

Computer Intensive Statistics in Ecology HW4

黃梓育

2018.03.28

<http://140.119.65.88/hw4/#1>

Computer Intensive Statistics in Ecology HW4

黃梓育

2018.03.28

Question 1

Compute the mean and stand error of the mean for the fish and copepod density (all data points) respectively using Jackknife. Plot the histogram of Jackknife means.



solution)

```
library(knitr) # plot table
library(readxl) # read excel
library(ggplot2) # plot figure
df <- data.frame(read_excel("./enviANDdensity.xls"))
```

(conti.)

```
DT::datatable(df[,11:12],fillContainer = F, options = list(
```

Show entries

	FishDensity..ind..1000m3. 	CopepodDensity.ind..m3. 
1	137.3233973	1119
2	20.96762377	1153
3	0	1719
4	0	855
5	180.7135569	1246

Showing 1 to 5 of 34 entries

Previous

1

2

3

4

5

6

7

Next

(conti.)

```
DT::datatable(df[,11:12],fillContainer = F, options = list(
```

Show entries

	FishDensity..ind..1000m3.	CopepodDensity.ind..m3.
1	137.3233973	1119
2	20.96762377	1153
3	0	1719
4	0	855
5	180.7135569	1246

Showing 1 to 5 of 34 entries

Previous

1234567

Next

So, we load these column

```
fish <- df$FishDensity..ind..1000m3.
cope <- df$CopepodDensity.ind..m3.
```

(conti.)

```
# resample the original sample by Jackknife  
jackknife.resample <- function(data){  
  return(t(sapply(1:length(data), function(i) data[-i])))  
}
```

(conti.)

```
# resample the original sample by Jackknife
jackknife.resample <- function(data){
  return(t(sapply(1:length(data), function(i) data[-i])))
}
```

Calculate Jackknife of mean of resample data

$$\hat{\theta}_{(.)} = \frac{\sum_{j=1}^n \hat{\theta}_{-j}}{n}$$

```
# calculate mean of Jackknife of resample data
jackknife.mean <- function(data){
  return(sum(data)/dim(data)[1]/dim(data)[2])
}
```


(conti.)

```
# resample the original sample by Jackknife
jackknife.resample <- function(data){
  return(t(sapply(1:length(data), function(i) data[-i])))
}
```

Calculate Jackknife of mean of resample data

$$\hat{\theta}_{(.)} = \frac{\sum_{j=1}^n \hat{\theta}_{-j}}{n}$$

```
# calculate mean of Jackknife of resample data
jackknife.mean <- function(data){
  return(sum(data)/dim(data)[1]/dim(data)[2])
}
```

Calculate Jackknife of stand error of resample data

$$S.R.(\hat{\theta}) = \sqrt{\frac{n-1}{n} \sum_{j=1}^n (\hat{\theta}_{-j} - \hat{\theta}_{(.)})^2}$$

```
jackknife.se <- function(data){
  m <- jackknife.mean(data)
  n <- dim(data)[1]
  return(sqrt(sum(sapply(1:n, function(i) (mean(data[i,])-m
```

(conti.)

```
fish.jackknife.sample <- jackknife.resample(fish)
cope.jackknife.sample <- jackknife.resample(cope)
fish.sample.mean <- data.frame(x=apply(fish.jackknife.saml
cope.sample.mean <- data.frame(x=apply(cope.jackknife.saml
```

(conti.)

```
kable(data.frame(Mean=jackknife.mean(fish.jackknife.sample)
                  SE_of_Mean=jackknife.se(fish.jackknife.sam
                  ,caption="Jackknife sample of fish", align="l", forma
```

