Stat 3202 Lab 5

Jane Weissberg.11

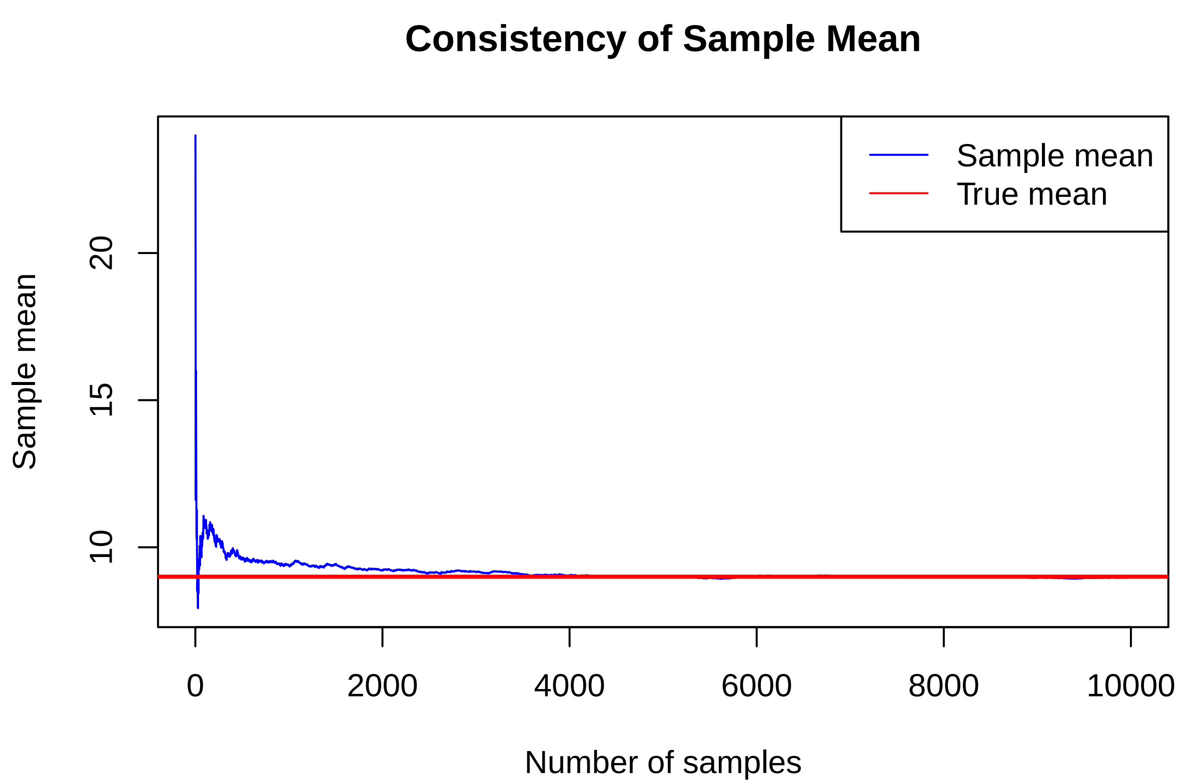
2024-06-22

1a.

```
p <- 0.10
mu <- (1-p) / p
n <- 10000
X <- rgeom(n, p)
x_bar <- cumsum(X) / (1:n)

plot(1:n, x_bar, type = "l", col = "blue", xlab = "Number of samples", ylab = "Sample mean", main = "Con
abline(h = mu, col = "red", lwd = 2)

legend("topright", legend = c("Sample mean", "True mean"), col = c("blue", "red"), lty = 1)
```



The plot shows that as the number of samples increases, the sample mean $\bar{x}$ converges to the true mean

