Lab 7

Jerry Xing

Question 1

X is the number of Aces obtained: X ~ H (52, 4, 8)

- Therefore:

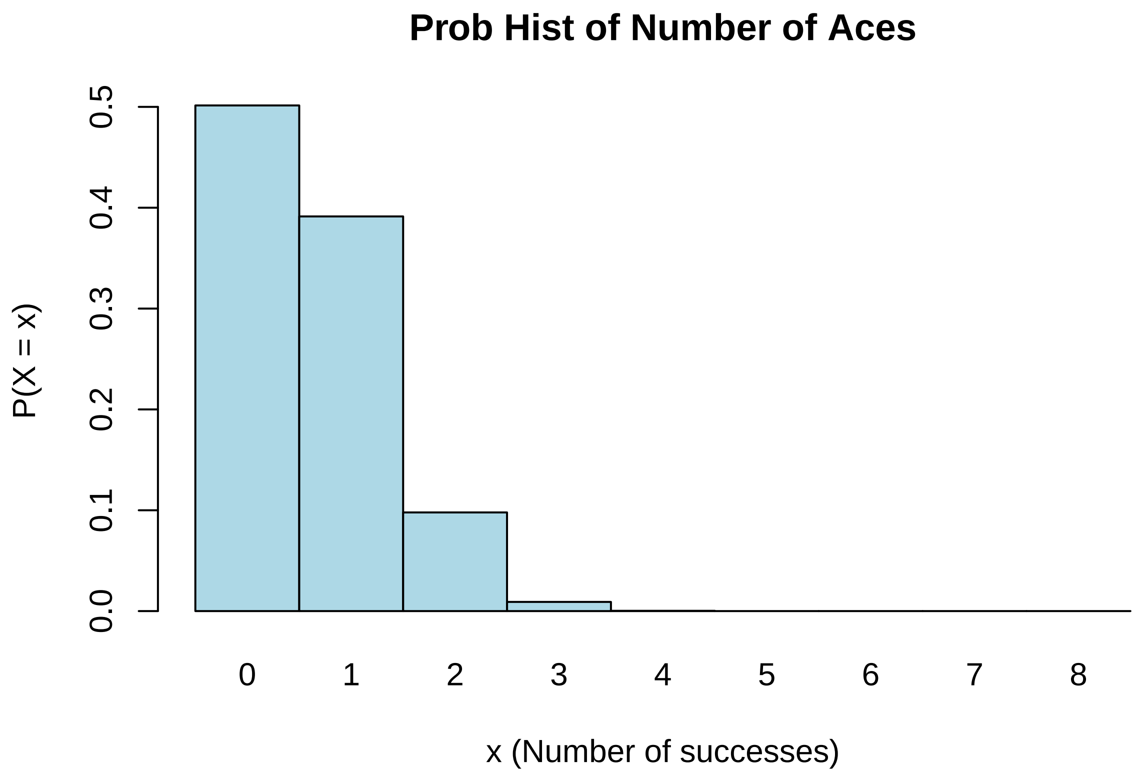
```
N <- 52
K <- 4
n <- 8
```

The probability distribution is as follows:

```
X.vals <- 0:8
d.vals <- as.table(dhyper(X.vals, K, N-K, n))
names(d.vals) <- X.vals
d.vals
```

```
##           0           1           2           3           4           5
## 0.501435036 0.391363930 0.097840983 0.009101487 0.000258565 0.000000000
##           6           7           8
## 0.000000000 0.000000000 0.000000000
```

The barplot is below:



Question 2

a.

