Assignment 2

Statistical analyses of whales data set

Name: Devosmita Chatterjee
Email Id: chatterjeedevosmita267@gmail.com

Statistical analyses of whales data set

1 Introduction

The whales data set consists of times (t_i, i=1, $\cdots$, 210) in hr(s) taken by the whales to swim 1 km.

2 Question 46a.

We make a histogram of the 210 values of t_i which is shown in figure 1.

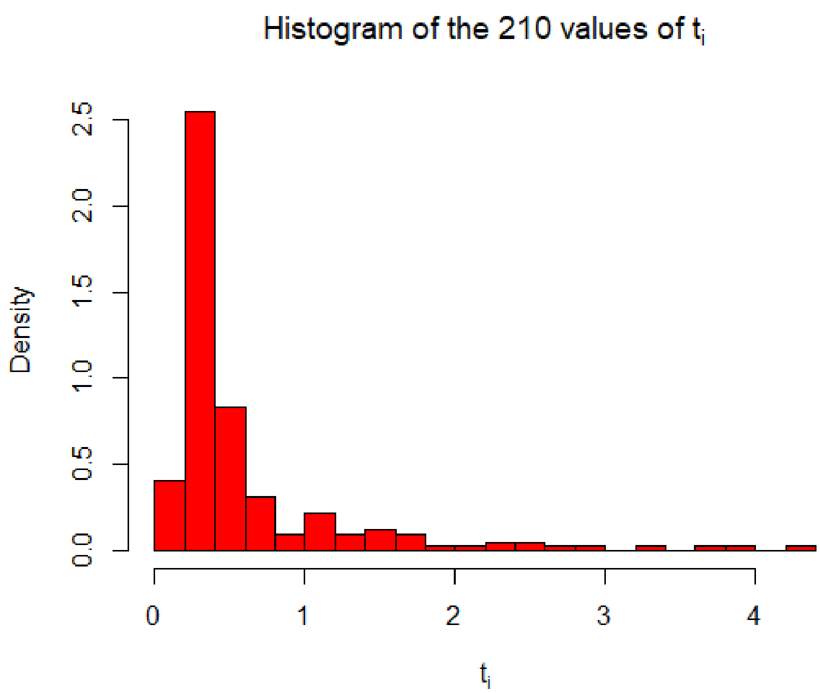

Figure 1: This figure shows the histogram of the 210 values of t_i.

Yes, gamma distribution is a plausible model to fit since the histogram looks like gamma distribution.

3 Question 46b.

We fit the parameters of the gamma distribution by the method of moments.
The method of moments estimates are given by

$$\hat{\alpha} = \frac{\hat{\mu}^2}{\hat{\sigma}^2}, \tag{1}$$

$$\hat{\lambda} = \frac{\hat{\mu}}{\hat{\sigma}^2} \tag{2}$$

where the estimates of μ and σ^2 from the sample moments are $\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$.
Therefore, the shape parameter fit of the gamma distribution by the method of moments is 0.7991741 and the scale parameter fit of the gamma distribution by the method of moments is 1.318771.

4 Question 46c.

We fit the parameters of the gamma distribution by maximum likelihood.
The density function of a gamma distribution is given by

$$f(x|\alpha,\lambda) = \frac{1}{\Gamma(\alpha)}\lambda^\alpha x^{\alpha-1}e^{-\lambda x}, \quad 0 \le x < \infty \tag{3}$$

Then, the log likelihood of an independent and identically distributed sample, $X_1, \cdots, X_n$ is given by

$$l(\alpha,\lambda) = \sum_{i=1}^{n}[\alpha \ln\lambda + (\alpha-1)\ln X_i - \lambda X_i - \ln\Gamma(\alpha)]$$

$$= n\alpha\ln\lambda + (\alpha-1)\sum_{i=1}^{n}\ln X_i - \lambda\sum_{i=1}^{n} X_i - n\ln\Gamma(\alpha) \tag{4}$$

