ST 502 Homework 3 Problem 8.43

Eric Warren

2024-02-06

Contents

1 Background **1**

2 Read in the data **1**

3 Part A **2**

4 Part B **3**

4.1 Method of Moments 3

4.2 Maximum Likelihood Estimation 4

5 Part C **4**

6 Part D **6**

6.1 Method of Moments 6

6.2 Maximum Likelihood Estimation 7

6.3 Comparison 7

1 Background

The file `gamma-arrivals` contains another set of gamma-ray data, this one consisting of the times between arrivals (interarrival times) of 3,935 photons (units are seconds).

2 Read in the data

Here we are going to read in our data for analysis purposes.

```
library(tidyverse)
library(MASS)

(df <- read_csv("gamma-arrivals.csv", col_names = "times"))
```

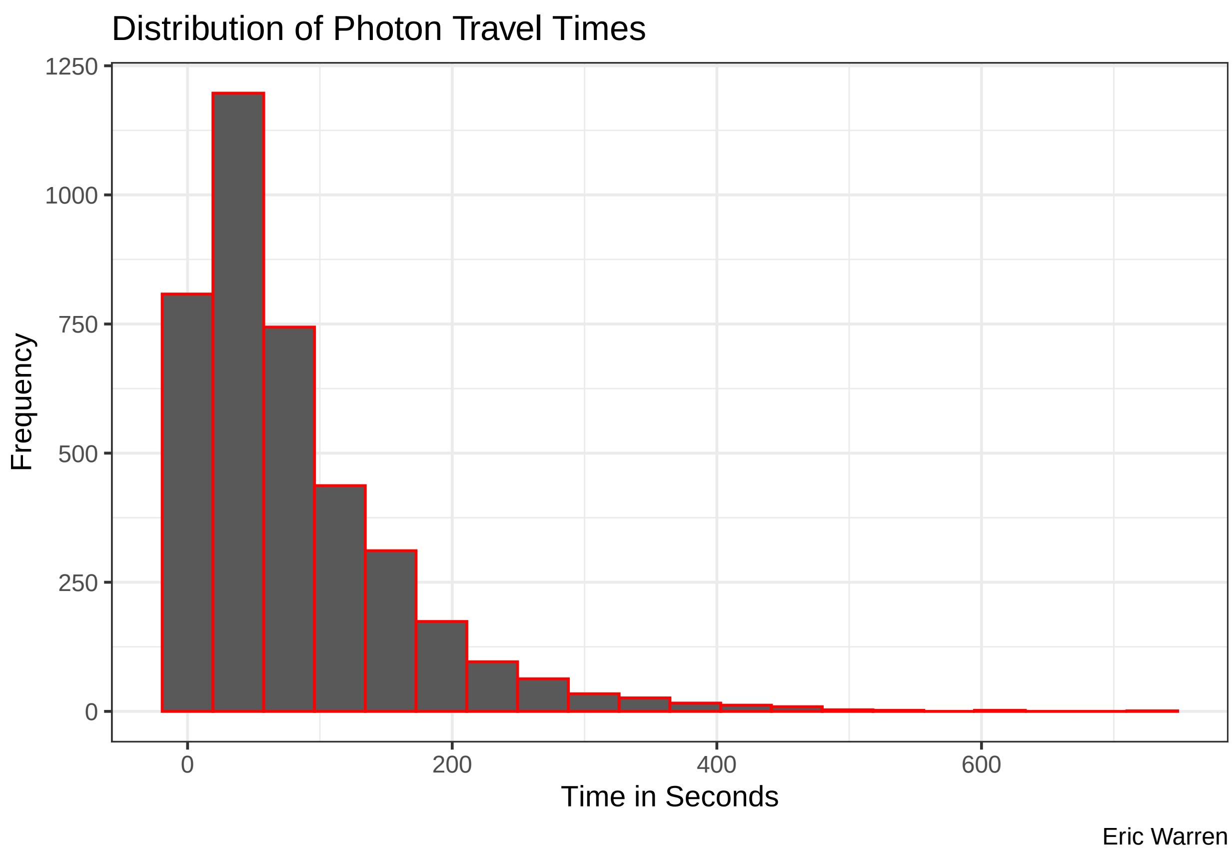
```
## # A tibble: 3,935 x 1
##    times
##    <dbl>
##  1  38.0
##  2 409.
##  3 272.
##  4  98.7
##  5  87.9
##  6  79.8
##  7 211.
##  8  99.3
##  9 181.
## 10 104.
## # i 3,925 more rows
```

3 Part A

Make a histogram of the interarrival times. Does it appear that a gamma distribution would be a plausible model?

First we are going to make our histogram.

```r
df %>%
  ggplot(aes(x = times)) +
  geom_histogram(color = "red",
                 bins = 20) +
  labs(title = "Distribution of Photon Travel Times",
       x = "Time in Seconds",
       y = "Frequency",
       caption = "Eric Warren") +
  theme_bw()
```



This "skewed right" distribution definitely looks like something that could be a gamma distribution. We have seen something like this before of pictures in the textbook of a shape similar to this. We will investigate more in-depth if that is an appropriate distribution.

