Estimation of Population Number of Female Bears in Yellowstone

Ruth Kristianingsih

October 23, 2017

Introduction

Studies about estimation of annual numbers of female bears with cub-of-the-year in the Yellowstone grizzly bear population was done by Keating (2002), from 1986 to 2001.For grizzly bears *(Ursus arctos horribilis)* in the Greater Yellowstone Ecosystem (GYE), minimum population size and allowable numbers of human-caused mortalities have been calculated as a function of the number of unique females with cubs-of-the-year seen during a 3-year period. (Keating et all, 2002) In the table below is the number of unique females with cubs-of-the-year that were seen exactly j times during year 1998.

Sights	Bears
1	11
2	13
3	5
4	1
5	1
6	0
7	2

The objective of this analysis is to estimate the total number of female bears and compute the confidence intervals (CI) using parametric bootstrap.

Figure 1: The Yellowstone Bears

Estimation

In order to estimate the number of female bears, there are some estimators that can be used. One of those estimators is Chao's estimator. There are two kind of Chao's estimator that have been introduced, Chao1 that was introduced in 1984 and Chao2 in 1989. (Keating et all, 2002)

1. Chao's Estimator

First of all, Chao1 estimator will be used to calculate the number of female bears. Suppose that, during a given year, after recording n independent random sightings of individuals from a closed population of size N (where N is unknown), there were observed m unique animals.(Keating, et all, 2002)

The number of different individuals observed exactly j times was f_j, and $f = (f_1, f_2, f_3..., fn)$, with the relationship that $m = \sum_{j=1}^{n} f_j$.

First Chao estimator (1984) was examined

$$\hat{N}_{Chao1} = m + \frac{f_1^2}{2f_2}$$

Now, by using Chao's first estimator, the estimation of the number of female bears can be calculated in the following.

```r
m <- sum(Bears)
f1 <- task3dat [1,2]
f2 <- task3dat [2,2]
N.chao1 <- m + (f1^2/(2*f2))
N.chao1
```

```
## [1] 37.65385
```

And from the result above, it is obtained the the estimation of the total number of female bears is around 38.

Then the second Chao's estimator (1989) which is the similar bias-corretcted of the previous estimator, and with the formula below, the number of female bears can be estimated.

$$\hat{N}_{Chao2} = m + \frac{f_1^2 - f_1}{2(f_2 + 1)}$$

And from data, we obtained the result below, the total number of female bears is around 37.

```r
N.chao2 <- m + (f1^2-f1)/(2*(f2+1))
N.chao2
```

```
## [1] 36.92857
```

The difference of estimation obtained from both estimator is only one individual female bear. However, in order to see the accuracy and the confidence interval of the estimation, parametric bootsrap need to be done. In the next step, simulation using parametric bootsrap using the Chao's second estimator will be examined.

2. Parametric Boostrap

By assuming that the counts follow a zero-truncated Poisson distribution, average value of the data is calculated.

```r
#average value of the data
y.tot <- sum(Sights*Bears)
y.tot
```

```
## [1] 75
```

```r
sum(Bears)
```

```
## [1] 33
```

```r
y.bar <- y.tot/sum(Bears)
y.bar
```

```
## [1] 2.272727
```

By using the average value calculated above, the value of λ can be found by expected value formula of zero-truncated poisson distribution.

