

Analiza rozkładu binomialnego na próbach losowych.

Próby losowe składające się z  $M = 1000$  próbek i wykres wartości wygenerowanych danych

```
M <- 1000          # Liczba prób
n1 <- 20           # Liczba prób dla każdej liczby
p1 <- 0.2          # Prawdopodobieństwo sukcesu w każdej próbie
n2 <- 20
p2 <- 0.8

proba_1 <- rbinom(M, size = n1, prob = p1)
proba_2 <- rbinom(M, size = n2, prob = p2)

head(proba_1)
```

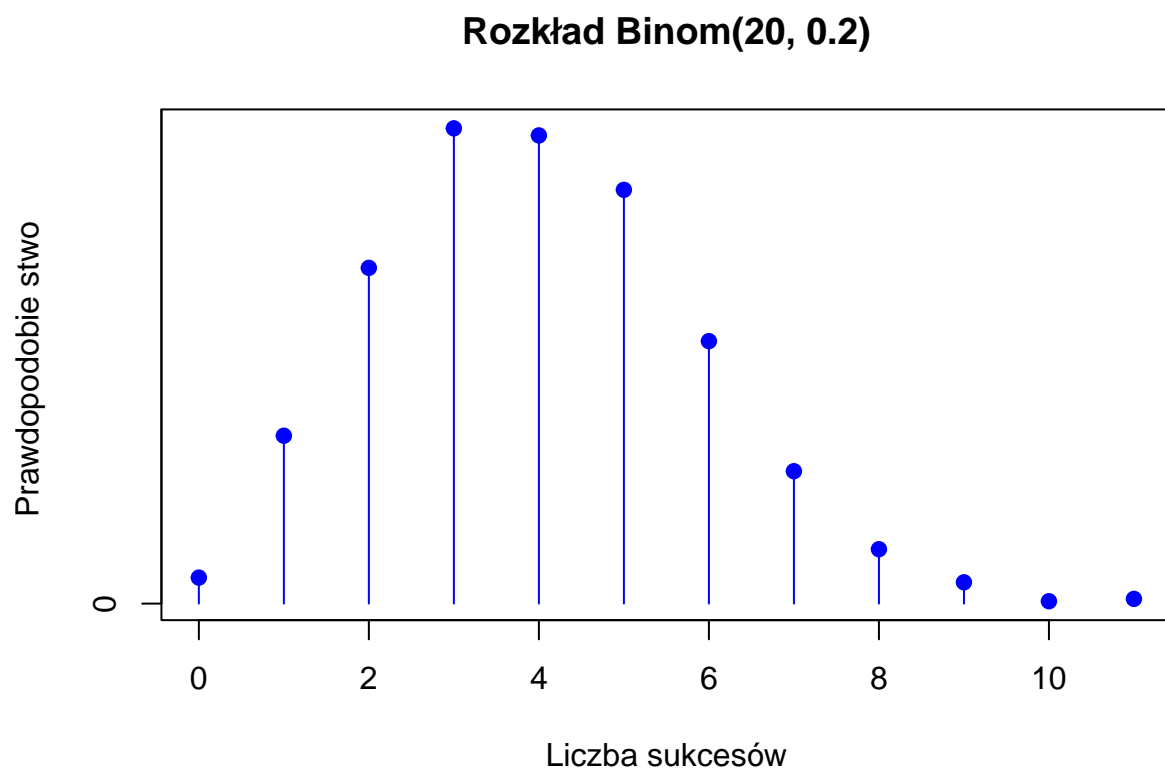
```
## [1] 5 6 5 4 7 2
```

```
head(proba_2)
```

```
## [1] 16 18 14 15 20 19
```