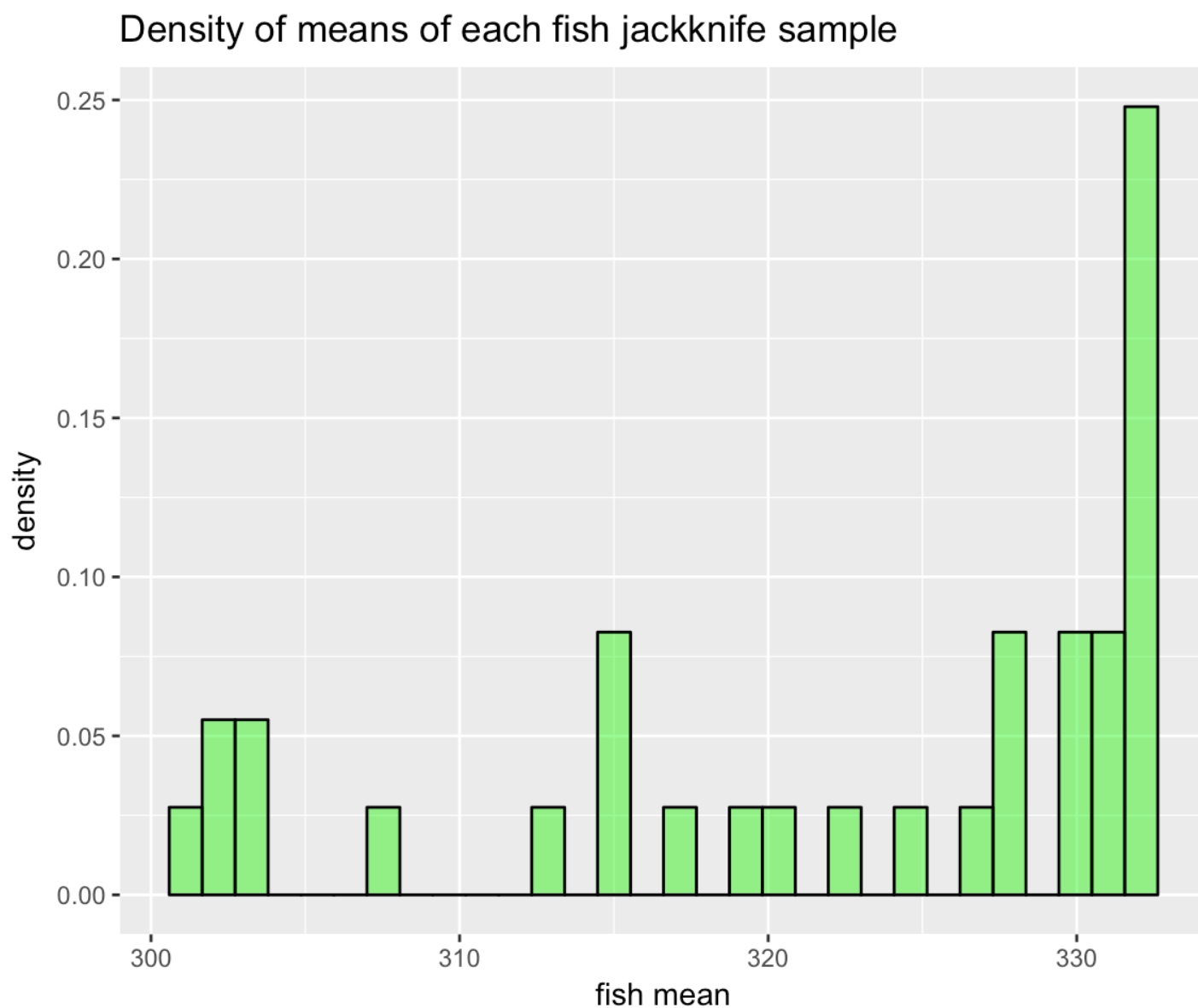
Jackknife sample of fish	
Mean	SE_of_Mean
322.4516	61.23107

```
kable(data.frame(Mean=jackknife.mean(cope.jackknife.sample)
                  SE_of_Mean=jackknife.se(cope.jackknife.sam
                  ,caption="Jackknife sample of copepod", align="l", fo
```

