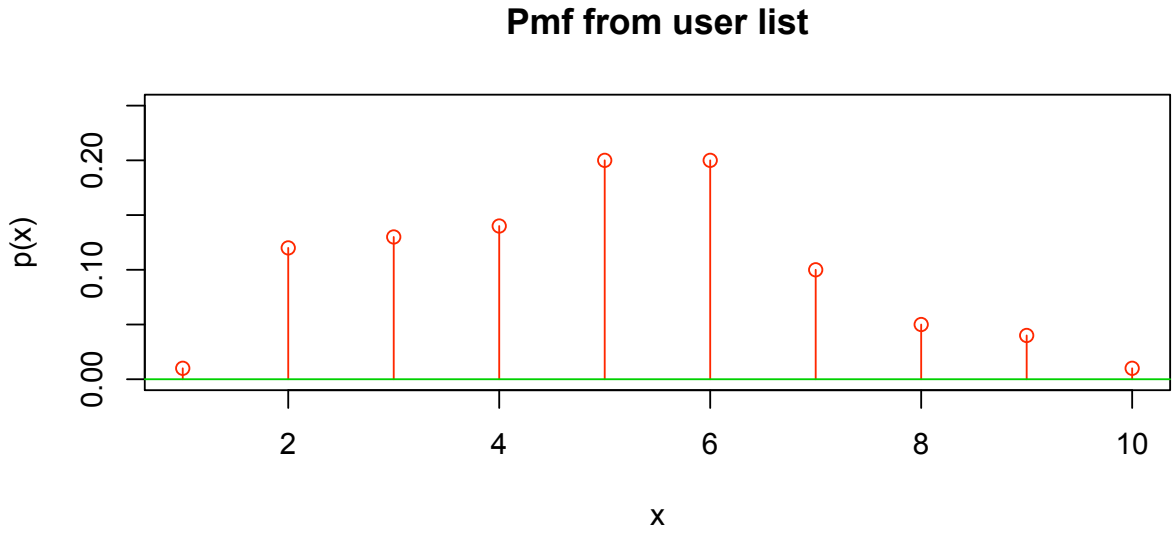
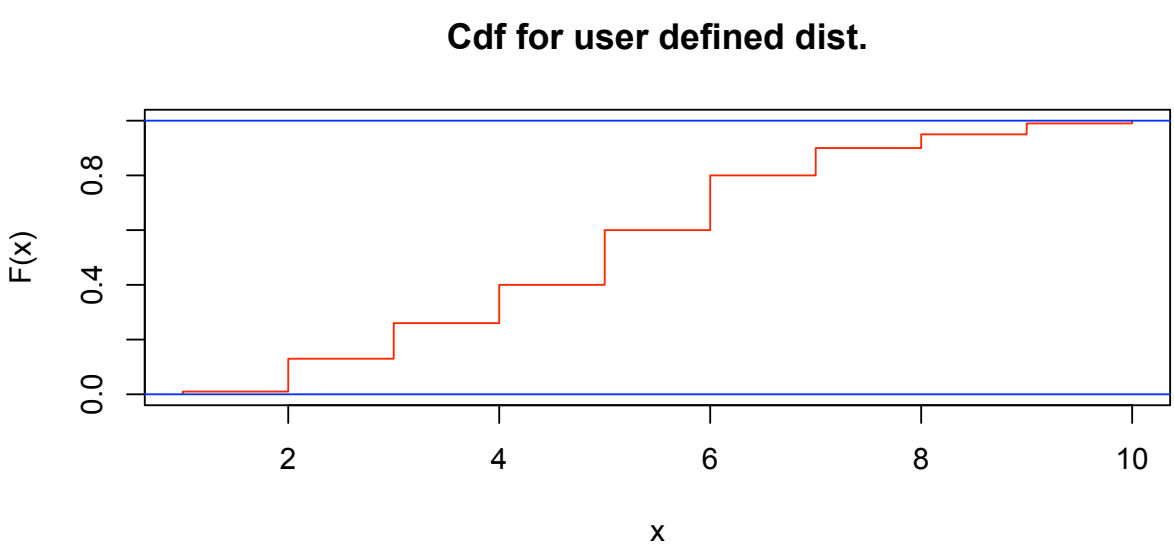
Math 3070 § 1. **PMF's and CDF's: User Specified,** Name: Example
Treibergs **Binary and Poisson Variables** October 6, 2010