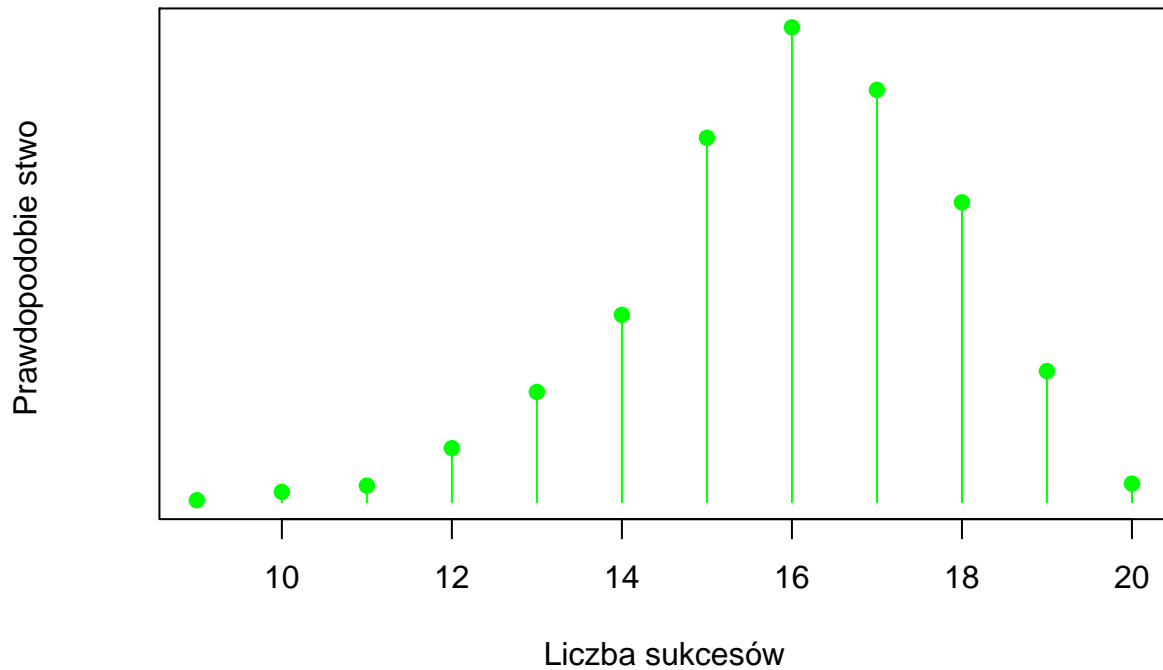
```
czestosc_1 <- table(proba_1)          # Liczenie ilości wystąpień danej liczby sukcesów
prawd_1 <- czestosc_1 / M              # Obliczanie prawdopodobieństwa dla każdej liczby sukcesów
czestosc_2 <- table(proba_2)
prawd_2 <- czestosc_2 / M

plot(as.numeric(names(prawd_1)), prawd_1, type = "h", col = "blue", xlab = "Liczba sukcesów",
     ylab = "Prawdopodobieństwo", main = "Rozkład Binom(20, 0.2)")
points(as.numeric(names(prawd_1)), prawd_1, col = "blue", pch = 19)
```



```
plot(as.numeric(names(prawd_2)), prawd_2, type = "h", col = "green", xlab = "Liczba sukcesów",  
     ylab = "Prawdopodobieństwo", main = "Rozkład Binom(20, 0.2)")  
points(as.numeric(names(prawd_2)), prawd_2, col = "green", pch = 19)
```

