# Exercise 1

Yongchao Terry Ma

9/19/2020

### Sample 100 samples from a standard normal distribution

```
set.seed(1234)

# 100 samples of 1000 sampled units
N <- 100
sims <- replicate(N, rnorm(5000, 0, 1))
```

### For each of these samples, calculate the following statistics for the mean

```
# mean
mean <- apply(sims, 2, mean)

# absolute bias
bias <- abs(mean - 0)

# standard error
se <- 1/sqrt(5000)

# lower bound of the 95% confidence interval
ci_lower <- mean - qt(.975, 4999)*se

# upper bound of the 95% confidence interval
ci_upper <- mean + qt(.975, 4999)*se
```

### Create a plot

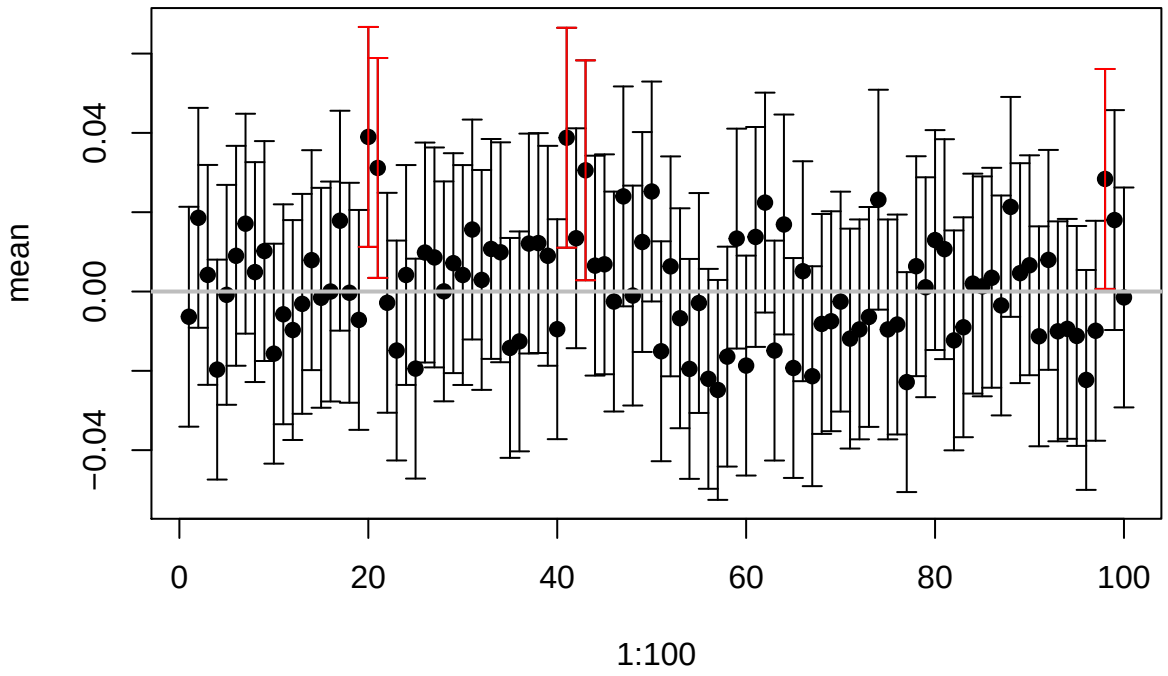
```
# find 5 CIs don't cover the true mean
index <- which(ci_lower > 0 | ci_upper < 0)

# plot all sample means with CI
plot(1:100, mean, ylim = range(c(ci_lower, ci_upper)), pch = 19)
arrows(1:100, ci_lower, 1:100, ci_upper,
       length = 0.05, angle = 90, code = 3)

# mark 5 CIs as red
```

```
arrows(index, ci_lower[index], index, ci_upper[index],
       length = 0.05, angle = 90, code = 3, col = "red")

# add a line of the true mean
abline(h = 0, lwd = 2, col = "grey")
```




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

```
tbl <- matrix(c(index, mean[index], bias[index], ci_lower[index], ci_upper[index]),
       nrow = 5)
colnames(tbl) <- c("index", "mean", "bias", "lower", "upper")
tbl
```

```
##      index       mean       bias        lower      upper
## [1,]    20 0.03898230 0.03898230 0.0112575122 0.06670709
## [2,]    21 0.03115910 0.03115910 0.0034343089 0.05888389
## [3,]    41 0.03877999 0.03877999 0.0110551983 0.06650478
## [4,]    43 0.03056728 0.03056728 0.0028424944 0.05829207
## [5,]    98 0.02838466 0.02838466 0.0006598667 0.05610945
```