Homework 2

Mengjiao Wu

February 4, 2018

Question 1

Part (a)

Log-likelihood function is:

$$\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} \frac{1}{\pi[1+(x_i-\theta)^2]} \\
&= (\frac{1}{\pi})^n \cdot [\frac{1}{1+(x_1-\theta)^2}] \cdot [\frac{1}{1+(x_2-\theta)^2}] \cdots [\frac{1}{1+(x_n-\theta)^2}] \\
&= \pi^{-n} \cdot [1+(x_1-\theta)^2]^{-1} \cdot [1+(x_2-\theta)^2]^{-1} \cdots [1+(x_2-\theta)^2]^{-1}
\end{aligned}$$

So,

$$\begin{aligned}
l(\theta) &= \ln\left\{\pi^{-n} \cdot [1+(x_1-\theta)^2]^{-1} \cdot [1+(x_2-\theta)^2]^{-1} \cdots [1+(x_2-\theta)^2]^{-1}\right\} \\
&= \ln\left(\pi^{-n}\right) + \sum_{i=1}^{n} \ln\left[1+(x_i-\theta)^2\right]^{-1} \\
&= -n\ln\pi - \sum_{i=1}^{n} \ln\left[1+(x_i-\theta)^2\right] \\
&= -n\ln\pi - \sum_{i=1}^{n} \ln\left[1+(\theta-x_i)^2\right]
\end{aligned}$$

First derivatives of log-likelihood function is:

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

Second derivatives of log-likelihood function is:

$$\begin{aligned}
l''(\theta) &= -2\sum_{i=1}^{n} \frac{1+(\theta-x_i)^2-(\theta-x_i)(2\theta-2x_i)}{[1+(\theta-x_i)^2]^2} \\
&= -2\sum_{i=1}^{n} \frac{1+(\theta-x_i)^2-2(\theta-x_i)^2}{[1+(\theta-x_i)^2]^2} \\
&= -2\sum_{i=1}^{n} \frac{1-(\theta-x_i)^2}{[1+(\theta-x_i)^2]^2}
\end{aligned}$$

For Fisher information, note that $I(\theta) = nI_{x_1}(\theta) = nI_x(\theta)$, so,

$$
\begin{aligned}
I(\theta) &= nE\left[\left(\frac{2(x-\theta)}{1+(x-\theta)^2}\right)^2\right] \\
&= n\int_{-\infty}^{\infty}\left(\frac{2(x-\theta)}{1+(x-\theta)^2}\right)^2\frac{1}{\pi\left[1+(x-\theta)^2\right]}dx \\
&= \frac{4n}{\pi}\int_{-\infty}^{\infty}\frac{(x-\theta)^2}{\left[1+(x-\theta)^2\right]^3}dx
\end{aligned}
$$

So, it can be derived that $I(\theta) = n\int \frac{[p'(x)]^2}{p(x)}dx$.
Letting $u = x-\theta$ and $du = dx$,

$$
I(\theta) = \frac{4n}{\pi}\int_{-\infty}^{\infty}\frac{u^2}{\left[1+u^2\right]^3}du = \frac{8n}{\pi}\int_{0}^{\infty}\frac{u^2}{\left[1+u^2\right]^3}du = \frac{8n}{\pi}\int_{0}^{\infty}\frac{u^2}{1+u^2}\left(\frac{1}{1+u^2}\right)^2du
$$

Substituting $x = 1/(1+u^2)$, $u = (1/x-1)^{1/2}$ and $du = \frac{1}{2}(1/x-1)^{-1/2}(-1/x^2)$, $I(\theta)$ can be expressed as:

$$
\begin{aligned}
I(\theta) &= \frac{8n}{\pi}\int_{0}^{\infty}\frac{u^2}{1+u^2}\left(\frac{1}{1+u^2}\right)^2du = \frac{8n}{\pi}\int_{0}^{1}(1-x)x^2\cdot(1/2)(1/x-1)^{-1/2}(-1/x^2)dx \\
&= \frac{4n}{\pi}\int_{0}^{1}x^{\frac{1}{2}}(1-x)^{\frac{1}{2}}dx = \frac{4n}{\pi}\int_{0}^{1}x^{\frac{3}{2}-1}(1-x)^{\frac{3}{2}-1}dx = \frac{4n}{\pi}\cdot\frac{(0.5\sqrt{\pi})^2}{2\times 1} = \frac{n}{2}
\end{aligned}
$$

Part (b)

Graph of log-likelihood function and Table containing results for different starting points are shown as:

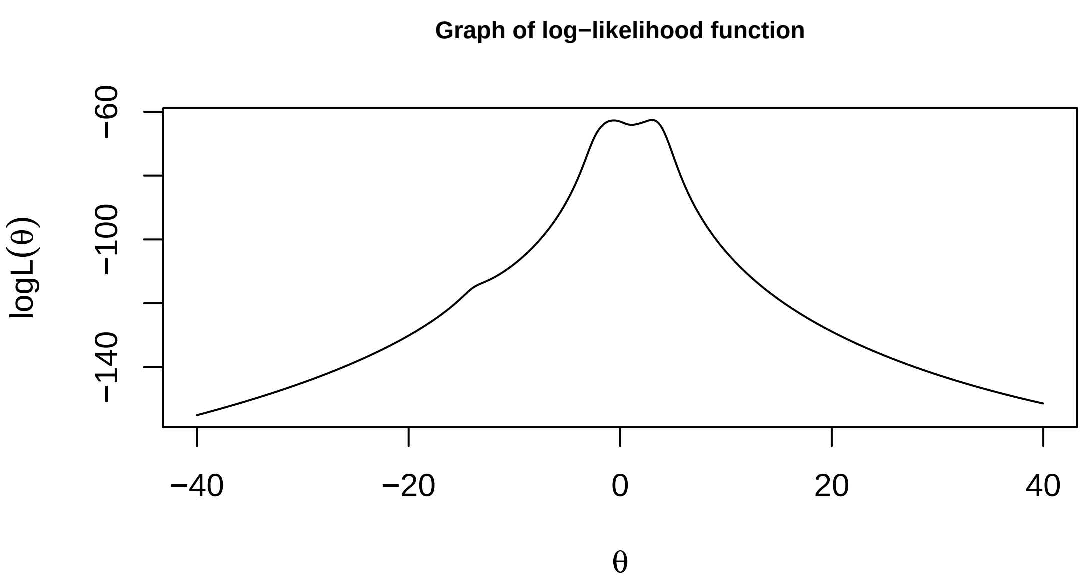


Table 1: MLE for theta using Newton-Raphson method

Initial value	MLE for theta
-11.00000	-6.200934e+09
-1.00000	-5.914700e-01
0.00000	-5.914700e-01
1.50000	1.092730e+00
4.00000	3.021350e+00
4.70000	-5.914700e-01
7.00000	5.175120e+09
8.00000	5.478284e+09
38.00000	4.807973e+09
3.25778	3.021350e+00

From the table, it can be observed that for initial values -1, 0 and 4.7, MLE for θ can be found to be -0.59147 through Newton-Raphson method. For initial values 1.5, 4 and $\overline{x}$, the results converge to the local maximum values. Therefore, 1.5, 4 and sample mean are not good starting points. In addition, for initial values -11, 7, 8 and 38, outputs of each step do not show the convergency. Hence, these values should not chosen as starting points either.

Part (c)

Results for different starting points are displayed in the following table with maximum iteration set to be 1000,

Table 2: MLE for theta using Fixed-point iterations

Initial value	alpha=1	alpha=0.64	alpha=0.25
-11	-0.591473515623568	-0.591473522524697	-0.591473525266997
-1	Maximum iterations reached	-0.59147352228533	-0.591473523438432
0	Maximum iterations reached	-0.591473519008384	-0.591473515739687
1.5	Maximum iterations reached	Maximum iterations reached	3.02134544306586
4	Maximum iterations reached	-0.591473519240639	3.02134544141267
4.7	Maximum iterations reached	-0.591473523492669	3.02134544310295
7	Maximum iterations reached	Maximum iterations reached	3.02134544282813
8	Maximum iterations reached	-0.591473523478013	3.02134544078601
38	Maximum iterations reached	Maximum iterations reached	3.02134544306369

Part (d)

Results for different starting points are displayed in the following table:

Table 3: MLE for theta using Fisher scoring and Newton method

Initial value	MLE for theta
-11.0	-0.59147
-1.0	-0.59147
0.0	-0.59147
1.5	3.02135
4.0	3.02135

Initial value	MLE for theta
4.7	3.02135
7.0	3.02135
8.0	3.02135
38.0	3.02135

Part (e)

For computation methods used in Part (a), Part (b) and Part(c), calculation time are respectively 0.184489 secs, 0.233618 secs and 0.2210901 secs. The difference of calculation time is relatively small among these three methods. However, by comparing the calculated results, it can be found that Newton method converges relatively slowly and choice of starting points has great influence on whether MLE values of θ can be derived correctly. For Fixed-point interation, it converges relatively quickly while the α value is important for convergency. If a proper α is chosen, Fixed-point interation can be an efficient computation method. The combination of Fisher scoring and Newton method converges most efficiently among these three methods and is considered to be the best choice for this question. Based on the log-likelihood graph, since there are two local maximum values, initial value of θ_0 are critical to all three methods so that proper initial values should be selected.