The partial derivatives are given by

$$\frac{\partial l}{\partial \alpha} = n\ln\lambda + \sum_{i=1}^{n}\ln X_i - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}, \tag{5}$$

$$\frac{\partial l}{\partial \lambda} = \frac{n\alpha}{\lambda} - \sum_{i=1}^{n} X_i \tag{6}$$

Setting the first partial derivative to zero, we get

$$\hat{\lambda} = \frac{\hat{\alpha}}{\overline{X}}, \tag{7}$$

Substituting the above equation in the second partial derivative and setting them to zero, we get

$$n\ln\hat{\alpha} - n\ln\overline{X} + \sum_{i=1}^{n}\ln X_i - n\frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} = 0 \tag{8}$$

This equation cannot be solved in closed form and therefore, we use an iterative method for finding the roots where we can use the initial value obtained by the method of moments.
Therefore, the shape parameter fit of the gamma distribution by maximum likelihood is 1.595458 and

the scale parameter fit of the gamma distribution by maximum likelihood is 2.632601.

We compare the shape parameter and the scale parameter obtained by the method of moments to those obtained by maximum likelihood which is represented in table 1.

Table 1: This table presents the parameter values (α, λ) obtained by the method of moments (MME) and by maximum likelihood (MLE).

Method	α	λ
MME	0.7991741	1.318771
MLE	1.595458	2.632601

5 Question 46d.

We plot the two gamma densities on top of the histogram which is shown in figure 2.

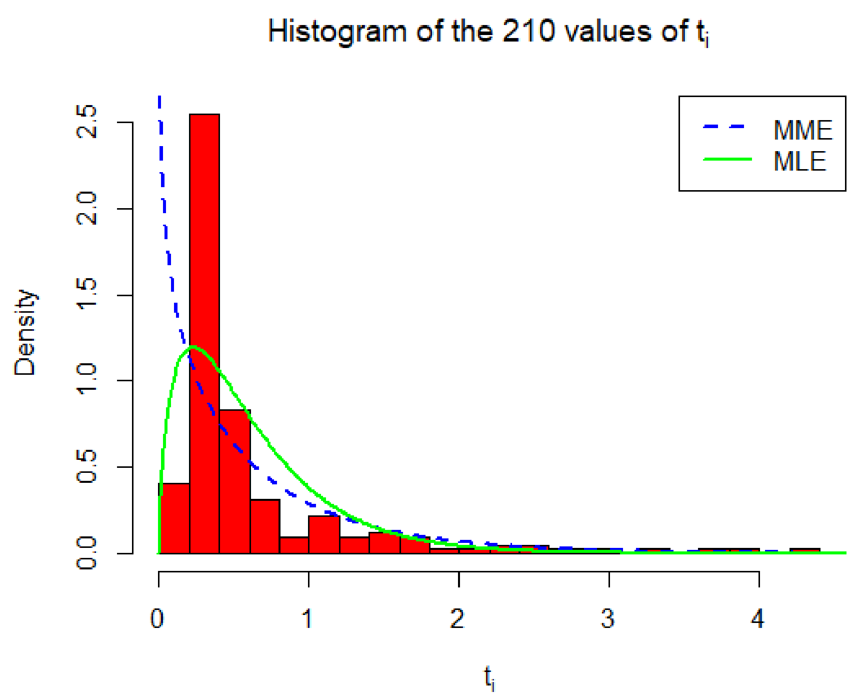

Figure 2: This figure shows the two gamma densities on top of the histogram.

Yes, both the fits look reasonable since they look like gamma distribution.

6 Question 46e.

We estimate the sampling distributions and the standard errors of the parameters fit by the method of moments by using the bootstrap.

In Bootstrap method, the standard deviation of the sampling distribution gives the standard error:

$$s_{\hat{\alpha}} = \sqrt{\frac{1}{B}\sum_{j=1}^{B}(\hat{\alpha}_j - \overline{\alpha})^2} \tag{9}$$

where $\overline{\alpha} = \frac{1}{B}\sum_{j=1}^{B}\hat{\alpha}_j$.
The sampling distributions of the parameters fit by the method of moments by using the bootstrap is shown in figure 3 and figure 4.

