.6416236	25	0.001
Fixed-Point (alpha = 1) (Normal)	TRUE	0.6416238	1	0.003
Fixed-Point (alpha = 0.64) (Normal)	TRUE	0.6416238	1	0.000
Fixed-Point (alpha = 0.25) (Normal)	TRUE	0.6416238	1	0.000
Secant (start = mean-1, mean+1)	TRUE	0.6416238	2	0.000
Secant (start = 0, 0)	FALSE	NA	NA	0.001

Problem 2.2

Part A

The log-likelihood function is:

$$\ell(\theta) = -n \log(2\pi) + \sum_{i=1}^{n} \log\left[1 - \cos(x_i - \theta)\right]$$



Part B

The method-of-moments estimator for θ is:

$$\hat{\theta}_{\mathrm{MM}} = \arctan 2\left(\frac{1}{n}\sum_{i=1}^{n}\sin x_i, \frac{1}{n}\sum_{i=1}^{n}\cos x_i\right)$$

```
[1] -3.078567
```

Part C

We need to compute the derivative of the log-likelihood,

$$\ell'(\theta) = \sum_{i=1}^{n} \frac{\sin(x_i - \theta)}{1 - \cos(x_i - \theta)}$$

Then solve for the second derivative as follows,

$$\ell''(\theta) = \sum_{i=1}^{n} \frac{\cos(x_i - \theta)}{1 - \cos(x_i - \theta)} - \left(\frac{\sin(x_i - \theta)}{1 - \cos(x_i - \theta)}\right)^2$$

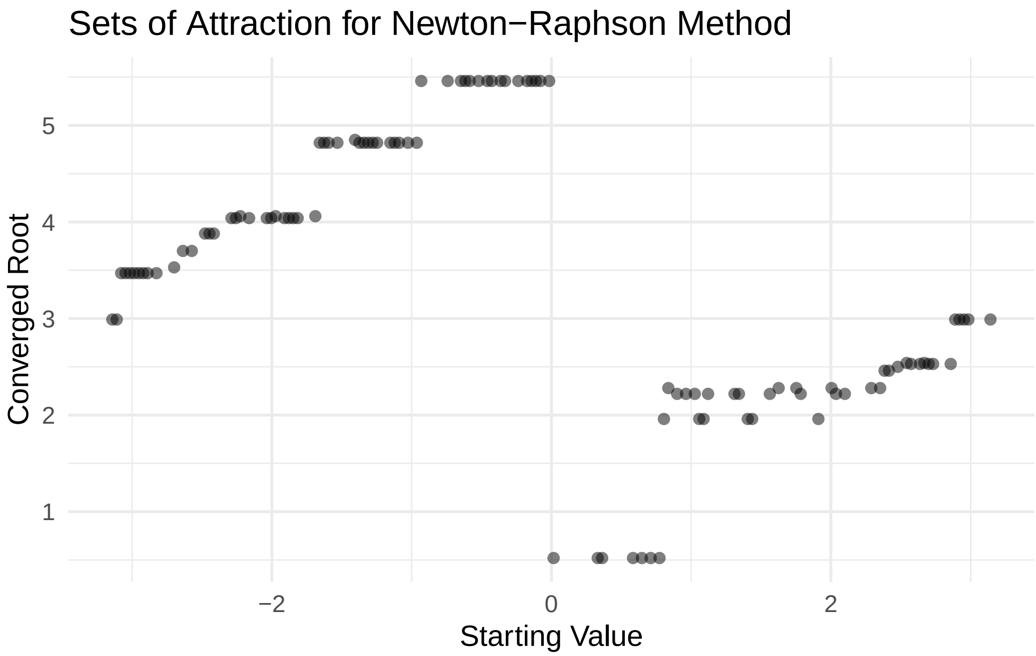
```
$theta
[1] -2.813185

$iterations
[1] 9
```

```
$theta
[1] -2.753185

$iterations
[1] 5
```

```
$theta
[1] 2.53

$iterations
[1] 5
```

Part D

We run the Newton–Raphson method with 200 equally spaced starting values between $-\pi$ and π.



The plot shows how different starting values lead to convergence to different roots. The interval is partitioned into regions where starting values converge to the same local maxima.

Part E

We search for two starting values close to each other that converge to different solutions:

```
startTheta1        root1
       1.50         2.22

startTheta2        root2
     1.5001       2.2200
```