R Session:

```
R version 2.10.1 (2009-12-14)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.31 (5538) powerpc-apple-darwin8.11.1]

[Workspace restored from /Users/andrejstreibergs/.RData]


################################################################################
#  USER DEFINED PROBABILITY MASS FUNCTION
#  SUPPOSE THE RANDOM VARIABLE CAN BE ALL INTEGERS 1 TO 10
#
> DD <- 1:10

#  ASSIGN PMF VALUES, EQ  p(1)=.01, p(2)=.12, etc.
PP <- c(.01,.12,.13,.14,.2,.2,.1,.05,.04,.01)

# CHECK THAT p(x)  SUMS TO ONE.
> sum(PP)
[1] 1


# THE FOLLOWING PUTS TWO GRAPHS ONE ON TOP OF THE OTHER ON EACH PAGE
> require(graphics)
> par(mfrow = c(2, 1))

# NOW PLOT THE PMF. type="h" MEANS THAT THE PLOT IS VERTICAL LINES.
# points ADDS CIRCLES. abline ADDS AXES.
> plot(DD,PP,type="h",col=2,main="Pmf from user list",xlab="x",ylab="p(x)")
> points(DD,PP,col=2);abline(h=0,col=3)
```

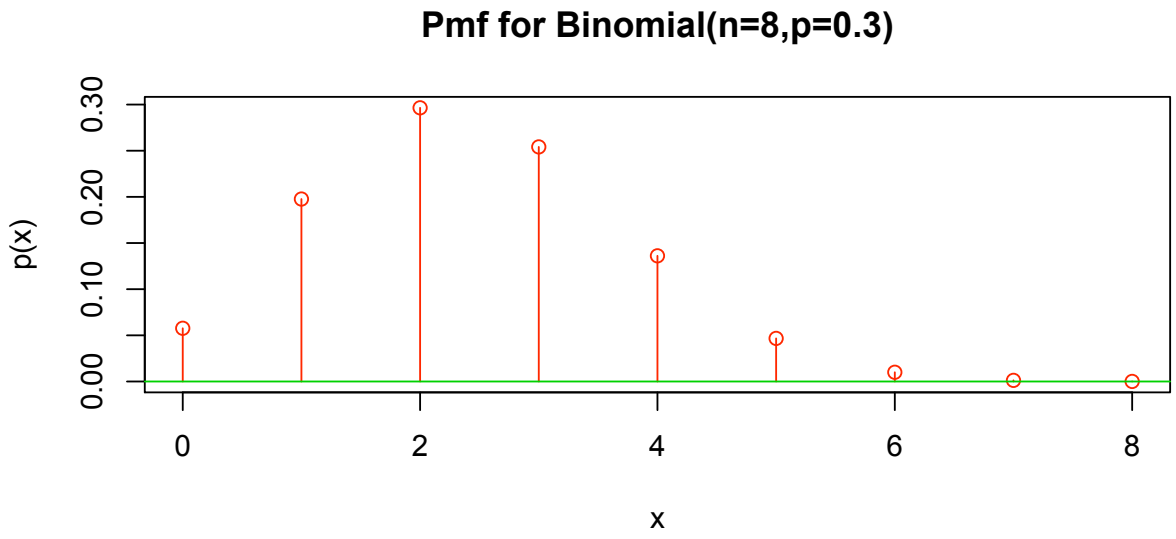
```
# TO GET THE CUMULATIVE DISTRIBUTION FUNCTION, WE NEED TO GET PARTIAL SUMS OF THE PDF.
> QQ <- cumsum(PP)

# SEE HOW THE cumulative sum QQ IS A LIST OF PARTIAL SUMS FROM PP.
> PP
 [1] 0.01 0.12 0.13 0.14 0.20 0.20 0.10 0.05 0.04 0.01
> QQ
 [1] 0.01 0.13 0.26 0.40 0.60 0.80 0.90 0.95 0.99 1.00


# NOW PLOT THE PMF. type="S" MEANS THAT THE PLOT IS STAIRCASE.  abline ADDS AXES.
# c(1,DD) AND c(0,QQ) ADDS A ZERO STARTING POINT TO MAKE PLOT START FROM ZERO.

> plot(c(1,DD),c(0,QQ),type="s", ylab="F(x)",col=2,xlab="x",
+ main="Cdf for user defined dist.");abline(h=0:1,col=4)
```

