Week 8 Lab handout

STAT 3021

In this lab handout, we will learn :
1. how to find expected value of discrete probability distribution when a table is given.
2. how to use built-in binomial function to find binomial probability
3. observe shape of binomial distribution for different values of n and p.
4. and how to create and knit R markdown file.

- To create your own R markdown file, select "File-New file - R Markdown"
- To insert R code chunk, click green button "[Insert]" then select "R".
- To generate a document, click "Knit". Use 'Knit to HTML' or 'Knit to Word' unless you have LaTex installed.

Part 1: Expected value in R for simple discrete random variable

Example 1 from Chapter 4.1

```
## x :enter x values
x<-c(4, 5, 6, 7, 8, 9)

## f(x) : enter probability mass function values
fx<-c(1/12, 1/12, 1/4, 1/4, 1/6, 1/6)

## x*f(x)
# x*fx

## E(X)=sum(x*f(x))
sum(x*fx)
```

```
## [1] 6.833333
```

```
## E(2X-3)=sum((2x-3)*f(x))
y<-2*x-1
sum(y*fx)
```

```
## [1] 12.66667
```

Part 2: Ch 5 binomial function

We use dbinom for P(X=x), pbinom for P(X<=x) where X~binom(n, p)

Candy machine example 1. 50% of the candies are orange, 25% are yellow, and 25% are brown. Let X be the number of yellows after drawing 10 candies. What is the probability you have 4 yellows?

$X \sim binom(n = 10, p = 0.25)$. $f(x) = P(X = x) = \binom{n}{x}p^x(1-p)^{n-x} =_n C_x p^x(1-p)^{n-x}$ $P(X = 4) = \binom{10}{4}(0.25^4)(0.75^6) = 0.145998$

```r
## by using ""choose" function
choose(10, 4)*.25^4*.75^6
```

```
## [1] 0.145998
```

We can also use built-in function "dbinom(x, n, p)" (density of binomial distribution)

```r
dbinom(4, 10, .25)
```

```
## [1] 0.145998
```

Candy machine Example 2. What is the probability you have at most 4 yellows? That is, what is P(X<=4)?

To find $P(X \leq x)$ where $X \sim binom(n, p)$, use the built-in cumulative binomial distribution function "pbinom(x, n, p)"

```r
pbinom(4, 10, .25)
```

```
## [1] 0.9218731
```

Candy machine Example 3. What is the probability you have at least 4 yellows? That is, what is P(X>=4)?

$P(X \geq 4) = P(X = 4, 5, 6, .., 9, 10) = 1 - P(X = 0, 1, 2, 3) = 1 - P(X \leq 3)$

