Tuto1_personal_assessment

Question 1

Question 1.a

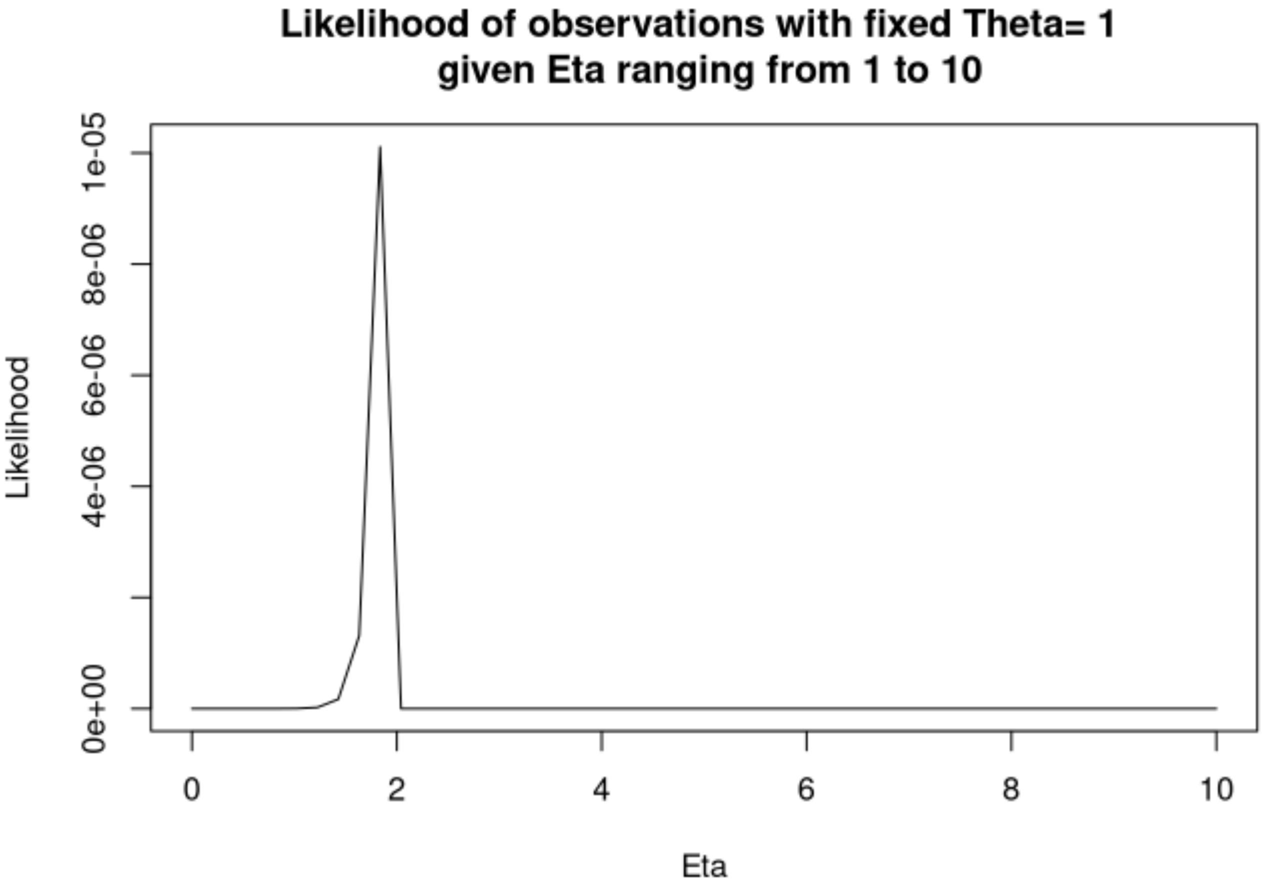
The likelihood could have also been rewritten with the sample mean in the exponent: $\bar{X} = \frac{1}{n} \sum_{i=1}{}^{n} X_i$

Notation mistake at the end of the likelihood function: $\theta^n e^{-\theta \sum_{i=1}^{n}(x_i - \eta)} 1_{min(x_i) \geq \eta}$ should have been $\theta^n e^{-\theta \sum_{i=1}^{n}(x_i - \eta)} 1_{min(X) \geq \eta}$.

Question 1.b

Though noted in the function `isi_generate` , the markdown could have been more explicit about the formula $X_i = \eta + Y_i$.

