Homework 7

Joanna Yu (Fall 2018 Tuesday 4pm)

10/23/2018

The Meat

1a) Do you expect the distribution of this measure (ground beef consumption per capita per month) to be approximately normal? Why or why not?
No, I do not expect the distribution of this to be normal because vegetarians consume zero meat. That will cause a spike at value 0. In a normal distribution, the tails should taper out.

1b) Do you expect the distribution of the sample mean to be approximately normal? Why or why not.
Yes, I expect the distribution of the sample mean to be approximately normal. Since the sample size is sufficiently large and I do not expect the population to be very skewed, the distribution of the sample mean should be approximately normal.

1c) What is the 95% confidence interval for Berkeley students?

$$(\bar{x} - 1.96 * \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 * \frac{\sigma}{\sqrt{n}})$$
$$(2.45 - 1.96 * \frac{2}{\sqrt{100}}, 2.45 + 1.96 * \frac{2}{\sqrt{100}})$$
$$(2.058, 2.842)$$

GRE Scores

2. What is the real confidence level for the interval we have made, if the sample size is 10? What if the sample is 200?

If n=10, df=9

```
1-2*pt(-1.96, df=9)
```

```
## [1] 0.9183556
```

The real confidence level is 91.84% when the sample size is 10.

If n=200, df=199

```
1-2*pt(-1.96, df=199)
```

```
## [1] 0.9486082
```

The real confidence level is 94.86% when the sample size is 200.

Maximum Likelyhood Estimation for an Exponential Distribution

3a) Write down the likelihood function, $L(\lambda)$.

$$\begin{aligned} L(\lambda) &= f(x_1, \ldots, x_n; \lambda) \\ &= \prod_{i=1}^{n} f(x_i; \lambda) \\ &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \end{aligned}$$

3b) Write down the log of the likelihood, and simplify it.

$$\sum_{i=1}^{n} (log(\lambda) - \lambda x_i)$$

3c) Take the derivative of the log of likelihood, set it equal to zero, and solve for λ. How is it related to the mean time between arrivals?

$$\sum_{i=1}^{n} \frac{1}{\lambda} - x_i = 0$$

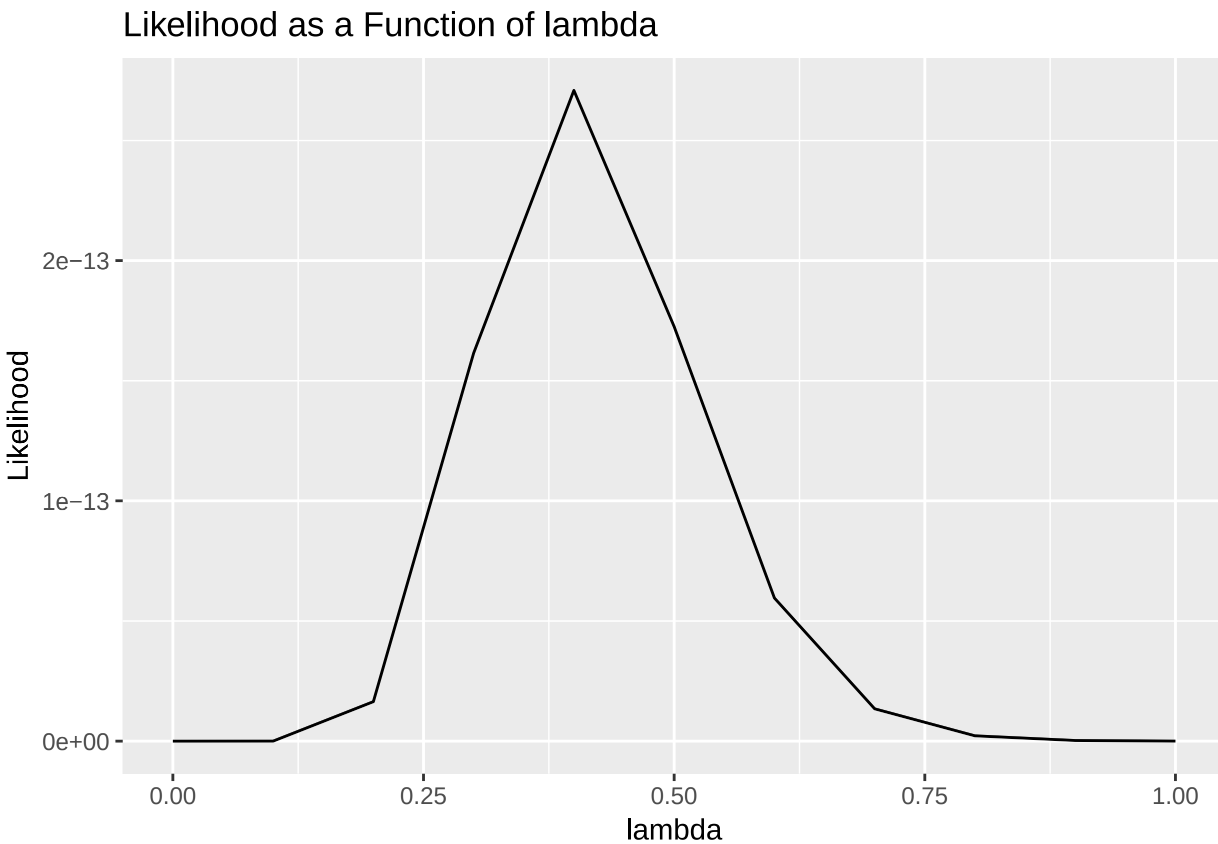
$$\frac{1}{\lambda} = \sum_{i=1}^{n} x_i$$

$$\lambda = \frac{1}{\sum_{i=1}^{n} x_i}$$

3d) Use R to plot the likelihood function. Then use optimize to approximate the maximum likelihood estimate for λ. How does your answer compare to your solution from part c?

```
times = c(2.65871285, 8.34273228, 5.09845548, 7.15064545,
0.39974647, 0.77206050, 5.43415199, 0.36422211,
3.30789126, 0.07621921, 2.13375997, 0.06577856,
1.73557740, 0.16524304, 0.27652044)
lambda_likelihood <- function(lambda) {
 return(lambda^length(times) * exp(-lambda * sum(times)))
}
lambda <- seq(0, 1, by = 0.1)

library(ggplot2)
qplot(lambda,
      sapply(lambda, function(lambda) {lambda_likelihood(lambda)}),
      geom = 'line',
      main = 'Likelihood as a Function of lambda',
      xlab = 'lambda',
      ylab = 'Likelihood')
```




```r
optimize(lambda_likelihood, interval = lambda, maximum = TRUE)
```

```
## $maximum
## [1] 0.3949072
##
## $objective
## [1] 2.712269e-13
```

```r
1/mean(times)
```

```
## [1] 0.3949269
```

Using optimize in R yields 0.3949072. The answer I have from c is 0.3949269.