**Worked out examples**

1. Suppose 10 rats are used in a biomedical study. These rats are injected with cancer cells and given a cancer drug that is designed to increase their survival rate. The survival times, in months, are shown below.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **x1** | **x2** | **x3** | **x4** | **x5** | **x6** | **x7** | **x8** | **x9** | **x10** |
| 14 | 17 | 27 | 18 | 12 | 8 | 22 | 13 | 19 | 12 |

Assume that the population from which the sample came from can be characterized by an exponential PDF so that



* 1. Find the maximum likelihood estimate for β and express the maximum likelihood estimate in terms of the xi’s.

The likelihood function is

.

The log likelihood function is

.

The first derivative with respect to β is

.

Setting this quantity equal to 0 and solving for β produces the MLE:

.

Thus, the maximum likelihood estimator for β is the sample mean!

* 1. Find the numerical value of the MLE.



* 1. What is the variance of the MLE?

The second derivative of the log likelihood function with respect to β is



Note that E(Xi) = β. The expected value of the second derivative of the log likelihood function is



Therefore, .

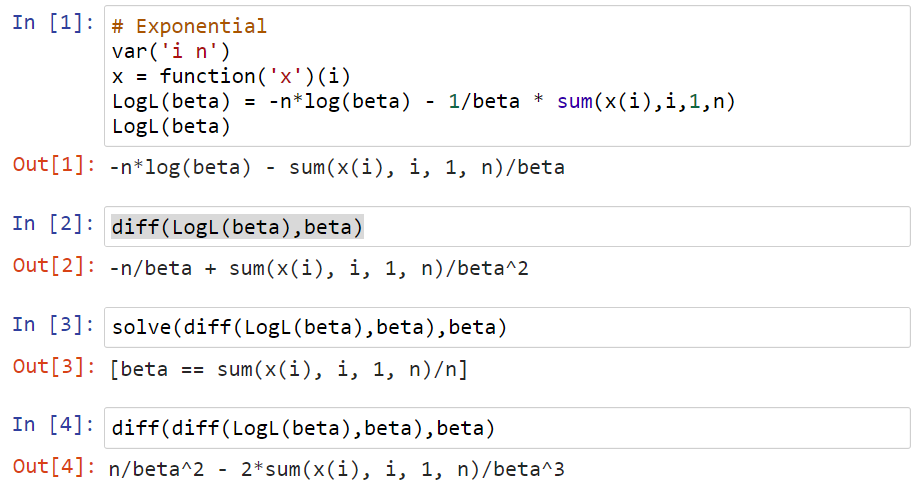
* 1. What is the estimated variance of the MLE?

.

* 1. What is the 95% Wald CI for β?

= (6.16, 26.24).

Sage:



R:

> x <- c(14, 17, 27, 18, 12, 8, 22, 13, 19, 12)

> sum(x)

[1] 162

> beta.hat <- mean(x)

> Var.beta.hat <- beta.hat^2 / length(x)

> beta.hat

[1] 16.2

> Var.beta.hat

[1] 26.244

> alpha <- 0.05

> lower <- beta.hat - qnorm(p = 1 - alpha/2, mean = 0, sd = 1) \* sqrt(Var.beta.hat)

> upper <- beta.hat + qnorm(p = 1 - alpha/2, mean = 0, sd = 1) \* sqrt(Var.beta.hat)

> data.frame(lower, upper)

lower upper

1 6.15932 26.24068

1. Suppose a random sample of size n = 50 is taken from a population which can be characterized by a one parameter beta PDF (see the beginning of the course notes for this section). The sample results in the observed values shown in beta\_sample.csv. Thus, y1 = 0.2805, …, y50 = 0.0285. Compete the following problems below.
   1. Find the MLE for β.

