

# Estimation Methods Application Solutions

Lt Col Ken Horton

Lt Col Pruitt

Professor Bradley Warner

14 October, 2020

## Exercises

1. In the Notes, we found that if we take a sample from the uniform distribution  $\text{Unif}(0, \theta)$ , the method of moments estimate of  $\theta$  is  $\hat{\theta}_{MoM} = 2\bar{x}$ . Suppose our sample consists of the following values:

0.2 0.9 1.9 2.2 4.7 5.1

a) What is  $\hat{\theta}_{MoM}$  for this sample?

```
x<-c(0.2,0.9,1.9,2.2,4.7,5.1)
thetamom<-2*mean(x)
thetamom
```

```
## [1] 5
```

b) What is wrong with this estimate?

For our distribution  $\theta$  is the upper bound or largest value for the random variable. The estimate for  $\theta$  is 5, which is an impossible value for  $\theta$ , since one of our observations (5.1) is beyond this value.

c) Show that this estimator is unbiased.

We need to show that

$$E(\hat{\theta}_{MoM}) = \theta$$

We proceed as follows

$$\begin{aligned} E(2\bar{X}) &= 2E\left(\sum \frac{X_i}{n}\right) = \frac{2}{n}E\left(\sum X_i\right) \\ &= \frac{2}{n} \sum E(X_i) = \frac{2}{n} \sum \frac{\theta}{2} = \frac{n\theta}{n} = \theta \end{aligned}$$

Notice that in performing this derivation, we treated  $X$  as a random variable and not as  $x$ , a data value.

- d) ADVANCED: Use simulation in R to find out how often the method of moment estimator is less the maximum observed value, ( $\hat{\theta}_{MoM} < \max x$ ). Report an answer for various sizes of samples. You can just pick an arbitrary value for  $\theta$  when you sample from the uniform. However, the minimum must be 0.

Let's start by writing code for one sample size and then generalize. This function will take as input the number of data points, sample that many points from a uniform with a max of 5, and then return a logical value comparing the method of moments estimate to the observed maximum. The choice of a maximum value for the uniform distribution is arbitrary.

```
check <- function(n=10){  
  temp<-runif(n,max=5)  
  2*mean(temp)<max(temp)  
}
```

Let's test the function.

```
set.seed(3030)  
temp<-runif(10,max=5)  
temp
```

```
## [1] 1.048619 2.370056 0.675677 2.434443 2.994015 3.930865 4.708367 1.645806  
## [9] 2.592365 3.664424
```

```
2*mean(temp)
```

```
## [1] 5.212927
```

Reset the seed and run the function, we should get FALSE.

```
set.seed(3030)  
check(10)
```

```
## [1] FALSE
```

Now let's repeat the test 10000 find the proportion of times the method of moments estimator is unrealistic.

```
(do(10000)*check(10)) %>%  
  summarize(mean(check)) %>%  
  pull()
```

```
## [1] 0.285
```

Let's make check a vectorized function for we can run for many sample sizes.

```
check <- Vectorize(check)
```

Run 1000 replicates for each sample size. The rest of the code gets my data frame in the proper shape.

```
my_data<-(do(1000)*check(seq(10,200,5))) %>%
  summarise_all(mean) %>%
  pivot_longer(everything(),names_to = "Sample",values_to = "Percent") %>%
  mutate(Sample=seq(10,200,5))
```

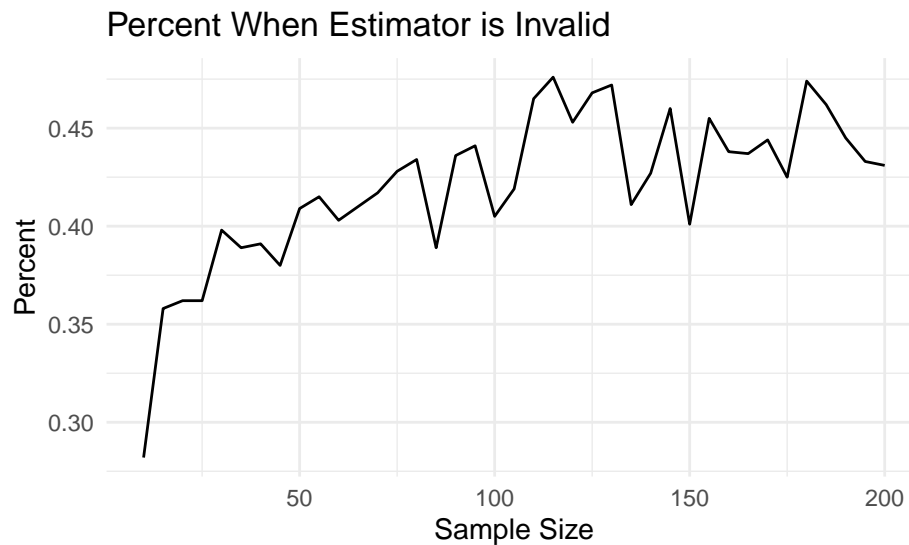
A quick look at the data.

```
head(my_data)
```

```
## # A tibble: 6 x 2
##   Sample Percent
##   <dbl>   <dbl>
## 1     10   0.282
## 2     15   0.358
## 3     20   0.362
## 4     25   0.362
## 5     30   0.398
## 6     35   0.389
```

