

Practical 1. Maximum Likelihood (ML) estimation

Ricard Garcia Isern & Antonio Arcas Pons

2023-10-21

Exercise 1: ML estimation of a one-parameter distribution. Let X be a random variable with probability density $f(x) = e^{-(x+1)}$ with $x > 1$. We consider a random sample of n observations of this distribution. First of all, before computing all the results, we must import the data:

```
data <- read.table('sample.dat')
data <- data.frame(data)
data <- data$V1
```

e) Download the file Sample.dat, which contains sample of observations from this probability distribution. Determine the sample size and calculate the value of the ML estimator for this sample.

We have computed in the pdf file how we have reached to the final expression of the MLE.

```
sample_size <- length(data)
MLE <- 1/(mean(data))
```

Our sample size is 10000.

The ML estimator (MLE) for this sample is 0.4736756.

f) Plot the log-likelihood function, and assess graphically if your ML estimate coincides with the maximum of this function.

First of all we must define the log-likelihood function:

```
log_likelihood_function <- function(beta,data){
  likelihood <- beta*exp(-beta*data+1)
  log_likelihood <- sum(log(likelihood))
  return(log_likelihood)
}

#different beta values
different_betas<-seq(0.1,3,by=0.01)

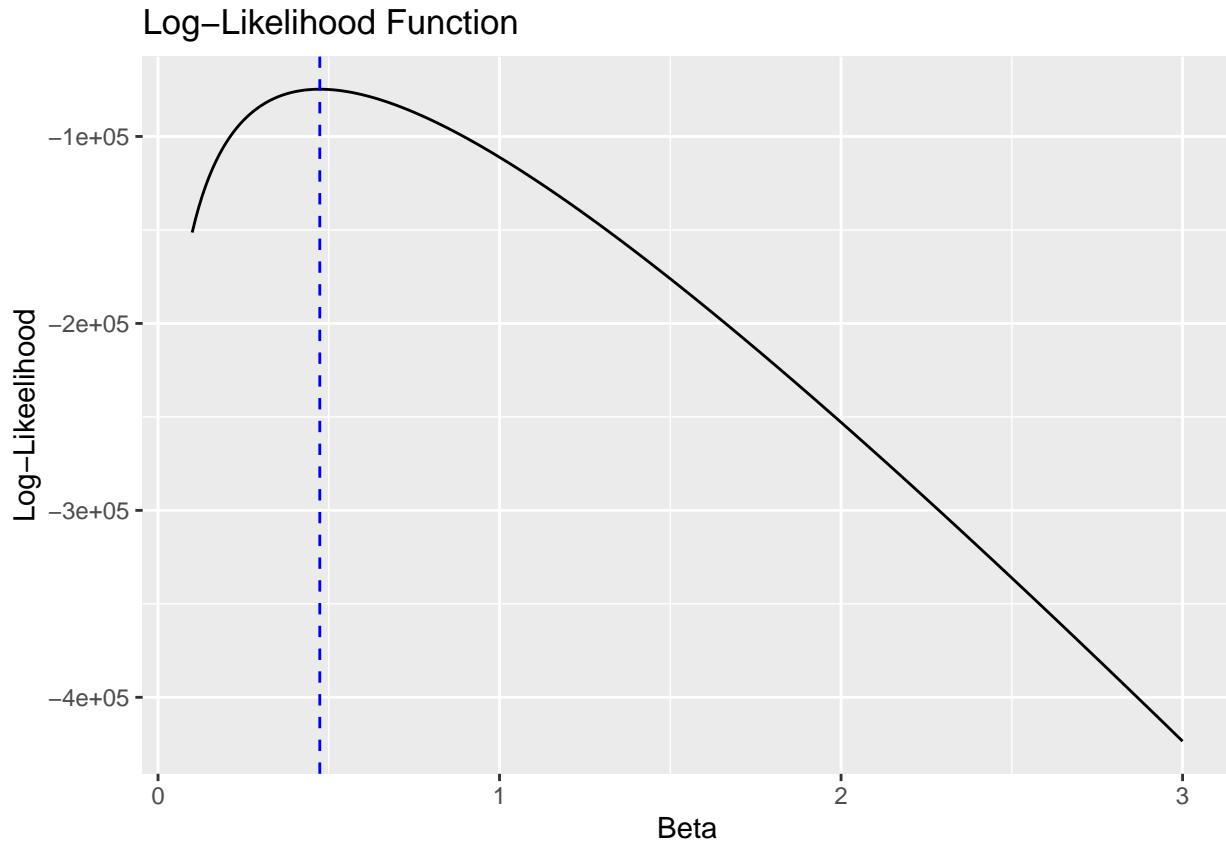
#compute log likelihood for each beta value
log_likelihood_values <- sapply(different_betas,function(beta) log_likelihood_function(beta,data))

#plot with ggplot2
library(ggplot2)
```

```

plotted_data <- data.frame(Beta=diferent_betas,LogLikelihood=log_likelihood_values)
ggplot(data=plotted_data,aes(x=Beta,y=LogLikelihood))+
  geom_line()+
  geom_vline(xintercept=MLE,color='blue',linetype='dashed')+
  xlab('Beta')+
  ylab('Log-Likeelihood')+
  ggtitle('Log-Likelihood Function')

```



Then we have to assign values to beta. In our case a range of values from 0.1 to 3:

```
diferent_betas<-seq(0.1,3,by=0.01)
```

Next, we must compute the log-likelihood for each beta value:

```
log_likelihood_values <- sapply(diferent_betas,function(beta) log_likelihood_function(beta,data))
```

Finally, we do the plot:

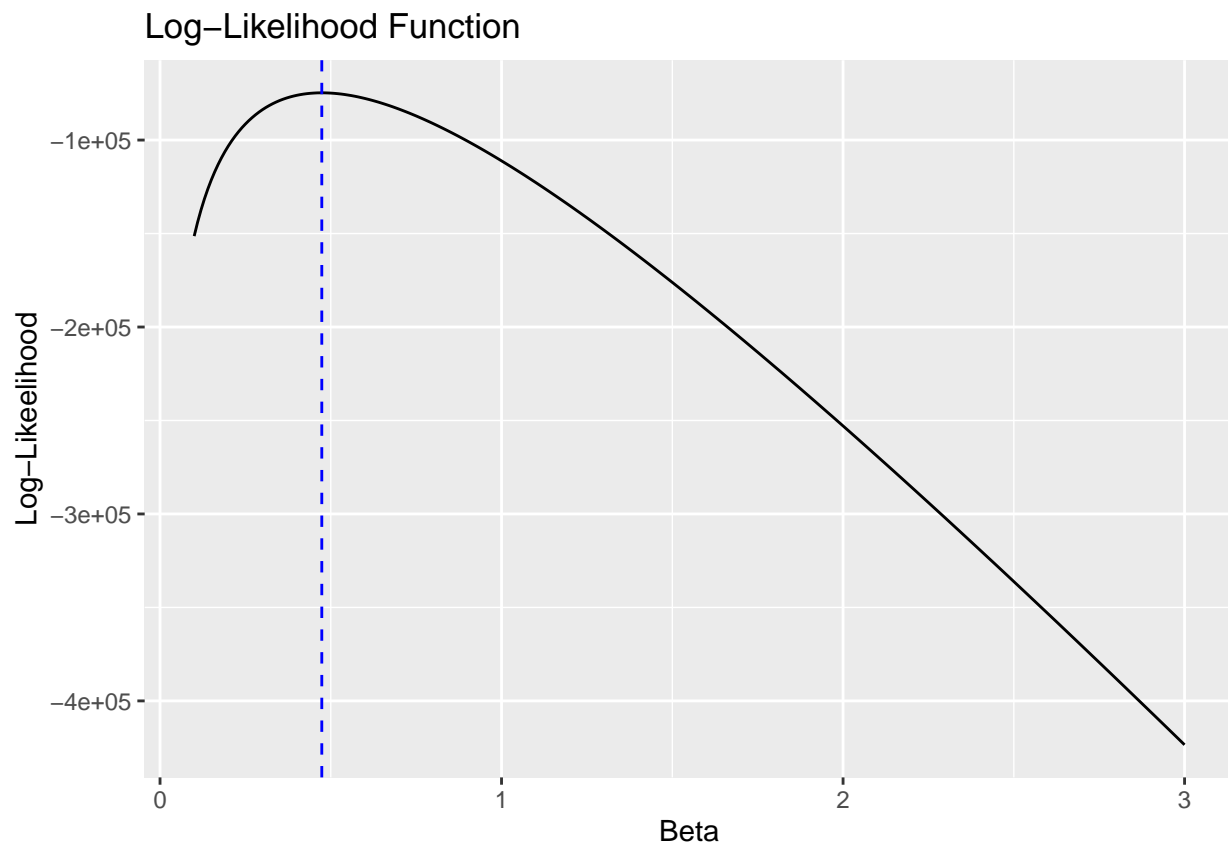
```

library(ggplot2)

plotted_data <- data.frame(Beta=diferent_betas,LogLikelihood=log_likelihood_values)
ggplot(data=plotted_data,aes(x=Beta,y=LogLikelihood))+
  geom_line()+
  geom_vline(xintercept=MLE,color='blue',linetype='dashed')+
  xlab('Beta')+

```

```
ylab('Log-Likeelihood')+
ggtitle('Log-Likelihood Function')
```



We can observe, in the plot, that the MLE previous computed coincides with the one of the graph.

g) Determine an expression for the Fisher information by calculating:

We have computed the formula the Fischer Information manually in the pdf file.

Here we compute the Fischer Information:

```
Fischer_information <- (10000/MLE^2)
```

The Fischer Information is 44569.52

h) Use the Fisher information for obtaining an expression for the variance of the maximum likelihood estimator $\hat{\theta}_{ML}$:

Since the variance of the maximum likelihood estimator is the inverse of the Fischer information we can compute it, as follows:

```
Variance <- 1/Fischer_information
```

The Variance of maximum likelihood estimator is 2.243686e-05.

j) Calculate a 95% confidence interval for parameter θ , using the dataset that you have downloaded:

```

a <- 0.05 #Alpha.
z <- qnorm(1 - a/2) #Computing the Z value of the normal distribution being alpha symmetrical.

lower <- MLE - z * (sqrt(1 /Fischer_information))
upper <- MLE +z* (sqrt(1/Fischer_information))

```

The CI is [0.4643917-0.4829595].

k) Do you think it is tenable that $\theta = 1$? Justify your answer:

Since $\theta = 1$ is outside of this confidence interval, it is not considered a tenable value for the parameter θ at a 95% confidence level. This means that, based on the data and the model assumptions, we have evidence to suggest that θ is different from 1.

l) Make a histogram of the data, using function `hist`, using the argument `freq=FALSE`. Overplot the histogram with the estimated probability density $f(x|)$, using the maximum likelihood estimate. What do you observe?

```

#First of all we create an histogram.
hist(data,freq=F,xlab='x',main='Histogram with estimated Probability Density Function')

#Here we define the function of the probability density function (PDF).
PDF <- function(x, beta) {
  beta * exp(-beta * x + 1)
}

#Here we overplot the curve of the PDF using the MLE to the histogram.
curve(PDF(x, MLE), col = "blue",add=T)

```

Histogram with estimated Probability Density Function