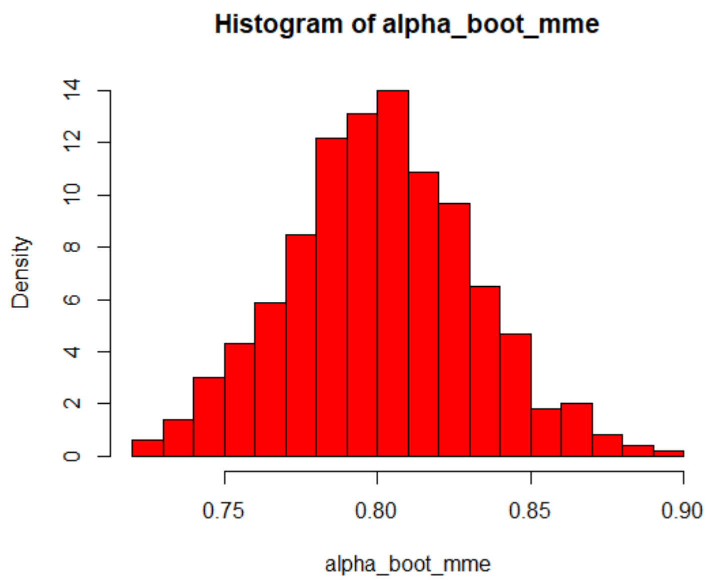

Figure 3: This figure shows the sampling distribution of the shape parameter fit by the method of moments by using the bootstrap.

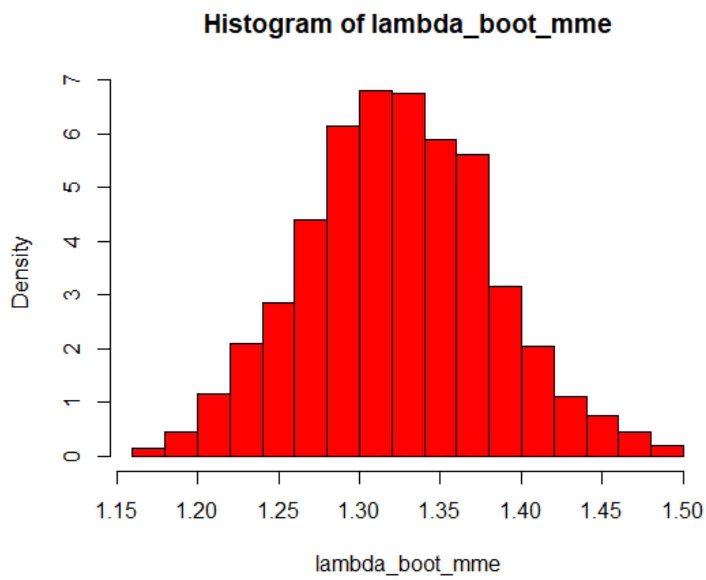

Figure 4: This figure shows the sampling distribution of the scale parameter fit by the method of moments by using the bootstrap.

Therefore, the standard error of the shape parameter fit by the method of moments by using the bootstrap is 0.03035455 and the standard error of the scale parameter fit by the method of moments by using the bootstrap is 0.05767604.

7 Question 46f.

We estimate the sampling distributions and the standard errors of the parameters fit by maximum likelihood by using the bootstrap.
The sampling distributions of the parameters fit by the method of moments by using the bootstrap is shown in figure 5 and figure 6.

