Math 3070 § 1. Treibergs | **Approximating the Binomial by Poisson Distribution Example.** | Name: Example June 13, 2011

The binomial distribution describes the probability that there will be x successes in a sample of size n, chosen with replacement from a population where the probability of success is p. In **R**, this is given by

$$P(X = x) = \mathrm{b}(x; n, p) = \mathtt{dbinom}(\mathtt{x}, \mathtt{n}, \mathtt{p}.$$

If $n \to \infty$, $p \to 0$ and $np \to \lambda$ with $\lambda > 0$, then $b(x, n, p) \to p(x; \lambda)$, where the Poisson pmf is given in **R** by

$$P(X = x) = p(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \mathtt{dpois}(x, \lambda)$$

Devore's rule of thumb, is that it's O.K. to approximate a binomial pmf by Poisson pmf if n is large and p is small.

> In a binomial experiment where each trial results in S or F and the sampling is with replacement from a population where the probability of success is p, if the sample size $n > 50$ and $np < 5$, then the experiment may be analyzed as if it were a Poisson experiment with $\lambda = np$.

We print and graph the binomial pmf and its Poisson approximation. We see how the errors compare as n increases and p decreases for several np's

---

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]
```

```
> ############## SUBPROGRAM TO LIST BINOMIAL VS POISSON PMF'S ###############
> # We use fix(lis)
> lis<- function(n,p,a,b){}
> fix(lis)
> lis <- function(n,p,a,b)
+               {
+                lambda <- n*p
+                x<- matrix(numeric((b-a+1)*3),ncol=3,
+                dimnames = list(a:b,c("Binomial","Poisson","Difference")))
+                x[,1]<-dbinom(a:b,n,p)
+                x[,2]<-dpois(a:b,lambda)
+                x[,3]<-x[,1]-x[,2]
+                error <- max(abs(x[,3]))
+                cat("Poisson Approx to Binomial\n n =", n, "\n p =", p,
+                "\n lambda =", lambda, "\n")
+                print (x)
+                cat("\n Error (Maximum Absolute Difference) = ",error,
+                "\n\n")
+               }

> lis(10,.3,0,10)
Poisson Approx to Binomial
 n = 10
 p = 0.3
 lambda = 3
      Binomial      Poisson    Difference
0  0.0282475249 0.0497870684 -0.0215395435
1  0.1210608210 0.1493612051 -0.0283003841
2  0.2334744405 0.2240418077  0.0094326328
3  0.2668279320 0.2240418077  0.0427861243
4  0.2001209490 0.1680313557  0.0320895933
5  0.1029193452 0.1008188134  0.0021005318
6  0.0367569090 0.0504094067 -0.0136524977
7  0.0090016920 0.0216040315 -0.0126023395
8  0.0014467005 0.0081015118 -0.0066548113
9  0.0001377810 0.0027005039 -0.0025627229
10 0.0000059049 0.0008101512 -0.0008042463

