HW5

2020324011
민동은

Problem 9.1. $X_1, \cdots, X_n \overset{iid}{\sim} Bernoulli(\theta) = Bin(1, \theta)$.

$R(X, F) = \bar{X} - \theta$ and $R^* = R(X^*, \hat{F})$.

$R(X, F) = \hat{\theta} - \theta$ since sample mean is mle of θ.

X^*: Bootstrap pseudo-dataset, $\hat{F}$: empirical distribution

Assume that the observed $X = x$. and sample mean of $x_i, \cdots, x_n$ is $\bar{x}_n$.

Bootstrap sample : $X_1^*, \cdots, X_n^* \sim Bernoulli(\bar{x}_n)$

Then, $R(X^*, \hat{F}) = \bar{X}_n^* - \bar{x}_n$, and $E(\bar{X}_n^* - \bar{x}_n) = 0$.

$$Var(\bar{X}_n^* - \bar{x}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i^* - \bar{x}_n\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i^*) = \frac{1}{n^2}\cdot n \cdot x_n(1-\bar{x}_n) = \frac{\bar{x}_n(1-\bar{x}_n)}{n}.$$

$$\therefore E(R^*) = 0, \quad Var(R^*) = \frac{\bar{x}_n(1-\bar{x}_n)}{n}.$$

Problem 9.4 (a) : R codes attached in next page.

$*$ Beverton-Holt model : $R = \dfrac{1}{\beta_1 + \beta_2/S}$.

$\rightarrow$ transform to $\frac{1}{R}, \frac{1}{S}$: $\frac{1}{R} = \beta_1 + \beta_2 \cdot \frac{1}{S}$.

Stable population when $R = S$: $\begin{cases} \frac{1}{R} = \beta_1 + \beta_2 \cdot \frac{1}{S} \\ \frac{1}{R} = \frac{1}{S} \end{cases} \rightarrow \begin{cases} y = \beta_1 + \beta_2 x \\ y = x. \end{cases}$

$\rightarrow \beta_1 + \beta_2 x = x$. $\beta_1 = (1-\beta_2)x$.

$$\therefore x = \frac{\beta_1}{1-\beta_2} = y.$$

Computing HW5

Dongeun Min

12/10/2020

Problem 9.4(a)

```r
# make dataset
R = c(68, 77, 299, 220, 142, 287, 276, 115, 64, 206, 222, 205, 233, 228, 188,
      132, 285, 188, 224, 121, 311, 166, 248, 161, 226, 67, 201, 267, 121, 301,
      244, 222, 195, 203, 210, 275, 286, 275, 304, 214)
S = c(56, 62, 445, 279, 138, 428, 319, 102, 51, 289, 351, 282, 310, 266, 256,
      144, 447, 186, 389, 113, 412, 176, 313, 162, 368, 54, 214, 429, 115, 407,
      265, 301, 234, 229, 270, 478, 419, 490, 430, 235)
table = data.frame(1/R, 1/S)
```

Observed data estimate

```r
set.seed(12345)
model = lm(table[,1]~table[,2])
theta = as.numeric(model$coefficients[1]/(1-model$coefficients[2]))
theta
```

```
## [1] 0.00666233
```

