title: “Chapter 4” author: “Akanksha” date: “2024-09-23” output: html\_document —

### 1. Expected value of

The expected value of in a binomial distribution is , where: - (the sample size) - is the proportion of Democrats in the population

N <- 25  
p <- 0.5   
E\_S <- N \* p  
E\_S

## [1] 12.5

### 2. Standard error of

The standard error (SE) of in a binomial distribution is .

SE\_S <- sqrt(N \* p \* (1 - p))  
SE\_S

## [1] 2.5

### 3. Expected value of

is the sample average, which is equivalent to . The expected value of is just because the sample average should estimate the proportion of Democrats.

E\_X\_bar <- p  
E\_X\_bar

## [1] 0.5

### 4. Standard error of

The standard error of is .

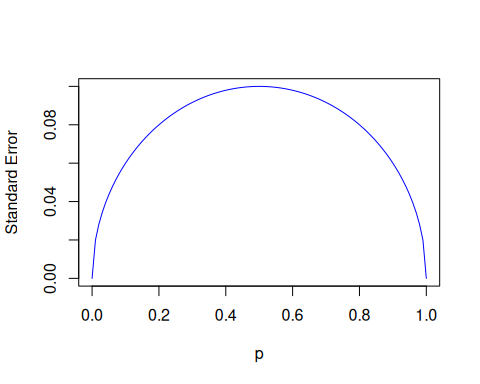
SE\_X\_bar <- sqrt(p \* (1 - p)) / sqrt(N)  
SE\_X\_bar

## [1] 0.1

### 5. Plotting the standard error for different values of

To plot the standard error for different values of , we can use a sequence of values ranging from 0 to 1 and plot versus .

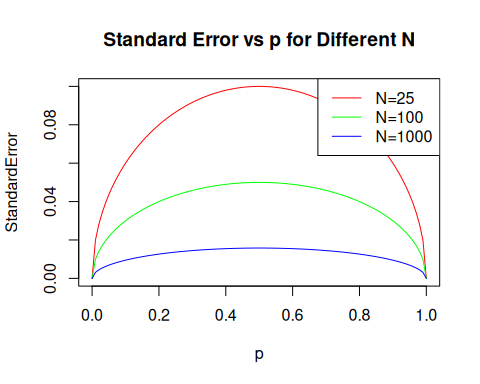
p\_vals <- seq(0, 1, length = 100)  
se\_vals <- sqrt(p\_vals \* (1 - p\_vals)) / sqrt(N)  
plot(p\_vals, se\_vals, type = "l", col = "blue", xlab = "p", ylab = "Standard Error")



### 6. Plotting for different values of

To extend the previous code to plot for , , and , you can use a loop.

# Define values for p  
p\_vals <- seq(0, 1, length = 100)  
  
# Define the values for N and their corresponding colors  
N\_vals <- c(25, 100, 1000)  
colors <- c("red", "green", "blue")  
  
# Create the initial plot with the first set of N = 25  
N <- N\_vals[1]  
se\_vals <- sqrt(p\_vals \* (1 - p\_vals)) / sqrt(N)  
plot(p\_vals, se\_vals, type = "l", col = colors[1], xlab = "p", ylab ="StandardError", ylim = c(0, max(se\_vals)), main = "Standard Error vs p for Different N")  
  
# Now, add the lines for N = 100 and N = 1000  
for (i in 2:length(N\_vals)) {  
 N <- N\_vals[i]  
 se\_vals <- sqrt(p\_vals \* (1 - p\_vals)) / sqrt(N)  
 lines(p\_vals, se\_vals, col = colors[i])  
}  
  
# Add a legend to the plot  
legend("topright", legend = c("N=25", "N=100", "N=1000"), col = colors, lty = 1)



### 7. Expected value of

The difference is , which simplifies to . The expected value of is .

E\_d <- 2 \* p - 1  
E\_d

## [1] 0

### 8. Standard error of

Since , the standard error of is .

SE\_d <- 2 \* SE\_X\_bar  
SE\_d

## [1] 0.2

### 9. Standard error with

If , we can compute the standard error for with .

p <- 0.45  
N <- 25  
SE\_X\_bar <- sqrt(p \* (1 - p)) / sqrt(N)  
SE\_d <- 2 \* SE\_X\_bar  
SE\_d

## [1] 0.1989975

### 10. Strategy for using

Given the standard error computed above, you can evaluate whether the standard error is large relative to the difference and decide if the sample size is sufficient.