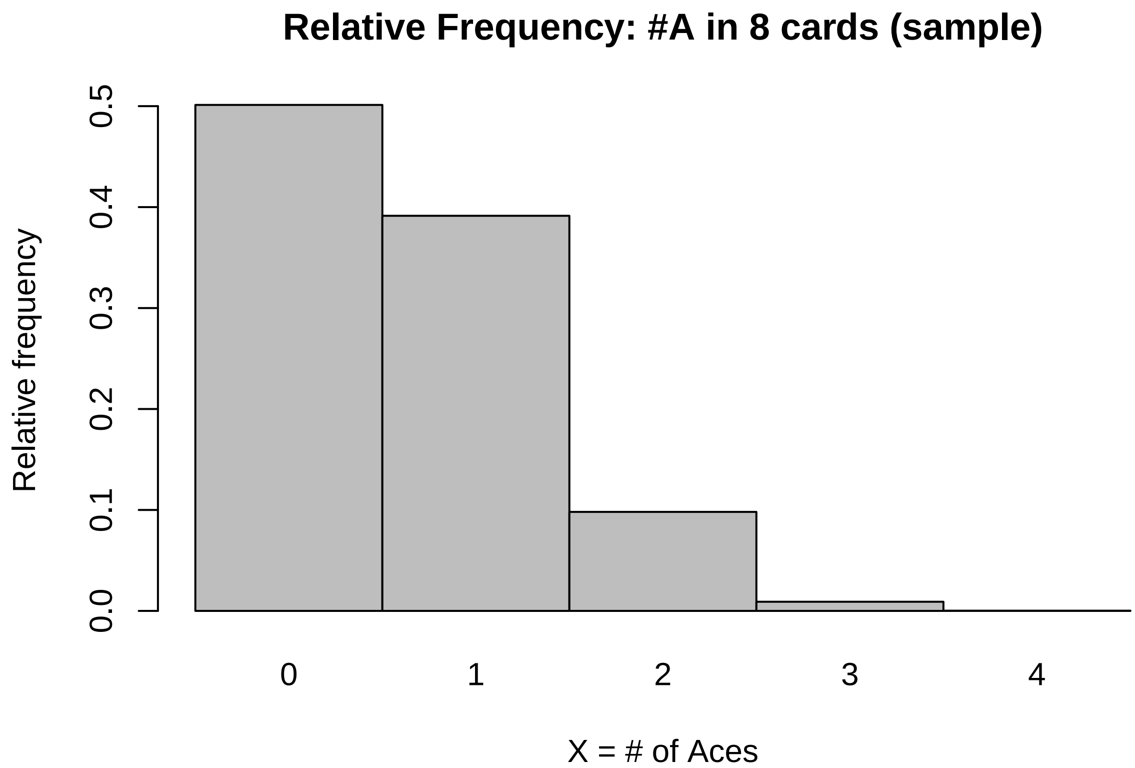
```
DrawCards <- function(n.cards, m.trials){
  X.vals <- integer(m.trials)
  deck <- c(rep("A", 4), rep("B", 48))
  for(i in seq_len(m.trials)){
    X.vals[i] <- sum(sample(deck, n.cards, replace = FALSE) == "A")
  }
  return(prop.table(table(X.vals)))
}

set.seed(69)
rel <- DrawCards(n.cards = 8, m.trials = 10^6)

print(rel)
```

```
## X.vals
##        0        1        2        3        4
## 0.501229 0.391402 0.098072 0.009078 0.000219
```