The likelihood function is



The log likelihood function is



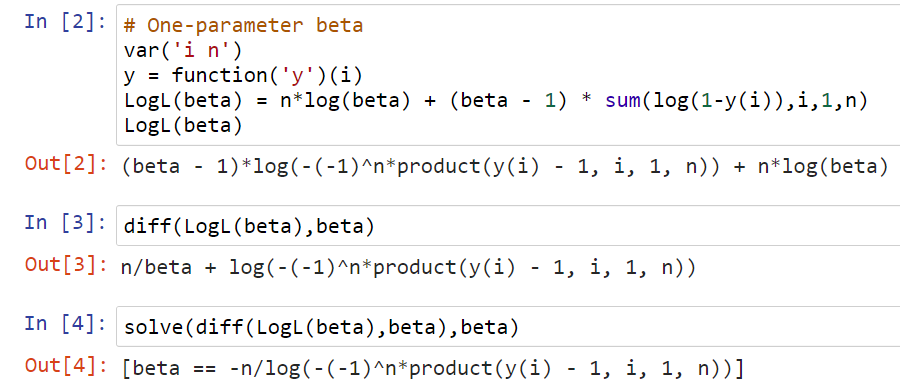
The first derivative with respect to β is



Setting this quantity equal to 0 and solving for β produces the MLE:



From Sage,



The expression given by Sage can be simplified further to result in . For example, the -(1)n part does not play a role once the log is distributed through the denominator.

Remember that 0 < β < ∞ for this beta PDF. To be very careful, we can examine a plot of the log likelihood function in part c) to show that the MLE does not occur when β = 0 or ∞ for the observed data in this problem.

Using the observed values results in .

> set1 <- read.csv("beta\_sample.csv")

> head(set1)

Observation.number y

1 1 0.2805

2 2 0.0593

3 3 0.0701

4 4 0.0135

5 5 0.1096

6 6 0.0543

> # MLE

> beta.hat <- -length(set1$y)/sum(log(1-set1$y))

> beta.hat

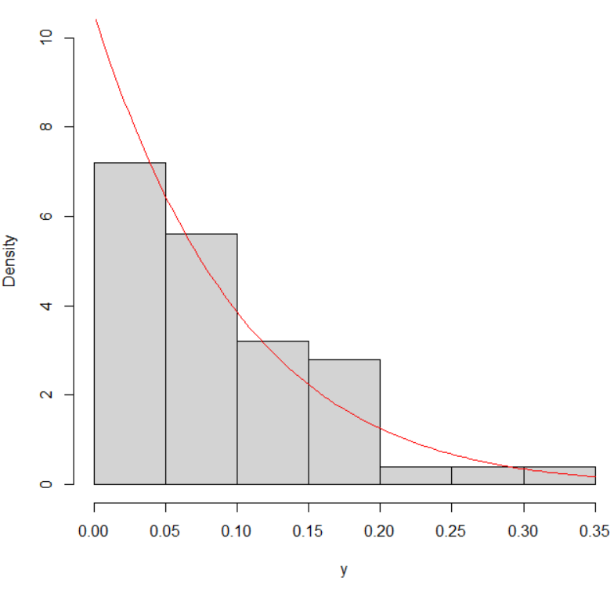
[1] 10.52628

Note that I used β = 10 here to actually sample the data. Thus, we would expect the sample to display similar characteristics to the PDF. As you can see, the MLE is approximately the true value of β. This shows the maximum likelihood estimation procedure works for this example!

* 1. Plot a histogram of the observed data and compare it to a PDF using  as β.

> hist(x = set1$y, freq = FALSE, xlab = "y", ylim = c(0,10), main = "")

> curve(expr = beta.hat \* (1 - x)^(beta.hat - 1), col = "red", add = TRUE)



