```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
set.seed(1929)
```

HW3

HW3.1

The distrubution of photons emitted by by X ($X \sim Poisson(100)$) and Y is the number of photons which pass through an object.
If we know the amount of protons emitted by the x-ray beam we can describe the number of photons phasing through the object by using the distrubution $Y|X \sim Binomial(X, 0.4)$

Calculate the probability mass function for (X,Y) [eg $p(k,l) = P((X = k, Y = l))$]

```r
set.seed(1929)
  n_reps = 1

#Creating a data frame containing random draws from the Poisson distrubution
photons_emitted.df <- data.frame(rpois(n_reps, 100))
  names(photons_emitted.df) <-c('Photons')
    #Number of photons emitted
    photons <- sum(photons_emitted.df$Photons)
      #Number of photons from one draw of the Poisson distrubution which passes through the material
      photons_pass_through_material.str <- rbinom(1, photons, 0.4)
      #Probability that a photon is emitted from the x-ray and passes through the object
      photons_pass_through_material.str/photons
```

```
## [1] 0.4059406
```

$p(k, l) = P(X = k, Y = l) = P(X = k) \ * \ P(Y = l|X = k)$

$P(X = k) * P(Y = l|X = k) = \frac{100^k e^{-100}}{k!} * kCl * (0.4)^l (0.6)^{k-l}$

HW3.3

You have n random number generators where the ith draw from the generator is from the interval [0, t_i] where t_i's are positive integers. Let X_i be the number drawn by the ith random number generator

What is the expected sum. $S = X_1 + ... + X_n$?

What is the expected value of Y, which is an indicator variable that counts for all X_i's <= 1?

HW3.4

Let X ~ Exponential(λ) and let t be a constant where: $0 < \text{t} < \lambda$

What is $E[e^{tX}]?

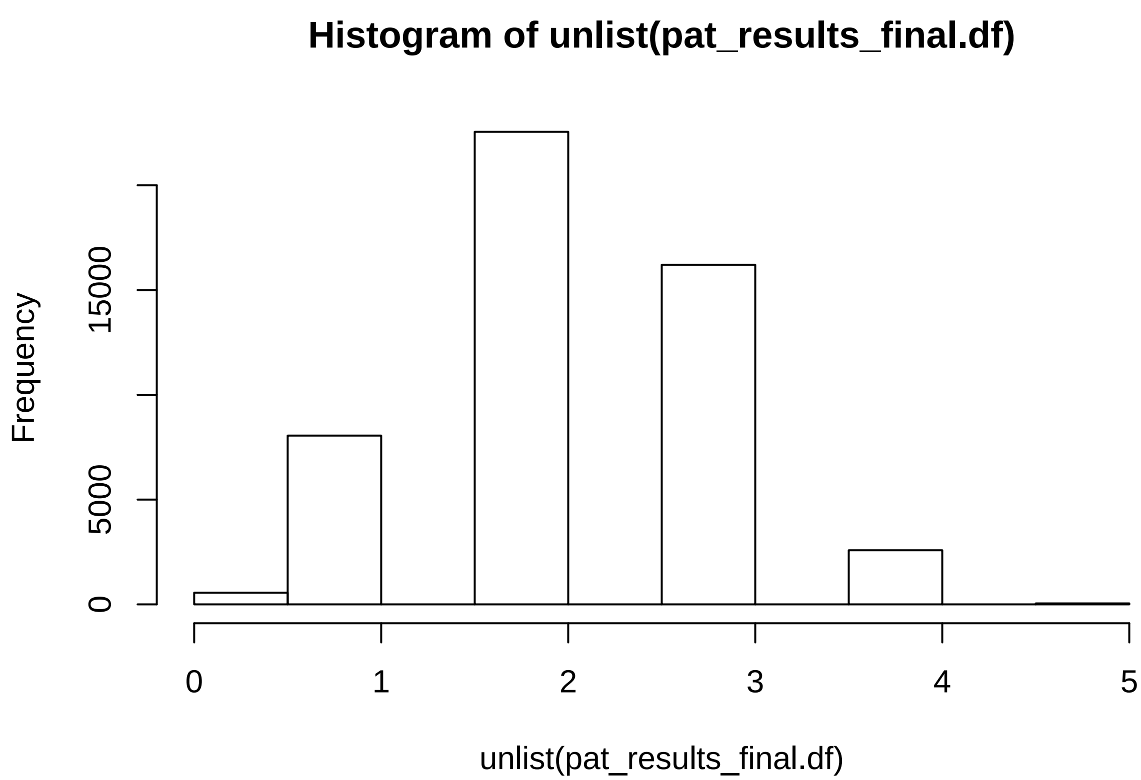
Use the Markov inequality to prove a bound on $P(e^{tx} => \alpha)$ *where* α *is a positive number*