## Rozkład Binom(20, 0.8)

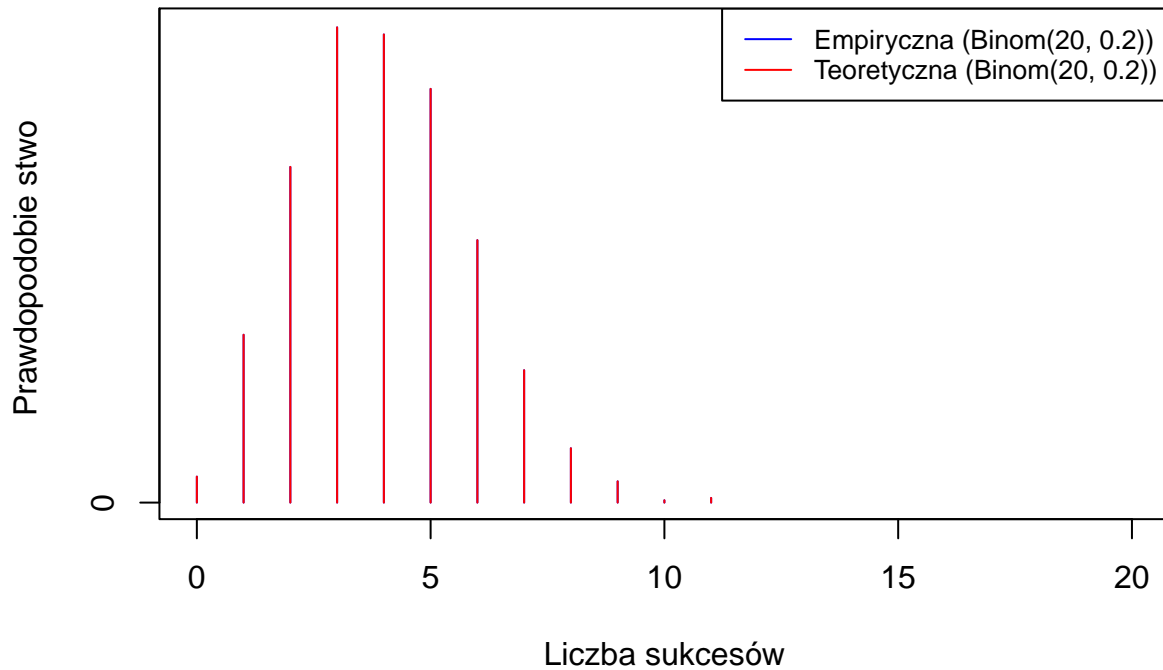


Empiryczne i teoretyczne funkcje prawdopodobieństwa

```
wyniki <- 0:n1
prawd_teor_1 <- dbinom(wyniki, size = n1, prob = p1)
wyniki <- 0:n2
prawd_teor_2 <- dbinom(wyniki, size = n2, prob = p2)

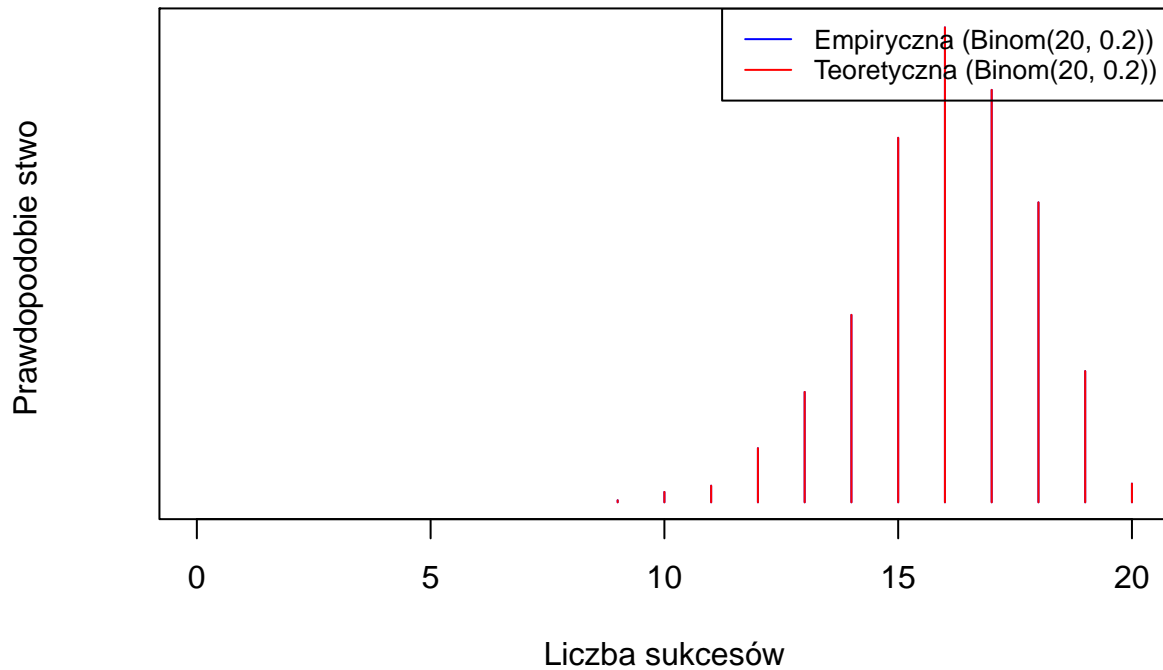
plot(as.numeric(names(prawd_1)), prawd_1, type = "h", col = "blue", xlab = "Liczba sukcesów", ylab = "Prawdopodobieństwo",
     main = "Empiryczna i teoretyczna funkcja prawdopodobieństwa",
     xlim = c(0, 20))
lines(as.numeric(names(prawd_1)), prawd_1, type = "h", col = "red")
legend("topright", legend = c("Empiryczna (Binom(20, 0.2))", "Teoretyczna (Binom(20, 0.2))"),
     col = c("blue", "red"), lty = 1, cex = 0.8)
```

## Empiryczna i teoretyczna funkcja prawdopodobieństwa



```
plot(as.numeric(names(prawd_2)), prawd_2, type = "h", col = "blue", xlab = "Liczba sukcesów", ylab = "Prawdopodobieństwo",
     main = "Empiryczna i teoretyczna funkcja prawdopodobieństwa",
     xlim = c(0, 20))
lines(as.numeric(names(prawd_2)), prawd_2, type = "h", col = "red")
legend("topright", legend = c("Empiryczna (Binom(20, 0.2))", "Teoretyczna (Binom(20, 0.2))"),
      col = c("blue", "red"), lty = 1, cex = 0.8)
```

