Final Project Part 1

Max Barshay

11/21/2020

Suppose I want to model the lifetimes (in years) of 10 batteries, $X_1, X_2, ..., X_{10}$, as i.i.d. random variables with p.d.f:

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & x > 0 \\ 0 & \text{otherwise} \end{cases}.$$

In order to model the lifetimes of the batteries we must generate an estimate for the unknown θ. One estimator is the MLE estimator which I will now calculate.

First, I will need to find the likelihood of θ.

Since the lifetimes of these batteries are independent, we find that:

$$L_X(\theta) = \frac{1}{\theta} e^{-X_1/\theta} \cdot \frac{1}{\theta} e^{-X_2/\theta} \cdot ... \cdot \frac{1}{\theta} e^{-X_{10}/\theta}$$

This simplifies to:

$$L_X(\theta) = \frac{1}{\theta^{10}} e^{-X_1/\theta - X_2/\theta - ... - X_{10}/\theta}$$

This further simplifies to:

$$L_X(\theta) = \frac{1}{\theta^{10}} e^{-\frac{1}{\theta} \sum_{i=1}^{10} X_i}$$

Now as we have done all quarter, lets take logs to make the math easier.

$$\ell_X(\theta) = log L_X(\theta) = -10 \log(\theta) - \frac{1}{\theta} \sum_{i=1}^{10} X_i$$

To make taking the derivative easier, I will rewrite this slightly.

$$\ell_X(\theta) = -10 \log(\theta) - (\sum_{i=1}^{10} X_i)\theta^{-1}$$

Now I will take the derivative of the log-likelihood.

$$\ell'_X(\theta) = -\frac{10}{\theta} + (\sum_{i=1}^{10} X_i)\theta^{-2}$$

I am going to rewrite the score in a slightly friendlier form.

$$\ell'_X(\theta) = -\frac{10}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{10} X_i$$

To find where the score crosses 0, or in other words where the log-likelihood is maximized, we set the score equal to 0 and solve for θ.

$$-\frac{10}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{10} X_i = 0$$

$$\frac{1}{\theta^2}\sum_{i=1}^{10} X_i = \frac{10}{\theta}$$

$$\frac{\theta}{\theta^2}\sum_{i=1}^{10} X_i = 10$$

$$\frac{1}{\theta}\sum_{i=1}^{10} X_i = 10$$

$$\sum_{i=1}^{10} X_i = 10\theta$$

We finally arrive at our MLE estimator of θ:

$$\hat{\theta} = \frac{\sum_{i=1}^{10} X_i}{10}$$

A natural question to ask is if $\hat{\theta}$ is the best unbiased estimator of θ.

By the Cramèr-Rao inequality, we know that variance of the best possible unbiased estimator is $\frac{1}{I(\theta)}$ where $I(\theta)$ is defined to be $Var[\ell'_X(\theta)]$. Note: There are other equivalent ways of calculting $I(\theta)$.

I will now calculate $I(\theta)$ to find the variance of the best possible unbiased estimator.

First recall:

$$\ell'_X(\theta) = -\frac{10}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{10} X_i$$

Now:

$$I(\theta) = Var[\ell'_X(\theta)] = Var[-\frac{10}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{10} X_i]$$

Recall that the score: $\ell'_X(\theta)$ is a random variable that depends on X and θ is just a contant. Using laws of variance we get:

$$Var[\ell'_X(\theta)] = Var[\frac{1}{\theta^2}\sum_{i=1}^{10} X_i]$$

We can pull the contant out, remembering to square it:

$$\frac{1}{\theta^4}Var[\sum_{i=1}^{10} X_i]$$

Since the $X_i's$ are independent random variables we can bring the variance inside of the sum:

$$\frac{1}{\theta^4}\sum_{i=1}^{10} Var[X_i]$$

Since the $X_i's$ are identically distributed random variables the variances of all of them are the same. We get:

$$\frac{1}{\theta^4}10Var[X_1]$$

To finish evaluating $Var[\ell'_X(\theta)]$, I must first evaluate $Var[X_1]$.

Recall from the shortcut formula for calculating variance:

$$Var[X_1] = E[X_1^2] - (E[X_1])^2$$

By LOTUS:

$$E[X_1^2] = \int_0^\infty x^2(\frac{1}{\theta}e^{-x/\theta})dx = 2\theta^2$$

Using the definition of expected value for continuous random variables we get:

$$E[X_1] = \int_0^\infty x(\frac{1}{\theta}e^{-x/\theta})dx = \theta$$

Therefore:

$$Var[X_1] = E[X_1^2] - (E[X_1])^2 = 2\theta^2 - (\theta)^2 = 2\theta^2 - \theta^2 = \theta^2$$

Returning to our previous calculation:

$$Var[\ell'_X(\theta)] = \frac{1}{\theta^4}10Var[X_1] = \frac{1}{\theta^4}10\theta^2$$

Simplifying, we get:

$$Var[\ell'_X(\theta)] = I(\theta) = \frac{10}{\theta^2}$$

This means, by the Cramèr-Rao inequality, that the variance of the best possible unbiased estimator of θ has variance $\frac{1}{\frac{10}{\theta^2}} = \frac{\theta^2}{10}$.