 Error (Maximum Absolute Difference) =  0.04278612

> # Devore's rule of thumb does not hold. Error = .04 is large.
```

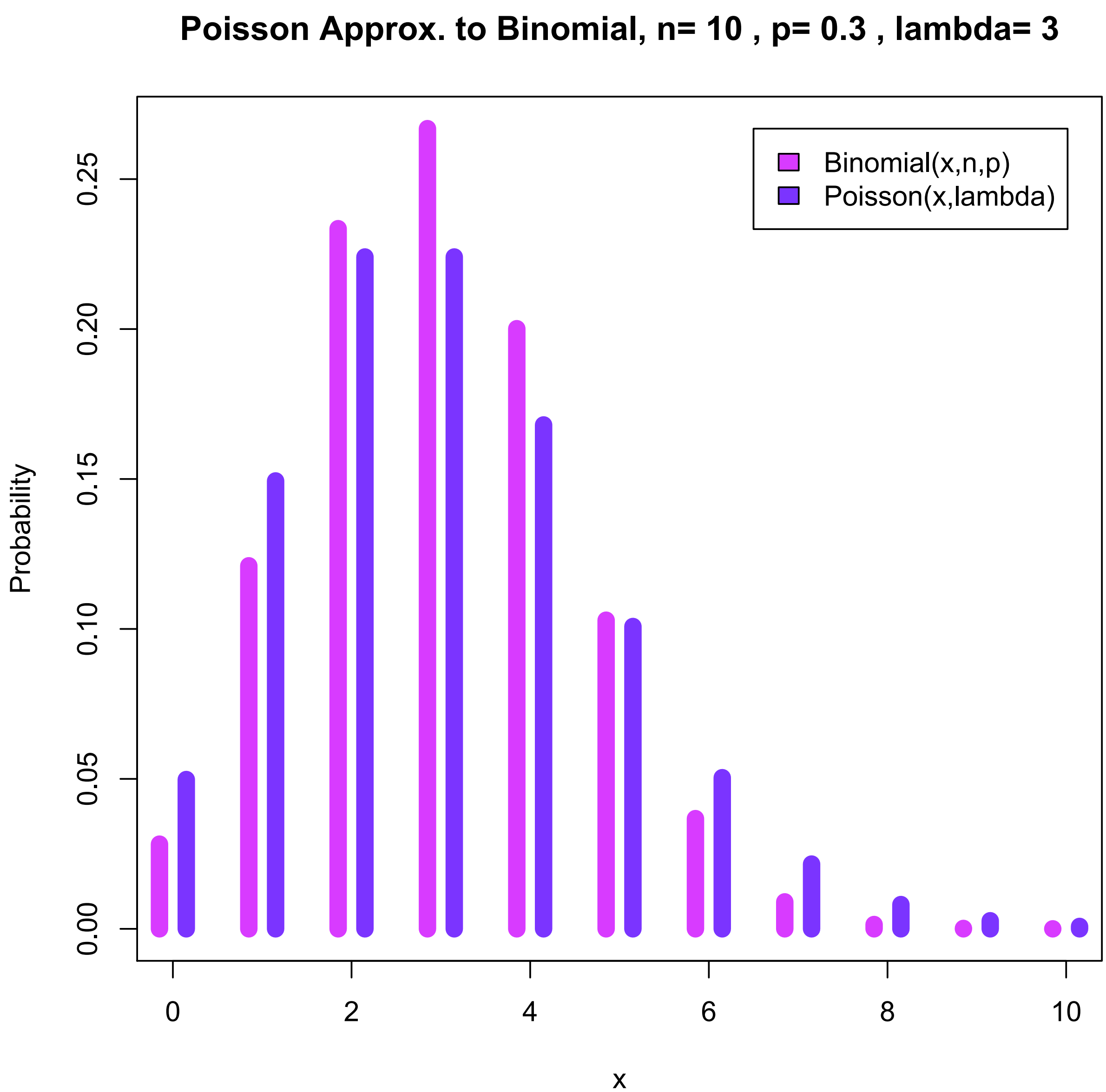
```
> lis(100,.03,0,10)
Poisson Approx to Binomial
 n = 100
 p = 0.03
 lambda = 3
       Binomial      Poisson    Difference
0  0.0475525079 0.0497870684 -0.0022345604
1  0.1470696121 0.1493612051 -0.0022915930
2  0.2251529629 0.2240418077  0.0011111553
3  0.2274741275 0.2240418077  0.0034323198
4  0.1706055956 0.1680313557  0.0025742399
5  0.1013080650 0.1008188134  0.0004892516
6  0.0496096195 0.0504094067 -0.0007997873
7  0.0206037006 0.0216040315 -0.0010003309
8  0.0074077738 0.0081015118 -0.0006937380
9  0.0023419766 0.0027005039 -0.0003585273
10 0.0006591336 0.0008101512 -0.0001510176

 Error (Maximum Absolute Difference) =  0.00343232


> # Devore's rule of thumb does hold. Error = .003 is smaller.


> lis(1000,.003,0,10)
Poisson Approx to Binomial
 n = 1000
 p = 0.003
 lambda = 3
       Binomial      Poisson    Difference
0  0.0495630828 0.0497870684 -2.239855e-04
1  0.1491366584 0.1493612051 -2.245467e-04
2  0.2241537439 0.2240418077  1.119363e-04
3  0.2243785721 0.2240418077  3.367645e-04
4  0.1682839291 0.1680313557  2.525734e-04
5  0.1008690833 0.1008188134  5.026984e-05
6  0.0503333690 0.0504094067 -7.603768e-05
7  0.0215065348 0.0216040315 -9.749664e-05
8  0.0080325937 0.0081015118 -6.891811e-05
9  0.0026641033 0.0027005039 -3.640064e-05
10 0.0007944212 0.0008101512 -1.573001e-05