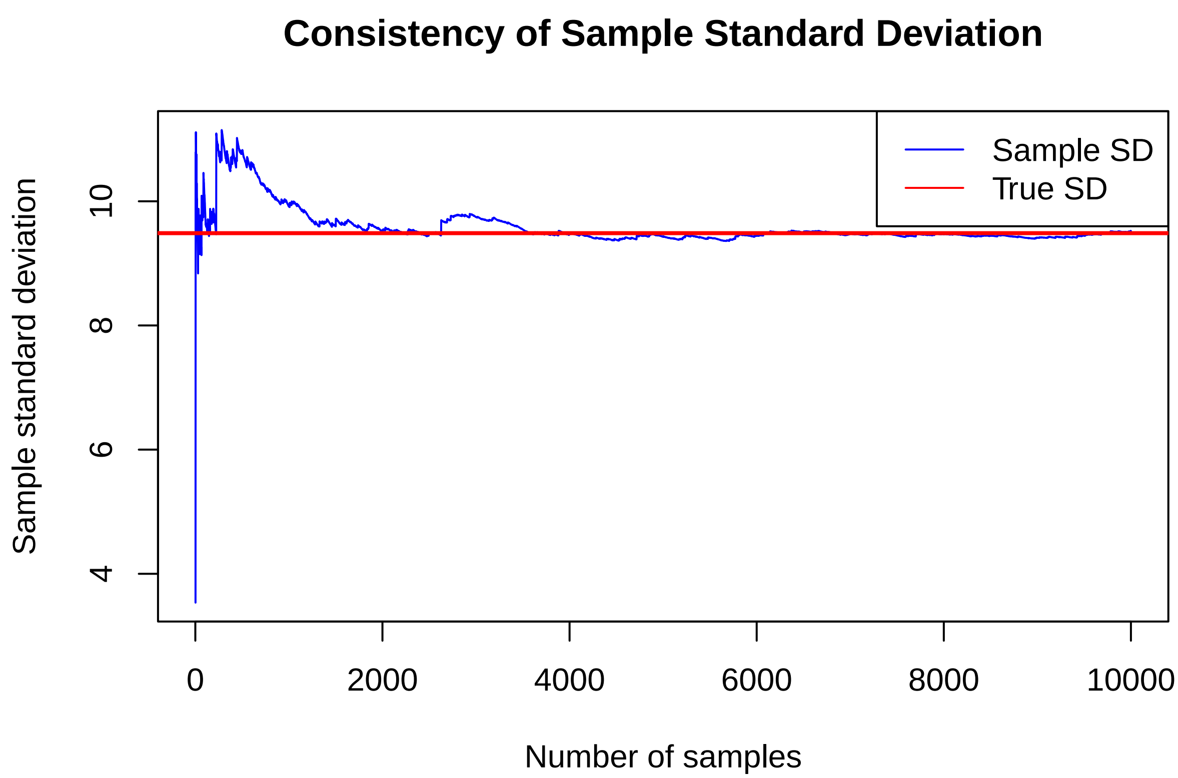
μ. This proves the second requirement of consistency, which is that the limit as n goes to infinity of the variance of the statistic over n must be zero. This also shows that $\bar{x}$ is an unbiased estimator for μ because the expected value of $\bar{x}$ would be equal to μ.

1b.

```
sigma <- sqrt((1 - p) / p^2)
sample_sd <- sapply(1:n, function(k) sd(X[1:k]))

plot(1:n, sample_sd, type = "l", col = "blue", xlab = "Number of samples", ylab = "Sample standard devi
abline(h = sigma, col = "red", lwd = 2)

legend("topright", legend = c("Sample SD", "True SD"), col = c("blue", "red"), lty = 1)
```



The plot shows that as the number of samples increases, the sample standard deviation s converges to the true standard deviation σ. This proves the second requirement of consistency, which is that the limit as n goes to infinity of the variance of the statistic over n must be zero. This also shows that s is an unbiased estimator for σ because the expected value of s would be equal to σ.

1c.

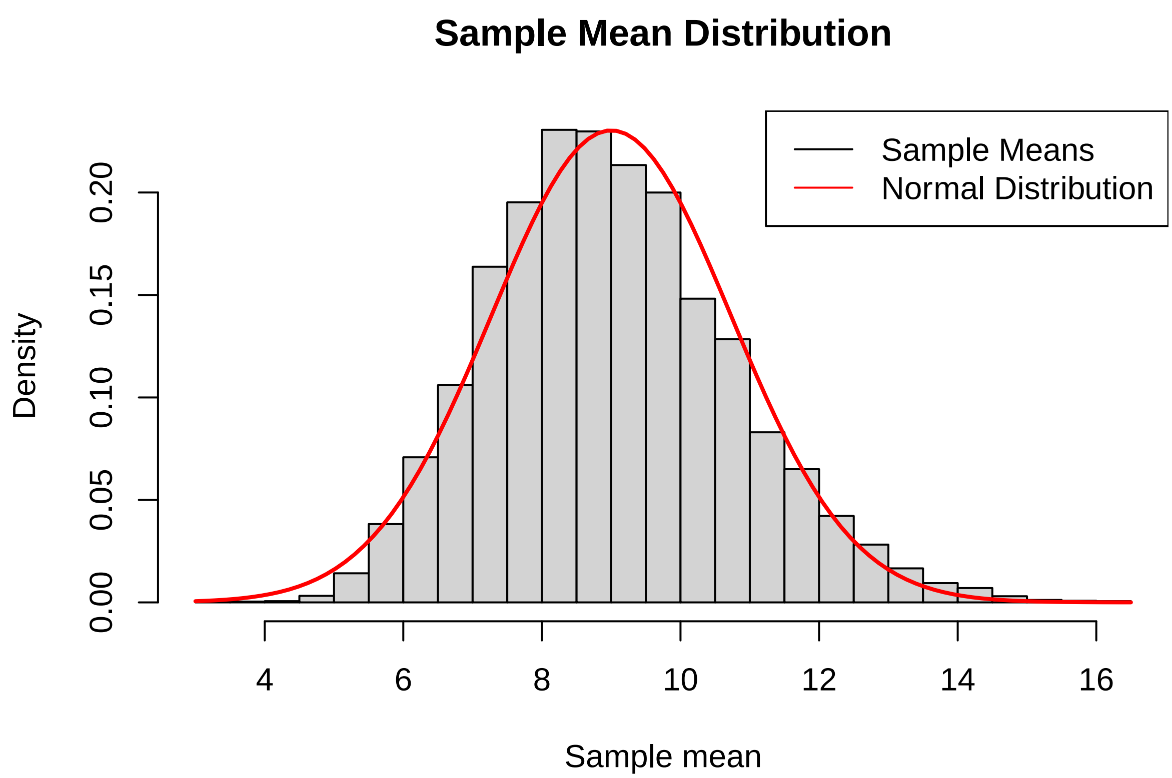
```
n <- 10000
sample_size <- 30

sample_means <- replicate(n, mean(rgeom(sample_size, p)))

hist(sample_means, probability = TRUE, breaks = 30, main = "Sample Mean Distribution", xlab = "Sample m
```

```r
curve(dnorm(x, mean = mu, sd = sigma / sqrt(sample_size)), col = "red", add = TRUE, lwd = 2)

legend("topright", legend = c("Sample Means", "Normal Distribution"), col = c("black", "red"), lty = 1)
```