Jackknife sample of copepod	
Mean	SE_of_Mean
1972.374	342.5549

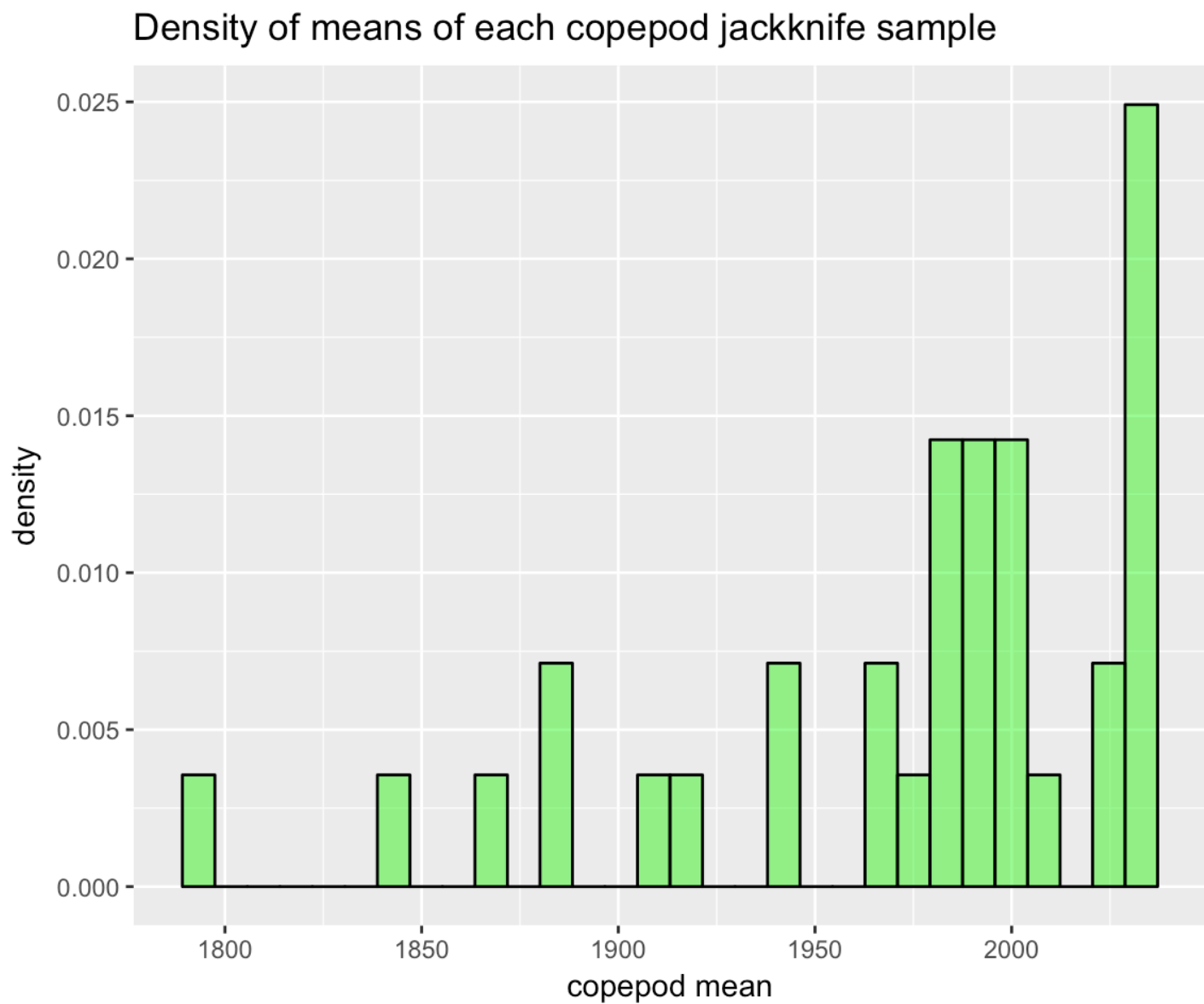
(conti.)

```
ggplot(fish.sample.mean,aes(x)) +  
  geom_histogram(aes(y=..density..), bins = 30, col="black",  
  labs(title="Density of means of each fish jackknife sampl
```



(conti.)

```
ggplot(cope.sample.mean, aes(x)) +  
  geom_histogram(aes(y=..density..), bins = 30, col="black",  
  labs(title="Density of means of each copepod jackknife sa
```



Question 2

Compute the regression coefficients for $\text{fish} = \beta_0 + \beta_1 * \text{copepod}$ and Jackknife SE of β_0 and β_1 .
Plot the histogram of Jackknife β_0 and β_1 .

solution)

Question 2

Compute the regression coefficients for fish = β_0 + β_1 * copepod and Jackknife SE of β_0 and β_1 .
Plot the histogram of Jackknife β_0 and β_1 .

solution) Calculate $\hat{\beta} = [\beta_0 \ \beta_1]'$ as following :

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Question 2

Compute the regression coefficients for fish = β_0 + β_1 * copepod and Jackknife SE of β_0 and β_1 .
Plot the histogram of Jackknife β_0 and β_1 .

solution) Calculate $\hat{\beta} = [\beta_0 \ \beta_1]'$ as following :

$$\hat{\beta} = (X'X)^{-1}X'Y$$

```
coeff.OLS <- function(X,Y){  
  # Y is dependent variable  
  # X is independent variable and intercept  
  X <- cbind(rep(1,length(X)),X)  
  return(solve(t(X) %*% X) %*% (t(X) %*% Y))  
}  
# compute coefficients of original sample  
b <- coeff.OLS(cope,fish)  
cat("Regression line : ", "β0 is", b[1], ", and β1 is", b[2])
```

