

Masterclass-bootstrap-eks-R

Harald.Weedon-Fekjar@medisin.uio.no (mailto:Harald.Weedon-Fekjar@medisin.uio.no)

2023

```
# OCBE Masterclass Bootstrap 2023 - R eksempel (https://www.r-project.org/)
#   https://www.med.uio.no/imb/english/research/centres/ocbe/
#   https://www.ous-research.no/ocbe/
#   ved Harald Weedon-Fekjær <harald.weedon-fekjar@medisin.uio.no>
```

```
# *** Bootstrap konfidens intervaller i R: ***
# Bootstrap i R gjøres via det innebygde bibloteket "boot".
# Først lager man en funksjon som regner ut estimatet man ønsker
# konfidens intervall for, og tar en indeks vektor (liste) som definer
# hva som skal og ikke skal med i det enkelte bootstrap utvalget.
# Siden kjører man "boot" og får konfidens intervall for estimatet.
```

```
# -----
# ----- Laster eksempel dataene: -----
# -----
# NB: Husk å forandre filområde til der du har filen!
setwd(file.path("M:\\OusBiostat\\Kurs-foredrag\\Masterclass",
               "Masterclass-Bootstap",
               "Eksempel\\Eksempel-programmer\\R"))
RoykeSlutt <- read.csv("RoykeSlutt.csv",header=T,sep=";")
```

```
# -----
# ----- 95% bootstrap konfidens intervall: -----
# -----
```

```
library(boot)
```

```
# Setter frøet til tilfeldig tall generator:
# (for å sikre at kjøringene er like hver gang)
set.seed(3)
```

```
# Bootstrap: Gjennomsnittlig røykeslutt motivasjon:
myfunc <- function(data,indices) { mean(data[indices],na.rm=T) }
results <- boot(data=RoykeSlutt$Motivasjon,statistic=myfunc,R=10000)
boot.ci(results, type="perc") # Persentil bootstrap
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = results, type = "perc")
##
## Intervals :
## Level      Percentile
## 95%      ( 7.982,  8.948 )
## Calculations and Intervals on Original Scale
```

```
boot.ci(results, type="bca")    # Bca bootstrap
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = results, type = "bca")
##
## Intervals :
## Level      BCa
## 95%      ( 7.914,  8.912 )
## Calculations and Intervals on Original Scale
```

```
# Bootstrap: Median røykeslutt motivasjon:
myfunc <- function(data,indices) { median(data[indices],na.rm=T) }
results <- boot(data=RoykeSlutt$Motivasjon,statistic=myfunc,R=10000)
boot.ci(results, type="perc")
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = results, type = "perc")
##
## Intervals :
## Level      Percentile
## 95%      (  8, 10 )
## Calculations and Intervals on Original Scale
```

```
# Bootstrap: Gjennomsnittlig forskjell i røykeslutt motivasjon mellom
#               de som har sluttet etter 6 måneder og de som ikke har sluttet:
myfunc <- function(data,indices) {
  bootdata <- data[indices,]
  m0 <- mean(bootdata$Motivasjon[bootdata$Sluttet6m==0],na.rm=T)
  m1 <- mean(bootdata$Motivasjon[bootdata$Sluttet6m==1],na.rm=T)
  return(m0-m1)
}
results <- boot(data=RoykeSlutt,statistic=myfunc,R=10000)
boot.ci(results, type="perc")
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = results, type = "perc")
##
## Intervals :
## Level      Percentile
## 95%      (-1.7679,  0.0455 )
## Calculations and Intervals on Original Scale
```

```
# Tilsvarende over, men via regresjon (mere generell metode):  
myfunc <- function(data,indices) {  
  bootdata <- data[indices,]  
  ret <- lm(Motivasjon~Sluttet6m,data=bootdata)$coefficients[2]  
  return(ret)  
}  
results <- boot(data=RoykeSlutt,statistic=myfunc,R=10000)  
boot.ci(results, type="perc")
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS  
## Based on 10000 bootstrap replicates  
##  
## CALL :  
## boot.ci(boot.out = results, type = "perc")  
##  
## Intervals :  
## Level      Percentile  
## 95%      (-0.0849,  1.7500 )  
## Calculations and Intervals on Original Scale
```