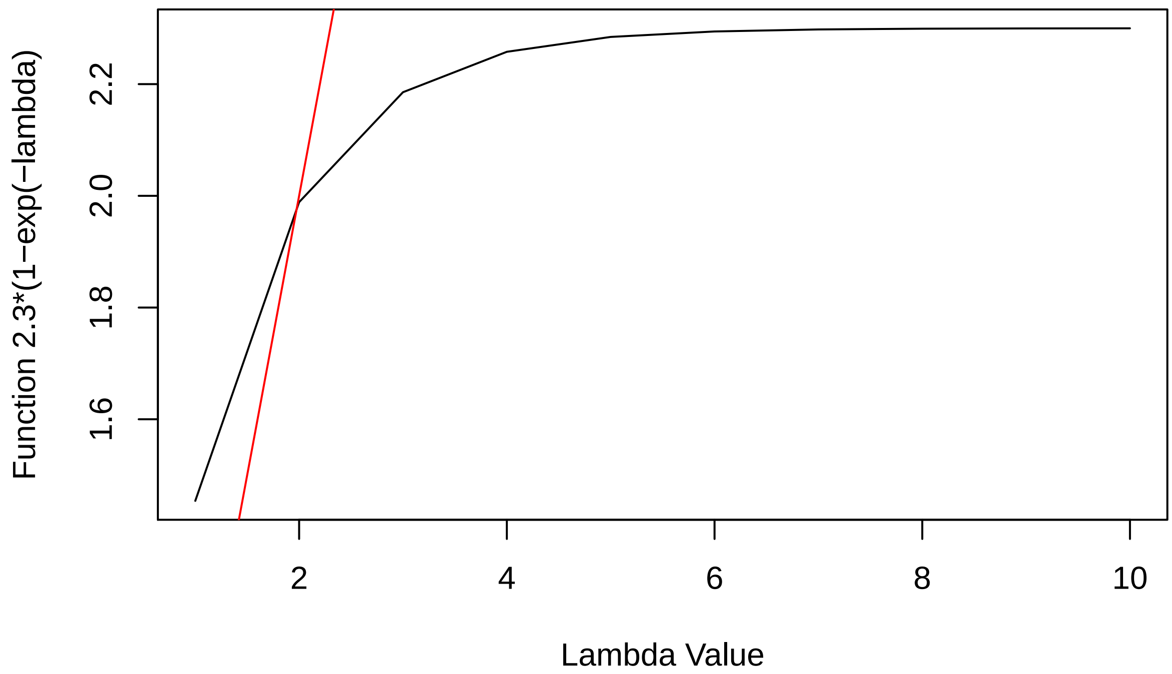
$$E(y) = \frac{\lambda}{1 - exp(-\lambda)}$$

By using the function inroot () in R, value of λ is estimated.

```
## $root
## [1] 1.983567
##
## $f.root
## [1] 4.739182e-06
##
## $iter
## [1] 8
##
## $init.it
## [1] NA
##
## $estim.prec
## [1] 6.103516e-05
```

The value of λ is approximately 1.98. Approximation using graph in R can be done as well to estimate the value of λ. The following graph shows that the intersection of two graph is approximately 1.98 which is same with what obtained above.

```r
lamda <- seq(1,10, by=1)
ylam <- 2.3*(1-exp(-lamda))
plot(ylam, type="l", xlab="Lambda Value", ylab="Function 2.3*(1-exp(-lambda)")
lines(lamda, col="red")
locator()
```



By using function locator (), it can be seen that the intersection of two graphs above is on $\lambda = 1.98$ that will be use to simulate the data.

Next step is simulating data using parametric boostrap by assuming that the data follows zero-truncated poisson.

```
simul <- matrix (0, 1000, 33)
for (i in 1:1000){
n <- 33; lambda <-1.98
Y<-rpois(n,lambda); Y0<-Y[Y>0]; r<-(n - length(Y0))
while(r>0){
  Y<-rpois(r,lambda); Y0<-c(Y0,Y[Y>0]); r<-(n - length(Y0))
}
simul[i,] <- Y0
}
```

Now, the total number of bears from the simulation done before is calculated. With the Chao's second estimator (1989) yields the estimation.

```
for (i in 1:1000){
  m[i] <- sum(simul[i,]==1)+sum(simul[i,]==2)+sum(simul[i,]==3)+sum(simul[i,]==4)+sum(simul[i,]==5) + su
  n [i] <- m[i] + ((sum(simul[i,]==1))^2-sum(simul[i,]==1))/(2*(sum(simul[i,]==2)+1))
}
n <- round(n,2)
n[1:10]
```

```
##  [1] 37.50 34.65 42.75 40.00 35.77 36.46 39.50 38.50 38.50 38.50
```

Then, after finding the number of total female bears form the simulation data, several calculation should be done in order to analyze the confidence intervals (CI), which are the average value of bootstrap estimate, expected bias, and estimated standard error.

