

Homework 3

1.

let's assume that two candidates are running in an election for Governor of California. This fictitious election pits Mr. Gubinator vs. Mr. Ventura. We would like to know who is winning the race, and therefore we conduct a poll of likely voters in California. If the poll gives the voters a choice between the two candidates, then the results can be reasonably modeled with the Binomial Distribution. States polls indicate that 62% of the population intend to vote for Mr. Gubinator.

Calculate the binomial probabilities, and CF (confidence interval):

a. in a random sample of 50 we find 10 likely voters b. in a random sample of 50 we find at least 3 likely voters c. What is the average and standard deviation of the number of voters that we will find in our random sample? d. plot the PMF and CMF from this random sample

```
In [8]: # a.  
a = dbinom(10,size=50,prob=.62)  
a  
library(binom)  
binom.confint(10, 50, conf.level = 0.95)
```

1.33945315982445e-09

method	x	n	mean	lower	upper
agresti-coull	10	50	0.2000000	0.11050245	0.3323061
asymptotic	10	50	0.2000000	0.08912769	0.3108723
bayes	10	50	0.2058824	0.10116836	0.3171207
cloglog	10	50	0.2000000	0.10318339	0.3196686
exact	10	50	0.2000000	0.10030224	0.3371831
logit	10	50	0.2000000	0.11113040	0.3332899
probit	10	50	0.2000000	0.10792338	0.3279450
profile	10	50	0.2000000	0.10570710	0.3242950
lrt	10	50	0.2000000	0.10568422	0.3242911
prop.test	10	50	0.2000000	0.10502158	0.3414368
wilson	10	50	0.2000000	0.11243750	0.3303711

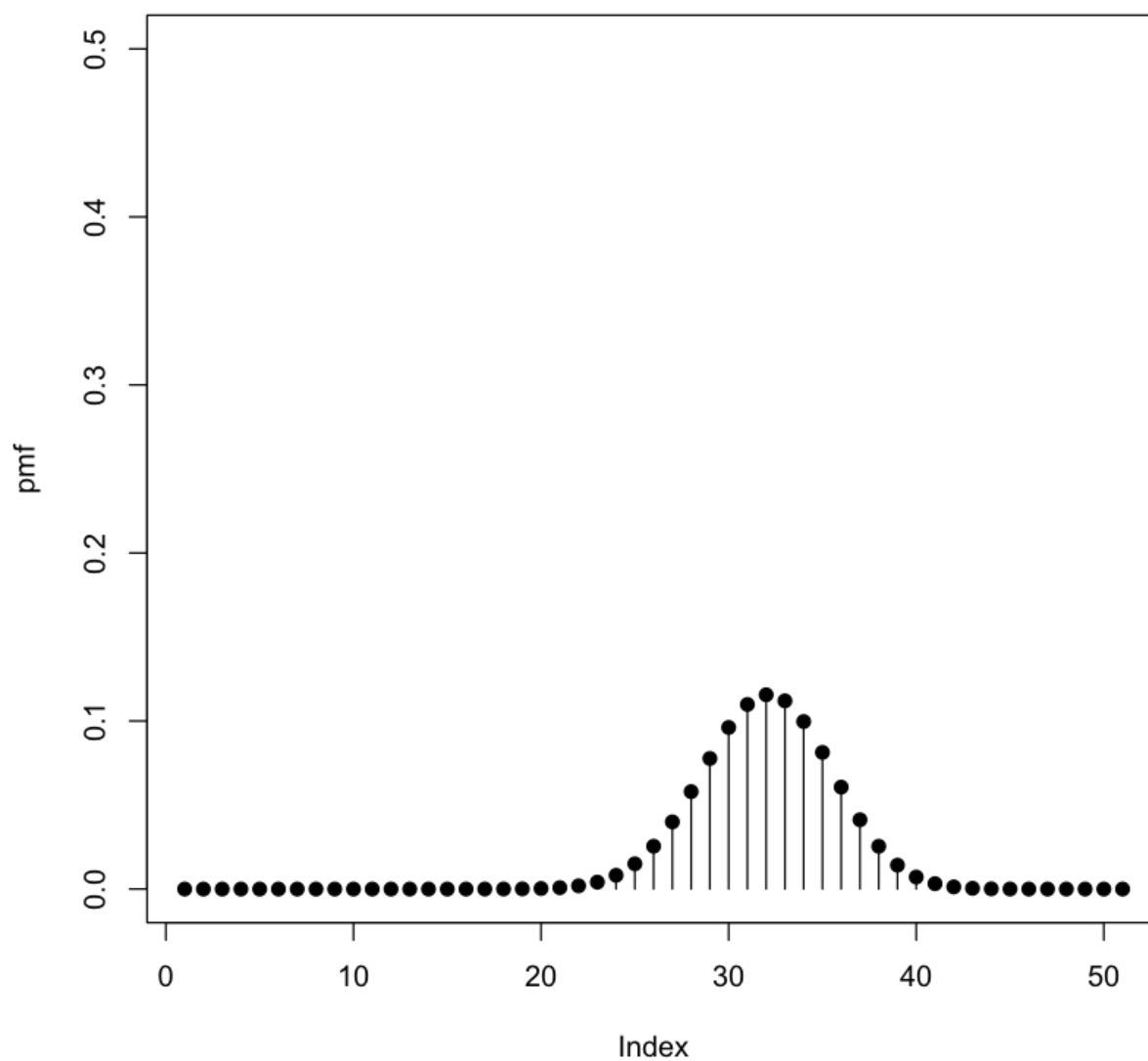
```
In [15]: # b.  
# in a random sample of 50 we find 1 likely voter, 2 likely voters, and no likely voters  
b1 = dbinom(1,size=50,prob=.62)  
b2 = dbinom(2,size=50,prob=.62)  
b3 = dbinom(0,size=50,prob=.62)  
b = 1-b1-b2-b3  
b  
  
b=0  
for (i in 2:50){  
  b=b+dbinom(i,size=50,prob=.62)  
}  
b
```