4 Part B

Fit the parameters by the method of moments and by maximum likelihood. How do the estimates compare?

4.1 Method of Moments

First let us take a look at Method of Moments. We know from our class lecture that the Method of Moment Estimators are $\hat{\alpha}_{MOM} = \frac{\bar{Y}^2}{S^2}$ and $\hat{\lambda}_{MOM} = \frac{\bar{Y}}{S^2}$. We can use that to get the estimates for our gamma distribution below.

```
# Find ybar (sample mean)
ybar <- mean(df$times)

# Find s2 (biased version of sample variance)
s2 <- mean(df$times^2) - ybar^2

# Now get alpha_hat = ybar^2 / s2
alpha_hat <- ybar^2 / s2

# Now get lambda_hat = ybar / s2
```

```r
lambda_hat <- ybar / s2

# Show the parameter estimates together
c(alpha_hat, lambda_hat)
```

```
## [1] 1.01235221 0.01266466
```

Through Method of Moments Estimation we can see that our estimators are $\hat{\alpha}_{MOM} = 1.0123522$ and $\hat{\lambda}_{MOM} = 0.0126647$.

4.2 Maximum Likelihood Estimation

Now let us take a look at Maximum Likelihood Estimation. We can find our Maximum Likelihood Estimators $\hat{\alpha}_{MLE}$ and $\hat{\lambda}_{MLE}$ by using the `MASS::fitdistr()` function in R. We can use that to get the estimates our shape and rate for our gamma distribution below.

```r
fitdistr(df$times, "gamma")[[1]]
```

```
##      shape       rate
## 1.02633211 0.01283955
```

```r
# Save the shape (alpha) estimate
alpha_mle <- fitdistr(df$times, "gamma")[[1]][[1]]

# Save the rate (lambda) estimate
lambda_mle <- fitdistr(df$times, "gamma")[[1]][[2]]
```

As we can see our shape also known as $\hat{\alpha}_{MLE}$ is equal to 1.0263321 and our rate also known as $\hat{\lambda}_{MLE}$ is equal to 0.0128396.

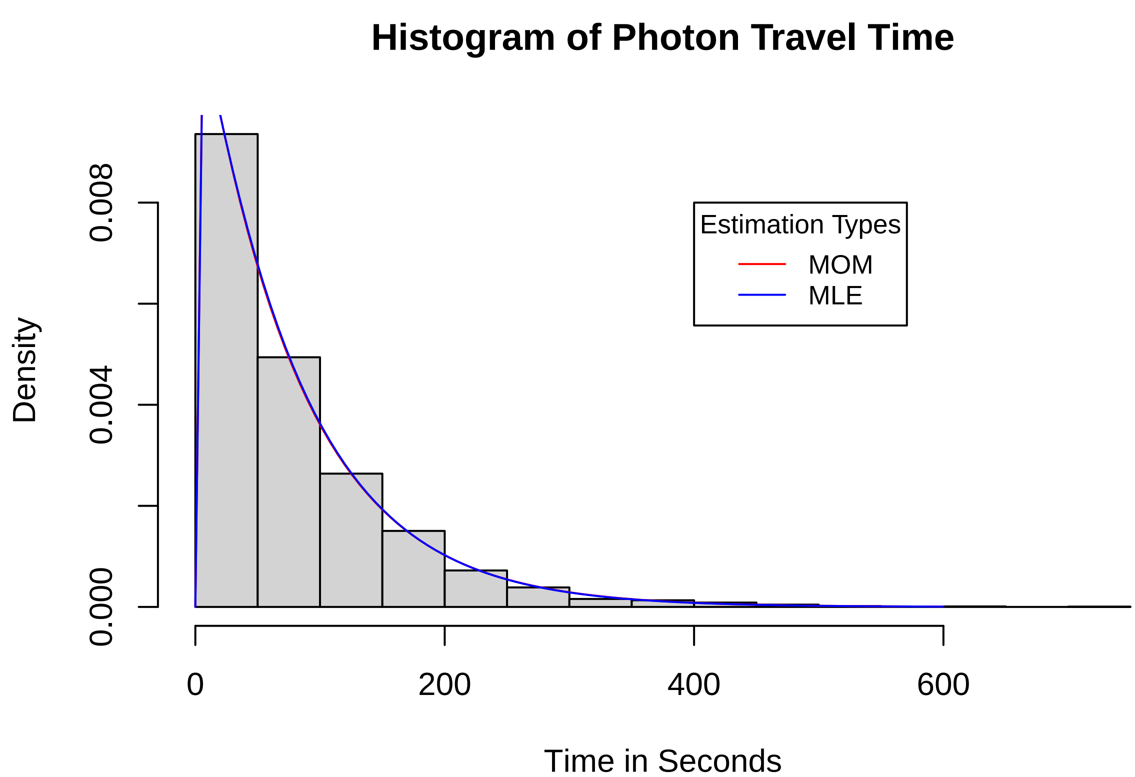
5 Part C

Plot the two fitted gamma densities on top of the histogram. Do the fits look reasonable?