Using the argument `type="l"` in the plotting function `isi_likelihood_plot_given_theta` could have given a continuous visualization of the likelihood of observations as a function of η with a fixed θ. The use of a finer sequence to define η could have given a smoother visualization as well (`etas = seq(0, 10, length.out=50)` was used.)



Question 1.c

The definition of $\hat{\eta}$ could have been more explicit, especially with regards to how it relates, in methodology, to the computation of a MLE for a uniform distribution (https://math.stackexchange.com/questions/411145/maximum-likelihood-estimation-of-a-b-for-a-uniform-distribution-on-a-b), plus the reference to Wasserman, L., *All of Statistics*, p125 (bottom right graph):

EM algorithm

multiparameter delta method

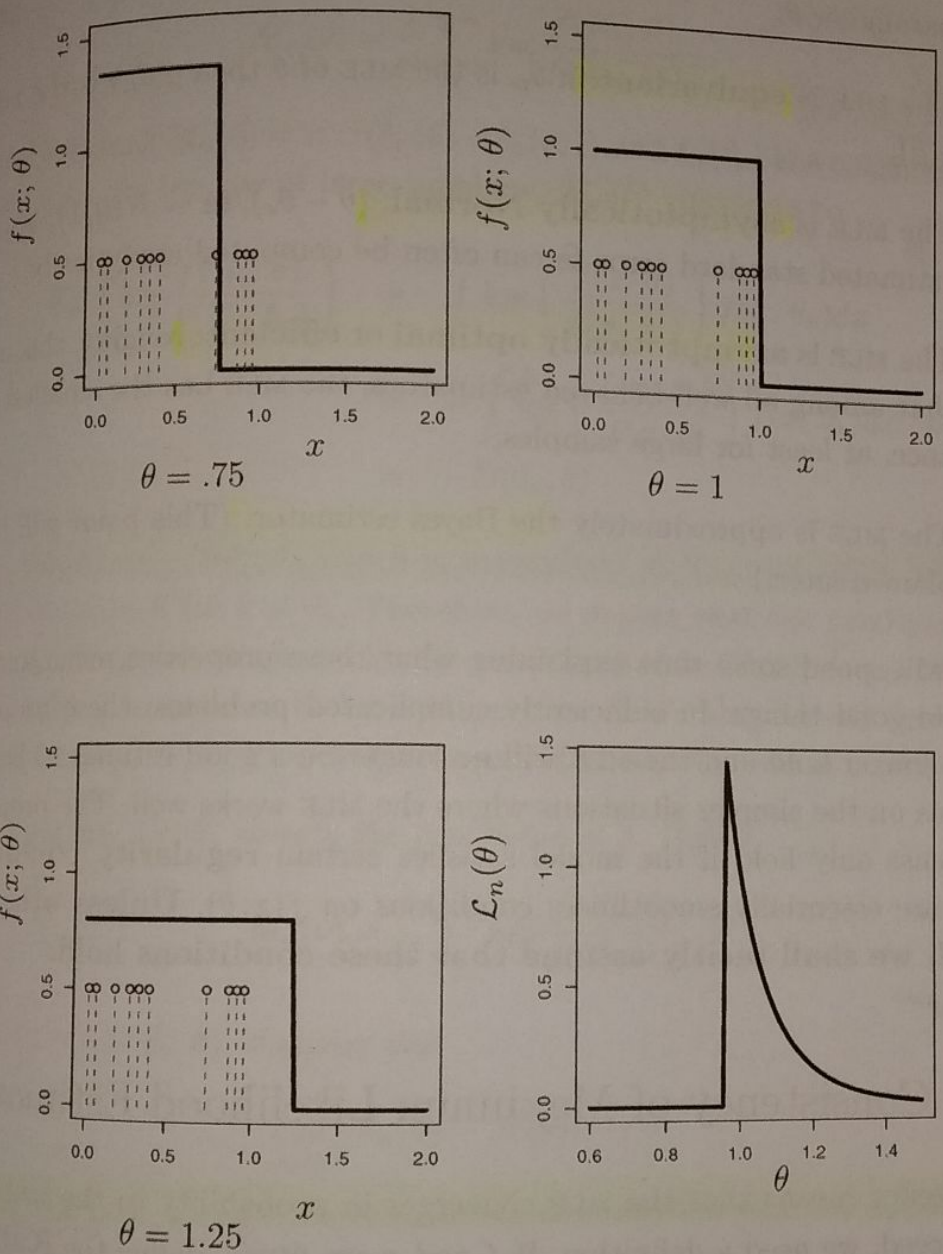


FIGURE 9.2. Likelihood function for Uniform $(0, \theta)$. The vertical lines show the observed data. The first three plots show $f(x;\theta)$ for three different values of θ. When $\theta < X_{(n)} = \max\{X_1, \ldots, X_n\}$, as in the first plot, $f(X_{(n)};\theta) = 0$ and hence $\mathcal{L}_n(\theta) = \prod_{i=1}^n f(X_i;\theta) = 0$. Otherwise $f(X_i;\theta) = 1/\theta$ for each i and hence $\mathcal{L}_n(\theta) = \prod_{i=1}^n f(X_i;\theta) = (1/\theta)^n$. The last plot shows the likelihood function.

