Quiz 1

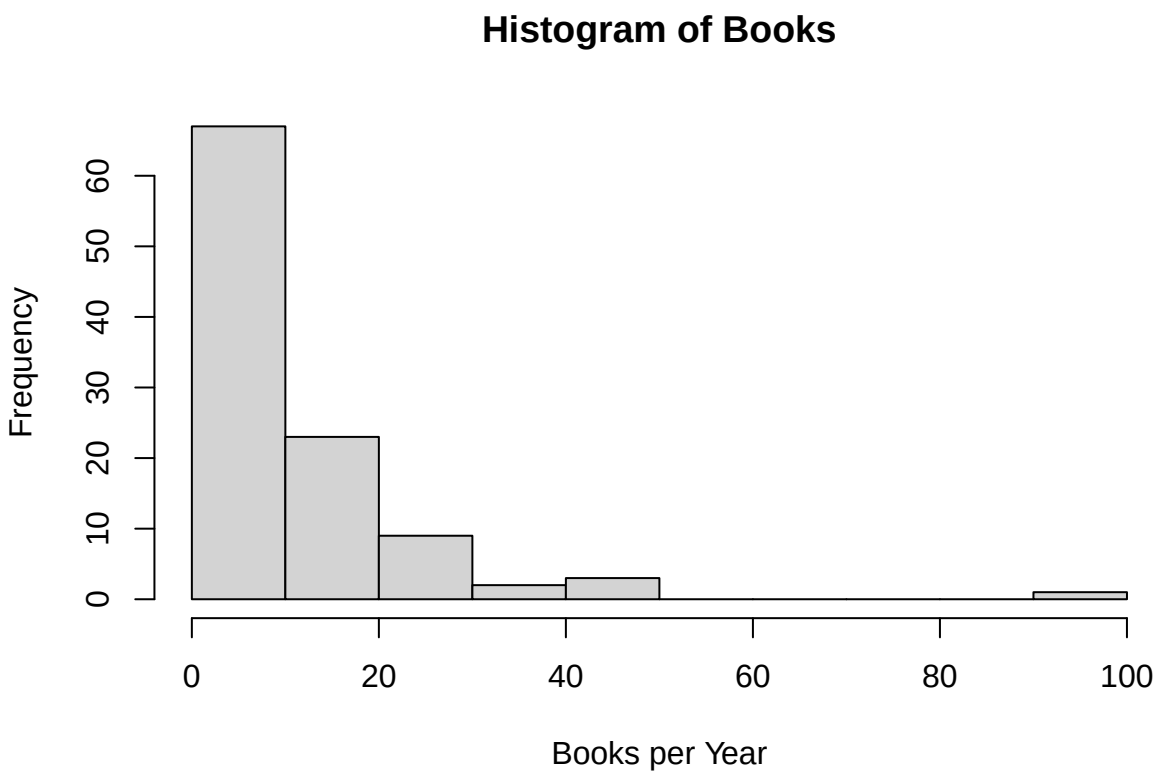
Anna Miletova, 89231151

2025-03-23

```
my.data <- read.table("data.txt", header = TRUE, sep = "\t")

x <- 1
z <- 5

my.data <- my.data[unique(c(seq(x,nrow(my.data),by=10),seq(z,nrow(my.data),by=10))),]
```

1. Histogram for Books..how.many.per.year

```
books <- my.data$Books..how.many.per.year

hist(books, main="Histogram of Books", xlab="Books per Year")
```



The histogram is positively asymmetric.

Since the graph is asymmetric to the right we expect median to be more than average and median to be a more appropriate measure.

Central tendency measures

```
modus <- function(v) {
  uniq_vals <- unique(v)
  uniq_vals[which.max(tabulate(match(v, uniq_vals)))]
}

cat(
  "Average: ", mean(books),
  "\nMedium: ", median(books),
  "\nModus: ", modus(books)
)
```

```
## Average:  12.73333
## Medium:  10
## Modus:  5
```

