Lab05_MC

Christophoros Spyretos

11/29/2021

Question 2

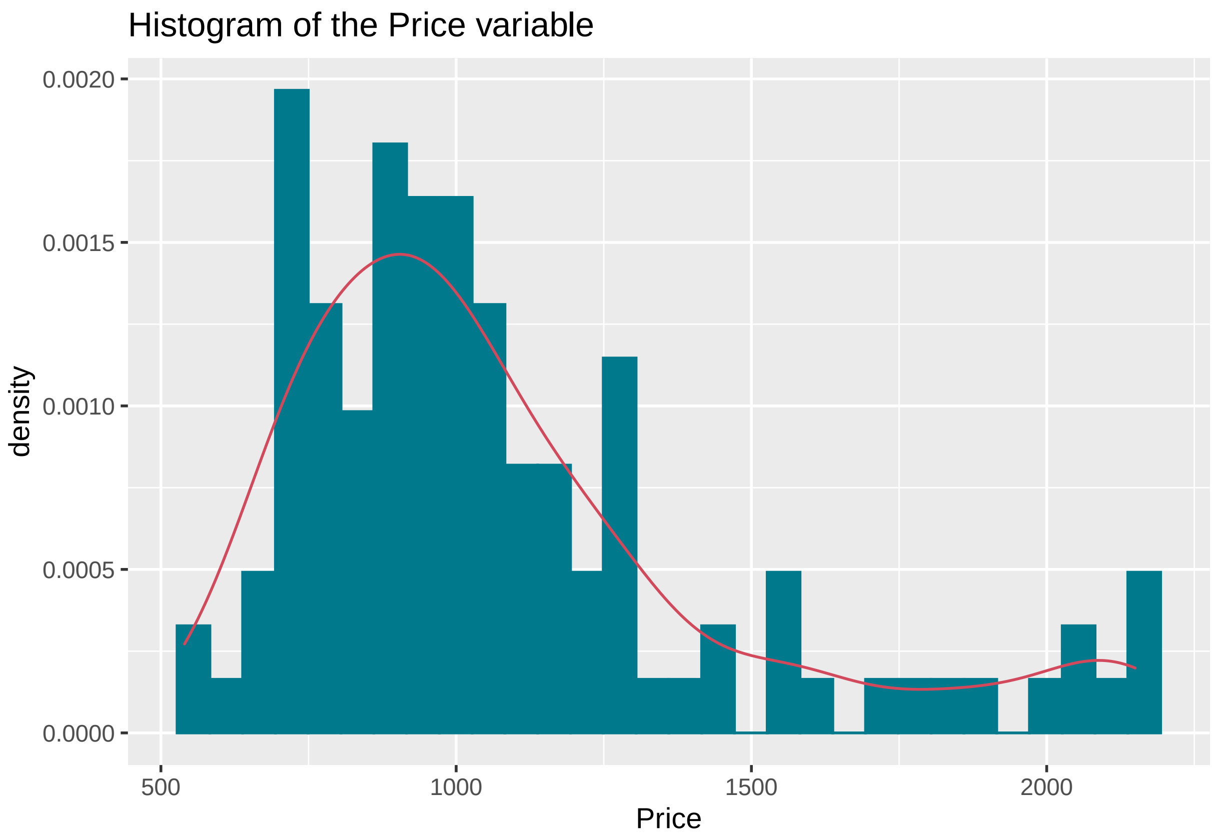
Task 1

```
data <- read.csv("prices1.csv", sep = ";")
```

```
library(ggplot2)
my_histogram <- ggplot(data, aes(x = Price)) + geom_histogram(bins = 30,
    color = "#00798c", fill = "#00798c", aes(y = ..density..)) +
    geom_density(colour = "#d1495b") + labs(title = "Histogram of the Price variable")
ylab("Density")
```

```
## $y
## [1] "Density"
##
## attr(,"class")
## [1] "labels"
```

```
my_histogram
```



```r
mean_value <- mean(data$Price)
mean_value
```

```
## [1] 1080.473
```

Task 2

The non-parametric bootstrap estimator of bias is given by:

$$\hat{T} = 2T(D) - \frac{1}{B}\sum_{i=1}^{B} T(D_i^*)$$

The variance of estimator is given by:

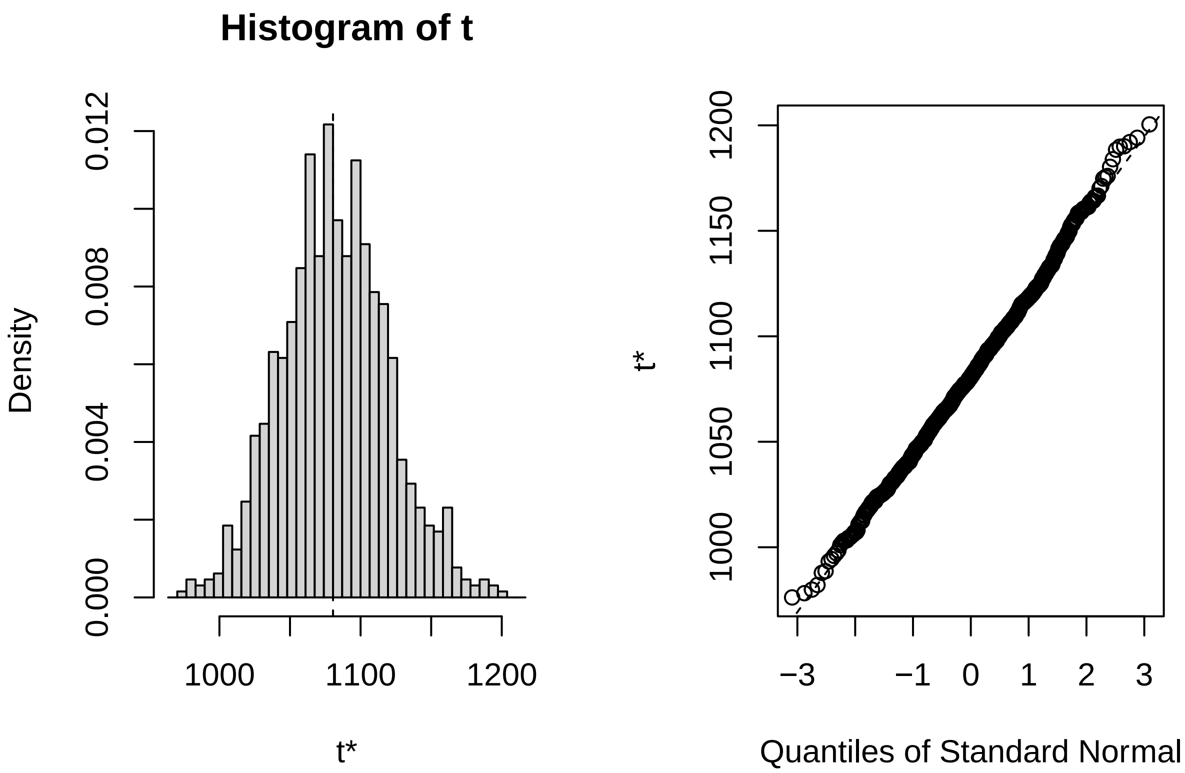
$$\widehat{Var[T(\cdot)]} = \frac{1}{B-1}\sum_{i=1}^{B}(T(D_i^*) - \overline{T(D^*)})^2$$

```r
library("boot")

# statistic
my_stat_fun <- function(data, index) {
    return(mean(data[index]))
}

# # typical bootstrap replicates 100-2000 B <- 100 B <- 200
# B <- 500
B <- 1000
# B <- 1500 B <- 2000

bootstrap <- boot(data = data$Price, statistic = my_stat_fun,
    R = B)
mean_estimator <- 2 * bootstrap$t0 - mean(bootstrap$t)
variance_estimator <- (1/(B - 1)) * sum((bootstrap$t - mean(bootstrap$t))^2)
ci <- boot.ci(boot.out = bootstrap, type = c("perc", "bca", "norm"))
plot(bootstrap)  # conf=0.95 default
```



Task 3

```
normal_low <- ci$normal[, 3] - ci$normal[, 2]
percent_low <- ci$percent[, 5] - ci$percent[, 4]
bca_low <- ci$bca[, 5] - ci$bca[, 4]

normal_upper <- ci$normal[, 3] - normal_low/2
percent_upper <- ci$percent[, 5] - percent_low/2
bca_upper <- ci$bca[, 5] - bca_low/2

intervals <- data.frame(Normal = c(normal_low, normal_upper),
    Percentile = c(percent_low, percent_upper), BCa = c(bca_low,
        bca_upper))

rownames(intervals) <- c("Low Interval", "Upper Interval")
knitr::kable(intervals)
```

	Normal	Percentile	BCa
Low Interval	147.2226	152.4158	149.1939
Upper Interval	1079.0048	1084.3497	1086.7890

Task 4

The variance of the mean price using the jackknife is given by $\widehat{Var[T(.)]} = \frac{1}{n(n-1)} \sum_{i=1}^{n} ((\mathrm{T_i^*}) - \mathrm{J(T)})^2$, where T_i^* is give by $ T_i^* = nT(D)-(n-1)T(D_i^*)$ and $J(T)$ is given by $J(T) = \frac{1}{n} \sum_{i=1}^{n} T_i^*$