.

Computing the MLE of η for the ISI exponential model is the reverse of the uniform distribution. The parameter η is upper-bounded by the lowest value among the observations X_i.

Question 1.d

The η used in the tutorial were integers compared to the small, floating point η found in the correction. This leads to a coarser representation in a graph.

Question 1.e

The different graphs should have been better commented to explain that increasing MLEs for θ should indicate an increasing spiking rate. Furthermore, the scale effect between θ and η results in the MLE of η being squished into a line (detail was lost when plotting both η and θ together). η should have been displayed with an inflated value to be visually interesting to look at.

Question 1.f

Instruction was misunderstood. The `non-parametric density estimator` was to be understood as `kernel density estimation function` , found in R’s standard library (https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/density).

Question 2

Question 2.a

Explanation is okay, just very verbose. It could have been condensed.

Question 2.b

Preliminary Note

Given the `question 2` is working with unit vectors, a notation shortcut was introduced in the question, which may cause incomprehension, where M is equated to θ with the following description:

$$\begin{aligned} M &= (x, y) \quad \text{with } x, y \in \mathbb{R} \\ e_1 &= (1, 0) \\ e_2 &= (0, 1) \\ < M, e_1 > &= x = cos(M) \\ < M, e_2 > &= y = sin(M) \end{aligned}$$

This should not have been done for the sake of clarity.

Comment

The main difference between the correction and the given tutorial is the representation of the B matrix as one-dimensional (given tutorial) instead of two-dimensional (correction). It was assumed that, given the instruction mentioned e_1 and e_2 as being unit vectors, the model representation was to be represented using sines and cosines in a single expression.

Though the expression differs, they should be equivalent in this question (the solution given as part of the assignment is $B\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$ evaluated.

Representing Y as a Gaussian expression was not explicitly given here, however:

$$\begin{aligned}
&\forall i \in \{1, \ldots, n_1 + n_2\} \\
a_i, b_i, \sigma_i &\in \mathbb{R} \\
\epsilon_i &\sim \mathcal{N}(0,1) \\
Y &= \begin{pmatrix} a_1 \\ \vdots \\ a_{n_1} \\ a_{n_1+1} \\ \vdots \\ a_{n_1+n_2} \end{pmatrix} + \begin{pmatrix} b_1 * cos(M) \\ \vdots \\ b_{n_1} * cos(M) \\ b_{n_1+1} * sin(M) \\ \vdots \\ b_{n_1+n_2} * sin(M) \end{pmatrix} + \sigma\epsilon \\
\sigma &= \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & & \ddots & \\ 0 & & & \sigma_{n_1+n_2} \end{bmatrix}; \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1+n_2} \end{pmatrix}
\end{aligned}$$

The given assignment relied on representing $B\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$ as the output of a function with input argument the angle M and output a $\mathbb{R}^{n_1+n_2}$ matrix $B = \begin{pmatrix} b_1 * cos(M) \\ \vdots \\ b_{n_1} * cos(M) \\ b_{n_1+1} * sin(M) \\ \vdots \\ b_{n_1+n_2} * sin(M) \end{pmatrix}$.

this representation carried over the next questions where Y and μ are also considered as functions of the angle M.

Question 2.c

Given the representation of B, Y, etc. as function of the angle M, it was possible to represent the log-likelihood as a function of the angle M.

This solution was a bit more involved and sidestepped `Question 2.d` slightly.

Question 2.d

To answer this question, and since a norm representation was already given in the previous question, it was resorted to show a correspondence between the log-likelihood and the OLS by using the definition of the OLS.

Question 2.e

Not completed

Note

The question was not answered because the intuition of developing the norm $||Y(M) - \mu(M)||^2$ into two separate sums $\sum^{n_1}{}_{i=1}[y_i - a_i - b_i\ cos(M)]^2$ and $\sum^{n_2}{}_{i=1}[y_i - a_i - b_i\ sin(M)]^2$ was missed.