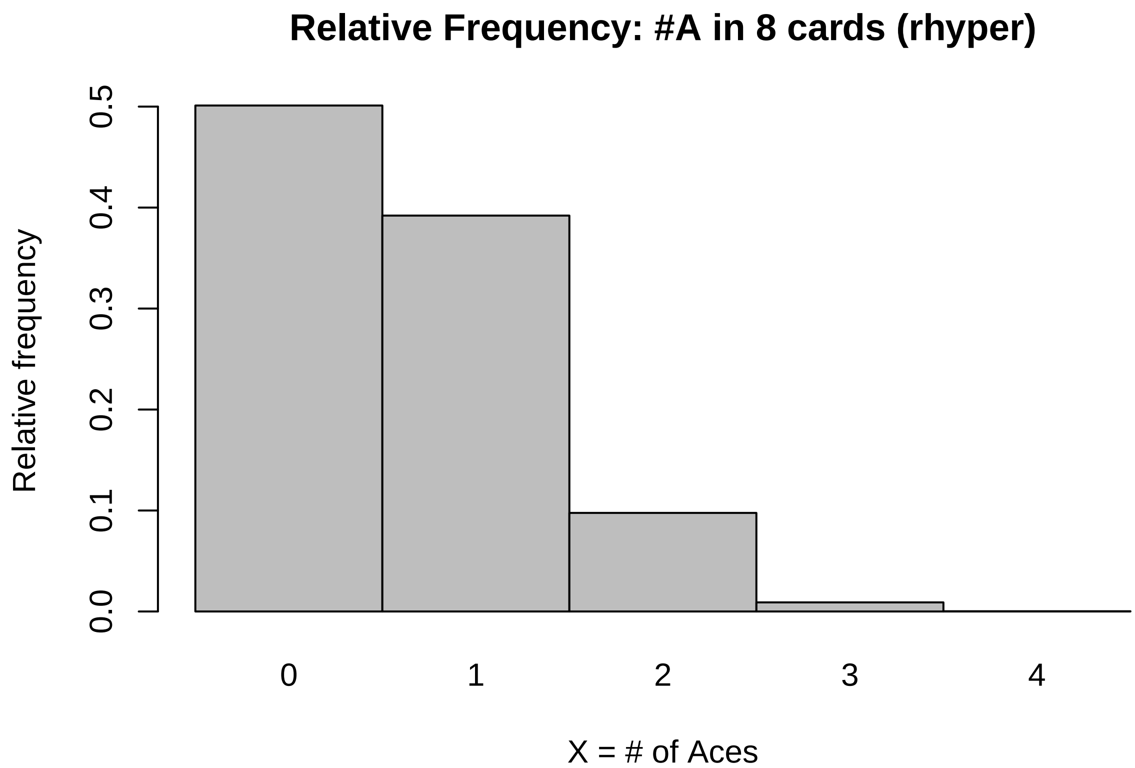
```r
barplot(rel, space=0,
        main = "Relative Frequency: #A in 8 cards (sample)",
        xlab = "X = # of Aces",
        ylab = "Relative frequency")
```



b.

```r
DrawCards_rhyper <- function(n.cards, m.trials){
  X <- rhyper(m.trials, m = 4, n = 52 - 4, k = n.cards)
  rel <- prop.table(table(X))
  barplot(rel, space = 0,
          main = "Relative Frequency: #A in 8 cards (rhyper)",
          xlab = "X = # of Aces", ylab = "Relative frequency")
  return(rel)
}

set.seed(1)
rel_rhyper <- DrawCards_rhyper(n.cards = 8, m.trials = 10^6)
```



Question 3

```r
m <- 333
n <- 999 - m
k <- 300

a_100_exact <- dhyper(100, m, n, k)
b_le_100    <- phyper(100, m, n, k)
c_lt_100    <- phyper(99,  m, n, k)
d_ge_110    <- 1 - phyper(109, m, n, k)
e_90_110    <- phyper(110, m, n, k) - phyper(89, m, n, k)

a_100_exact
```

```
## [1] 0.05834763
```

```r
b_le_100
```

```
## [1] 0.5304624
```

```r
c_lt_100
```

```
## [1] 0.4721148
```

```r
d_ge_110
```

```
## [1] 0.08253631
```

```r
e_90_110
```

```
## [1] 0.8759101
```

Question 4

X is the number of tickets the student has bought when they first win a prize: X ~ G(p)