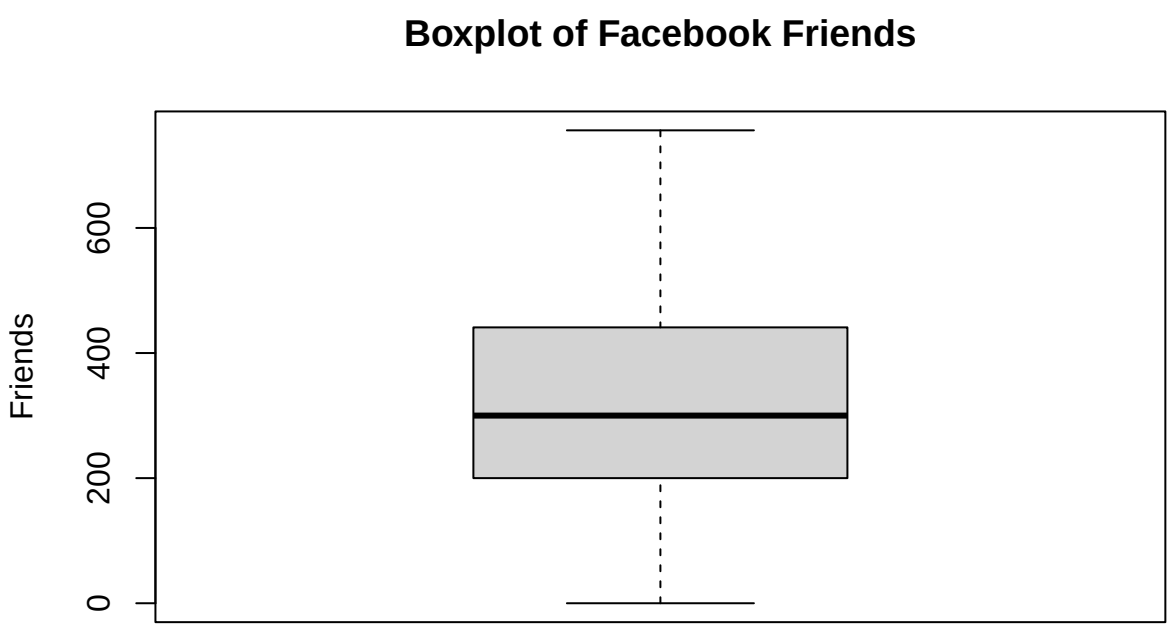
Measures of variability

```
# cat("IQR: ", quantile(books, 0.75) - quantile(books, 0.25))
cat(
  "Range: ", max(books) - min(books),
  "\nIQR: ", IQR(books),
  "\nStandard deviation: ", sd(books)
)
```

```
## Range:  100
## IQR:  10
## Standard deviation:  13.70939
```

2. Boxplot for Friends.on.Facebook

```
friends <- my.data$Friends.on.Facebook

boxplot(friends, main="Boxplot of Facebook Friends", ylab="Friends")
```




```
summary(friends) # Min, Max, Mean, Median, Quartiles
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##     0.0   200.0   300.0   304.8   441.0   756.0      56
```

```
IQR(friends, na.rm=TRUE)
```

```
## [1] 241
```

The outliers, points outside of the "whiskers", are not observed in the boxplot.

3. Circumference of a circle with a radius r

```
circumference <- function(r) {
  return(2 * pi * r)
}
cat("Circumference of the circle, where r = x =", x ,"is", circumference(x))
```

```
## Circumference of the circle, where r = x = 1 is 6.283185
```

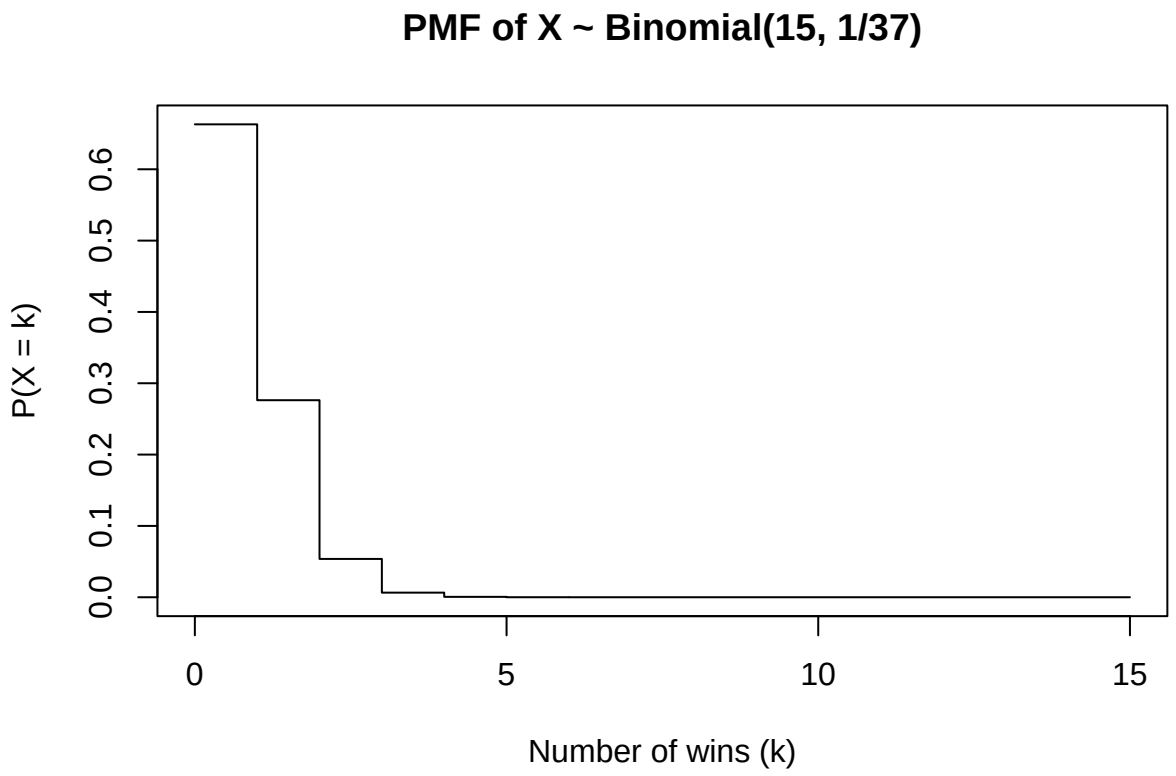
4. Roulette probability

```
times <- as.numeric(paste0(x, 5))
win_probability <- 1/37

P_X_equals_z <- dbinom(z, size = times, prob = win_probability)
cat("P(X =", z, ") =", P_X_equals_z)
```

```
## P(X = 5 ) = 3.292718e-05
```