Now reformulate this into a bound on $P(X >= b)$ *How does this compare to the true value of the probability P(X >= b)*

HW3.5

Flip a coin 10 times and let X be the sequence 'HT'. What is $P(X = 0)$ *?*

```
library(stringr)
  set.seed(1929)
#Assigning variables
coin_sides <- c('H', 'T')
  pat <- setNames(nm = "HT") #Pattern which we are looking for
    prob <- c(0.5, 0.5) #The coin is fair
      n_reps <- 50000 #Replicaitons of the Experiment
        all_flips.df <- 0 #Zeroing out the data frame


#A simple program which will toss the coin 'flips' times and record the results, creates a
# string from that data, and returns the number of times the pattern 'HT' is seen in the data
coin_pattern_toss <-function(flips){
all_flips.df <- data.frame(sample(coin_sides, flips, replace = TRUE, prob), stringsAsFactors = TRUE)
  names(all_flips.df) <- c('Coin')

all_flips.str <- paste(unlist(all_flips.df),collapse = '')
  pat_results.df <- data.frame(lapply(pat, str_count, string=all_flips.str))
return(pat_results.df)
}

#A large amount of 10 flip trials
pat_results_final.df <- data.frame(replicate(n_reps, coin_pattern_toss(10)))

#Histogram of the distrubuiton of pattern sucesses
hist(unlist(pat_results_final.df))
```




```
#Tabulation of the number of successes in the pattern
pat_results_final.ls <- data.frame(table(unlist(pat_results_final.df)))
  #Probability that P(X = 0)
  (pat_results_final.ls[1,2]/n_reps)
```

```
## [1] 0.01114
```

What is the expectation [$E[X]$]?

The probability that the pattern 'HT' comes up once in the coin tosses is;
$P(Heads\ and\ Tails) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4} = E_i$ Where i is 1-9 and if one was to calculate the expected value by weighing each coin toss by the probability and itterations we would see; $E(X) = E_1 + E_2 + ... + E_9 = \sum_{i=1}^{9} E_i = 9 * E_i = 9 * \frac{1}{4} = 2.25$

```
#Can be found using the weighted.mean function
x <- c(0, 1, 2, 3, 4, 5) #Number of patterns seen in ten flips
  wt <- (pat_results_final.ls$Freq)/n_reps #Weighted rate of patterns seen

#Expected number of patterns
weighted.mean(pat_results_final.df)
```

```
## [1] 2.24682
```

If you were to draw one card at a time from a full deck wihtout replacement what is the probability that you will NOT have the pattern 'RB'?

Use a hypergeometric distrubution in order to determine the chance of drawing only black cards first.

$P(X = 0) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$
Where;
k = 26 | Successes
n = 26 | Draws