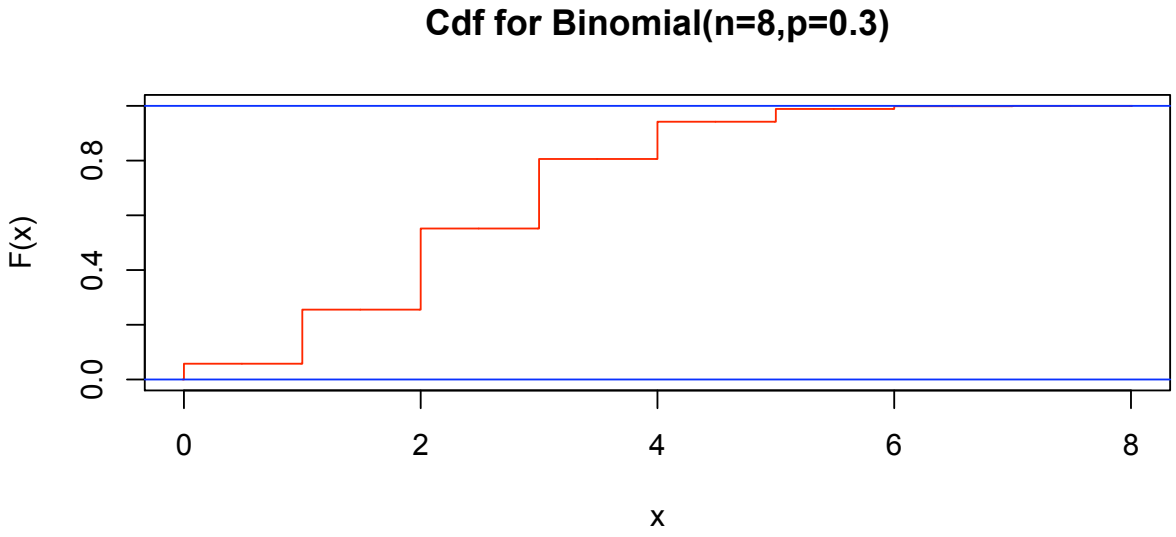

Pmf from user list
p(x)
0.00
0.10
0.20
2
4
6
8
10
x


Cdf for user defined dist.
F(x)
0.0
0.4
0.8
2
4
6
8
10
x

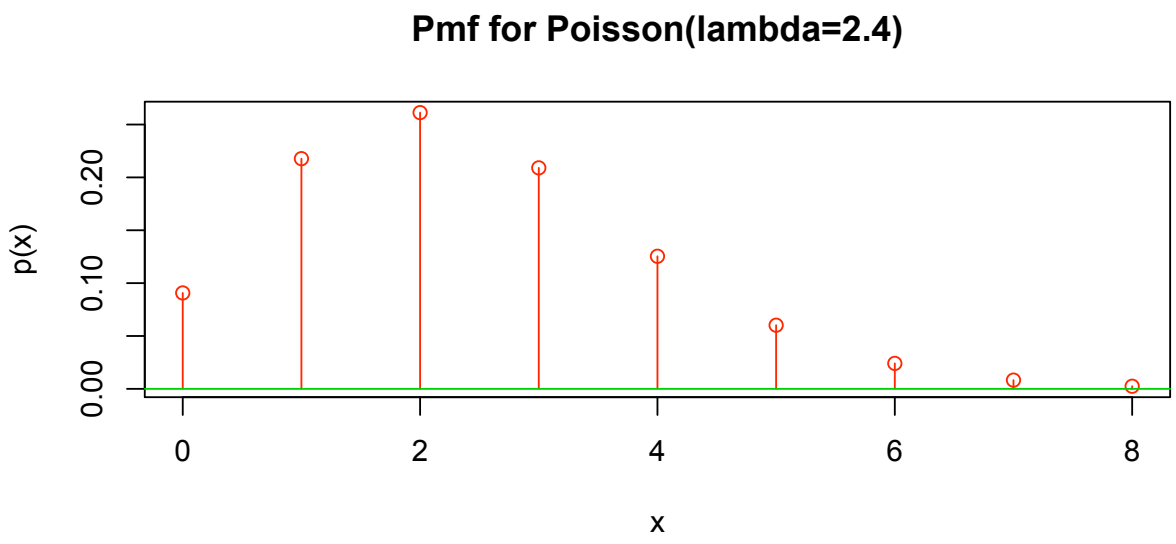
```
################################################################################
#  BINOMIAL AND POISSON VARIABLES ARE CANNED IN R

# FOR EXAMPLE, IF lambda=2 THEN PMF FOR p(x)=P(X=x) where
# X ~ Poisson(lambda=2) is by hand, p(10)=
> exp(-8)*8^10/factorial(10)
[1] 0.09926153

# IT IS DESCRIBED BY
> help(poisson)

# THE CANNED POISSON PDF IS CALLED dpois(x,lambda)
> dpois(10,8)
[1] 0.09926153

# THE CANNED BINOMIALS INCLUDE PDF dbinom(x,x,p) and the CDF pbinom(x,n,p)
# YOU CAN READ ABOUT IT IN
> help(dbinom)

# E.G., IF n-5 AND p=0.5 WE LIST OUT ALL x FROM 0 TO 5
> dbinom(0:5,5,.5)
[1] 0.03125 0.15625 0.31250 0.31250 0.15625 0.03125
> pbinom(0:5,5,.5)
[1] 0.03125 0.18750 0.50000 0.81250 0.96875 1.00000

