

Materi 14
“Bootstrap”

Bootstrap adalah metode inferensi tentang populasi menggunakan data sampel. Bradley Efron pertama kali memperkenalkannya dalam makalah ini pada tahun 1979. Bootstrap mengandalkan pengambilan sampel dengan penggantian dari data sampel. Teknik ini dapat digunakan untuk memperkirakan kesalahan standar dari setiap statistik dan untuk memperoleh interval kepercayaan (CI) untuk itu. Bootstrap sangat berguna ketika CI tidak memiliki formulir tertutup, atau memiliki yang sangat rumit.

Jika kita ingin menghitung CI untuk lebih dari satu statistik sekaligus maka dilakukan tahap-tahap berikut ini dengan mengembalikannya sebagai vektor tunggal.

#1

```
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1      5.1         3.5        1.4         0.2  setosa
2      4.9         3.0        1.4         0.2  setosa
3      4.7         3.2        1.3         0.2  setosa
4      4.6         3.1        1.5         0.2  setosa
5      5.0         3.6        1.4         0.2  setosa
6      5.4         3.9        1.7         0.4  setosa

> library(boot)
> foo = function(data, indices){
+   dt = data[indices,]
+   c(
+     cor(dt[,1],dt[,2],method = 's'),
+     median(dt[,1]),
+     median(dt[,2])
+   )
+ }
> foo <- function(data, indices, cor.type){
+   dt<-data[indices,]
+   c(
+     cor(dt[,1], dt[,2], method=cor.type),
+     median(dt[,1]),
+     median(dt[,2])
+   )
+ }
> set.seed(12345) (#set.seed digunakan untuk reproduksibilitas)
> myBootstrap = boot(iris, foo, R = 1000, cor.type = 's')
> head(myBootstrap$t)
      [,1] [,2] [,3]
```

```
[1,] -0.20526295 5.70 3.1
[2,] -0.23549118 5.80 3.0
[3,] -0.18780154 5.80 3.1
[4,] -0.07147936 5.70 3.0
[5,] -0.25135181 5.80 3.0
[6,] -0.21180913 5.75 3.0
```

```
> myBootstrap$t0
[1] -0.1667777 5.8000000 3.0000000
> myBootstrap
```

ORDINARY NONPARAMETRIC BOOTSTRAP

Call:

```
boot(data = iris, statistic = foo, R = 1000, cor.type = "s")
```

Bootstrap Statistics :

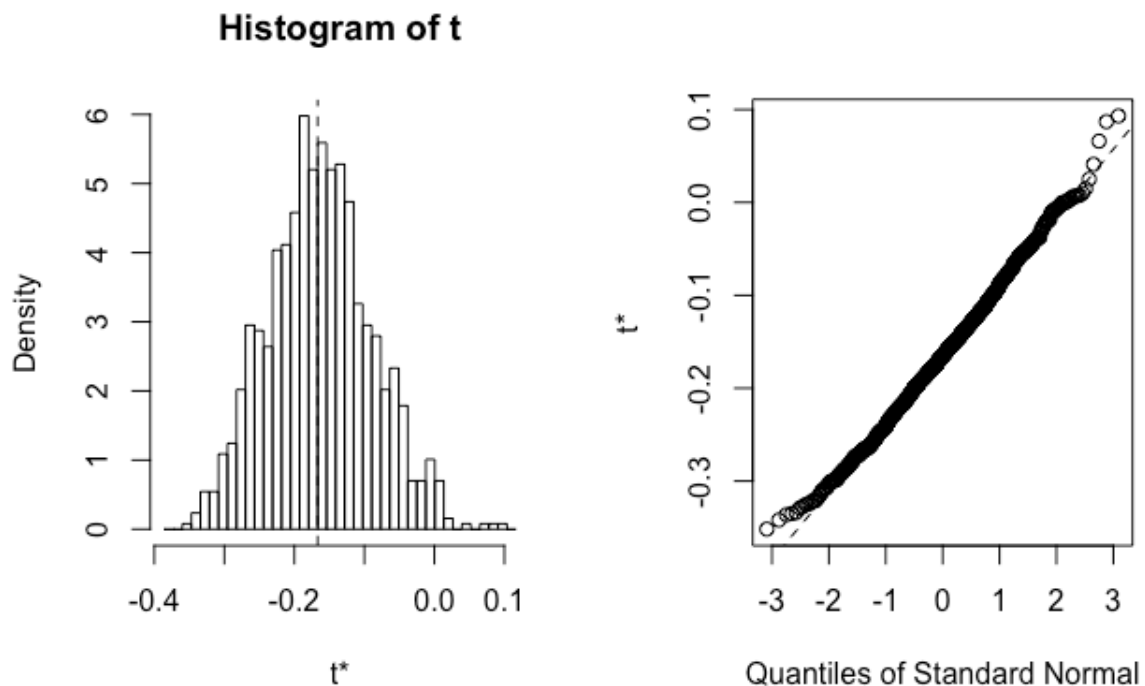
```
      original    bias std. error
t1* -0.1667777 0.003003139 0.07370122
t2*  5.8000000 -0.011300000 0.10678817
t3*  3.0000000 0.010000000 0.02925501
> colMeans(myBootstrap$t)-myBootstrap$t0
[1] 0.003003139 -0.011300000 0.010000000
> apply(myBootstrap$t,2,sd)
[1] 0.07370122 0.10678817 0.02925501
```