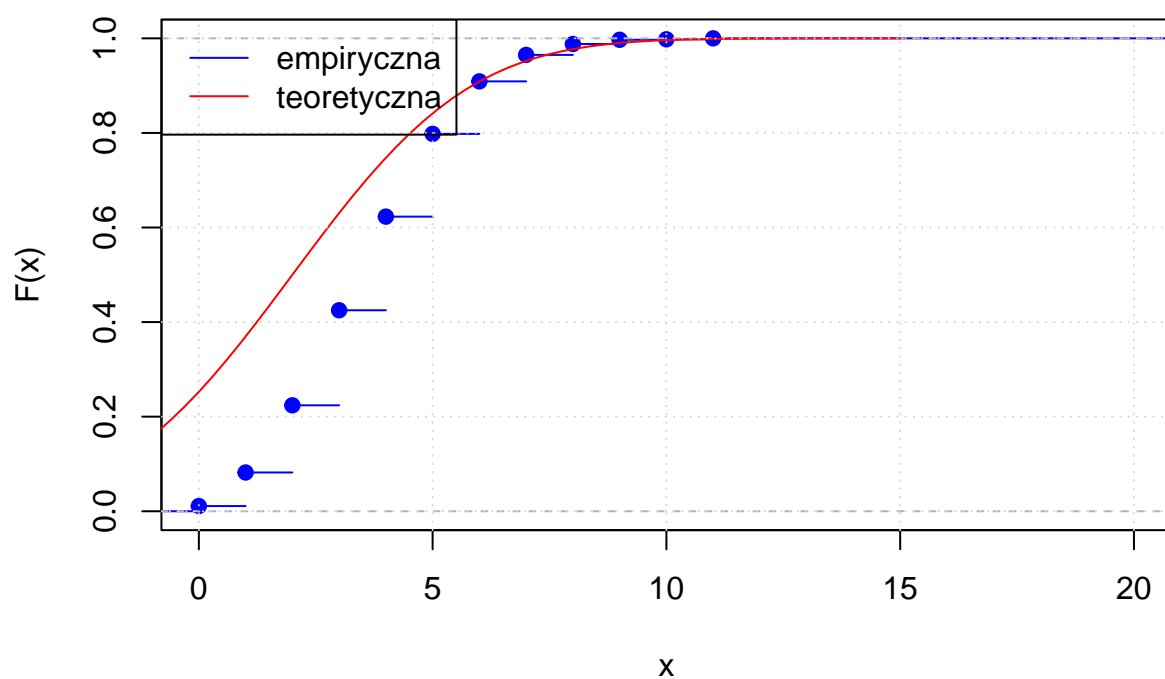
## Empiryczna i teoretyczna funkcja prawdopodobieństwa



Empiryczne i teoretyczne dystrybuanty

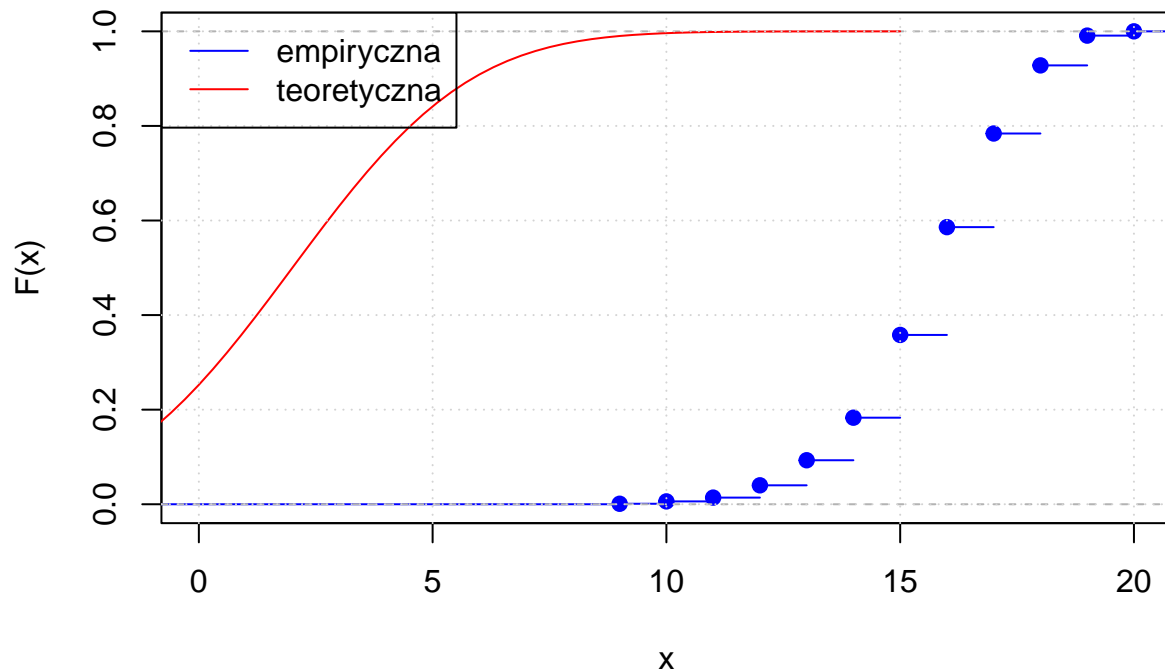
```
plot(ecdf(proba_1), xlab = 'x', ylab = 'F(x)', main = 'Dystrybuanta', col = 'blue', xlim = c(0, 20))
curve(pnorm(x, mean = 2, sd = 3), add = T, col = 'red', -15, 15)
grid()
legend('topleft', c('empiryczna', 'teoretyczna'),
      col = c('blue', 'red'), lwd = 1)
```

## Dystrybuanta



```
plot(ecdf(proba_2), xlab = 'x', ylab = 'F(x)', main = 'Dystrybuanta', col = 'blue', xlim = c(0, 20))
curve(pnorm(x, mean = 2, sd = 3), add = T, col = 'red', -15, 15)
grid()
legend('topleft', c('empiryczna', 'teoretyczna'),
      col = c('blue', 'red'), lwd = 1)
```