The average value of bootstrap estimate can be calculated simply by taking the average of the number of bears obtained above, by function mean () in R, the average is around 38.

```
mean(n)
```

```
## [1] 38.20768
```

In order to know the expected bias, taking the difference of the average of bootstrap estimate and the results of calculating the number of female bears directly from actual data have to be done. And the expected bias is shown below

```
mean(n)-37
```

```
## [1] 1.20768
```

And the standard error of the data simulation is around 4.

```
sd(n)
```

```
## [1] 3.838571
```

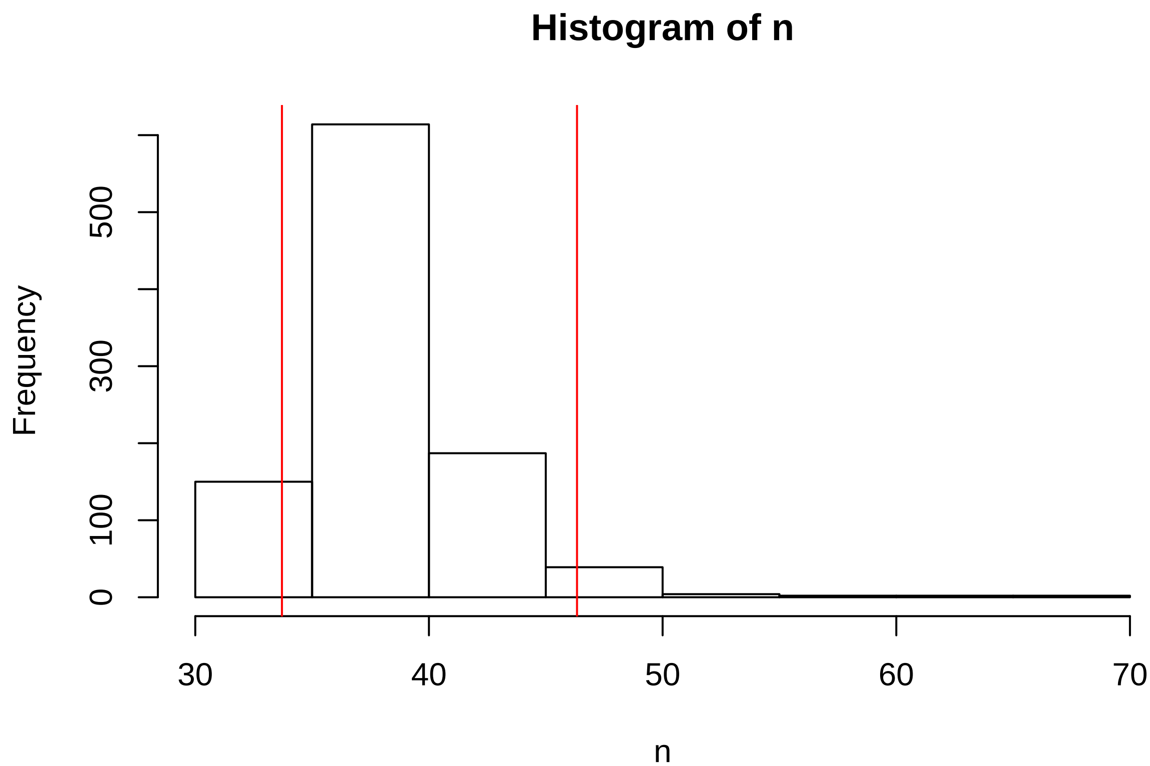
Therefore, the population estimate using bootstrap is around 38 with standard error 4.015. And confidence interval (CI) can be calculated which is 38 ± 8. Confidence Intervals (CI) using the quantiles of the bootstrap replicates can be obtained as well which yields the results below

```
quantile(n, probs=c(0.025, 0.975))
```

```
##     2.5%    97.5%
## 33.71000 46.33675
```

Therefore, we can say that the procedure using bootsrap has result in the interval [34,48]: 41 ± 7. The following histogram can also show where the range of number population of female bears are.

```
qts <- quantile(n,probs=c(.025,.975))
hist(n)
abline(v=qts[1],col="red")
abline(v=qts[2],col="red")
```



Conclusion

Based on the results obtained above, it can be seen that the estimation process using parametric boostrap is better than using the classical procedure, especially by using the quantiles of the bootstrap replicates, it also gives the results to the interval [34,48].

References

Keating, Kim A., et all. 2002. Estimating Numbers of Females with Cubs-of-Year in the Yellowstone Grizzly Bear Population. Ursus 13:161-174.