Now we can plot the results of sample size versus

```
my_data %>%
  gf_line(Percent~Sample,xlab="Sample Size",title="Percent When Estimator is Invalid") %>%
  gf_theme(theme = theme_minimal())
```



Here is more traditional, old school R code.

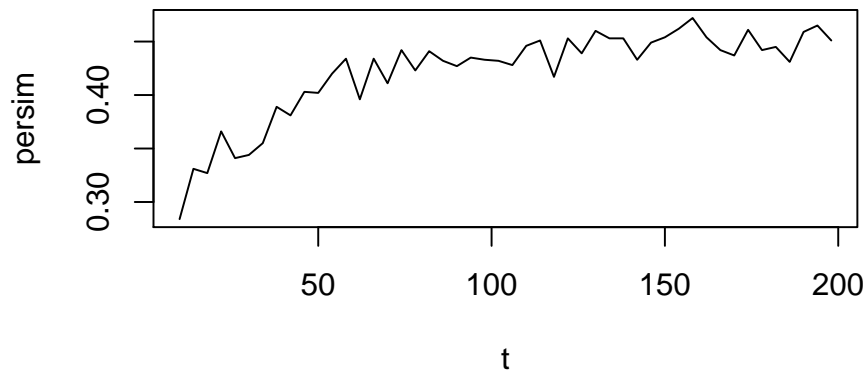
```
simn<-function(n){
  y<-replicate(1000,{
    x<-runif(n)
    (2*mean(x))<max(x)
  })
}
```

```

    mean(y)
}

t<-seq(10,200,4)
persim<-sapply(t,simn)
plot(t,persim,type="l")

```



2. Let  $x_1, x_2, \dots, x_n$  be a simple random sample from an exponentially distributed population with parameter  $\lambda$ . Find  $\hat{\lambda}_{MoM}$ .

Recall that  $E(X) = \frac{1}{\lambda}$ . Setting this equal to the sample moment  $\bar{x}$  and solving for  $\lambda$  yields the method of moment estimator. Thus,

$$\hat{\lambda}_{MoM} = \frac{1}{\bar{x}}$$

3. Let  $x_1, x_2, \dots, x_n$  be an iid random sample from an exponentially distributed population with parameter  $\lambda$ . Find  $\hat{\lambda}_{MLE}$ .

Recall that

$$f_X(x; \lambda) = \lambda e^{-\lambda x}$$

So the likelihood function is:

$$L(\lambda; \mathbf{x}) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum x_i}$$

And the log-likelihood function is:

$$l(\lambda; \mathbf{x}) = n \log \lambda - \lambda \sum x_i$$

Taking the derivative with respect to  $\lambda$  and setting equal to 0:

$$\frac{dl(\lambda; \mathbf{x})}{d\lambda} = \frac{n}{\lambda} - \sum x_i = 0$$

Note that  $\frac{d^2 l(\lambda; \mathbf{x})}{d\lambda^2} = -\frac{n}{\lambda^2} < 0$ , so  $l$  is always concave down. Thus, any optimum found is a maximum.

So,

$$\hat{\lambda}_{MLE} = \frac{n}{\sum x_i} = \frac{1}{\bar{x}}$$

4. It is mathematically difficult to determine if the estimators found in questions 2 and 3 are unbiased. Since the sample mean is in the denominator; mathematically we may have to work with the joint pdf. So instead, use simulation to get an sense of whether the method of moments estimator for the exponential distribution is unbiased.