## Dystrybuanta



Próby losowe składające się z  $M = 100, 1000$  i  $10000$  próbek

```
M_values <- c(100, 1000, 10000)
n <- 20
p <- 0.2
```

Funkcja do generowania prób, rysowania wykresów i obliczeń

```
analize_sample <- function(M) {
  # Generowanie próby
  sample <- rbinom(M, size = n, prob = p)

  # Empiryczna funkcja prawdopodobieństwa
  x_vals <- 0:max(sample)
  #freq <- table(factor(sample, levels = x_vals)) / M
  freq <- as.vector(table(factor(sample, levels = x_vals))) / M

  # Wykres funkcji prawdopodobieństwa
  plot(freq ~ x_vals, type = "h", col = "blue", xlab = "x", ylab = "f(x)",
       main = paste("Funkcja prawdopodobieństwa dla M =", M))
  grid()
  points(freq ~ x_vals, col = "blue")

  # Teoretyczna funkcja prawdopodobieństwa
  prob_theoretical <- dbinom(x_vals, size = n, prob = p)
  lines(prob_theoretical ~ x_vals, type = "h", col = "red")
  points(prob_theoretical ~ x_vals, col = "red")
}
```

```

legend("topright", c("empiryczna", "teoretyczna"), col = c("blue", "red"), lwd = 1)

# Empiryczna dystrybuanta
cum_freq <- cumsum(freq)
plot(cum_freq ~ x_vals, type = "s", col = "blue", xlab = "x", ylab = "F(x)",
     main = paste("Dystrybuanta dla M =", M))
grid()
points(cum_freq ~ x_vals, col = "blue")

# Teoretyczna dystrybuanta
cdf_theoretical <- pbinom(x_vals, size = n, prob = p)
lines(cdf_theoretical ~ x_vals, type = "s", col = "red")
points(cdf_theoretical ~ x_vals, col = "red")
legend("topleft", c("empiryczna", "teoretyczna"), col = c("blue", "red"), lwd = 1)

# Obliczenia statystyk empirycznych
mean_empirical <- mean(sample)
variance_empirical <- var(sample)

list(mean_empirical = mean_empirical, variance_empirical = variance_empirical)
}

```

Teoretyczne wartości średniej i wariancji

```

mean_theoretical <- n * p
variance_theoretical <- n * p * (1 - p)