- Therefore:

```r
p <- 1/3
```

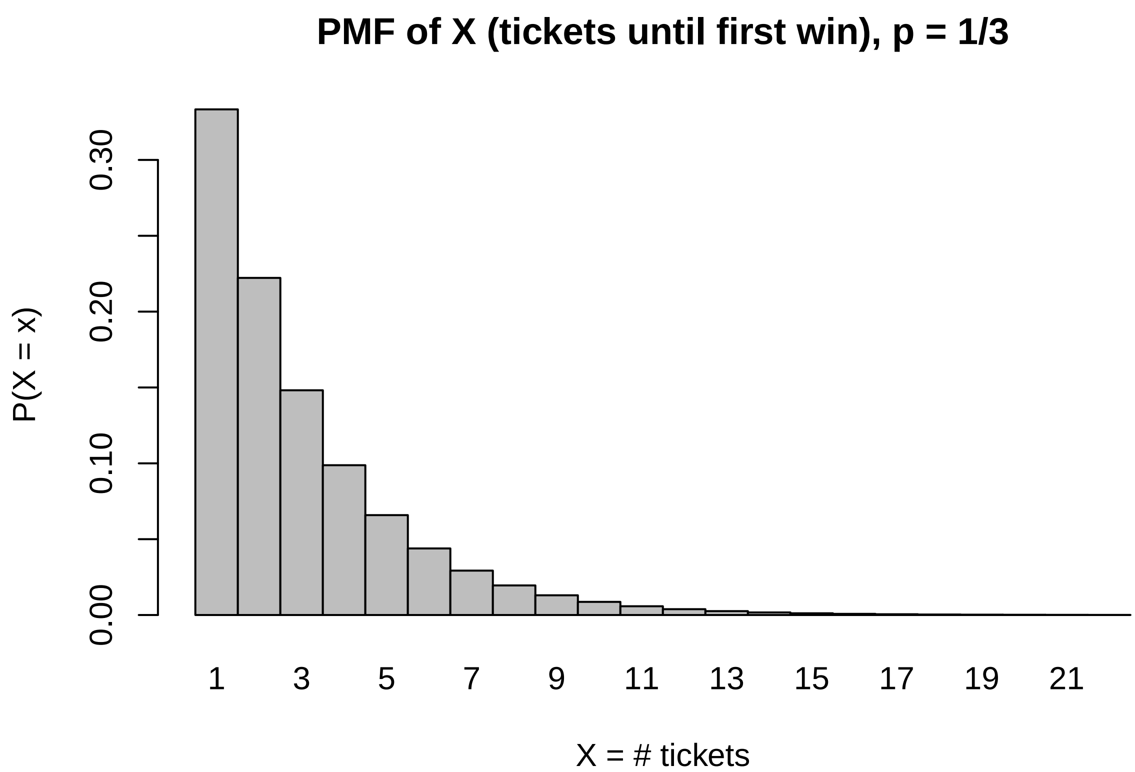
The probability distribution is as follows:

```r
cutoff <- 0.00005
X <- c()
P <- c()
k <- 1L
repeat{
  pk <- dgeom(k - 1L, prob = p)
  if(pk < cutoff) break
  X <- c(X, k); P <- c(P, pk)
  k <- k + 1L
}

pmf <- setNames(P, X)
pmf
```

```
##            1            2            3            4            5            6
## 3.333333e-01 2.222222e-01 1.481481e-01 9.876543e-02 6.584362e-02 4.389575e-02
##            7            8            9           10           11           12
## 2.926383e-02 1.950922e-02 1.300615e-02 8.670765e-03 5.780510e-03 3.853673e-03
##           13           14           15           16           17           18
## 2.569116e-03 1.712744e-03 1.141829e-03 7.612194e-04 5.074796e-04 3.383197e-04
##           19           20           21           22
## 2.255465e-04 1.503643e-04 1.002429e-04 6.682859e-05
```

```r
barplot(height = pmf, space = 0,
        main = "PMF of X (tickets until first win), p = 1/3",
        xlab = "X = # tickets", ylab = "P(X = x)")
```