Regression line : β0 is 93.06466 , and β1 is 0.1162999

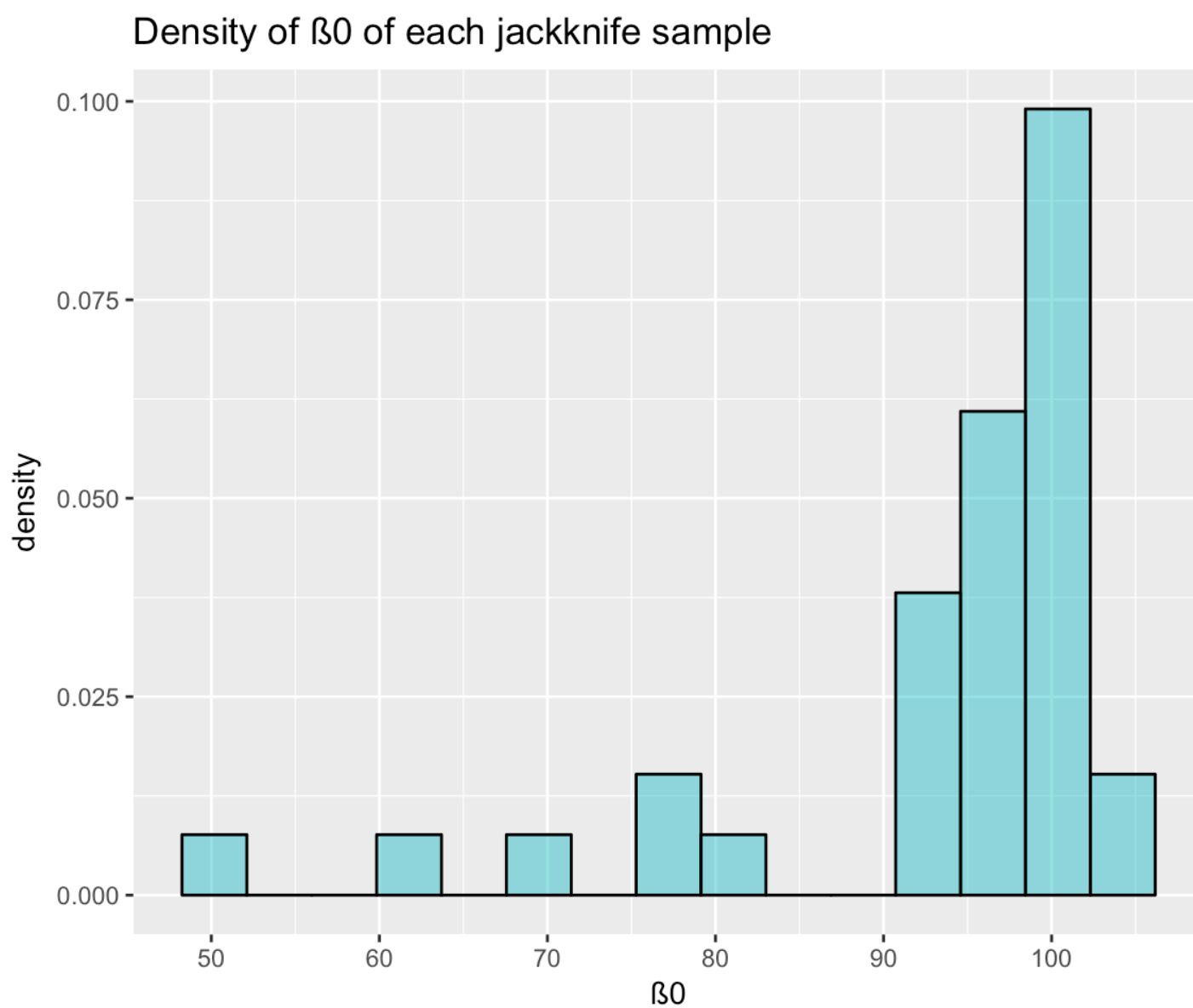
(conti.)

```
# compute coefficients of jackknife sample
jackknife.Coeff <- data.frame(t(
  sapply(1:length(fish),
    function(i) coeff.OLS(cope.jackknife.sample[i,], fi
  colnames(jackknife.Coeff) <- c("beta0", "beta1")
kable(head(jackknife.Coeff), format = "html")
```

beta0	beta1
96.81970	0.1157231
101.96731	0.1149709
103.09380	0.1157187
102.30830	0.1146001
95.45166	0.1159731
100.10574	0.1165966