```

Analiza dla każdej liczby prób i podsumowanie wyników

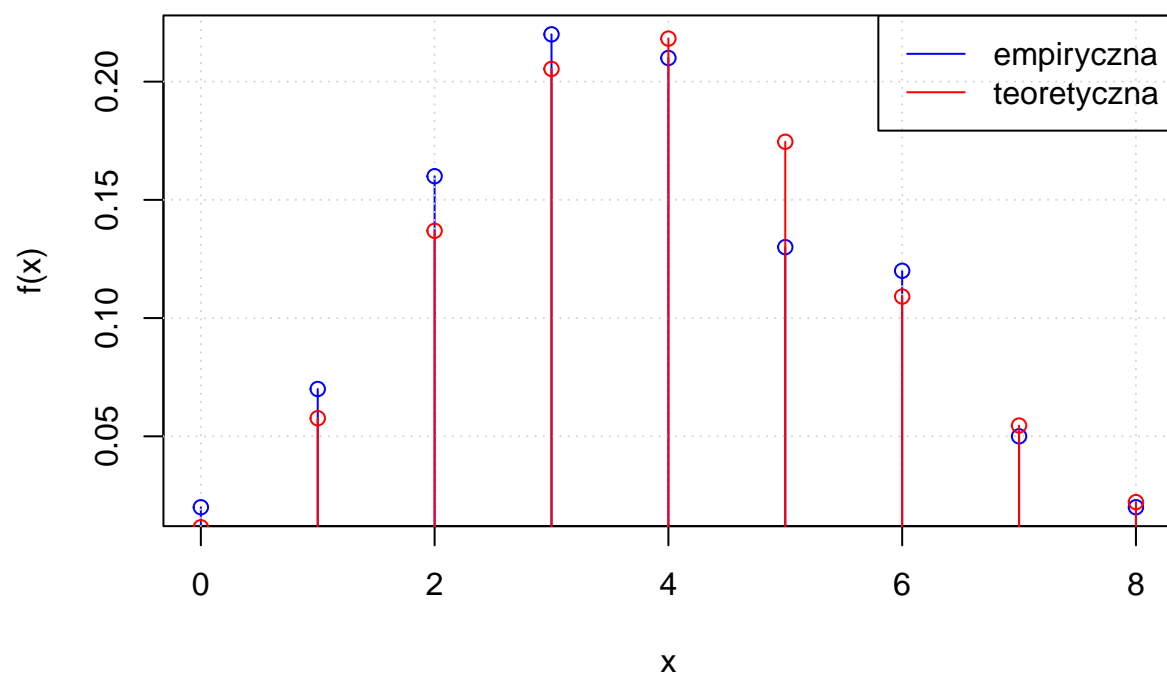
```

results <- lapply(M_values, analyze_sample)

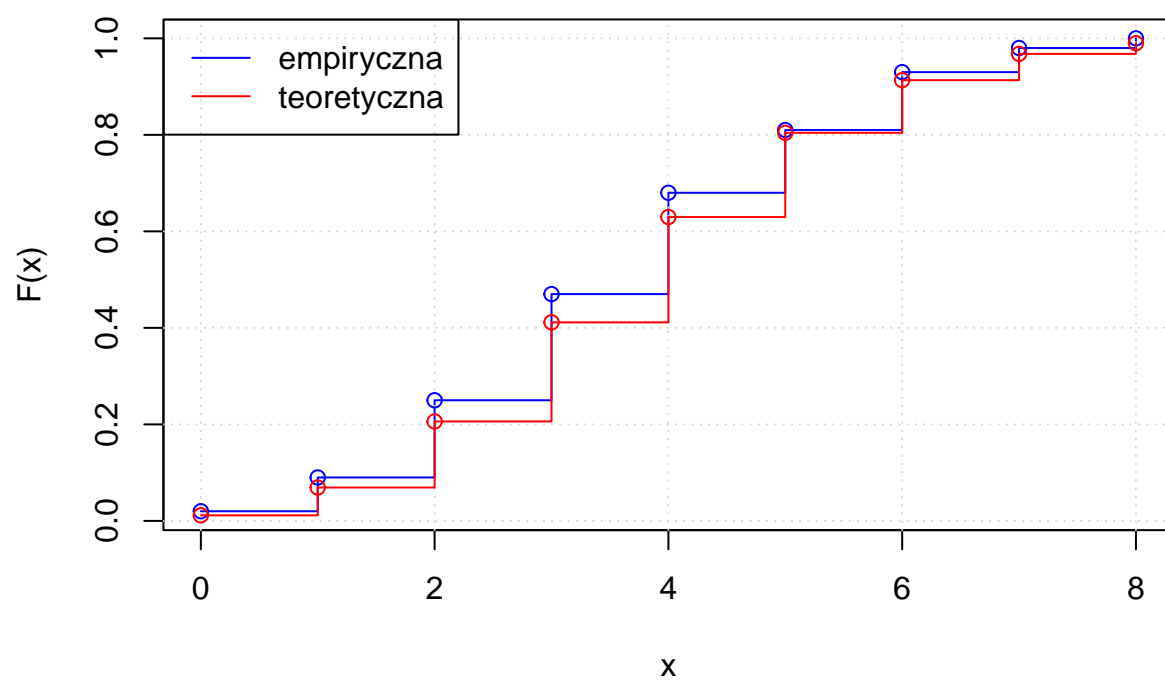
```



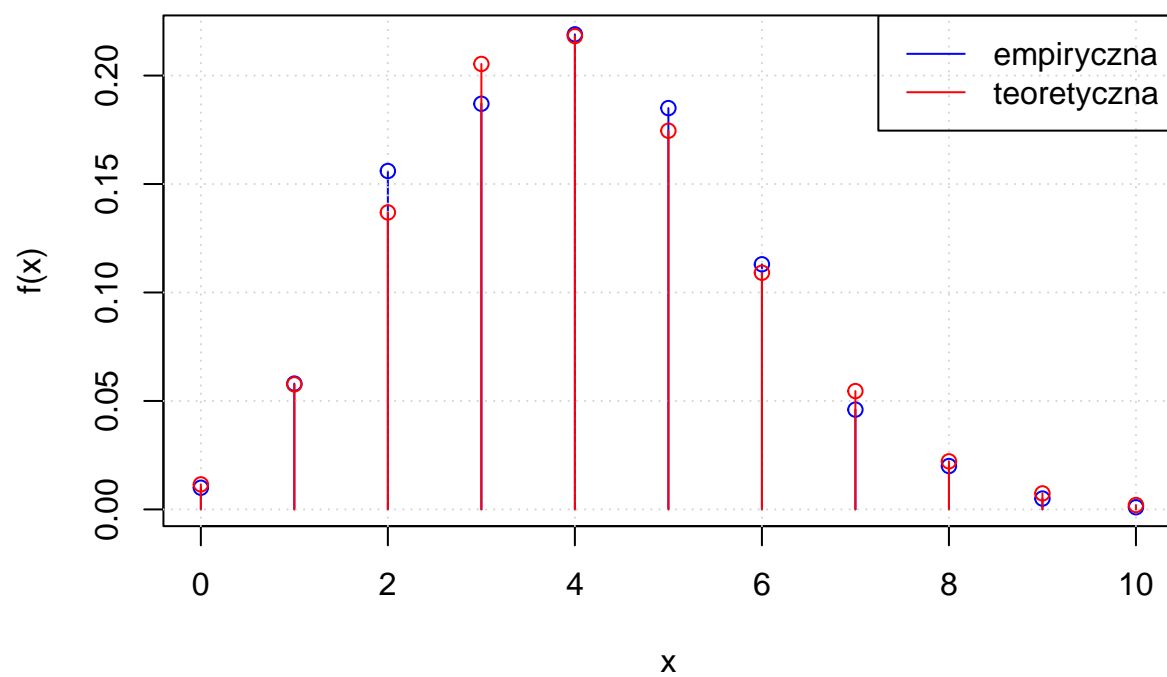
### Funkcja prawdopodobieństwa dla $M = 100$