Keterangan :

- foo memilih elemen yang kita inginkan yang nomornya disimpan dari data dan menghitung koefisien korelasi dari dua kolom pertama.
- method = 's' berfungsi sebagai koefisien spearman.

#2

```
> plot(myBootstrap, index=1)
```



```
> boot.ci(myBootstrap, index=1)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 1000 bootstrap replicates
```

CALL :

```
boot.ci(boot.out = myBootstrap, index = 1)
```

Intervals :

Level	Normal	Basic
95%	(-0.3142, -0.0253)	(-0.3246, -0.0327)

Level	Percentile	BCa
95%	(-0.3008, -0.0090)	(-0.3006, -0.0089)

Calculations and Intervals on Original Scale

Warning message:

In boot.ci(myBootstrap, index = 1) :

bootstrap variances needed for studentized intervals

```
> boot.ci(myBootstrap, index=1, type=c('basic','perc'))
```

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS

Based on 1000 bootstrap replicates

CALL :

```
boot.ci(boot.out = myBootstrap, type = c("basic", "perc"), index = 1)
```

Intervals :

```
Level   Basic      Percentile
95%    (-0.3246, -0.0327 ) (-0.3008, -0.0090 )
Calculations and Intervals on Original Scale
> boot.ci(myBootstrap, index=1, type='norm')$norm
  conf
[1,] 0.95 -0.3142325 -0.02532907
> boot.ci(myBootstrap, index=1, type='basic')$basic
  conf
[1,] 0.95 975.98 25.03 -0.3245646 -0.03272326
```

Keterangan :

Index 1 adalah koefisien korelasi spearman antara sepal.length dan width.

Index 2 adalah median sepal.length

Index 3 adalah median sepal width

Boot.ci digunakan untuk menciptakan objek dari kelas

#3

```
> boot.ci(myBootstrap, index=3)
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 1000 bootstrap replicates
```

CALL :

```
boot.ci(boot.out = myBootstrap, index = 3)
```

Intervals :

```
Level   Normal      Basic
95%    ( 2.933, 3.047 ) ( 2.900, 3.000 )
```

```
Level   Percentile      BCa
```

```
95%    ( 3.0, 3.1 ) ( 2.9, 2.9 )
```

Calculations and Intervals on Original Scale

Warning : BCa Intervals used Extreme Quantiles

Some BCa intervals may be unstable

Warning messages:

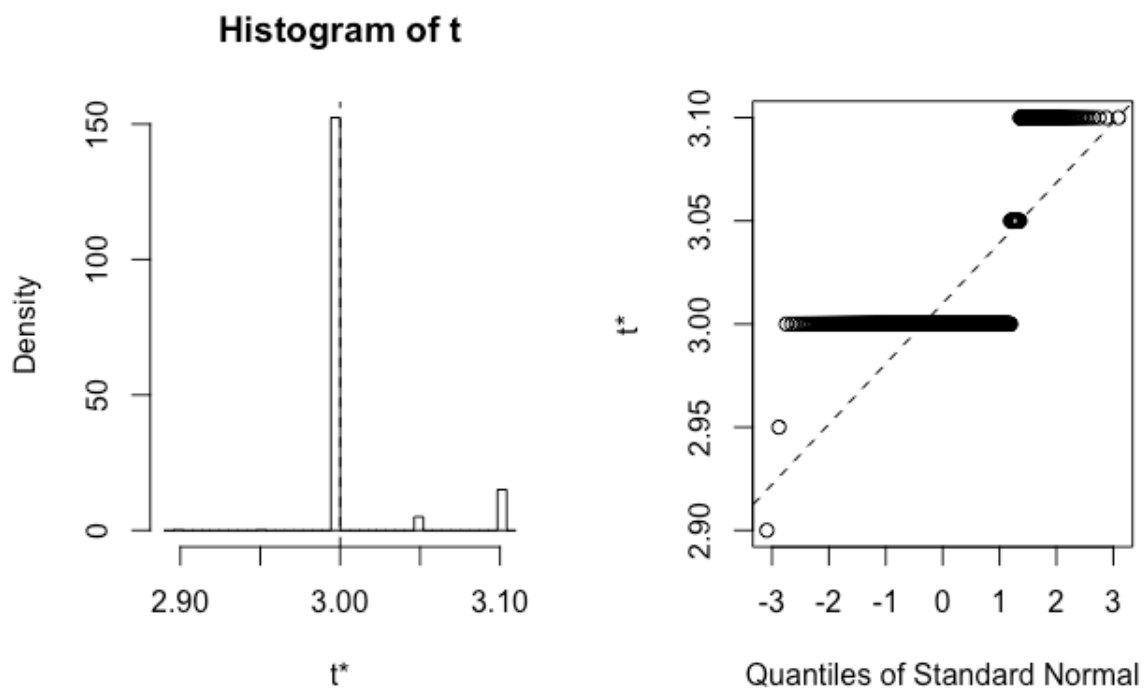
1: In boot.ci(myBootstrap, index = 3) :

bootstrap variances needed for studentized intervals

2: In norm.inter(t, adj.alpha) :

extreme order statistics used as endpoints

```
> plot(myBootstrap, index=3)
```



```
> table(myBootstrap$t[,3])
```

```
2.9 2.95 3 3.05 3.1
1 1 882 29 87
```

#4

```
> tableOfIndices<-boot.array(myBootstrap, indices=T)
```

```
> tableOfIndices[1,]
```

```
[1] 142 31 28 43 6 26 42 92 30 3 34 73 68 46 118 25 4 29 48 12 146
[22] 64 19 102 22 136 98 11 90 109 22 89 56 113 60 6 146 126 97 62 45 41
[43] 73 10 148 59 53 133 37 106 102 102 2 125 142 23 132 92 135 17 118 4 8
[64] 87 43 79 49 23 83 70 63 149 5 82 65 19 55 145 104 25 21 19 98 49
[85] 8 52 12 43 101 30 59 80 146 99 120 98 135 19 12 84 13 26 57 51 30
[106] 141 149 104 7 104 82 101 20 13 107 80 33 24 79 75 64 49 95 84 119 46
[127] 126 80 78 141 45 131 14 20 9 113 6 44 98 81 93 110 29 110 123 38 118
[148] 4 141 3
```

```
> tableOfAppearances<-boot.array(myBootstrap)
```

```
> tableOfAppearances[1,]
```

```
[1] 0 1 2 3 1 3 1 2 1 1 1 3 2 1 0 0 1 0 4 2 1 2 2 1 2 2 0 1 2 3 1 0 1 1 0 0 1 1 0 0 1 1
[43] 3 1 2 2 0 1 3 0 1 1 1 0 1 1 1 0 2 1 0 1 1 2 1 0 0 1 0 1 0 0 2 0 1 0 0 1 2 3 1 2 1 2
[85] 0 0 1 0 1 1 0 2 1 0 1 0 1 4 1 0 2 3 0 3 0 1 1 0 1 2 0 0 2 0 0 0 0 3 1 1 0 0 1 0 1 2
[127] 0 0 0 0 1 1 1 0 2 1 0 0 0 0 3 2 0 0 1 3 0 1 2 0
```

```
> onceAgain<-apply(tableOfIndices, 1, foo, data=iris, cor.type='s')
```

```
> head(t(onceAgain))
```

```
      [,1] [,2] [,3]
[1,] -0.20526295 5.70 3.1
[2,] -0.23549118 5.80 3.0
```

```
[3,] -0.18780154 5.80 3.1
[4,] -0.07147936 5.70 3.0
[5,] -0.25135181 5.80 3.0
[6,] -0.21180913 5.75 3.0
> head(myBootstrap$t)
      [,1] [,2] [,3]
[1,] -0.20526295 5.70 3.1
[2,] -0.23549118 5.80 3.0
[3,] -0.18780154 5.80 3.1
[4,] -0.07147936 5.70 3.0
[5,] -0.25135181 5.80 3.0
[6,] -0.21180913 5.75 3.0
```