

Exercise2

Lina Kramer

1/11/2023

Monte Carlo simulation

```
set.seed(15)
library(dplyr)
library(tidyr)
library(plotrix)

data <- replicate (100, rnorm(200)) %>% as.data.frame()# 100 samples from a standard normal distribution

data <- data %>% pivot_longer(cols=everything(), names_to = "set", values_to = "val")

sumdata <- data %>% # means, bias, standard errors, CIs
  group_by(set) %>%
  summarise_each(funs(mean,
    bias= 0-mean,
    se=sd())/sqrt(n()),
    lower = mean - qnorm(0.975)*sd()/sqrt(n()),
    upper= mean + qnorm(0.975)*sd()/sqrt(n()))

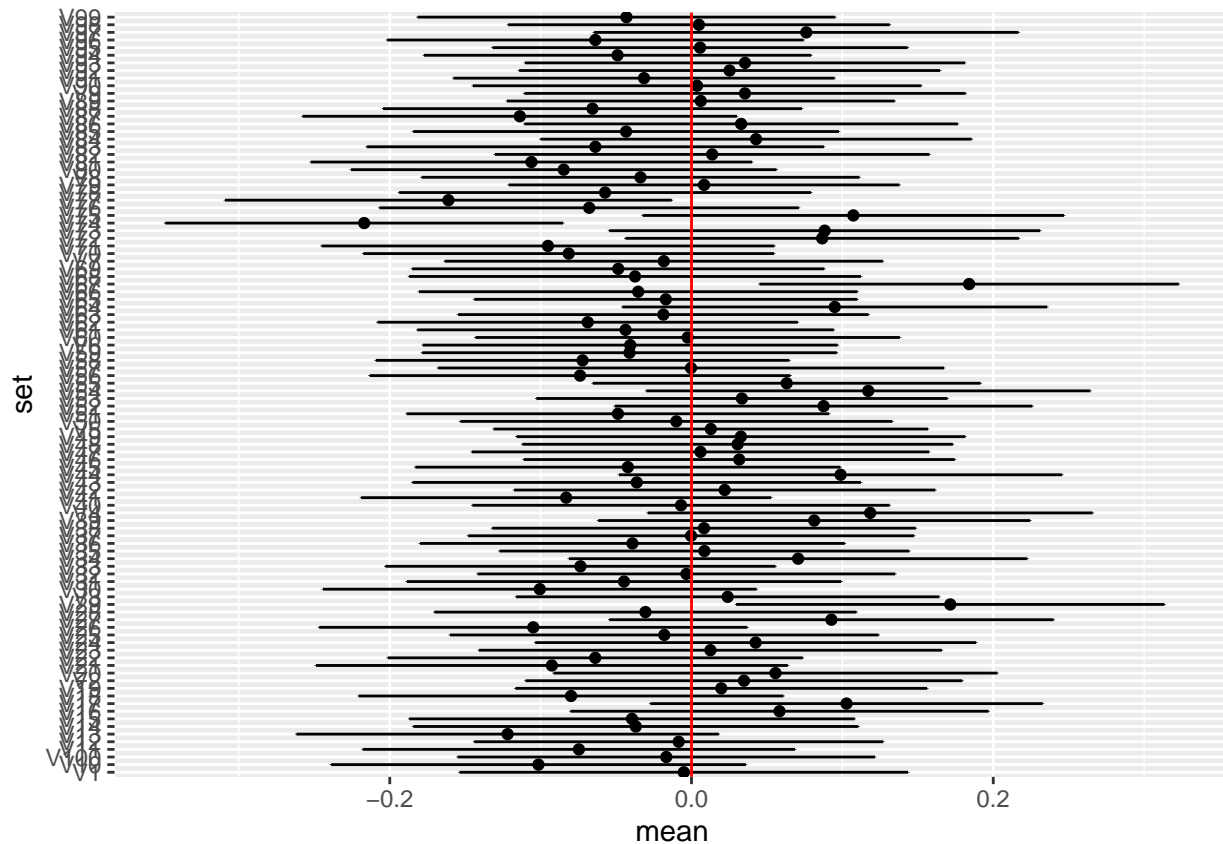
sumdata

## # A tibble: 100 x 6
##   set      mean      bias      se  lower upper
##   <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl>
## 1 V1     -0.00510  0.00510 0.0743 -0.151  0.141
## 2 V10    -0.101     0.101  0.0687 -0.236  0.0333
## 3 V100   -0.0167     0.0167 0.0691 -0.152  0.119
## 4 V11    -0.0746     0.0746 0.0716 -0.215  0.0657
## 5 V12    -0.00851  0.00851 0.0678 -0.141  0.124
## 6 V13    -0.122     0.122  0.0700 -0.259  0.0153
## 7 V14    -0.0369     0.0369 0.0738 -0.181  0.108
## 8 V15    -0.0394     0.0394 0.0738 -0.184  0.105
## 9 V16     0.0583   -0.0583 0.0693 -0.0774 0.194
## 10 V17    0.103    -0.103  0.0650 -0.0247 0.230
## # ... with 90 more rows

library(ggplot2)
library(gghighlight)
```

```
pd <- position_dodge(0.78)

ggplot(sumdata, aes(x=mean, y = set)) +
  #draws the means
  geom_point(position = pd) +
  #draws the CI error bars
  geom_errorbar(data=sumdata, aes(xmin=mean-2*se, xmax=mean+2*se), width=.1, position = pd)+
  geom_vline(xintercept = 0, color ="red")
```



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

```
## # A tibble: 4 x 6
##   set    mean  bias    se  lower  upper
##   <chr> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 V29   0.172 -0.172 0.0710 0.0324 0.311
## 2 V67   0.184 -0.184 0.0695 0.0479 0.320
## 3 V74  -0.217  0.217 0.0659 -0.346 -0.0878
## 4 V77  -0.161  0.161 0.0740 -0.306 -0.0160
```