Formative Assessment 6

Elisha Sophia Borromeo and Zyann Lynn Mayo

2025-03-05

I. Geometric Distribution

1. Set the probability of success: p <- 0.2

```
p <- 0.2
```

2. Generate 1000 random variables from the geometric distribution.

To generate the 1000 random variables, we can use rgeom function which generates a random sample from a geometric distribution. We add 1 because this function returns the number of failures before the first success.

```
x <- rgeom(1000, prob = p) + 1
```

3. Calculate some basic statistics:

```
mean_x <- mean(x)
var_x <- var(x)
sd_x <- sd(x)
```

4. Print the results in item 3 with the following output (string):

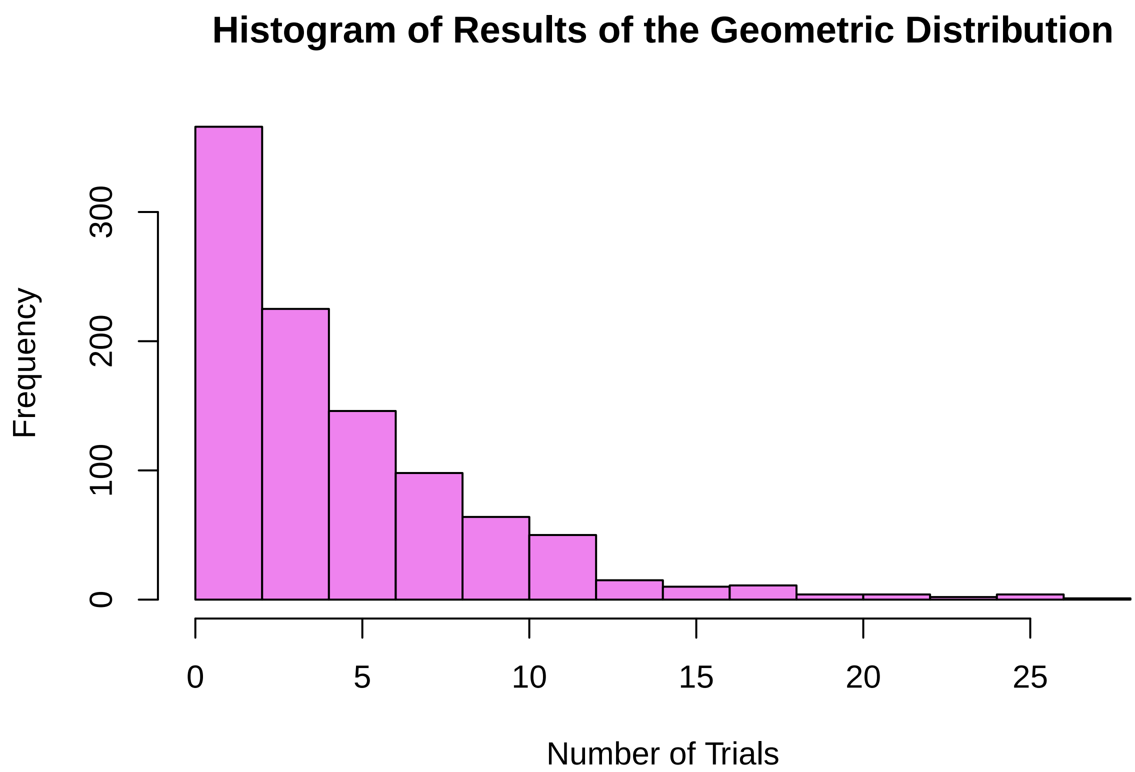
For the number of trials required to achieve the first success, we assume that this asks for how many trials it took for the first simulation to be successful. These results vary.

```
cat("Number of trials required to achieve first success:", x[1], "\n",
   "Mean (in 2 decimal places):", round(mean_x, 2), "\n",
    "Variance (in 2 decimal places):", round(var_x, 2), "\n",
    "Standard deviation (in 2 decimal places):", round(sd_x, 2), "\n")
```

```
## Number of trials required to achieve first success: 3
##  Mean (in 2 decimal places): 4.87
##  Variance (in 2 decimal places): 17.99
##  Standard deviation (in 2 decimal places): 4.24
```

5. Plot the histogram of the results.

```
hist(x, breaks = 10, col = "violet", main = "Histogram of Results of the Geometric Distribution",
     xlab = "Number of Trials", ylab = "Frequency", border = "black")
```



II. Hypergeometric Distribution

We need to find

$$P(X > 1) = 1 - P(X = 0) - P(X = 1)$$

For both cases, we will follow the Hypergeometric Probability formula:

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Case 1: A sample of 10 is selected from a box of 40

Insert values into variables

```
# k = selected number of defective chips
k1 <- 0.10
# n = sample size
```

```r
n1 <- 10
# N = population from given case
N1 <- 40
# K = defective
K1 <- N1*k1

# Compute the probabilities P(X = 0) and P(X = 1)
P_X0_case1 <- dhyper(x = 0, m = K1, n = N1 - K1, k = n1)
P_X1_case1 <- dhyper(x = 1, m = K1, n = N1 - K1, k = n1)

# Compute P(X > 1) = 1 - P(X = 0) - P(X = 1)
P_greater1_case1 <- 1 - P_X0_case1 - P_X1_case1
```

We get that the probability that the sample contains more that 10% defectives when a sample of 10 is selected from a box of 40 is **25.59%**

Case 2: A sample of 10 is selected from a box of 5000

Insert values into variables

```r
# k = selected number of defective chips
k2 <- 0.10
# n = sample size
n2 <- 10
# N = population from given case
N2 <- 5000
# K = defective
K2 <- N2*k2

# Compute the probabilities P(X = 0) and P(X = 1)
P_X0_case2 <- dhyper(x = 0, m = K2, n = N2 - K2, k = n2)
P_X1_case2 <- dhyper(x = 1, m = K2, n = N2 - K2, k = n2)

# Compute P(X > 1) = 1 - P(X = 0) - P(X = 1)
P_greater1_case2 <- 1 - P_X0_case2 - P_X1_case2
```

We get that the probability that the sample contains more that 10% defectives when a sample of 10 is selected from a box of 5000 is **26.39%**