Question 5

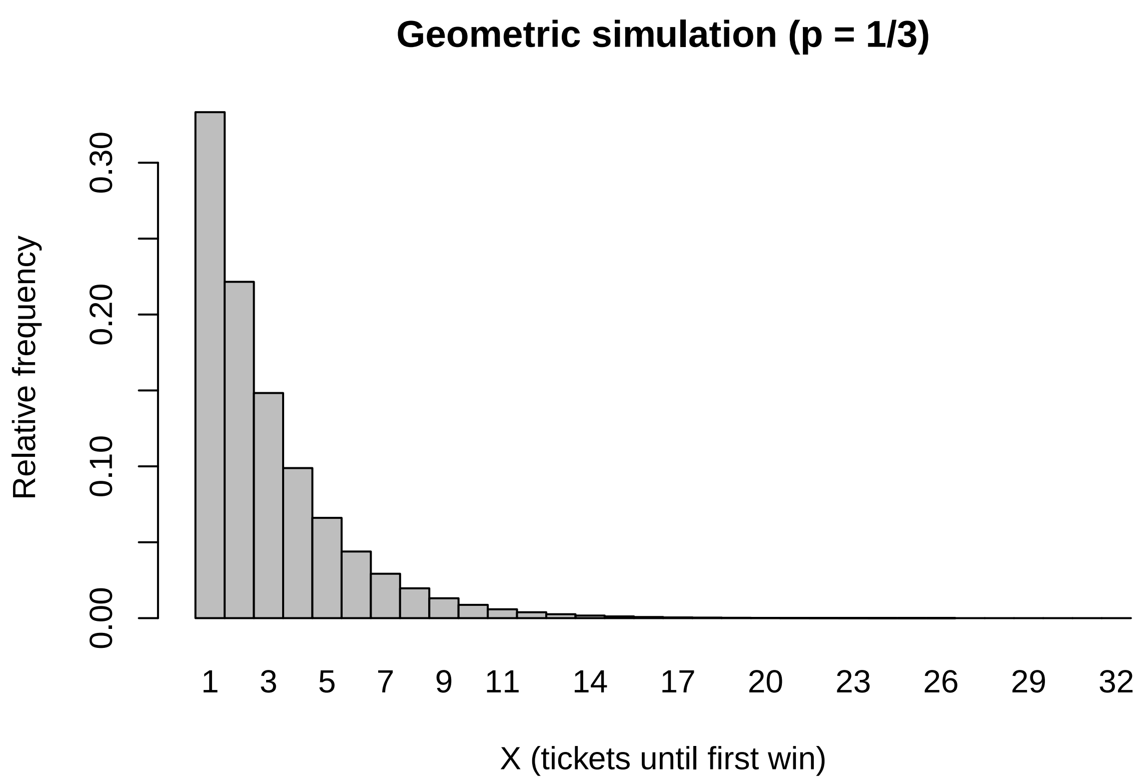
a.

```r
SimulateLotteryTickets <- function(p.win, m.trials){
  X.vals <- integer(m.trials)
  for(i in seq_len(m.trials)){
    have.won  <- FALSE
    n.tickets <- 0L
    while(!have.won){
      ticket <- sample(c("Win", "Lose"), 1, prob = c(p.win, 1 - p.win))
      n.tickets <- n.tickets + 1L
      if (ticket == "Win"){
        have.won <- TRUE
      }
    }
```

```r
    X.vals[i] <- n.tickets
  }
  X.vals
}
set.seed(1)
res <- SimulateLotteryTickets(1/3, 10^6)
rel_res <- prop.table(table(res))
print(rel_res)
```

```
## res
##        1        2        3        4        5        6        7        8
## 0.333305 0.221542 0.148294 0.098846 0.066051 0.043893 0.029254 0.019644
##        9       10       11       12       13       14       15       16
## 0.013096 0.008751 0.005835 0.003876 0.002588 0.001701 0.001134 0.000737
##       17       18       19       20       21       22       23       24
## 0.000468 0.000346 0.000224 0.000136 0.000084 0.000070 0.000046 0.000026
##       25       26       27       28       29       30       31       32
## 0.000014 0.000015 0.000007 0.000005 0.000004 0.000004 0.000003 0.000001
```