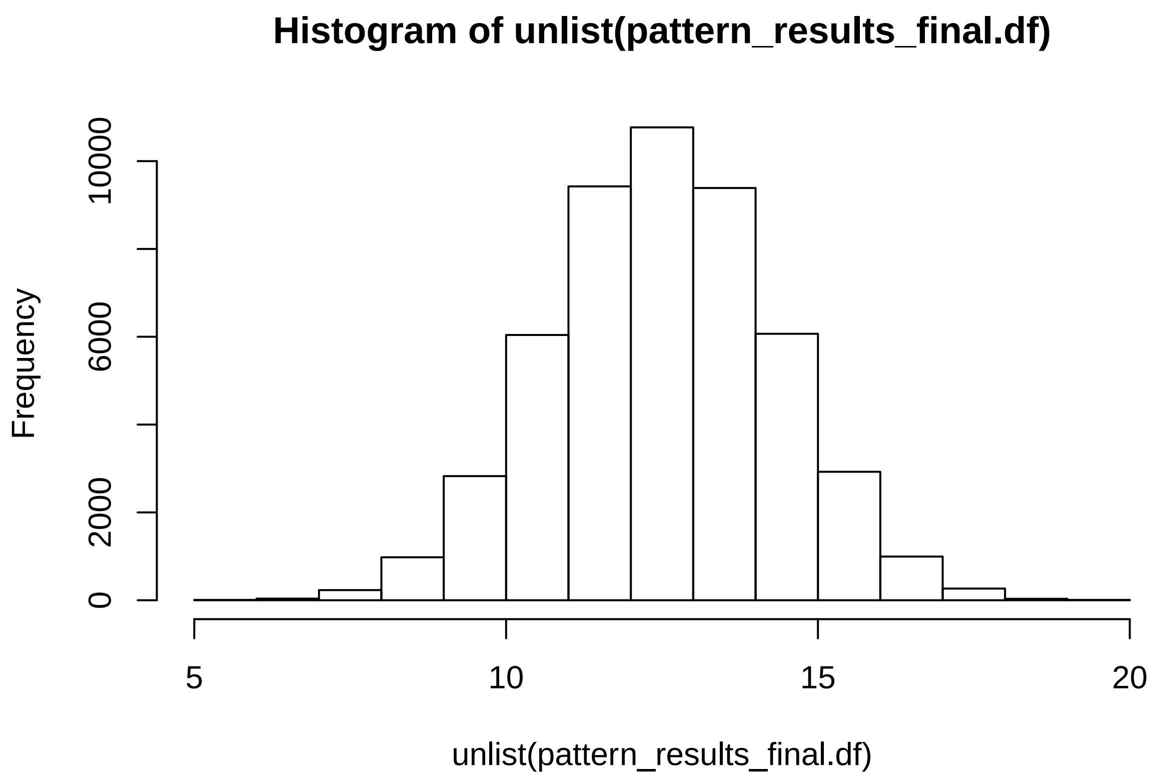
N = 52 | Number of objects in the Urn K = 26 | Number of Desired Object in the 'Urn'[Black Cards]

$\therefore \frac{\binom{26}{26}\binom{26}{0}}{\binom{52}{26}} = 2.0164e^{-15}$

Pseudocode; Perform this drawing action 50000 times using a string of R's and B's, do the same pattern searching as above, and then find the proability by density.

Notes: There is only one way for $P(Y = 0)$ and that is when all Black cards are drawn first.

```
library(stringr)
  set.seed(1929)
    pat <- setNames(nm = 'RB')  #Pattern that we wish to observe
      deck_attributes <- c(rep(c('R', 'B'), 26)) #Creaation of a variable containing the card colours
        n_reps <- 50000 #Number of reps for the simulation
          cards_to_draw <- 52 #Number of cards to draw from the deck

#Creation of a colour deck data frame [Red and Black]
full_deck.df <- data.frame(deck_attributes)

#Function to draw cards from the deck and to record the number of times the pattern is seen
card_draw <- function(cards_to_draw){
draw_results.df <-
  data.frame(
  sample(full_deck.df$deck_attributes, cards_to_draw, replace = FALSE), stringsAsFactors = TRUE)
names(draw_results.df) <- c('Cards')

draw_results.str <- paste(unlist(draw_results.df), collapse = '')

pattern_results.df <- data.frame(lapply(pat, str_count, string = draw_results.str))

return(pattern_results.df)
}

#Drawing the entire deck a large number of times
pattern_results_final.df <- data.frame(replicate(n_reps, card_draw(cards_to_draw)))

#Histogram of the results
hist(unlist(pattern_results_final.df))
```




```
#Transforming the data frame into a list
pattern_results_final.ls <- data.frame(table(unlist(pattern_results_final.df)))
  #Probability that P(Y = 0) --- must have one trial come up as P(Y = 0) in the sim to use this
  (pattern_results_final.ls[1,2]/n_reps)
```

```
## [1] 2e-05
```

As stated above using the hypergeometric distrubution the chance of this being true is very low $(2.0164\ e^{-15})$ so it may take 10^{15} trials until one comes up as $\mathrm{P(Y = 0)}$.

What is E(Y)? Since the probability of achieving a 'RB' pattern is computed without replace one needs take this into account. From the data obtained above from the simulation we can see that the weighted mean (according to the number of trials performed [50,000]) sits around 14 – so we should expected to see a similar or comprobable result from using the rules for discrete estimation.

$E[Y] = \sum_{i=1}^{51} 51 \frac{26}{52} \frac{26}{51} = 13$

This answer is close to mean of 14 from the experiment above