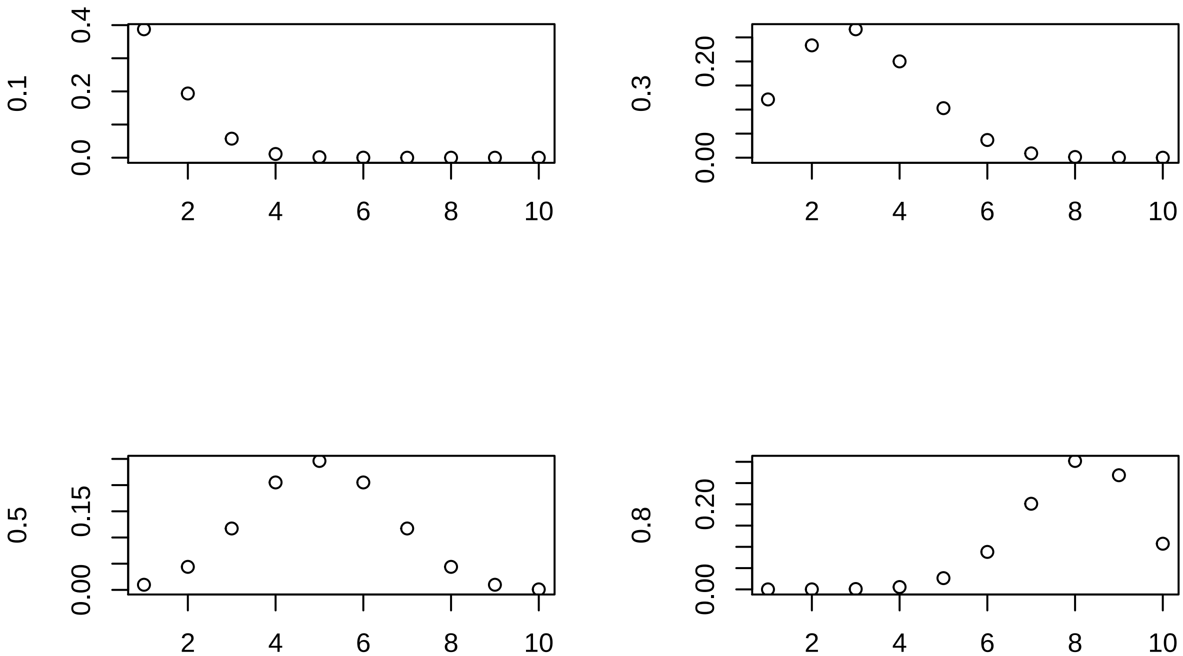
```r
1-pbinom(3, 10, 0.25)
```

```
## [1] 0.2241249
```

Part 3: Shape of Binomial distriution with different n, p

Round 1: n=10 & different probabiliy of successes (p).

```r
n=10 #number of trials
p=c(0.1, 0.3, 0.5, 0.8) #probability of success
x<-c(1:n)
par(mfrow=c(2,2)) ## setting a 2 by 2 frame
for (i in c(1:4)){
  plot(x, dbinom(x, n, p[i]), xlab="", ylab=p[i])}
```



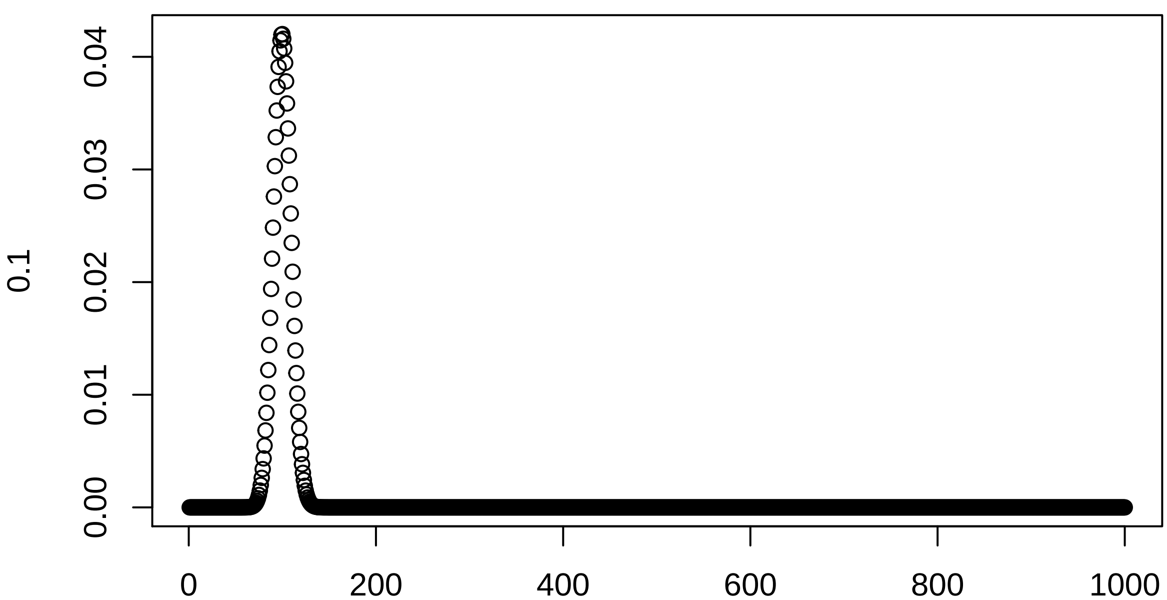
For $n = 10$, if $p = 0.1$ or $p = 0.3$, the distribution of $X \sim binom(n, p)$ is skewed to the right.