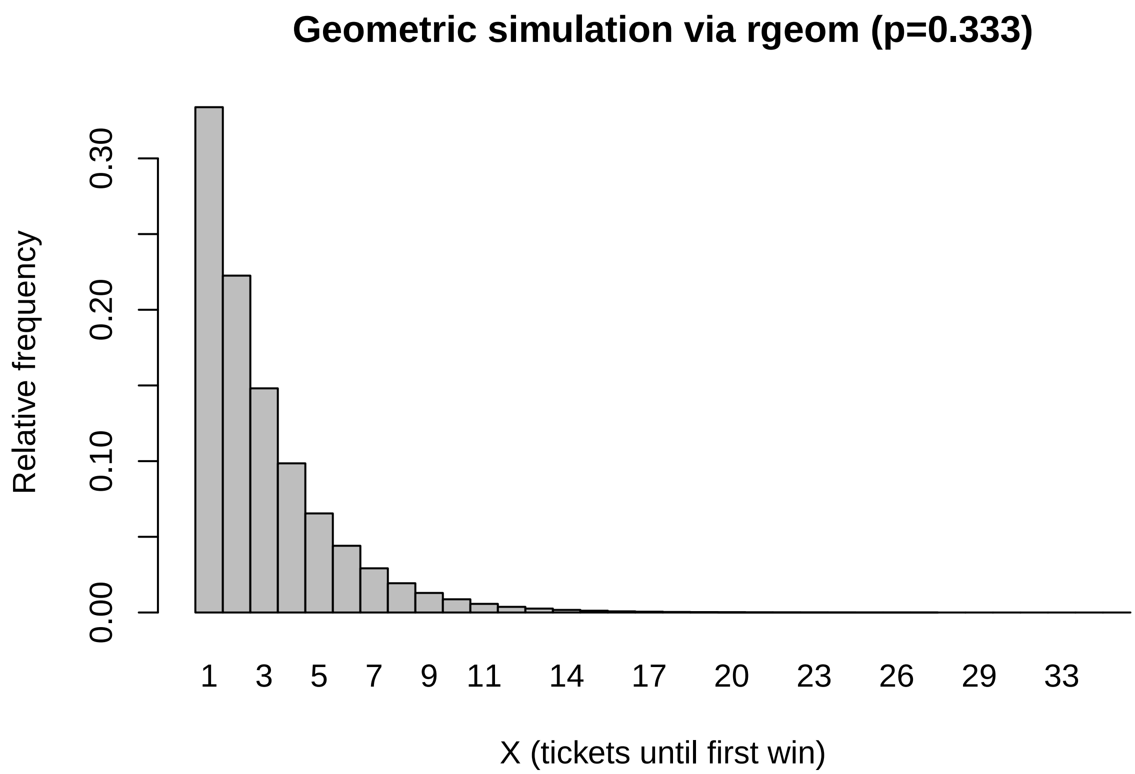
```r
barplot(rel_res, space = 0,
        xlab = "X (tickets until first win)",
        ylab = "Relative frequency",
        main = "Geometric simulation (p = 1/3)")
```



b.

```r
SimulateLottery_rgeom <- function(p.win, m.trials){
  X <- rgeom(m.trials, prob = p.win) + 1L
  rel <- prop.table(table(X))
  barplot(rel, space = 0,
          main = sprintf("Geometric simulation via rgeom (p=%.3f)", p.win),
          xlab = "X (tickets until first win)", ylab = "Relative frequency")
  invisible(rel)
}

set.seed(1)
rel_sim_rgeom <- SimulateLottery_rgeom(1/3, 1e6)
```



Question 6

```r
p_fail <- 0.001

a_by_500    <- pgeom(499,   p_fail)
```

```r
b_atleast_1200 <- 1 - pgeom(1198, p_fail)
c_1000_2000 <- pgeom(1999, p_fail) - pgeom(998, p_fail)

a_by_500
```

```
## [1] 0.3936211
```

```r
b_atleast_1200
```

