Homework3

Gifty Osei

2024-09-30

Problem 3c

```r
# Parameters
n <- 1000

# Case 1: Small p
p1 <- 0.001
lambda1 <- n * p1  # lambda = 1
k1 <- floor(lambda1 + 0.5 * sqrt(lambda1))  # k = 1

# Case 2: Larger p
p2 <- 0.01
lambda2 <- n * p2  # lambda = 10
k2 <- floor(lambda2 + 0.5 * sqrt(lambda2))  # k = 11



# Binomial CDFs
cdf_binom_case1 <- pbinom(k1, size = n, prob = p1)
cdf_binom_case2 <- pbinom(k2, size = n, prob = p2)

# Poisson CDFs
cdf_poisson_case1 <- ppois(k1, lambda = lambda1)
cdf_poisson_case2 <- ppois(k2, lambda = lambda2)
```

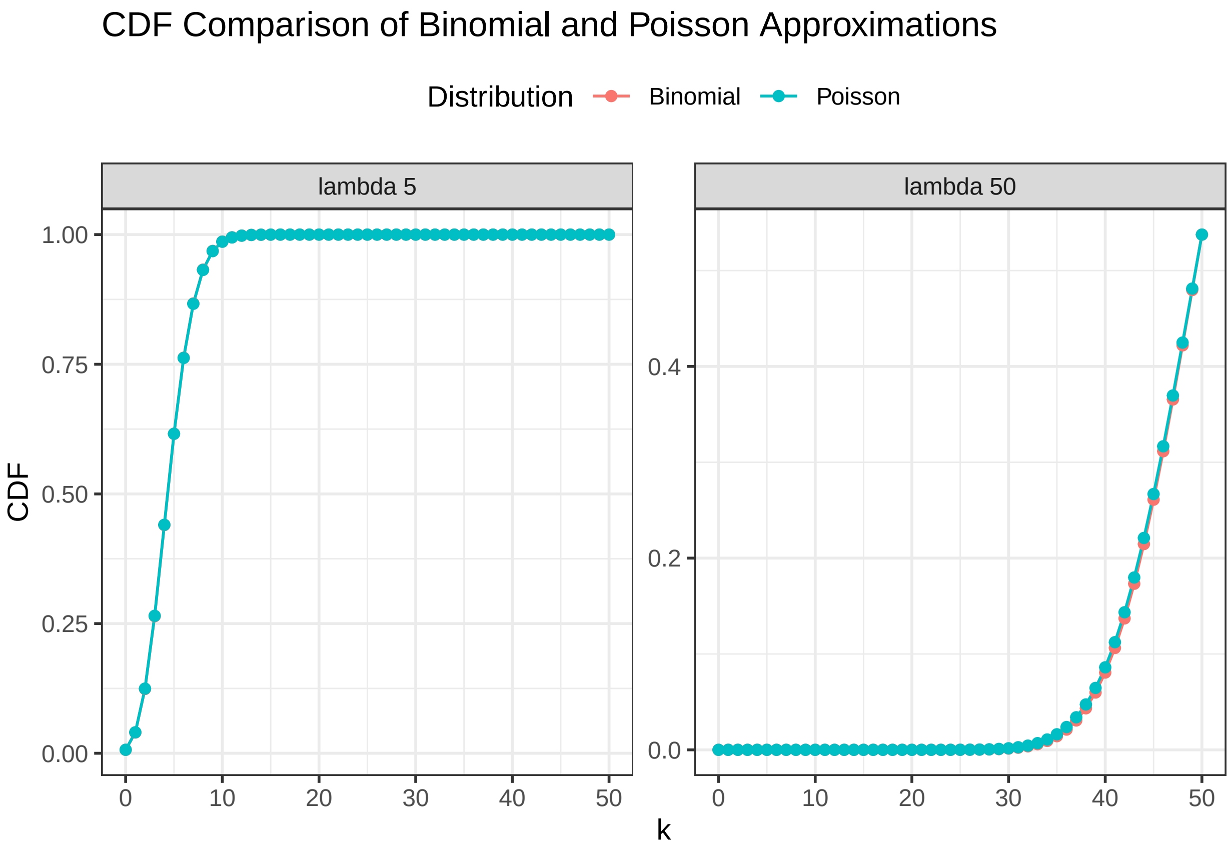
For Case 1:Small $p(p = 0.001, \lambda = 1)$, Binomial CDF at $k = 1$ is 0.7357589 and Poisson CDF at $k = 1$ is 0.7357589.

For Case 2: Larger $p(p = 0.01, \lambda = 10)$, Binomial CDF at $k = 2$ is 0.6973501 and Poisson CDF at $k = 2$ is 0.6967761.

Plot to show Difference

```r
library(ggplot2)

# Set parameters for Binomial and Poisson distributions
n <- 1000
lambda_values <- c(5, 50)  # Small and large values of lambda
```

```r
k_values <- c(floor(5 + 0.5 * sqrt(5)), floor(50 + 0.5 * sqrt(50)))  # Values for k

# Create a data frame to store CDF values
results <- data.frame(k = integer(),
                      CDF = numeric(),
                      Distribution = character(),
                      Lambda = numeric())

# Calculate CDF for Binomial and Poisson distributions
for (lambda in lambda_values) {
  p <- lambda / n  # Probability of success for Binomial
  k_range <- 0:50  # Range of k values to calculate CDF

  # Calculate Binomial CDF
  binom_cdf <- pbinom(k_range, n, p)
  results <- rbind(results,
                   data.frame(k = k_range,
                              CDF = binom_cdf,
                              Distribution = "Binomial",
                              Lambda = lambda))

  # Calculate Poisson CDF
  poisson_cdf <- ppois(k_range, lambda)
  results <- rbind(results,
                   data.frame(k = k_range,
                              CDF = poisson_cdf,
                              Distribution = "Poisson",
                              Lambda = lambda))
}

# Plotting the CDF comparison
ggplot(results, aes(x = k, y = CDF, color = Distribution)) +
  geom_line() +
  geom_point() +
  facet_wrap(~ Lambda, scales = "free_y",
             ncol = 2, labeller =
               labeller(Lambda = function(x) paste(expression(lambda), x))) +
  labs(title = "CDF Comparison of Binomial and Poisson Approximations",
       x = "k", y = "CDF") +
  theme_bw() +
  theme(legend.position = "top")
```



The Poisson approximation to the Binomial distribution is more accurate when λ (or p) is small, as illustrated in the left plot. When λ is large, the Poisson approximation is still useful, but its accuracy diminishes as p increases or λ increases. Therefore, The distribution Y_n converges to the distribution of $Y \sim Poisson(\lambda)$.