# THE SOLUTION TO PROBLEM 125[88]b.  IS
> 1-pbinom(3,200,.01)
[1] 0.1419660

# THE POISSON APPROXIMATION IS
> 1-ppois(3,2)
[1] 0.1428765

################################################################################
# PLOT THE BINOMIAL PMF AND CDF FOR n=8 AND p=0.3

> DD <- 0:8
> PP <- dbinom(DD,8,0.3); PP
[1] 0.05764801 0.19765032 0.29647548 0.25412184 0.13613670 0.04667544
[7] 0.01000188 0.00122472 0.00006561
> pbinom(DD,8,0.3)
[1] 0.05764801 0.25529833 0.55177381 0.80589565 0.94203235 0.98870779
[7] 0.99870967 0.99993439 1.00000000

> plot(DD,PP,type="h",col=2,main="Pmf for Binomial(n=8,p=0.3)",xlab="x",ylab="p(x)");
> points(DD,PP,col=2); abline(h=0,col=3)


# THE EASIEST WAY TO PLOT THE CDF FOR THE SAME BINOMIAL VARIABLE IS TO VIEW
# pbinom(x,8,0.3) AS A FUNCTION OF THE REAL VARIABLE x PLOTTED FOR -0.01 < x < 8.01.
# XX HOLDS REAL NUMBERS AT INCREMENTS OF .01. type="s" MEANS STEP FUNCTION
```