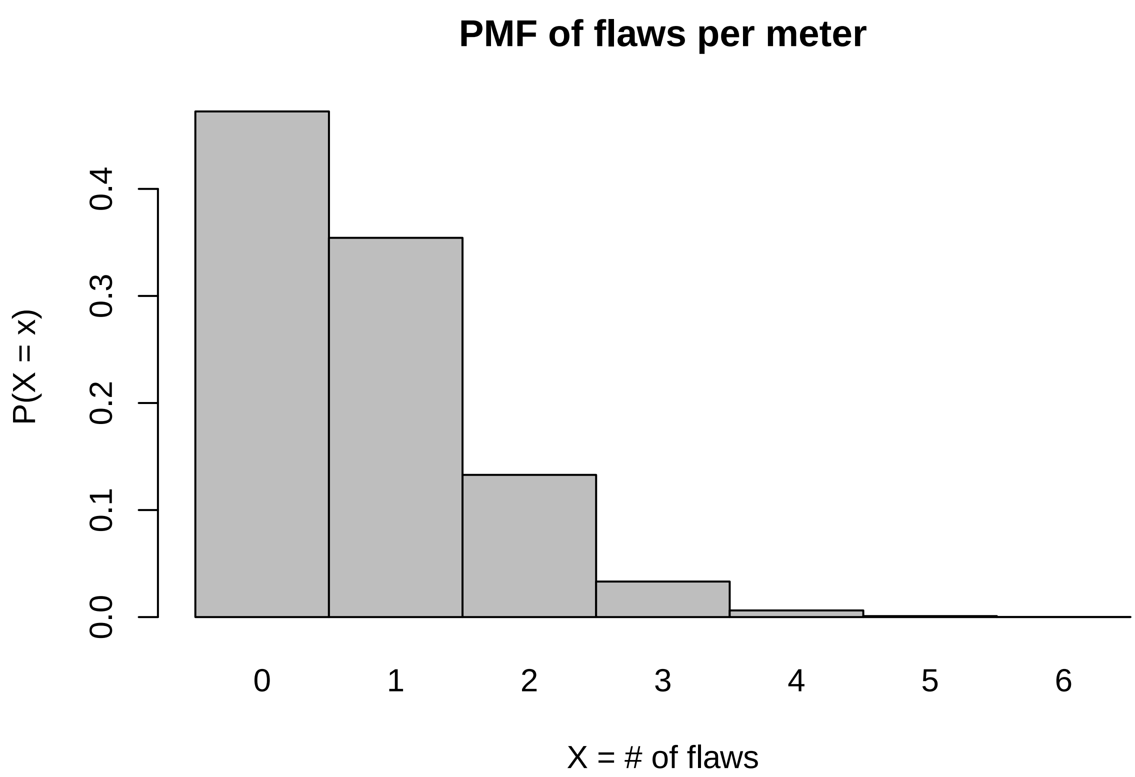
```
## [1] 0.3013147
```

```r
c_1000_2000
```

```
## [1] 0.2328636
```

