**Задача 1**

sample\_sizes <- c(30, 50, 100, 200, 500)

size <- length(sample\_sizes)

num\_tests <- 10000

independent\_sample\_test <- function(x, y) {

result <- t.test(x, y)

return(result$p.value > 0.05)

}

paired\_sample\_test <- function(x, y) {

result <- t.test(x, y, paired = TRUE)

return(result$p.value > 0.05)

}

test\_independent <- function(n) {

result <- replicate(10000, {

x <- rnorm(n, mean = 7, sd = 1)

e <- rnorm(n, mean = 0.2, sd = 1)

y <- x + e

independent\_sample\_test(x, y)

})

return(result)

}

test\_paired <- function(n) {

result <- replicate(10000, {

x <- rnorm(n, mean = 7, sd = 1)

e <- rnorm(n, mean = 0.2, sd = 1)

y <- x + e

paired\_sample\_test(x, y)

})

return(result)

}

results\_independent <- rep(0, times = size)

results\_paired <- rep(0, times = size)

for (i in 1:size) {

independent <- test\_independent(sample\_sizes[i])

paired <- test\_paired(sample\_sizes[i])

results\_independent[i] <- (sum(independent)/length(independent))

results\_paired[i] <- (sum(paired)/length(paired))

}

plot(sample\_sizes, results\_independent, type="b", pch=19, col="blue",

ylim = c(0, 1), xlab = "N", ylab = "Ratio", main = "Plot of Two Vectors")

lines(sample\_sizes, results\_paired, col = "red", type = "b", pch = 19)

legend("topright", legend = c("Independent", "Paired"), col = c("blue", "red"), lty = 1, pch = 19)

**Резултат:**

За n = 30; 50; 100; 200; 500

Процентът на верни заключения за две независими извадки и за двойки наблюдения са съответно:

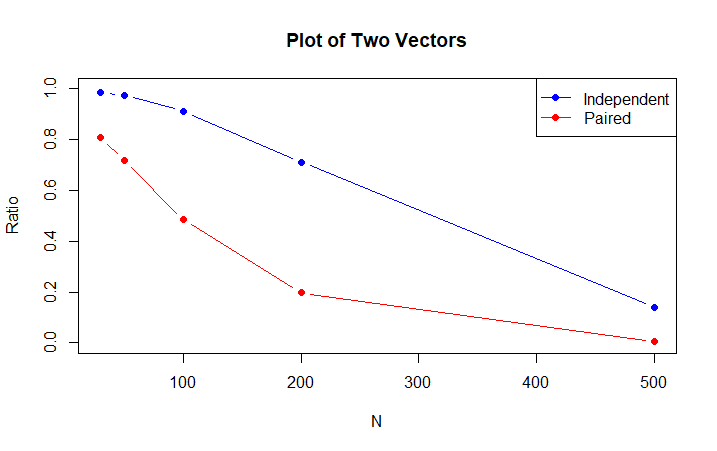
N = 30; Independent = 98.53 % ; Paired = 80.84 %

N = 50; Independent = 97.51 % ; Paired = 71.77 %

N = 100; Independent = 91.21 % ; Paired = 48.44 %

N = 200; Independent = 71.17 % ; Paired = 19.67 %

N = 500; Independent = 13.97 % ; Paired = 0.057 %



**Задача 2**

num\_simulations <- 10000

n\_values <- c(5, 10, 20, 30, 50, 100)

prob <- c(0.1, 0.1, 0.1, 0.1, 0.1, 0.5)

simulate\_rolls <- function(n, probabilities) {

sides <- length(probabilities)

rolls <- sample(1:sides, n, replace = TRUE, prob = probabilities)

return(rolls)

}

test\_hypothesis <- function(rolls) {

r <- as.numeric(table(factor(rolls, levels = 1:6)))

result <- chisq.test(r)

return(result$p.value > 0.05)

}

test\_sample <- function(n) {

result <- replicate(10000, {

rolls <- simulate\_rolls(n, prob)

test\_hypothesis(rolls)

})

return(result)

}

results <- c(1:6)

for (i in 1:length(n\_values)) {

rolls <- test\_sample(n\_values[i])

results[i] <- sum(rolls)/num\_simulations

}

plot(n\_values, results, type="b", pch=19, col="blue",

ylim = c(0, 1), xlab = "N", ylab = "Ratio", main = "Plot of Vector")

**Резултат:**

За n = 5; 10; 20; 30; 50; 100

Процентът вярно заключение на теста е съответно:

N = 5; SameProb = 81.01%

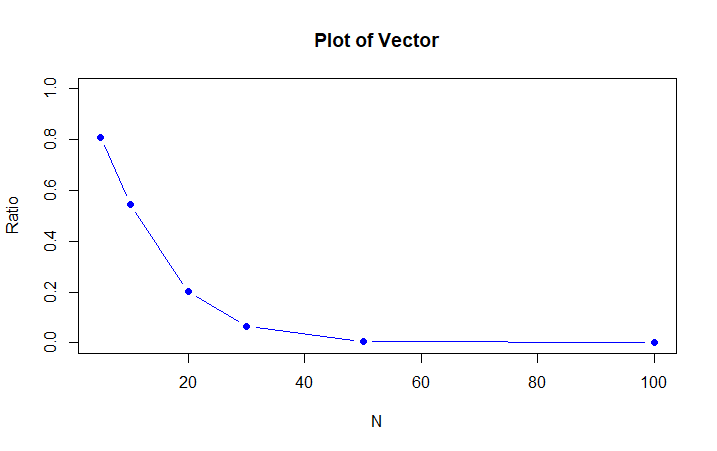
N = 10; SameProb = 54.59 %

N = 20; SameProb = 20.37 %

N = 30; SameProb = 6.62 %

N = 50; SameProb = 0.4 %

N = 100; SameProb = 0 %

Имаме условие, което изисква за най-голямата стойност на n честотата на вярно заключение да е 98%, но очевидно с увеличаването на броя хвърляния все по-рядко ще имаме разпределение на цифрите с еднаква вероятност, затова моята графика изглежда по този начин.