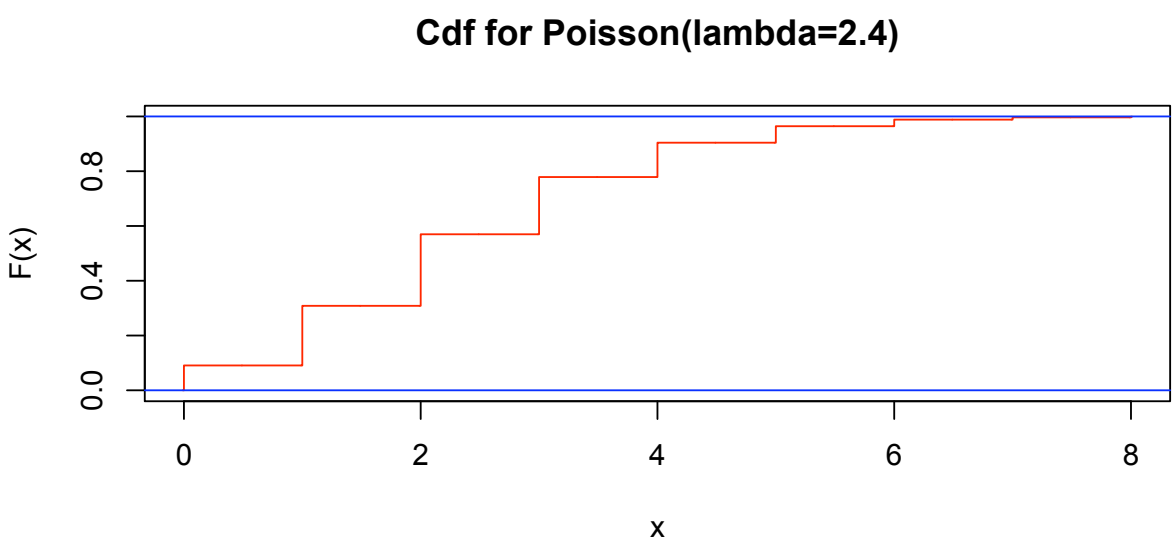
```
> XX <- seq(-0.01, 8.01, 0.01)
> plot(XX, pbinom(XX, 8, 0.3),type="s", ylab="F(x)",col=2,xlab="x",
main="Cdf for Binomial(n=8,p=0.3)"); abline(h=0:1,col=4)
```





```
################################################################################
# TO COMPARE BINOMIAL AND POISSON, USE SAME EXPECTED VALUE. lambda = pn =2.4

# VALUES OF PMF AND CDF  FROM 0 TO 6 FOR POISSON(lambda=2.4)
> DD <- 0:8
> PP <- dpois(DD,2.4); PP
[1] 0.090717953 0.217723088 0.261267705 0.209014164 0.125408499 0.060196079
[7] 0.024078432 0.008255462 0.002476639
> ppois(DD,2.4)
[1] 0.09071795 0.30844104 0.56970875 0.77872291 0.90413141 0.96432749
[7] 0.98840592 0.99666138 0.99913802

> plot(DD,PP,type="h",col=2,main="Pmf for Poisson(lambda=2.4)",xlab="x",ylab="p(x)")
> points(DD,PP,col=2); abline(h=0,col=3)
> XX <- seq(-0.01, 8.01, 0.01)
> plot(XX, ppois(XX, 2.4),type="s", ylab="F(x)",col=2,xlab="x",
    main="Cdf for Poisson(lambda=2.4)");abline(h=0:1,col=4)
```


Pmf for Poisson(lambda=2.4)
p(x)
0.00
0.10
0.20
0
2
4
6
8
x


Cdf for Poisson(lambda=2.4)
F(x)
0.0
0.4
0.8
0
2
4
6
8
x