1

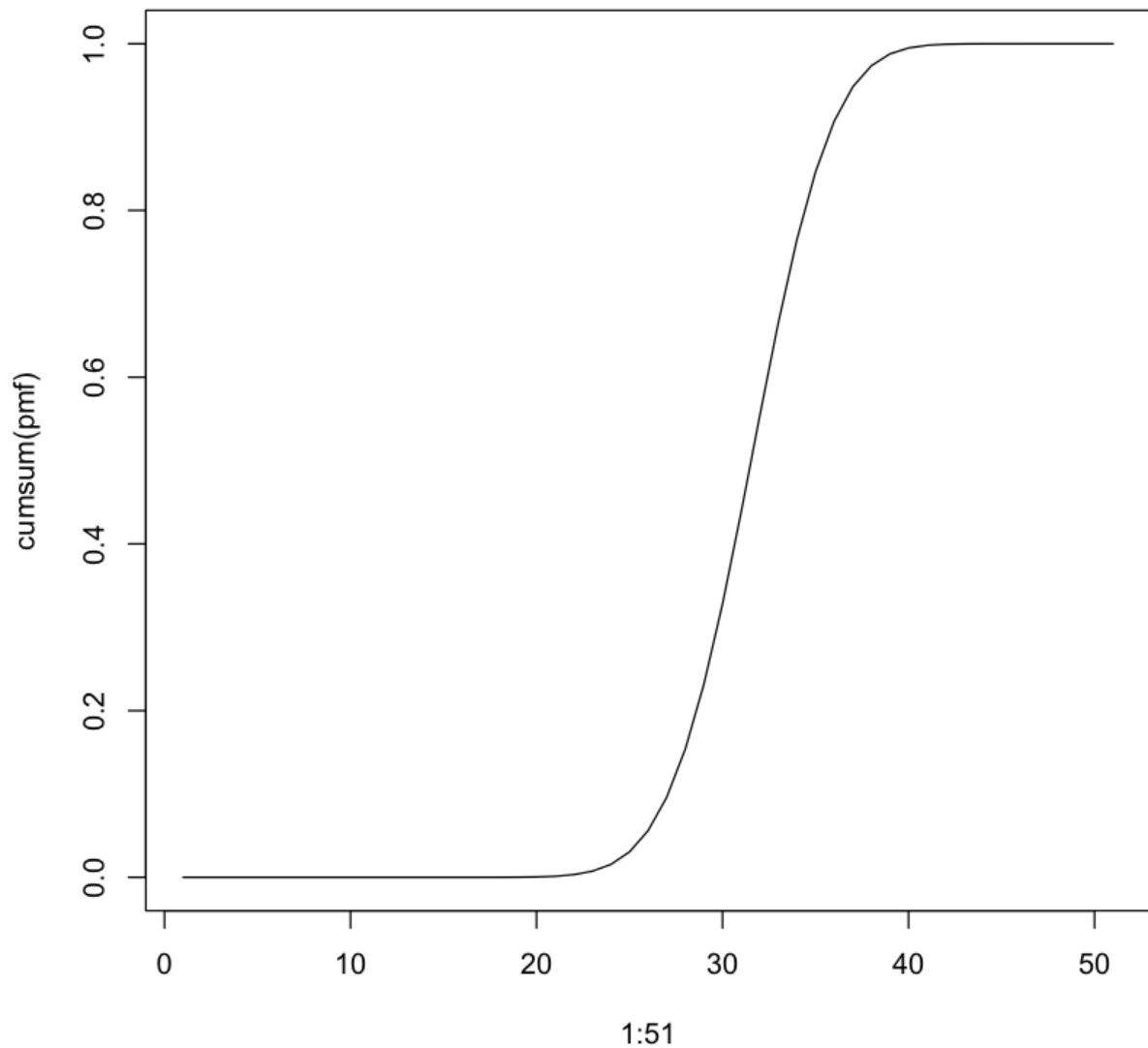
1

```
In [16]: # c.  
# The mean of the distribution is  $50 \times 0.62 = 31$   
# standard deviation is  $\sqrt{50 \times 0.62 \times (1 - 0.62)} = 3.43$ 
```

```
In [20]: # d.  
pmf <- dbinom(0:50, size = 50, prob = 0.62)  
plot(pmf, type = "h", ylim = c(0, 0.5))  
points(pmf, pch = 19)
```



```
In [21]: plot(1:51,cumsum(pmf), type = 'l')
```



2.

What are all the possible codon arrangements available during translation in genetics? explain

```
In [ ]: 4^3 = 64
```

We have 4 types (A,U,C,G). For each position **in** codon, we can have 4 choices. Then, the possible codon arrangements is 64.

```
In [24]: library(gtools)
subun = c("A","U","C","G")
permutations(4,3, subun,repeats.allowed = T)
```

A	A	A
A	A	C
A	A	G
A	A	U
A	C	A
A	C	C
A	C	G
A	C	U
A	G	A
A	G	C
A	G	G
A	G	U
A	U	A
A	U	C
A	U	G
A	U	U
C	A	A
C	A	C
C	A	G
C	A	U
C	C	A
C	C	C
C	C	G
C	C	U
C	G	A
C	G	C
C	G	G
C	G	U
C	U	A
C	U	C
⋮	⋮	⋮
G	A	G
G	A	U
G	C	A
G	C	C
G	C	G
G	C	U
G	G	A
G	G	C
G	G	G
G	G	U
G	U	A

G	U	C
G	U	G
G	U	U
U	A	A
U	A	C
U	A	G
U	A	U
U	C	A
U	C	C
U	C	G
U	C	U
U	G	A
U	G	C
U	G	G
U	G	U
U	U	A
U	U	C
U	U	G
U	U	U

3.