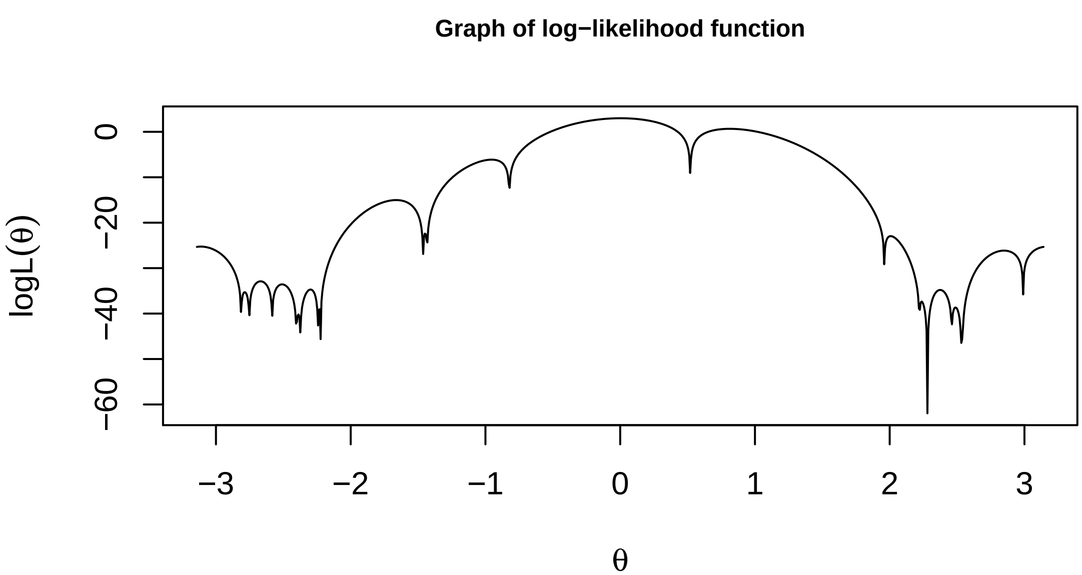
Question 2

Part (a)

Log-likelihood function is:

$$l(\theta) = \log \frac{1 - \cos(x - \theta)}{2\pi} = \log\left[1 - \cos(x - \theta)\right] - \log(2\pi)$$

Graph of log-likelihood function is shown as:



Part (b)

According to the question, $\mathbb{E}[X \mid \theta] = \overline{x}$. Since $\mathbb{E}[X \mid \theta] = \int x f(x) dx$,

$$\overline{x} = \frac{1}{2\pi} \int_0^{2\pi} x \left[1 - \cos(x - \theta)\right] dx = \pi - \frac{1}{2\pi} \int_0^{2\pi} x \cos(x - \theta) dx$$

Using integration by parts, it can be obrained that,

$$\int_0^{2\pi} x \cos(x - \theta) dx = x \sin(x - \theta)\Big|_{x=0}^{2\pi} - \int_0^{2\pi} \sin(x - \theta) dx = 2\pi \sin(2\pi - \theta)$$

Since $\sin(-x) = \sin(x)$ and $\sin(2\pi + x) = \sin(x)$, $\overline{x}$ can be calculated and θ is:

$$\overline{x} = \pi - \frac{1}{2\pi} \int_0^{2\pi} x \cos(x - \theta) dx = \pi + \sin\theta \iff \theta = \arcsin(\overline{x} - \pi)$$

Therefore,

$$\hat{\theta}_{moment} = \arcsin(\overline{x} - \pi) = 0.0954$$

Part (c) and (d)

MLEs for θ at $\theta_0 = \hat{\theta}_{moment}, -2.7, 2.7$ are presented as:

Table 4: MLE for theta

Initial value	**MLE for theta**
0.09539	0.00312
-2.70000	-2.66886
2.70000	2.84842

Part (e)

Interval $[-\pi, \pi]$ can be partitioned into several subintervals A_i based on the standard that $\theta \in A_i$ if and only if $\theta_0 = \theta$ suggests $\theta_t = \hat{\theta}_i$. Estimations of the upper and lower bounds of the corresponding A_i at different $\hat{\theta}_i$ are demonstrated in the forrlowing table:

Table 5: Partition of the interval

	Lower	**Upper**
-3.1125	-3.14159	-2.82585
-2.7866	-2.79428	-2.76271
-2.6689	-2.73113	-2.60484
-2.5094	-2.57326	-2.41540
-2.3883	-2.38382	-2.38382
-2.2979	-2.35225	-2.25753
-2.2322	-2.22595	-2.22595
-1.6627	-2.19438	-1.46818
-1.4475	-1.43661	-1.43661

	Lower	Upper
-0.9544	-1.40503	-0.83671
0.0031	-0.80513	0.48939
0.8126	0.52097	1.94179
2.0072	1.97336	2.19438
2.237	2.22595	2.25753
2.3747	2.28910	2.44697
2.4884	2.47854	2.51012
2.8484	2.54169	2.98372
3.1707	3.01530	3.14159

Question 3

Part (a)

The main idea of Gauss-Newton is to approximate the function $f_\theta(t)$ by alinear function. In this question, partial derivatives of f_θ with respect to K and r are calculated to derive the gradient vector:

$$\frac{\partial\, f_\theta(t)}{\partial\, K} = \frac{N_0^2(1-e^{-rt})}{[N_0(1-e^{-rt})+Ke^{-rt}]^2}$$

$$\frac{\partial\, f_\theta(t)}{\partial\, r} = \frac{tKN_0(K-N_0)e^{-rt}}{[N_0(1-e^{-rt})+Ke^{-rt}]^2}$$

So, the gradient vector g_i is:

$$g_i = \left[\frac{1}{N_0(1-e^{-rt_i})+Ke^{-rt_i}}\right]^2 \left(N_0^2(1-e^{-rt_i}), t_iKN_0(K-N_0)e^{-rt_i}\right)^{\mathrm{T}}$$

and the residual at each iteration is:

$$z_i = N_i - f_\theta(t_i)$$

Let G be the matrix of g_i as its $i-th$ row, θ_{t+1} can be expressed as:

$$\theta_{t+1} = \theta_t + (G_t^{\mathrm{T}}G_t)^{-1}G_t^{\mathrm{T}}z_t$$

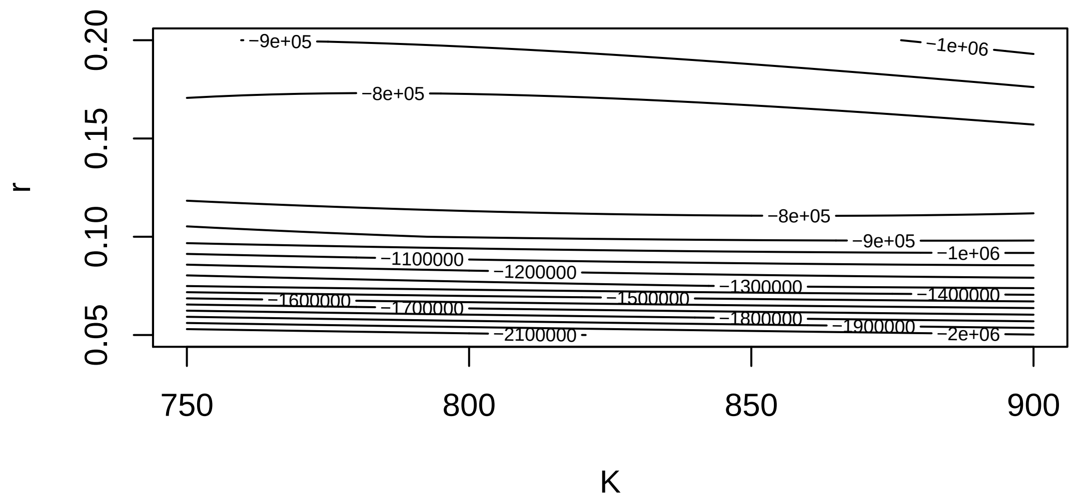
Least-squares estimates of K and r with initial values $(800, 0.1)$ are:

Table 6: LSE values of K and r

LSE of K	LSE of r
828.8996	0.13406

Part (b)

Contour plot is presented as:



Part (c)

According to the question, it can be derived that:

$$\log(N_t) \sim (\log\left[\frac{KN_0}{N_0 + (K - N_0)e^{-rt}}\right], \sigma^2)$$

So, using the bbmle.mle2() function in R, maximum likelihood estimators of K, r, σ^2 with initial value $(500, 0.5, 0.5)$ and the variance of these estimations are:

Table 7: MLE and the variance summary

	Estimate	Std. Error	z value	Pr(z)
K	500.00272	130.4146	3.83395	0.00013
r	0.40495	0.1046	3.87123	0.00011
var	0.55168	0.2468	2.23536	0.02539