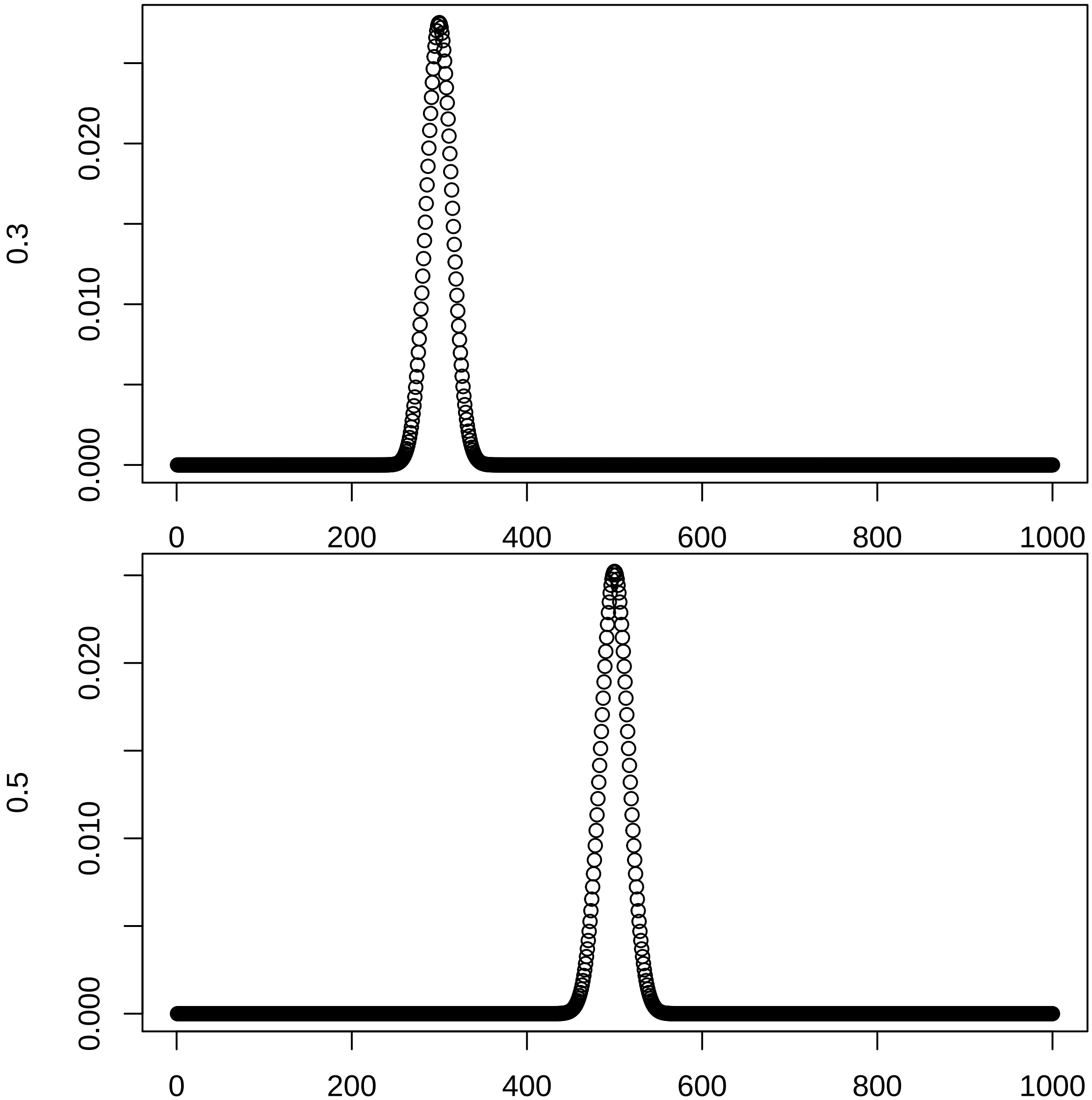
If $p = 0.5$, the distribution of X is symmetric.

