Monte Carlo Simulation

Emilia Löscher

2022-09-14

8

##Monte Carlo simulation exercise

Perform a small simulation that does the following: a. Sample 100 samples from a standard normal distribution.

```
library(ggplot2)
library(dplyr)
```

```
##
## Attache Paket: 'dplyr'

## Die folgenden Objekte sind maskiert von 'package:stats':
##
##     filter, lag

## Die folgenden Objekte sind maskiert von 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
set.seed(1409)
N <- 100 #Number of samples
n <- 1000 #Size of each sample
samples <- matrix(rnorm(n*N), nrow = n, ncol = N, byrow = FALSE)
#samples in column of size 1000 (n)
```

b. For each of these samples, calculate the following statistics for the mean:

- absolute bias
- standard error
- lower bound of the 95% confidence interval
- upper bound of the 95% confidence interval

```
sample_mean <-  abs_bias <-  std_error <-  low_CI <-  up_CI <- length(N)
for(i in 1:ncol(samples)){
  sample_mean[i] <- mean(samples[,i])
  abs_bias[i] <- abs(sample_mean[i] - 0)
  std_error[i] <- sd(samples[,i])/sqrt(n)
```

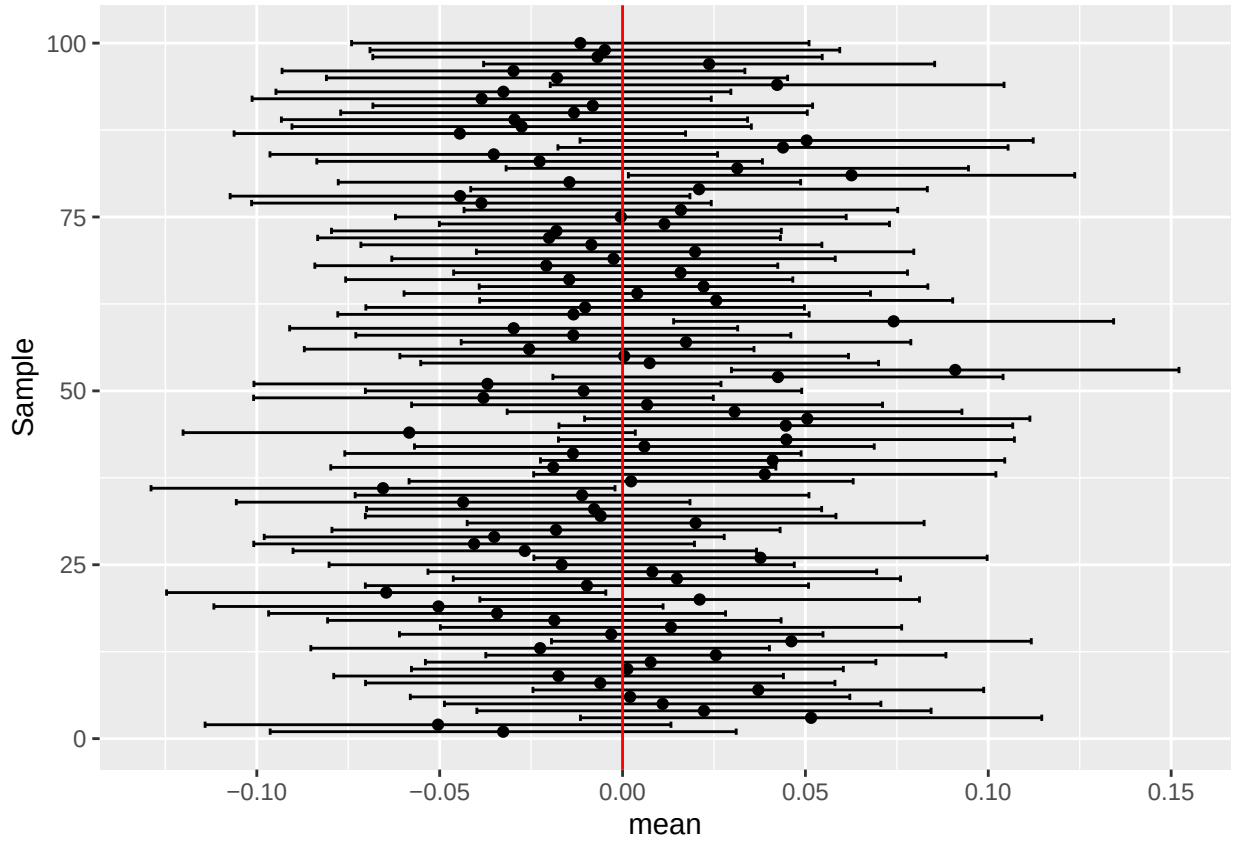
```
  low_CI[i] <- sample_mean[i] - 1.96 * std_error[i]
  up_CI[i] <- sample_mean[i] + 1.96 * std_error[i]
}

mcs_list <- list(sample_mean, abs_bias, std_error, low_CI, up_CI)
```

c. Create a plot that demonstrates the following: "A replication of the procedure that generates a 95% confidence interval that is centered around the sample mean would cover the population value at least 95 out of 100 times" (Neyman, 1934)

```
data <- data.frame(Sample = 1:N, mean= sample_mean, lower = low_CI, upper = up_CI)

ggplot(data, aes(mean, Sample)) +        # ggplot2 plot with confidence intervals
  geom_point() +
  geom_errorbar(aes(xmin = lower, xmax = upper))+
  geom_vline(xintercept = 0,color = "red", size = 0.5)
```



d. Present a table containing all simulated samples for which the resulting confidence interval does not contain the population value.

```
data %>% filter(lower > 0 | upper < 0)
```

```
##   Sample        mean        lower        upper
```

```
## 1     21 -0.06462372 -0.124675493 -0.004571948
## 2     36 -0.06549365 -0.128930587 -0.002056709
## 3     53  0.09095593  0.029770987  0.152140879
## 4     60  0.07411772  0.013961615  0.134273826
## 5     81  0.06260920  0.001594079  0.123624330
```

5 samples as expected when using a 95% confidence interval.