Computational Statistics II

Homework 5

Subash Kharel

April 07, 2020

1. We will now consider the Boston housing data set, from the MASS library.
   a. Based on this data set, provide an estimate for the population mean of medv. Call this estimate ^μ

```
mew = mean(medv)
mew
```

```
Result:
[1] 22.53281
```

   b. Provide an estimate of the standard error of ^μ. Interpret this result.
   Hint: We can compute the standard error of the sample mean by dividing the sample standard deviation by the square root of the number of observations.

```
standard_error = sqrt(var(medv)/nrow(Boston))
standard_error
```

```
Result:
[1] 0.4088611
```

The standard error of the sample mean is ~0.409 using the standard formula. This means the sample mean is ~0.409 far from the population mean.

   c. Now estimate the standard error of ^μ using the bootstrap. How does this compare to your answer from (b)?

```
mean_function = function(data, index){
  return (mean(data[index]))
}
boot(medv ,mean_function, R=1000)
```

```
Result:
ORDINARY NONPARAMETRIC BOOTSTRAP
Call:
boot(data = medv, statistic = mean_function, R = 1000)
Bootstrap Statistics:      original       bias              std.error
t1* 22.53281    0.01249901    0.3994601
```

The standard error calculated using bootstrap is 0.399 which is almost equal to the value calculated using the standard formula.

   d. Based on this data set, provide an estimate, ^μmed, for the median value of medv in the population.

```
med = median(medv)
med
```

```
Result:
[1] 21.2
```

e. We now would like to estimate the standard error of ^μmed. Unfortunately, there is no simple formula for computing the standard error of the median. Instead, estimate the standard error of the median using the bootstrap. Comment on your findings.

```
median_function = function(data, index){
  return (median(data[index]))
}

boot(medv, median_function, R=1000)
```

```
Result:
ORDINARY NONPARAMETRIC BOOTSTRAP

Call:
boot(data = medv, statistic = median_function, R = 1000)

Bootstrap Statistics:
     original          bias        std. error
t1*   21.2         -0.00425    0.3766182
```

The standard error of the median is 0.3766 using the bootstrap. This is the measure of how far the sample median is from the population median.