Let us take a look to see how our estimates (both Method of Moments and Maximum Likelihood Estimation) are looking compared to our histogram from before.

```r
#Show histogram
hist(df$times,
     xlab = "Time in Seconds",
     ylab = "Density",
     main = "Histogram of Photon Travel Time",
     freq = FALSE
     )
# Show MOM estimation
curve(dgamma(x,
             shape = alpha_hat,
             rate = lambda_hat
             ),
```

```
      from = 0,
      to = 600,
      add = TRUE,
      col = "red"
      )
# Show MLE estimation
curve(dgamma(x,
             shape = alpha_mle,
             rate = lambda_mle
             ),
      from = 0,
      to = 600,
      add = TRUE,
      col = "blue"
      )
# Make legend
legend(400, 0.008,
       legend=c("MOM", "MLE"),
       col=c("red", "blue"),
       lty = 1, cex = 0.8,
       title = "Estimation Types"
       )
```



After plotting our Method of Moments and Maximum Likelihood Estimation estimated parameters for a gamma distribution on top of our histogram of data, we can see that the fit for both is pretty good. While both estimated distributions extend much higher vertically at their point, it seems the data mirrors the distribution pretty well. Note how the Method of Moment (MOM on the graph legend) and Maximum Likelihood Estimation (MLE on the graph legend) both provide very similar estimators for this gamma distribution which is why it might be hard to see the difference between them. All in all, both seem like good fits for our data.

6 Part D

For both maximum likelihood and the method of moments, use the bootstrap to estimate the standard errors of the parameter estimates. How do the estimated standard errors of the two methods compare?

Please note for both maximum likelihood and the method of moments we are going to use the bootstrap with a $N = 10000$.

6.1 Method of Moments

First we will use the bootstrap to find the estimated standard errors of the estimated parameters from the Method of Moments estimation.

```
# Put down our estimates from MOM
alpha0 <- alpha_hat
lambda0 <- lambda_hat

# Number of samples to create
N <- 10000

# Sample size
n <- length(df$times)

# Get our estimates
set.seed(999)
estimates <- replicate(N, {
  sim_data <- rgamma(n, shape = alpha0, rate = lambda0)
  ybar <- mean(sim_data)
  s2 <- mean(sim_data^2) - ybar^2
  alpha_hat_sim <- ybar^2 / s2
  lambda_hat_sim <- ybar / s2
  return(c("alpha_hat" = alpha_hat_sim, "lambda_hat" = lambda_hat_sim))
})

# SE(alpha_hat)
se_mom_alpha_hat <- sd(estimates[1, ])

# SE(lambda_hat)
se_mom_lambda_hat <- sd(estimates[2, ])

# Show them together
c(se_mom_alpha_hat, se_mom_lambda_hat)
```

```
## [1] 0.0319333512 0.0004482477
```

Using Method of Moments Estimation we can see that our standard error of $\hat{\alpha}_{MOM}$ is equal to 0.0319334 and our standard error of $\hat{\lambda}_{MOM}$ is equal to 0.0004482.

6.2 Maximum Likelihood Estimation

First we will use the bootstrap to find the estimated standard errors of the estimated parameters from the Maximum Likelihood Estimation.

```
# Put down our estimates from MLE
alpha0 <- alpha_mle
lambda0 <- lambda_mle

# Number of samples to create
N <- 10000

# Sample size
n <- length(df$times)

# Get our estimates
set.seed(999)
estimates <- replicate(N, {
  sim_data <- rgamma(n, shape = alpha0, rate = lambda0)
  return(fitdistr(sim_data, "gamma")[[1]])
})

# SE(alpha_hat)
se_mle_alpha_hat <- sd(estimates[1, ])

# SE(lambda_hat)
se_mle_lambda_hat <- sd(estimates[2, ])

# Show them together
c(se_mle_alpha_hat, se_mle_lambda_hat)
```

```
## [1] 0.0204185408 0.0003263877
```

Using Maximum Likelihood Estimation we can see that our standard error of $\hat{\alpha}_{MLE}$ is equal to 0.0204185 and our standard error of $\hat{\lambda}_{MLE}$ is equal to 0.0003264.

6.3 Comparison

Here we can compare both estimation methods via a table to look at their respective standard errors.

Estimation	$\hat{SE}(\hat{\alpha})$	$\hat{SE}(\hat{\lambda})$
Method of Moments	0.0319334	0.0004482
Maximum Likelihood	0.0204185	0.0003264

As we can see here, the standard errors for both $\hat{\alpha}$ and $\hat{\lambda}$ are lower when using Maximum Likelihood Estimation over Method of Moments Estimation. Therefore, we can say that Maximum Likelihood Estimation is better to use for estimating parameters if using this gamma distribution for this specific data. However we must decide if this precision by Maximum Likelihood Estimation outweighs its computational cost (as it takes longer to bootstrap values).