Bootstrapping residuals

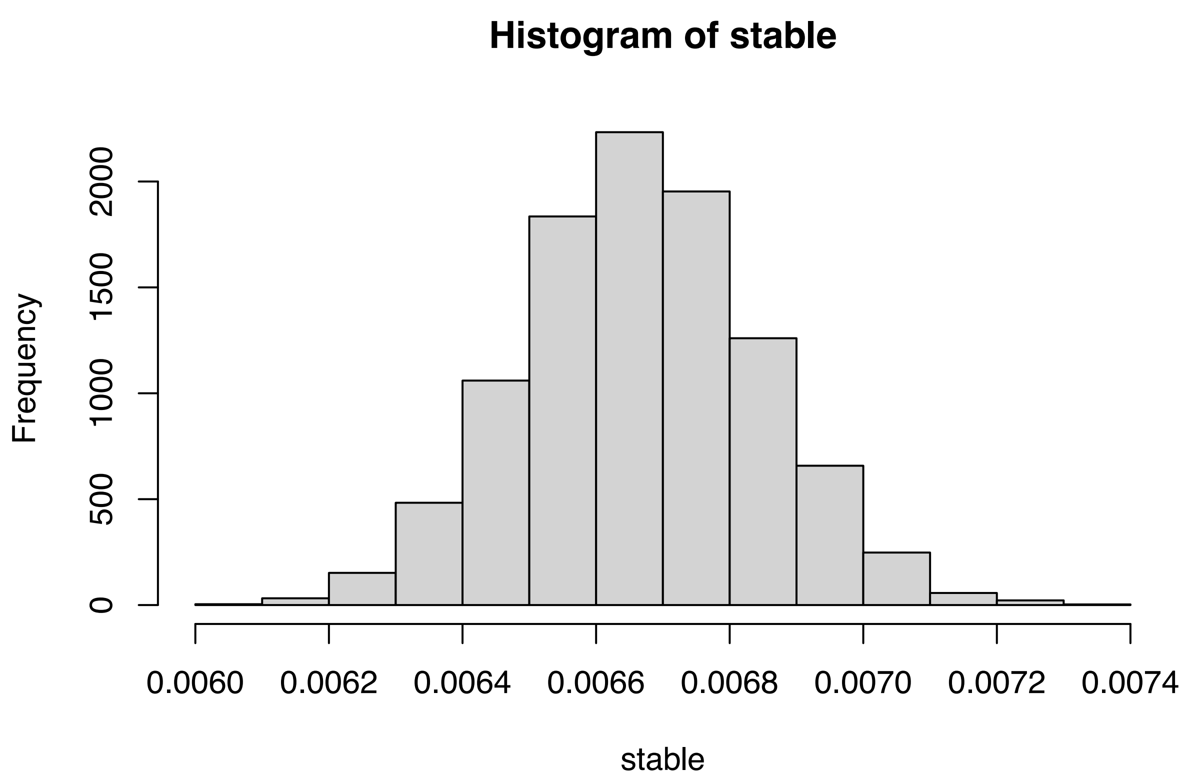
```r
iter = 10000
stable = rep(0,iter)

model = lm(table[,1]~table[,2])
fitted = as.numeric(model$fitted.values)
res = data.frame(model$residuals)

# Bootstrap residuals
for(i in 1:iter){
      newres = res[sample(1:length(res[,1]),replace=T),]
      newY = fitted + newres
      model1 = lm(newY~table[,2])
      # when (1/R) = beta1 + beta2*(1/S) meets with (1/R) = (1/S) (45 degree line)
      stable[i] = model1$coefficients[1] / (1-model1$coefficients[2])
}
ci = quantile(stable, c(0.025,0.975), na.rm=T)
se = sd(stable)/sqrt(length(stable))
ci  #confidence interval of estimate
```

```
##        2.5%       97.5%
## 0.006318194 0.007028899
```

```r
se  #standard error of estimate
```

```
## [1] 1.79193e-06
```

```r
hist(stable)  #histogram of point estimate for stable population level (R=S)
```