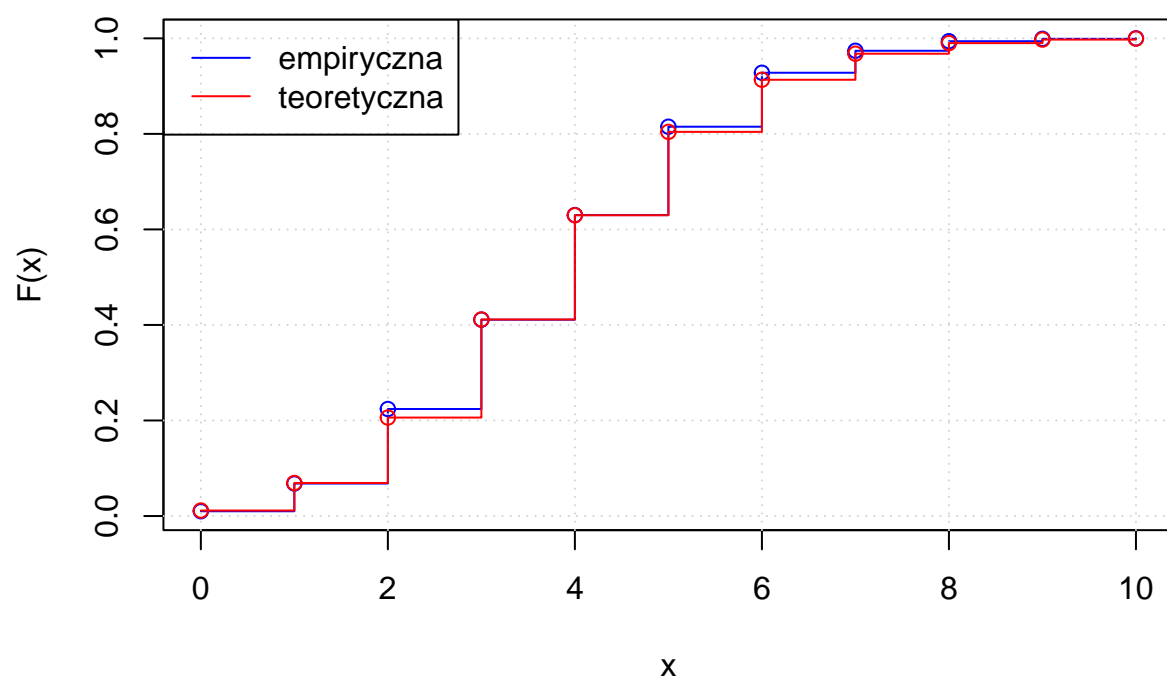
### Dystrybuanta dla $M = 100$



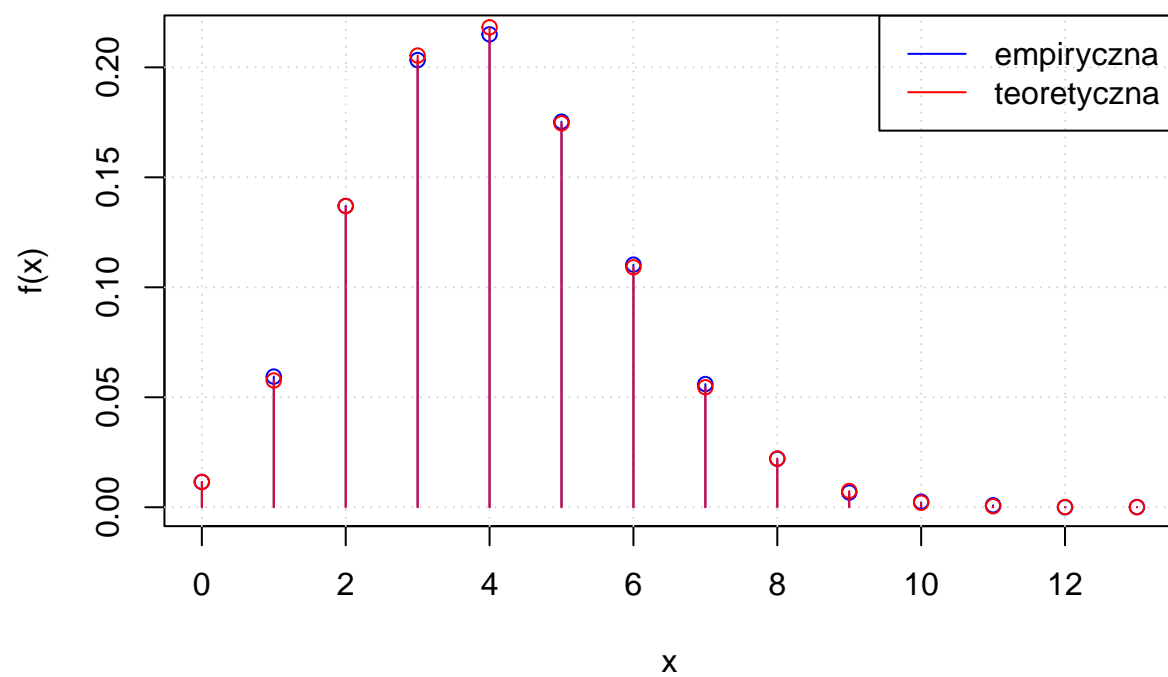
### Funkcja prawdopodobieństwa dla $M = 1000$