Let me first confirm that $\hat{\theta}$ is an unbiased estimator of θ. We must check that $E[\hat{\theta}] - \theta = 0$.

$$E[\hat{\theta}] - \theta = E[\frac{\sum_{i=1}^{10} X_i}{10}] - \theta$$

Using the fact that the X_i are i.i.d. we get:

$$\text{bias} = E[\frac{\sum_{i=1}^{10} X_i}{10}] - \theta = \frac{1}{10}10E[X_1] - \theta$$

We know from before that $E[X_1] = \theta$ thus:

$$\text{bias} = E[\hat{\theta}] - \theta = \frac{10\theta}{10} - \theta = \theta - \theta = 0$$

Therefore $\hat{\theta}$ is an unbiased estimator of θ.

Now that we know $\hat{\theta}$ is unbiased. We can check its variance to see if its the best possible unbiased estimator.

$$Var[\hat{\theta}] = Var[\frac{\sum_{i=1}^{10} X_i}{10}]$$

We can pull out the $\frac{1}{10}$ and remember to square it as well as use the fact the $X_i's$ are i.i.d. to get:

$$\frac{1}{10^2}10Var[X_1] = \frac{1}{10}Var[X_1]$$

We already found $Var[X_1] = \theta^2$ though. Plugging that in we get:

$$Var[\hat{\theta}] = \frac{1}{10}Var[X_1] = \frac{\theta^2}{10}$$

Since the variance of the unbiased estimator $\hat{\theta}$ is the same as the best possible variance of an unbiased estimator we know that $\hat{\theta}$ is the best possible estimator of θ. The reason being that we want estimators with as low of variance and bias as possible. Since this esimator has no bias and has the lowest possible variance of estimators of θ it is the best estimator of θ.

We might also want to know what the MSE of $\hat{\theta}$ is.

We know $\text{bias}(\hat{\theta}) = 0$ and $\text{variance}(\hat{\theta}) = \frac{\theta^2}{10}$.

Also, we know $\text{MSE}(\hat{\theta}) = \text{variance}(\hat{\theta}) + (\text{bias}(\hat{\theta}))^2$.

It clearly follows that $\text{MSE}(\hat{\theta}) = \text{variance}(\hat{\theta}) = \frac{\theta^2}{10} = .1\theta^2$.

Now, consider another possible estimator: $\tilde{\theta} = \alpha\bar{X}$ where $\alpha = \frac{9}{10}$.

We have that $\tilde{\theta} = \frac{9}{10}\bar{X}$.

We can calculate $\text{bias}(\tilde{\theta})$ as follows:

$$\text{bias}(\tilde{\theta}) = E[\tilde{\theta}] - \theta$$

$$E[\tilde{\theta}] = E[\frac{9}{10}\bar{X}] = \frac{9}{10}E[\bar{X}] = \frac{9}{10}E[\frac{X_1 + ... + X_{10}}{10}] = \frac{9}{10^2}E[X_1 + ... + X_{10}]$$

Using the fact the X_i are identically distributed we get:

$$E[\tilde{\theta}] = \frac{9}{10^2}10E[X_1] = \frac{9}{10^2}10\theta = \frac{9\theta}{10}$$

This means that the bias is:

$$\text{bias}(\tilde{\theta}) = E[\tilde{\theta}] - \theta = \frac{9\theta}{10} - \frac{10\theta}{10} = -\frac{\theta}{10}$$

To calculate the variance we get:

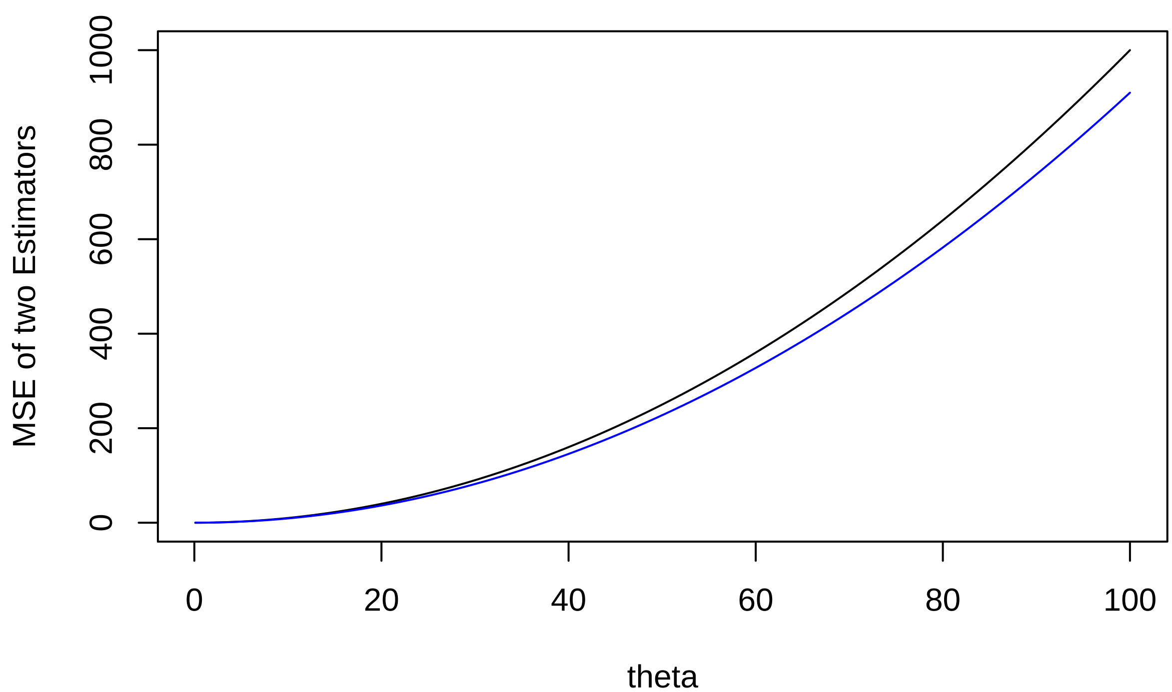
$$Var[\tilde{\theta}] = Var[\frac{9}{10}\bar{X}] = (\frac{9}{10})^2 Var[\bar{X}] = (\frac{9}{10})^2 Var[\frac{X_1 + ... + X_{10}}{10}]$$