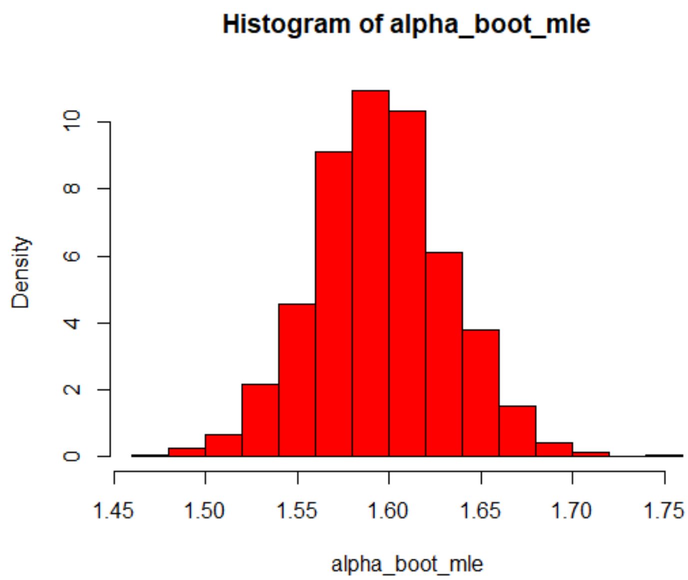

Figure 5: This figure shows the sampling distribution of the shape parameter fit by maximum likelihood by using the bootstrap.

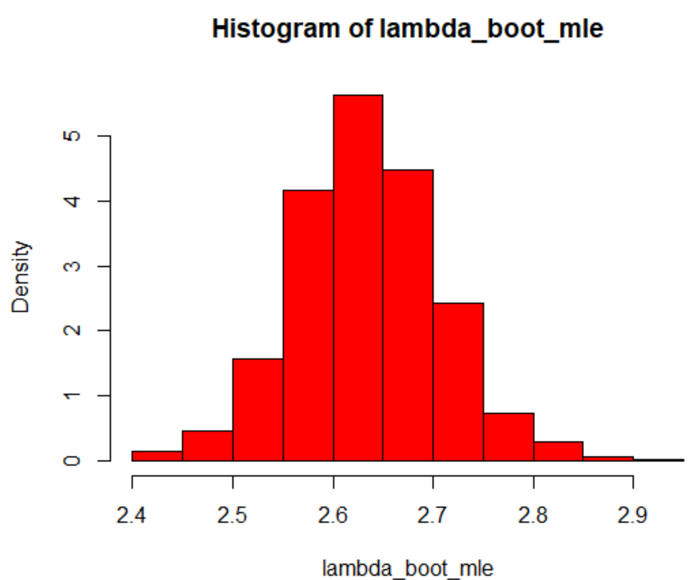

Figure 6: This figure shows the sampling distribution of the scale parameter fit by maximum likelihood by using the bootstrap.

Therefore, the standard error of the shape parameter fit by maximum likelihood by using the bootstrap is 0.03645341.
Therefore, the standard error of the scale parameter fit by maximum likelihood by using the bootstrap is 0.07305184.

We compare these results to the results found previously in table 2.

Table 2: This table presents the mean ($\overline{\alpha}$, $\overline{\lambda}$) and the standard error ($s_{\hat{\alpha}}$, $s_{\hat{\lambda}}$) of parameter values (α, λ) obtained by the method of moments (MME) and maximum likelihood (MLE).

Method	$\overline{\alpha}$	$s_{\hat{\alpha}}$	$\overline{\lambda}$	$s_{\hat{\lambda}}$
MME	0.8010712	0.03035455	1.323598	0.05767604
MLE	1.59583	0.03645341	2.634662	0.07305184

8 Question 46g.

We find approximate confidence intervals for the parameters estimated by maximum likelihood.
Approximate 100(1-a)% confidence interval for the shape parameter is given by

$$I_\alpha = \overline{\alpha} \pm z_{a/2} s_{\hat{\alpha}}. \tag{10}$$

Approximate 100(1-a)% confidence interval for the scale parameter is given by

$$I_\lambda = \overline{\lambda} \pm z_{a/2} s_{\hat{\lambda}}. \tag{11}$$

Therefore, 95% confidence interval for the shape parameter estimated by maximum likelihood is (1.524382, 1.667279) and 95% confidence interval for the scale parameter estimated by maximum likelihood is (2.49148, 2.777843).
Therefore, 95% confidence interval for the shape parameter estimated by the method of moments is (0.7415762, 0.8605661) and 95% confidence interval for the scale parameter estimated by the method of moments is (1.210553, 1.436643).