Bootstrapping cases

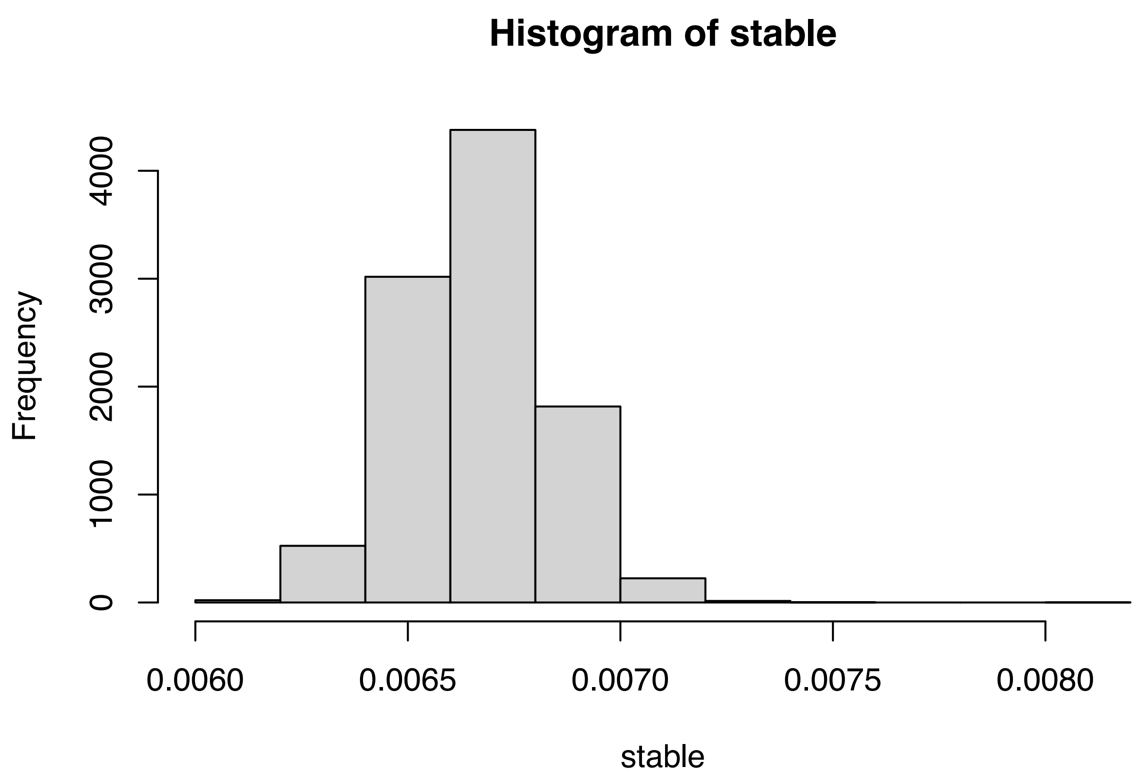
```r
# Bootstrap cases (pairs)
stable = rep(0,iter)

for(i in 1:iter){
      boot = table[sample(1:nrow(table),replace=T),]
      model2 = lm(boot[,1]~boot[,2])
      # when (1/R) = beta1 + beta2*(1/S) meets with (1/R) = (1/S) (45 degree line)
      stable[i] = model2$coefficients[1] / (1-model2$coefficients[2])
}
ci = quantile(stable, c(0.025,0.975), na.rm=T)
se = sd(stable)/sqrt(length(stable))
ci  #confidence interval of estimate
```

```
##        2.5%       97.5%
## 0.006343359 0.006997631
```

```r
se  #standard error of estimate
```

```
## [1] 1.685017e-06
```

```r
hist(stable)  #histogram of point estimate for stable population level (R=S)
```



The point estimate obtained from bootstrapping residuals has a larger standard error and a wider confidence interval than the point estimate obtained from bootstrapping cases. As the point estimate from the observed data is 0.00666233, the histogram of the point estimate obtained from bootstrapping cases shows that it is slightly closer to the point estimate from the observed data (than the point estimate obtained from bootstrapping residuals). However, both estimates are close enough to the result of the observed data.

```r
model1$coefficients
```

```
## (Intercept)  table[, 2]
##  0.00196964  0.70932045
```

```r
model2$coefficients
```

```
## (Intercept)   boot[, 2]
## 0.001996538 0.689638232
```

Also, the Beverton-Holt model has the constraint of beta_1 and beta_2 being non-negative, so the intercept and coefficient of model1 (bootstrapped with residuals) and model2 (bootstrapped cases) may also be checked. Since it is shown that both the intercept (beta_1) and coefficient (beta_2) are non-negative, no other constraint is needed.