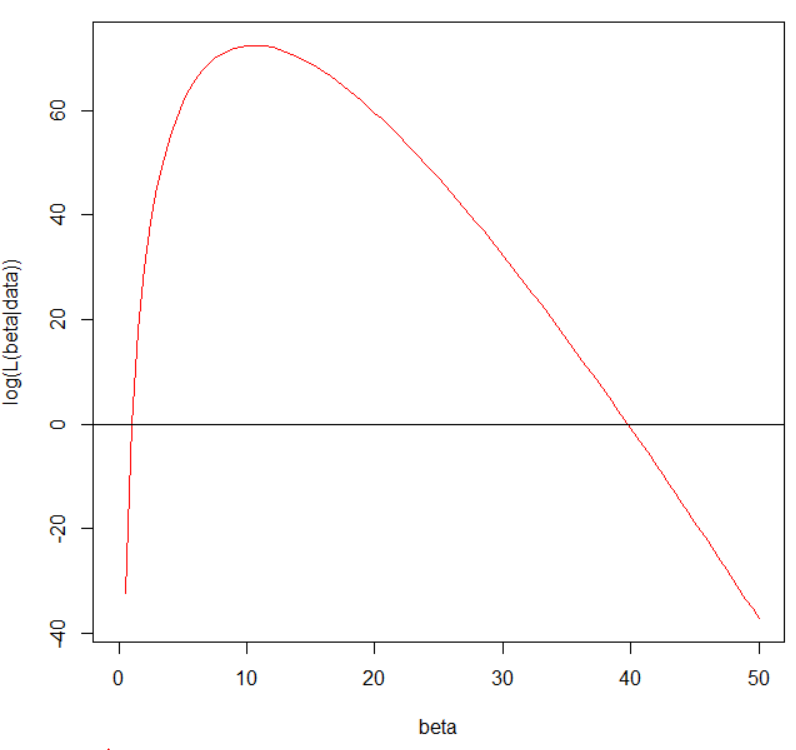
As discussed many times already in this course, we would expect the histogram to be similar in shape to the PDF!

* 1. Plot the log likelihood function using the observed values and inspect where the maximum value occurs.

> curve(expr = 50\*log(x) + (x - 1)\*sum(log(1-set1$y)), col = "red", xlim =

c(0,50), xlab = "beta", ylab = "log(L(beta|data))")

> abline(h = 0)



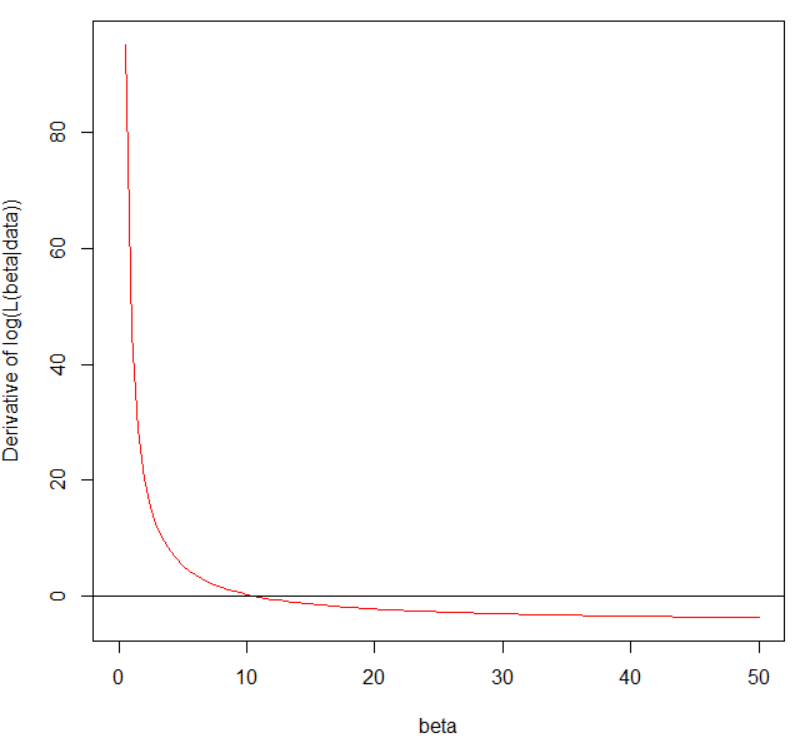
We can see the maximum does occur when β = 10.53.

* 1. Plot the derivative of the log likelihood function and discuss how the MLE value can be found from it.

> curve(expr = 50/x + sum(log(1-set1$y)), col = "red", xlim = c(0,50),

xlab = "beta", ylab = "Derivative of log(L(beta|data))")

> abline(h = 0)



Because the function crosses the x-axis at 10.53, this is the root and the MLE for β.