

# Assignment 4

Anant Vishwakarma

5/4/2022

Suppose that we sample 10 numbers randomly *without* replacement from the population above. What would the standard error for  $\hat{p}$  be? Is this larger or smaller than the corresponding value when sampling with replacement? Using a simulation, estimate the mean and the SD of the estimated standard error for  $\hat{p}$  when sampling without replacement. Would you say that, in this case, the estimated standard error also tends to be fairly close to the (true) standard error?

## Discussion of Results

The standard error for  $\hat{p}$  is 0.1059903. When sampling with replacement, in the above literature we saw the standard error for  $\hat{p}$  was 0.1449138, so the value we obtained without replacement was lower than the one obtained with replacement. Thus sampling without replacement seems to have lower error.

We then used a simulation to calculate the estimated mean and standard deviation of the estimated standard error. The estimated standard error mean is equal to 0.1020322 and the estimated standard error standard deviation is equal to 0.0134216.

The simulated estimated SD of the standard error is a relatively small number i.e the observations do not differ too much from the central value, thus we can say that due to the small value of the SD of the standard error, the calculated standard error is fairly close to the true standard error.

## Appendix: R Script

```
knitr::opts_chunk$set(echo = FALSE)
knitr::opts_chunk$set(echo = FALSE)
pop <- c(1,0,1,1,0,1,1,1,0,1,0,1,1,0,1,1,1,0)
N <- length(pop)

simulatePhat <- function(n, population){ # this simulates a p hat value
  phat <- mean(sample(population, n, replace=FALSE))
  phat # outputs phat
}
simulatedPhats <- replicate(1000, # 1000 simulations
  simulatePhat(10,pop)
)
simulatedPhatsFreqs <- table(simulatedPhats) # get frequencies for phat simulations
simulatedPhatsMean <- mean(simulatedPhats)
simulatedPhatsSD <- sd(simulatedPhats)

#the function we want to use
#but we want to sample without replacement so what do we need to change?
#hint: we only need to change one word
simulateEstSE <- function(n, population){
  phat <- mean(sample(population, n, replace=FALSE))
  estSE <- sqrt((phat*(1-phat)/n)*(N-n)/(N-1))
  estSE
}

# 1000 simulations
simulatedEstSEs <- replicate(1000, simulateEstSE(10,pop))

#our average standard error in 1000 simulations
simulatedEstSEsMean <- mean(simulatedEstSEs)
#the standard deviation (spread) of our errors in 1000 simulations
simulatedEstSEsSD <- sd(simulatedEstSEs)
```