Find the number of three-digit integers that have either 2 or 7 in the tens digit and 2 in the units digit?

In []: For the hundred digit, we have 1~9 9 choices. For tens digit we have 2 choices. And the units digit only has 1 choice.
 $9 \times 2 \times 1 = 18$

4.

Write a function in R that calculates nPk and nCk, and answer the following questions keeping in mind the type of arrangement asked (Permutations or combinations):

- a. How many ways can we give awards to the first 3 places (gold,silver and bronze) among eight contestants?
- b. Claire took 12 photos. She wants to put 5 of them on Facebook. How many groups of photos are possible?

```
In [25]: perm <- function(n, r){
  return(factorial(n)/factorial(n - r))
}

comb <- function(n, r){
  return( factorial(n) / (factorial(n-r) * factorial(r)) )
}
```

```
In [28]: a. perm(8,3) =336

336
```

```
In [30]: b. comb(12,5) =792

792
```

5.

Seven friends went to watch a game. They were supposed to occupy seat numbers 50 to 57. In how many different ways can they do it?

```
In [ ]: They have  $8*7*6*5*4*3*2 = 8!$  choices.
```

6.

A famous example of a markov chain is the recruitment probabilities for males in some ivy league schools in the past, where Harvard, Dartmouth, and Yale admitted only male students. Assume that, at that time, 80 percent of the sons of Harvard men went to Harvard and the rest went to Yale, 40 percent of the sons of Yale men went to Yale, and the rest split evenly between Harvard and Dartmouth; and of the sons of Dartmouth men, 70 percent went to Dartmouth, 20 percent to Harvard, and 10 percent to Yale.

Draw the transition matrix, plot the transition diagram and answer:

Calculate the probability that a grandson from Yale went to Harvard Draw the transition matrix after 5 generations

```
In [1]: library(markovchain)

#Lets define our states
universityStates <- c("Harvard", "Dartmouth", "Yale")
byRow <- TRUE

##lets create the transitional matrix
admitMatrix <- matrix(data = c(0.8, 0, 0.2,
                              0.2, 0.7, 0.1,
                              0.3, 0.3, 0.4), byrow = byRow, nrow = 3,
                      dimnames = list(universityStates, universityStates))

##make the matrix a markovchain class
mcUniversity <- new("markovchain", states = universityStates, byrow = byRow,
                   transitionMatrix = admitMatrix, name = "University")

##define the initial state
initialState <- c(0, 0, 1)

Package: markovchain
Version: 0.6.9.8-1
Date: 2017-08-15
BugReport: http://github.com/spedygiorgio/markovchain/issues
```

```
In [2]: admitMatrix
```

	Harvard	Dartmouth	Yale
Harvard	0.8	0.0	0.2
Dartmouth	0.2	0.7	0.1
Yale	0.3	0.3	0.4

```
In [3]: mcUniversity
```

```
University
A 3 - dimensional discrete Markov Chain defined by the following states:
Harvard, Dartmouth, Yale
The transition matrix (by rows) is defined as follows:
      Harvard Dartmouth Yale
Harvard      0.8      0.0  0.2
Dartmouth     0.2      0.7  0.1
Yale          0.3      0.3  0.4
```

```
In [4]: grandson <- initialState * (mcUniversity ^2)
grandson
```

Harvard	Dartmouth	Yale
0.42	0.33	0.25

```
In [ ]: the probability that a grandson from Yale went to Harvard: 0.42
```

```
In [5]: after5gen <-(mcUniversity ^5)
```

```
In [6]: after5gen
```

University^5

A 3 - dimensional discrete Markov Chain defined by the following states:

Harvard, Dartmouth, Yale

The transition matrix (by rows) is defined as follows:

	Harvard	Dartmouth	Yale
Harvard	0.59006	0.17778	0.23216
Dartmouth	0.49883	0.29620	0.20497
Yale	0.52602	0.25935	0.21463