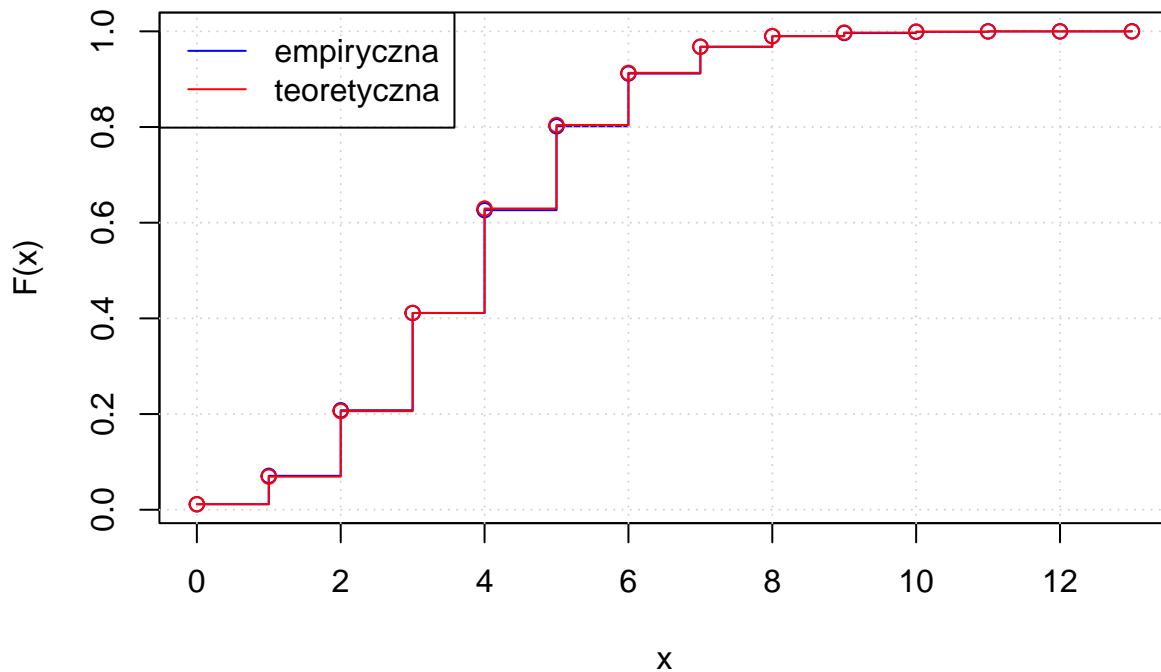
### Dystrybuanta dla $M = 1000$



### Funkcja prawdopodobieństwa dla $M = 10000$



## Dystrybuanta dla M = 10000



```
for (i in seq_along(M_values)) {
  cat("Dla M =", M_values[i], ":\n")
  cat("  Empiryczna średnia:", results[[i]]$mean_empirical, "\n")
  cat("  Empiryczna wariancja:", results[[i]]$variance_empirical, "\n")
  cat("  Różnica średnich:", abs(mean_theoretical - results[[i]]$mean_empirical), "\n")
  cat("  Różnica wariancji:", abs(variance_theoretical - results[[i]]$variance_empirical), "\n\n")
}
```

```
## Dla M = 100 :
##   Empiryczna średnia: 3.77
##   Empiryczna wariancja: 3.168788
##   Różnica średnich: 0.23
##   Różnica wariancji: 0.03121212
##
## Dla M = 1000 :
##   Empiryczna średnia: 3.947
##   Empiryczna wariancja: 3.025216
##   Różnica średnich: 0.053
##   Różnica wariancji: 0.1747838
##
## Dla M = 10000 :
##   Empiryczna średnia: 4.0061
##   Empiryczna wariancja: 3.252388
##   Różnica średnich: 0.0061
##   Różnica wariancji: 0.05238803
```