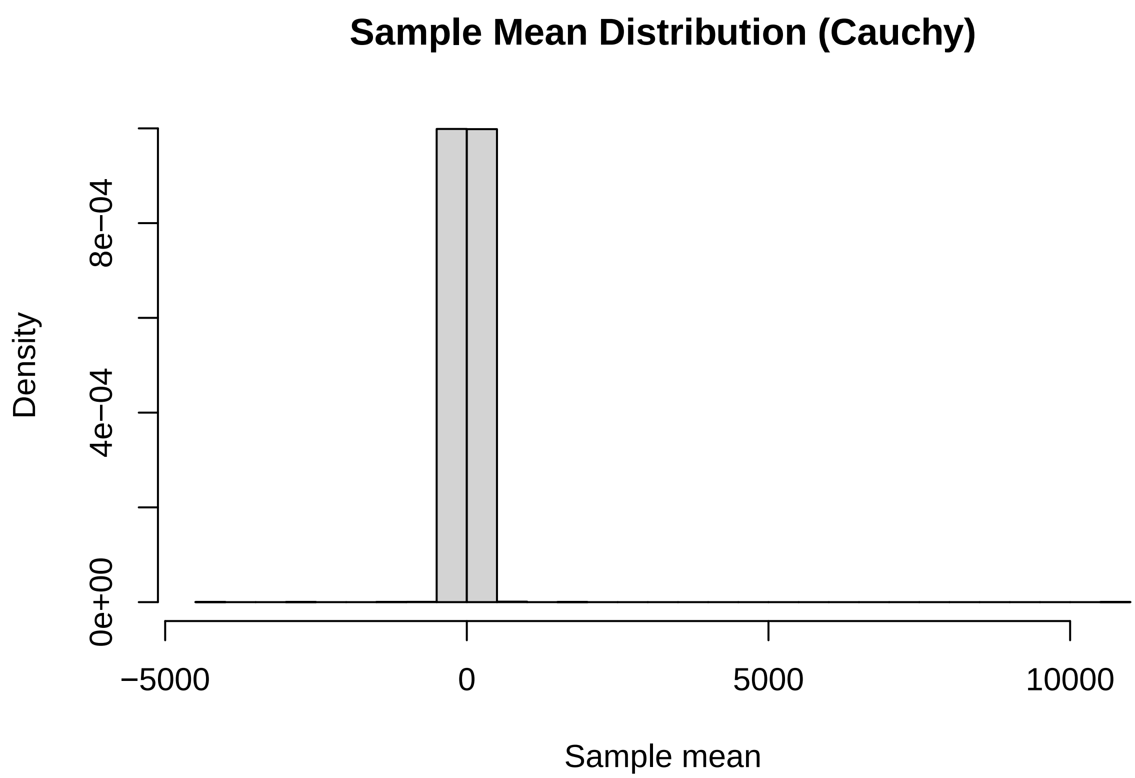
The histogram of the sample means closely follows the red curve, which represents the normal distribution. This demonstrates that the sample mean $\bar{x}$ adheres to the Central Limit Theorem for the geometric distribution because with a high enough sample size and with many samples taken, the histogram of many samples of $\bar{x}$ is close to the normal distribution.

2.

```r
x0 <- 0
gamma <- 1

n <- 10000
sample_size <- 30

sample_means <- replicate(n, mean(rcauchy(sample_size, location = x0, scale = gamma)))
hist(sample_means, probability = TRUE, breaks = 30, main = "Sample Mean Distribution (Cauchy)", xlab =
```



The histogram of the sample means for the Cauchy distribution shows that the sample means do not converge to a normal distribution. This is due to the heavy tails and undefined variance of the Cauchy distribution, which prevents the Central Limit Theorem from applying.

3.

```
n <- 10000
compute_K <- function() {
  sum <- 0
  count <- 0
  while (sum <= 1) {
    sum <- sum + runif(1)
    count <- count + 1
  }
  return(count)
}

K_values <- replicate(n, compute_K())
E_K <- mean(K_values)

E_K
```

```
## [1] 2.711
```

The expected value E(K) appears to be approximately 2. This makes sense because the mean of the Uniform(0, 1) distribution is 0.5, so it takes about 2 samples on average for their sum to exceed 1.