hw04

Zeng Geduo 2020302121048

2022/4/27

7.2

Refer to the law data (bootstrap). Use the jackknife-after-bootstrap method to estimate the standard error of the bootstrap estimate of se(R).

Solution

```
data_law <- as_tibble(bootstrap::law)

# Bootstrap
n <- 10
boot_ <- data_law %>%
  bootstrap(n)

grp_boot <- boot_$strap %>% map(as_tibble)

# define function of R
r <- function(tib, col1=1, col2=2) {
  return(cor(tib[, col1], tib[, col2])[[1]])
  }

# jackknife
r_boot <-  grp_boot %>% map_dbl(r)

jknf <- r_boot %>% jackknife(mean)

#result
jknf$jack.se
```

```
## [1] 0.03048046
```

7.3

Obtain a bootstrap t confidence interval estimate for the correlation statistic in Example 7.2 (law data in bootstrap).

Solution

First compute the $\hat{R}$.

```
r_mu <- r(data_law)
r_mu
```

```
## [1] 0.7763745
```

Then use bootstrap to compute the $\hat{se}\widehat{R^{(b)}}$ for every sample in 7.2.

```
# bootstrap for every sample
boot_grp_boot <- list()
grp_boot_grp_boot <- list()
sd_r_boot <- numeric(0)

for (i in 1:n){
  boot_grp_boot[i] <- bootstrap(grp_boot[[i]] %>%
                                  as.data.frame(), n)

  grp_boot_grp_boot[[i]] <- boot_grp_boot[[i]] %>%
    map(as_tibble) # convert to tibble

  sd_r_boot[[i]] <- grp_boot_grp_boot[[i]] %>%
      map_dbl(r) %>%
      sd()
}
```

Then Compute the t-statistics for every $\hat{se}\widehat{R^{(b)}}$ and compute the quantile of them.

```
t_boot <- (r_boot - r_mu) / sd_r_boot

alpha <- 0.05

Qt <- quantile(t_boot, c(alpha/2, 1-alpha/2), type = 1)
```

In the end, compute the sample standard deviation $\hat{se}\hat{R}$ in the first resampling and compute the Bootstrap t CI.

```
se_boot <- sd(r_boot)

r_mu + Qt * se_boot
```

```
##      2.5%     97.5%
## 0.6915912 1.0729670
```

```
#
# boot::boot.ci()
```

7.4

Refer to the air-conditioning data set *aircondit* provided in the *boot* package. The 12 observations are the times in hours between failures of airconditioning equipment [63, Example 1.1]:

3, 5, 7, 18, 43, 85, 91, 98, 100, 130, 230, 487.

Assume that the times between failures follow an exponential model Exp(λ). Obtain the MLE of the hazard rate λ and use bootstrap to estimate the bias and standard error of the estimate.

Solution

```
data_air <- boot::aircondit

vec_air_time_diff <- data_air %>%
  as_vector() %>%
  diff()
```

```r
vec_air_time_diff
```

```
##  hours2  hours3  hours4  hours5  hours6  hours7  hours8  hours9 hours10 hours11 
##       2       2      11      25      42       6       7       2      30     100 
## hours12 
##     257
```

Because the times between failures follow an exponential model Exp(λ), so the likelihood function is

$$L(\lambda) = \prod \lambda e^{\lambda X_i}$$

.

Considering that

$$lnL(\lambda) = nln\lambda - \lambda \sum X_i$$

$$\frac{\partial lnL}{\partial \lambda} = \frac{n}{\lambda} - \sum X_i = 0$$

so that

$$MLE(\lambda) = \frac{1}{\overline{X_i}}$$

```r
MLE_exp <- function(data, i){
  return(1 / mean(data[i]))
}
boot::boot(vec_air_time_diff, MLE_exp, n)
```

```
## 
## ORDINARY NONPARAMETRIC BOOTSTRAP
## 
## 
## Call:
## boot::boot(data = vec_air_time_diff, statistic = MLE_exp, R = n)
## 
## 
## Bootstrap Statistics :
##       original       bias    std. error
## t1* 0.02272727 -0.001535735  0.01180792
```