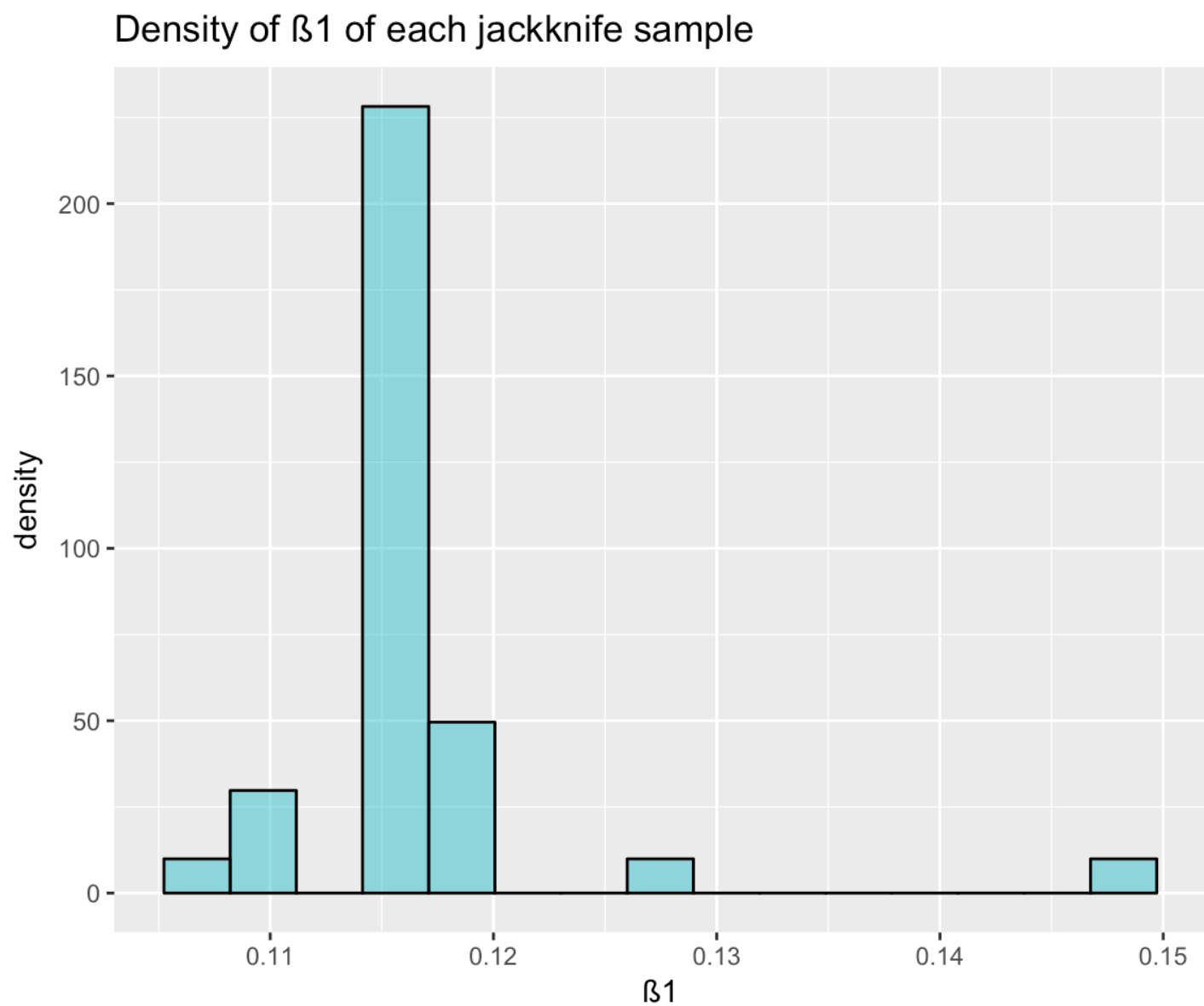
(conti.)

```
ggplot(jackknife.Coeff,aes(beta0)) + theme(legend.position=  
  geom_histogram(aes(y=..density..), bins = 15, col="black"  
  labs(title="Density of  $\beta_0$  of each jackknife sample", x=" $\beta$ 
```