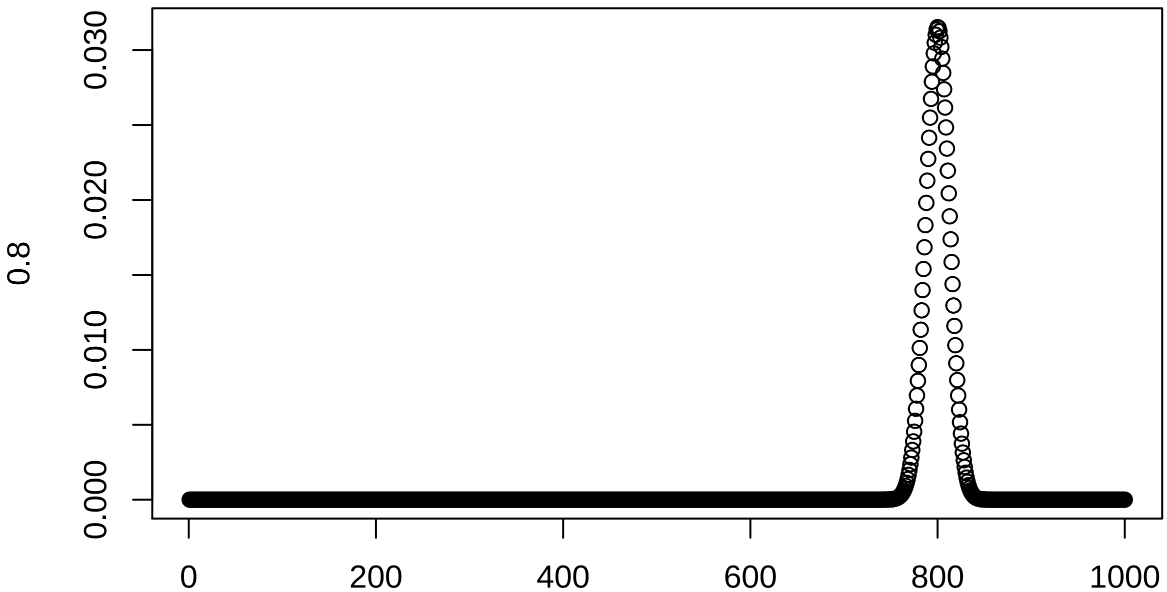
If $p = 0.8$, the distribution of X is skewed to the left.

Round 2: With n=1000 and different probability values.

```
n=1000
p=c(0.1, 0.3, 0.5, 0.8)
x<-c(1:n)
par(mfrow=c(1,1))
for (i in c(1:4)){
  plot(x, dbinom(x, n, p[i]), xlab="", ylab=p[i])}
```




0.3
0.000
0.010
0.020
0
200
400
600
800
1000
0.5
0.000
0.010
0.020
0
200
400
600
800
1000



For binomial distribution, if n is large (n=1000 in this example), then the distributions are asymptotically symmetric.

You zoom the range of x-axis to better observe this. You can adjust the range of x-axis using xlim=c(a, b) option.

```
n=1000
p=0.1
x<-c(1:n)
par(mfrow=c(1,1))
plot(x, dbinom(x, n, p), xlab="", ylab=p, xlim=c(50, 150), main="Shape of binom(n=1000, p=0.1)")
```