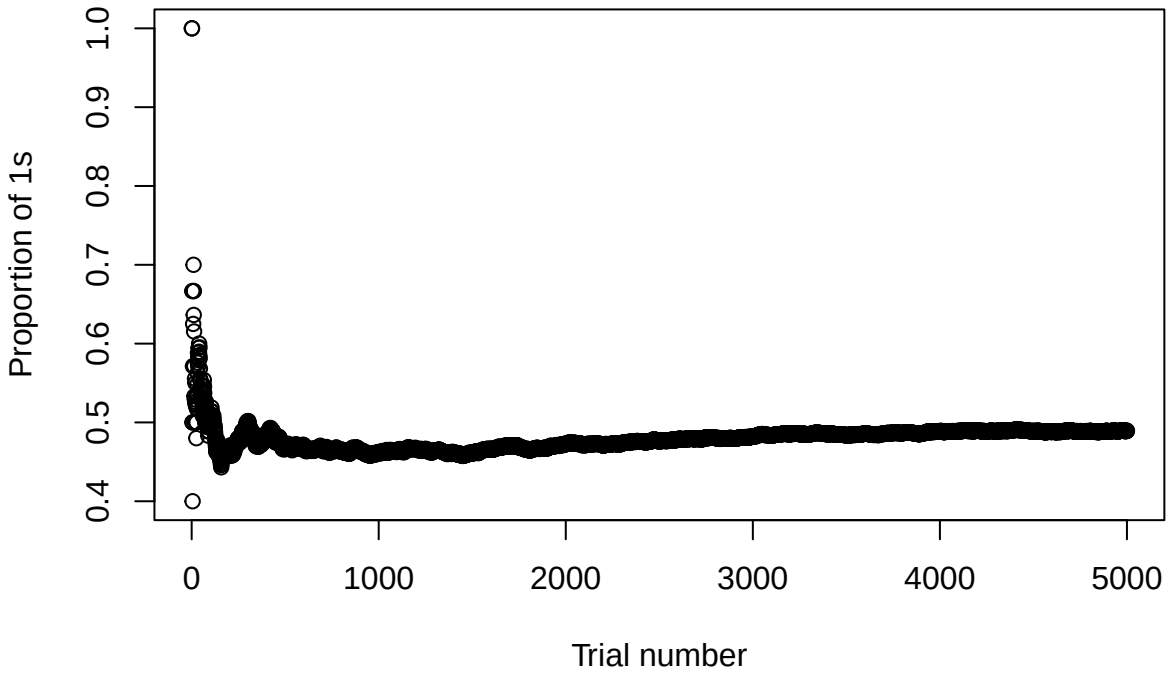
BB503/BB602 - R Training - Week IV

Ege Ulgen

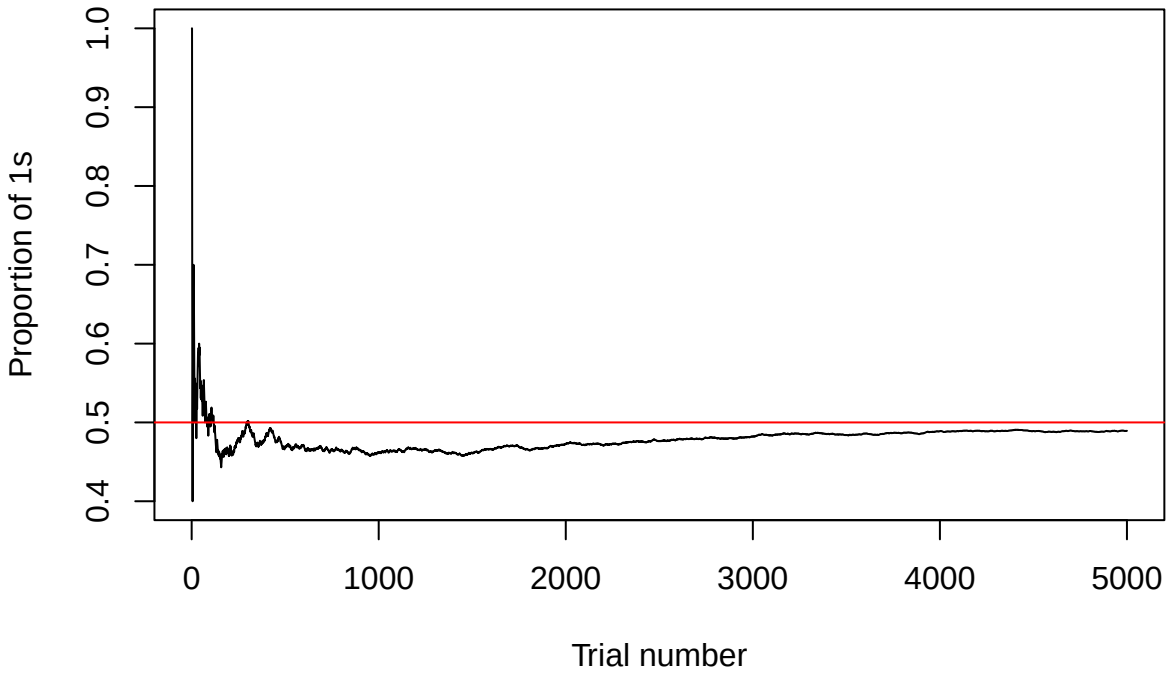
Probability

Basic Simulation - Flipping a fair coin

```r
# by this you get the same answer all the time
set.seed(321)

# number of trials
n <- 5000

# tossing coins
results <- c() # to store results
average_at_i <- c() # to store averages (will approximate 0.5)
for (i in 1:n) {
    current_res <- sample(c(0, 1), 1)
    results <- c(results, current_res)
    average_at_i <- c(average_at_i, mean(results))
}

# plot resulting averages by each trial
plot(average_at_i, xlab = "Trial number", ylab = "Proportion of 1s")
```




```
plot(average_at_i, xlab = "Trial number", ylab = "Proportion of 1s", type = "l")
abline(h = 0.5, col = "red")
```



Bayes Theorem

```
bayesTheorem <- function(pA, pB, pBA) {
  pAB <- pA * pBA / pB
  return(pAB)
}
```

Given the following statistics, what is the probability that a woman over 50 has cancer if she has a positive mammogram result?

- One percent of women over 50 have breast cancer
- Ninety percent of women who have breast cancer test positive on mammograms
- Eight percent of women will have false positives

```
bayesTheorem(0.01, 0.9 * 0.01 + 0.08 * 0.99, 0.90)
```

```
## [1] 0.10204
```