(conti.)

```
ggplot(jackknife.Coeff,aes(beta1)) + theme(legend.position=  
  geom_histogram(aes(y=..density..), bins = 15, col="black"  
  labs(title="Density of  $\beta_1$  of each jackknife sample", x=" $\beta$ 
```



Question 3

Compare the estimates for Q1 and Q2 obtained from normal theory, bootstrap, and jackknife.

solution)

```
# normal theory
normal.se <- function(data){
  data <- matrix(data)
  n <- length(t(data))
  return(sqrt(sum((data-mean(data))^2)/(n*(n-1))))
}
```

Question 3

Compare the estimates for Q1 and Q2 obtained from normal theory, bootstrap, and jackknife.

solution)

```
# normal theory
normal.se <- function(data){
  data <- matrix(data)
  n <- length(t(data))
  return(sqrt(sum((data-mean(data))^2)/(n*(n-1))))
}
```

```
# resample the original sample by bootstrap
bootstrap.resample <- function(data, times){
  n <- length(data)
  return(t(replicate(times, data[ceiling(runif(n,0,n))])))
}
```

Question 3

Compare the estimates for Q1 and Q2 obtained from normal theory, bootstrap, and jackknife.

solution)

```
# normal theory
normal.se <- function(data){
  data <- matrix(data)
  n <- length(t(data))
  return(sqrt(sum((data-mean(data))^2)/(n*(n-1))))
}
```

```
# resample the original sample by bootstrap
bootstrap.resample <- function(data, times){
  n <- length(data)
  return(t(replicate(times, data[ceiling(runif(n,0,n))])))
}
```

```
# calculate bootstrap of stand error of resample data
bootstrap.se <- function(data){
  return(sd(apply(data, 1, mean)))
}
```

(conti.)

```
fish.bootstrap.sample <- bootstrap.resample(fish,1e3)  
cope.bootstrap.sample <- bootstrap.resample(fish,1e3)
```

(conti.)

```
fish.bootstrap.sample <- bootstrap.resample(fish,1e3)
cope.bootstrap.sample <- bootstrap.resample(fish,1e3)

compare.fish <- data.frame(cbind(
  c(mean(fish),normal.se(fish)),
  c(mean(fish.bootstrap.sample), bootstrap.se(fish.bootstrap.sample)),
  c(jackknife.mean(fish.jackknife.sample), jackknife.se(fish.jackknife.sample))
))
rownames(compare.fish) <- c("Mean", "S.E")
colnames(compare.fish) <- c("Normal_Theory", "Bootstrap", "Jackknife")
kable(compare.fish, caption="Comparison of fish", align="l")
```

Comparison of fish			
	Normal_Theory	Bootstrap	Jackknife
Mean	322.45163	323.10264	322.45163
S.E	61.23107	57.18685	61.23107

(conti.)

```
compare.cope <- data.frame(cbind(
  c(mean(cope),normal.se(cope)),
  c(mean(cope.bootstrap.sample), bootstrap.se(cope.bootstrap.sample)),
  c(jackknife.mean(cope.jackknife.sample), jackknife.se(cope.jackknife.sample))
))
rownames(compare.cope) <- c("Mean", "S.E")
colnames(compare.cope) <- c("Normal_Theory", "Bootstrap", "Jackknife")
kable(compare.cope, caption="Comparison of copepod", align="center")
```

Comparison of copepod			
	Normal_Theory	Bootstrap	Jackknife
Mean	322.45163	323.10264	322.45163
S.E	61.23107	57.18685	61.23107

(conti.)

```
n <- length(fish)
btIndex <- sapply(1:1e3,function(ignore) c(1:n)[ceiling(run
btCoeff <- apply(btIndex, 2, function(x) coeff.OLS(cope[x],

beta.cope <- data.frame(cbind(
  c(b),
  c(mean(jackknife.Coeff[,1]),mean(jackknife.Coeff[,2])),
  c(mean(btCoeff[,1]),mean(btCoeff[,2]))))
rownames(beta.cope) <- c("β0", "β1")
colnames(beta.cope) <- c("Normal_Theory", "Bootstrap", "Jac
kable(beta.cope, caption="Comparison of β", align="l", form
```

Comparison of β			
	Normal_Theory	Bootstrap	Jackknife
β_0	93.0646559	92.7309256	83.887058
β_1	0.1162999	0.1165492	0.123219

Thanks you