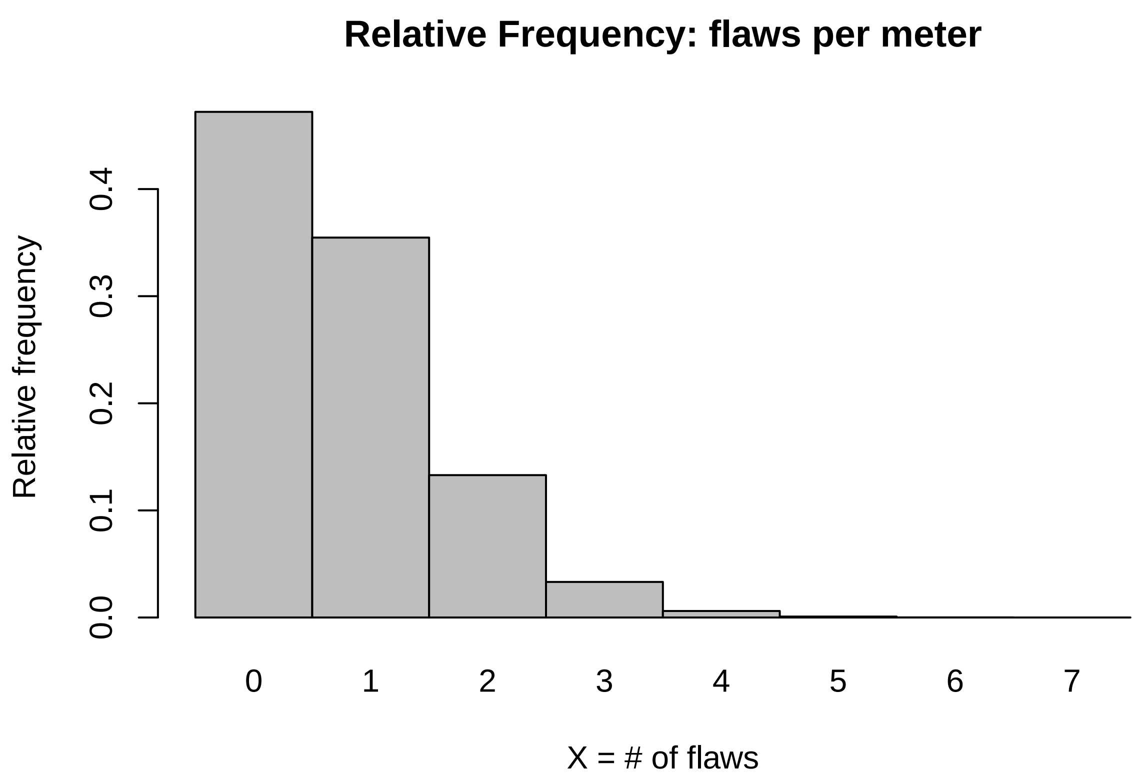
Question 7

```
##            0            1            2            3            4            5
## 0.4723665527 0.3542749146 0.1328530930 0.0332132732 0.0062274887 0.0009341233
##            6
## 0.0001167654
```



Question 8

```
## X
##        0        1        2        3        4        5        6        7
## 0.472090 0.354683 0.132927 0.033210 0.006082 0.000880 0.000111 0.000017
```



Question 9

```r
lam <- 34

a_exact_34 <- dpois(34, lam)
b_le_30    <- ppois(30, lam)
c_lt_30    <- ppois(29, lam)
d_gt_38    <- 1 - ppois(38, lam)
e_ge_38    <- 1 - ppois(37, lam)
f_30_40    <- ppois(40, lam) - ppois(29, lam)
```

```r
g_two_years_each_le_34 <- ppois(34, lam)^2
h_two_years_total_le_68 <- ppois(68, 2 * lam)

a_exact_34
```

```
## [1] 0.06825056
```

```r
b_le_30
```

```
## [1] 0.2803502
```

```r
c_lt_30
```

```
## [1] 0.2235049
```

```r
d_gt_38
```

```
## [1] 0.2166179
```

```r
e_ge_38
```

```
## [1] 0.2681011
```

```r
f_30_40
```

```
## [1] 0.6429105
```

```r
g_two_years_each_le_34
```

```
## [1] 0.297506
```

```r
h_two_years_total_le_68
```

```
## [1] 0.532192
```

Question 10

a.

$$
\begin{aligned}
X &\sim Poisson(5) \\
P(X = 4) &= \frac{e^{-5} \cdot 5^4}{4!} \\
&= \frac{e^{-5} \cdot 625}{24} \\
&\approx 0.1754674
\end{aligned}
$$

b.

$$
\begin{aligned}
Y &\sim B(10^6, \frac{5}{10^6}) \\
P(X=4) &= \binom{10^6}{4} \cdot (\frac{5}{10^6}))^4 \cdot (1 - (\frac{5}{10^6}))^{10^6-4} \\
&\approx 0.1754676
\end{aligned}
$$