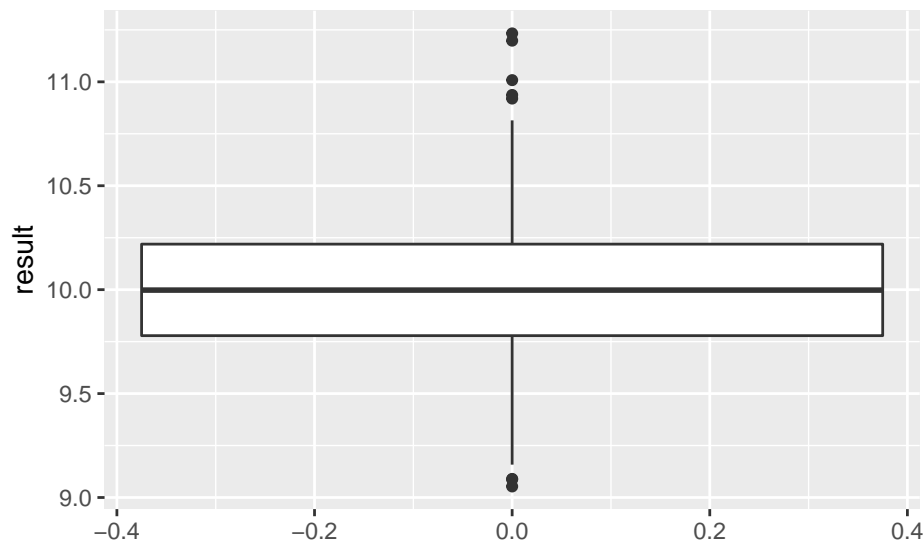
We need to sample data from an exponential and then compare the the reciprocal of the mean to the parameter.

```
set.seed(630)
1/mean(rexp(1000,rate=10))
```

```
## [1] 10.09422
```

This is close, maybe we just got lucky. Let's repeat many times.

```
(do(1000)*(1/mean(rexp(1000,rate=10)))) %>%
  gf_boxplot(~result)
```



Looks like it has the potential to be unbiased. We would need to investigate other values for  $\lambda$ .

5. Find a maximum likelihood estimator for  $\theta$  when  $X \sim \text{Unif}(0, \theta)$ . Compare this to the method of moments estimator we found. Hint: Do not take the derivative of the likelihood function.

$$L(\theta; \mathbf{x}) = \frac{1}{\theta^n}, \quad \text{only if all } x_i \leq \theta$$

A better way to write this is so as to see the answer is to let  $M = \max(x_i)$ , then:

$$L(\theta; \mathbf{x}) = \begin{cases} 0, & \theta < M = \max(x_i) \\ \frac{1}{\theta^n}, & \theta \geq M = \max(x_i) \end{cases}$$

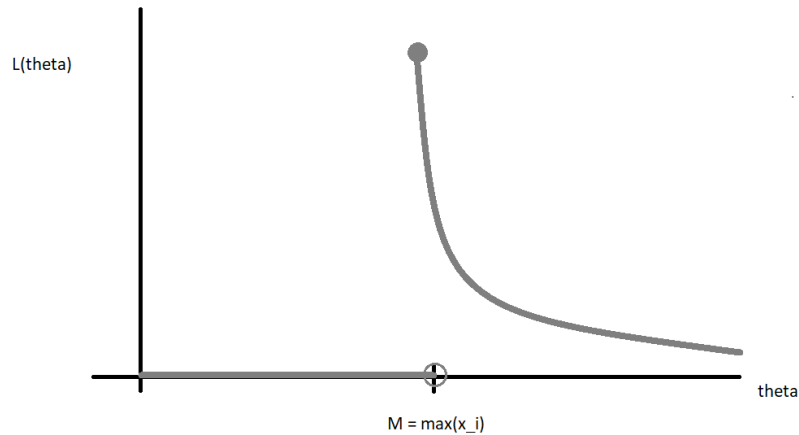


Figure 1: A graph of the likelihood function

Figure 1 is a plot of this likelihood function and you can see that the maximum occurs at the maximum observed data point.

Recall that  $f(x_i; \theta) = \frac{1}{\theta}$  if  $x_i \in [0, \theta]$  and 0 elsewhere. And since the likelihood function is simply the product of these pdfs, if any  $x_i$  is beyond  $\theta$ , then the entire likelihood function is 0.

You can picture  $L$  as a decreasing function of  $\theta$ , but remembering that  $L$  takes the value 0 if  $\theta$  is smaller than at least one  $x_i$ . Thus,  $L$  achieves its maximum at  $\theta = \max x_i$ .

This estimate is more intuitive than the method of moments estimate ( $2\bar{x}$ ). The method of moments estimate is sometimes not feasible. Meanwhile, the MLE ( $\max x_i$ ) is always feasible.

## File Creation Information

- File creation date: 2020-10-14
- Windows version: Windows 10 x64 (build 18362)
- R version 3.6.3 (2020-02-29)
- **mosaic** package version: 1.7.0
- **tidyverse** package version: 1.3.0
- **fastR2** package version: 1.2.1