Now using properties of variance and the fact that these random variables are i.i.d. we get:

$$Var[\tilde{\theta}] = (\frac{9}{10})^2 \frac{1}{10^2} 10 Var[X_1] = (\frac{9}{10})^2 \frac{1}{10}\theta^2 = \frac{9^2}{10^3}\theta^2$$

Now, to calculate the MSE of $\tilde{X}$:

$$\text{MSE}(\tilde{\theta}) = \text{variance}(\tilde{\theta}) + (\text{bias}(\tilde{\theta}))^2 = \frac{9^2}{10^3}\theta^2 + (-\frac{\theta}{10})^2 = \frac{91}{1000}\theta^2 = .091\theta^2$$



The MSE of $\tilde{\theta}$ is smaller for all values of θ. Therefore $\tilde{\theta}$ has better MSE than $\hat{\theta}$.

This does not contradict what I found earlier because $\tilde{\theta}$ is not an unbiased estimator of θ and the bound given by the Cramèr-Rao inequality only applies to unbiased estimators.

A natural question that might arise is what value of α minimizes the MSE of $\tilde{\theta}$.

When we define $\tilde{\theta}$ in terms of an unknown alpha we get:

$$\text{bias}(\tilde{\theta}) = E[\tilde{\theta}] - \theta = E[\alpha \bar{X}] - \theta = \theta(\alpha - 1)$$

Similarly:

$$Var[\tilde{\theta}] = Var[\alpha \bar{X}] = \frac{\theta^2}{10}\alpha^2$$

This means that the formula for the MSE of $\tilde{\theta}$ is:

$$\text{MSE}(\tilde{\theta}) = \frac{\theta^2}{10}\alpha^2 + (\theta(\alpha - 1))^2$$

We can rewrite this as:

$$\text{MSE}(\tilde{\theta}) = \frac{\theta^2}{10}\alpha^2 + \theta^2\alpha^2 - 2\theta^2\alpha + \theta^2$$

We now take the derivative to see where this function of alpha is maximized. Note that we are treating θ as a constant here (in the sense that it does not depend on the variable we are integrating with respect to).

$$\frac{d}{d\alpha}(\text{MSE}(\tilde{\theta})) = \frac{\theta^2 + 10\theta^2}{5}\alpha - 2\theta^2$$

Setting this equal to 0 and solving for alpha we get $\alpha = \frac{10}{11}$.

Thus, $\alpha = \frac{10}{11}$ is the optimal value of α in terms of the MSE.

Now, I will run some empirical tests just to confirm our math.

```
theta <- 1
alphas <- seq(0, 2, .0001)

mse <- function(theta, alpha){
  val <- theta^2/10*alpha^2 + (theta * (alpha - 1))^2
  return (val)
}

resp <- mse(theta, alphas)

alphas[which.min(resp)]
```

```
## [1] 0.9091
```

```
theta <- 10
alphas <- seq(0, 2, .0001)

mse <- function(theta, alpha){
  val <- theta^2/10*alpha^2 + (theta * (alpha - 1))^2
```

```r
  return (val)
}

resp <- mse(theta, alphas)

alphas[which.min(resp)]
```

```
## [1] 0.9091
```

```r
theta <- 1931293.12312312
alphas <- seq(0, 2, .0001)

mse <- function(theta, alpha){
  val <- theta^2/10*alpha^2 + (theta * (alpha - 1))^2
  return (val)
}

resp <- mse(theta, alphas)

alphas[which.min(resp)]
```

```
## [1] 0.9091
```

This illustrates that $\alpha = \frac{10}{11}$ is optimal in the case where $\theta = 1$, $\theta = 10$, and $\theta = 1931293.12312312$. Note that I am only considering $\alpha's$ ranging from 0 to 2 here. The calculus proves this in the general case.