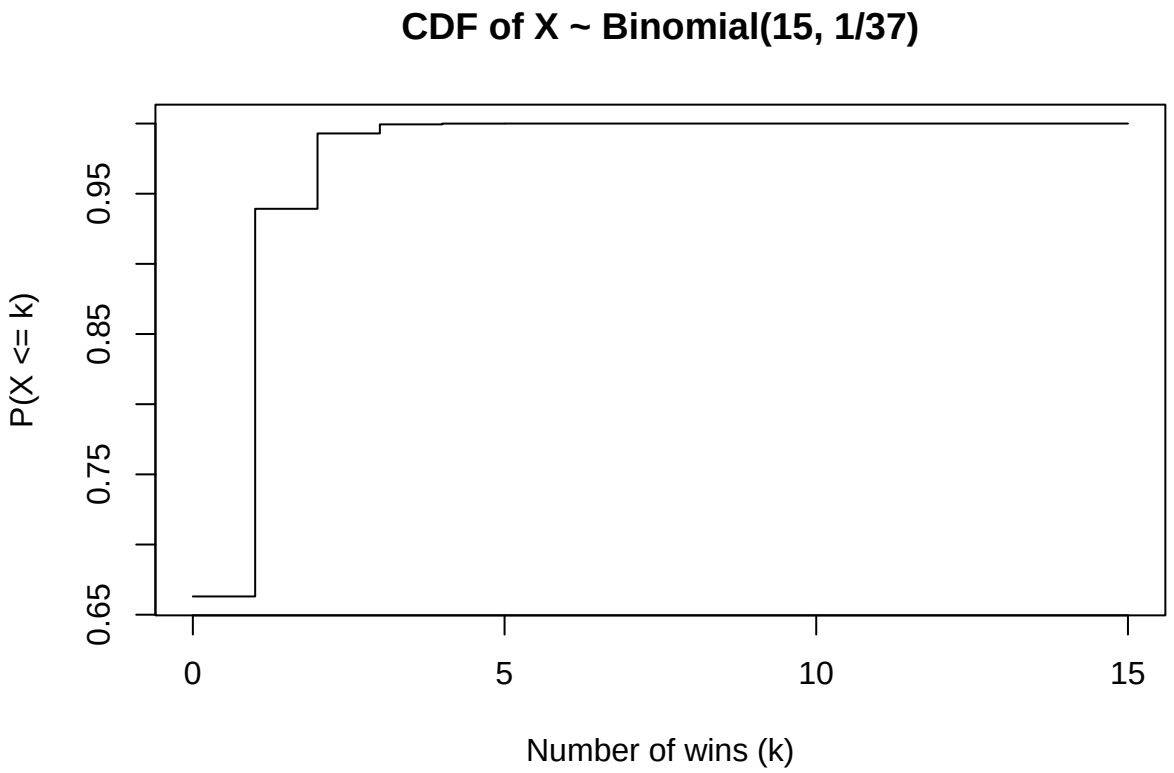
PDF

```
k_values <- 0:times
pmf_values <- dbinom(k_values, size = times, prob = win_probability)

plot(
  x    = k_values,
  y    = pmf_values,
  type = "s",
  main = "PMF of X ~ Binomial(15, 1/37)",
  xlab = "Number of wins (k)",
  ylab = "P(X = k)"
)
```



CDF

```
cdf_values <- pbinom(k_values, size = times, prob = win_probability)

plot(
  x    = k_values,
  y    = cdf_values,
  type = "s",
  main = "CDF of X ~ Binomial(15, 1/37)",
  xlab = "Number of wins (k)",
  ylab = "P(X <= k)"
)
```



Probability of winning less than 5 times

```
P_less_than_5 <- pbinom(4, size = times, prob = win_probability)
cat("P(X < 5) =", P_less_than_5)
```

```
## P(X < 5) = 0.9999655
```