 Error (Maximum Absolute Difference) =  0.0003367645

 > # Devore's rule of thumb does hold. Error = .0003 is small.
```

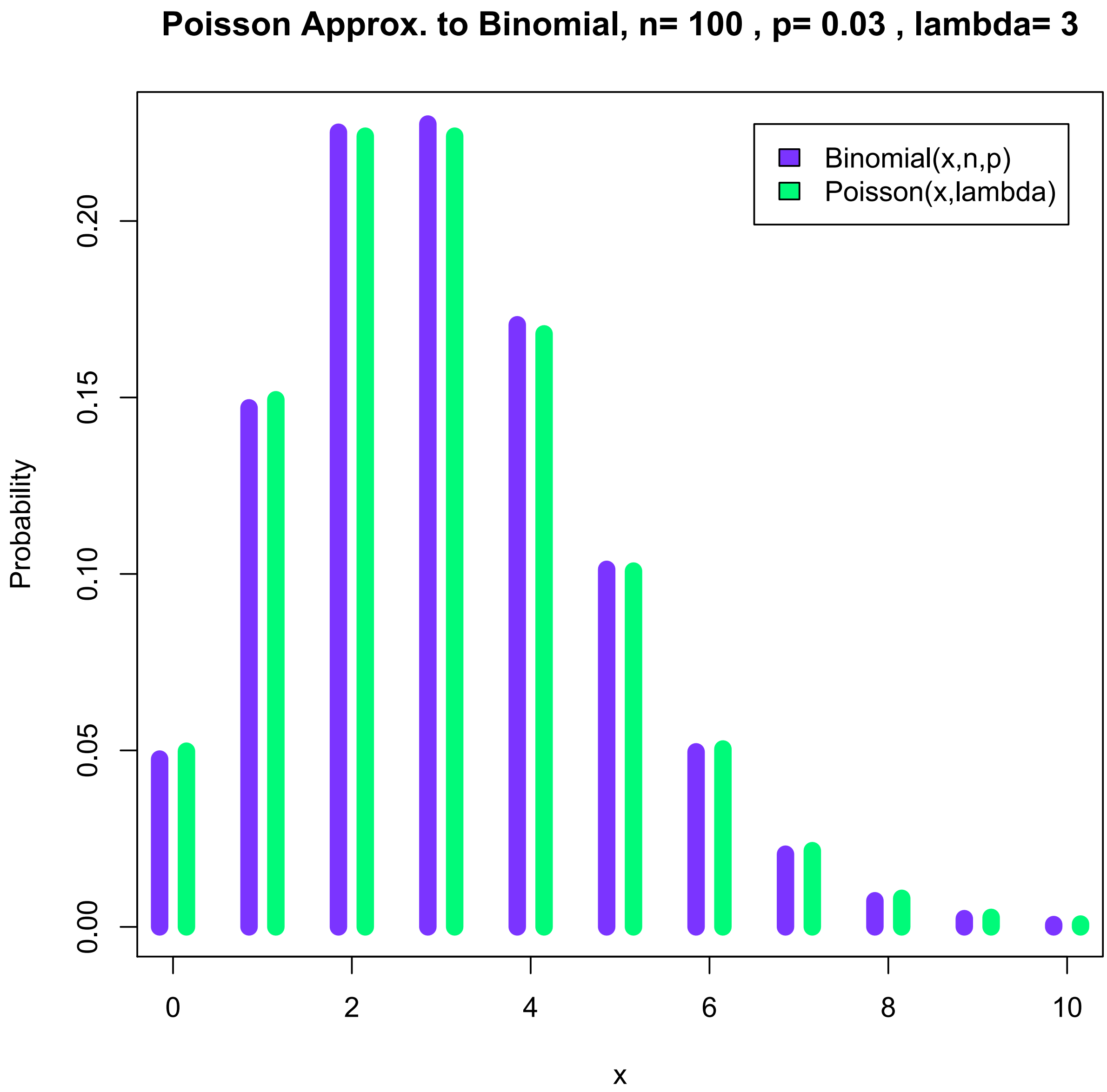
```
> lis(10000,.0003,0,10)
Poisson Approx to Binomial
 n = 10000
 p = 3e-04
 lambda = 3
      Binomial      Poisson    Difference
0  0.049764665 0.0497870684 -2.240362e-05
1  0.149338796 0.1493612051 -2.240922e-05
2  0.224053009 0.2240418077  1.120125e-05
3  0.224075421 0.2240418077  3.361327e-05
4  0.168056566 0.1680313557  2.520996e-05
5  0.100823853 0.1008188134  5.039553e-06
6  0.050401841 0.0504094067 -7.565634e-06
7  0.021594307 0.0216040315 -9.724596e-06
8  0.008094625 0.0081015118 -6.886842e-06
9  0.002696859 0.0027005039 -3.645123e-06
10 0.000808572 0.0008101512 -1.579117e-06

 Error (Maximum Absolute Difference) =  3.361327e-05


 > # Devore's rule of thumb does hold. Error = .00003 is small.
```

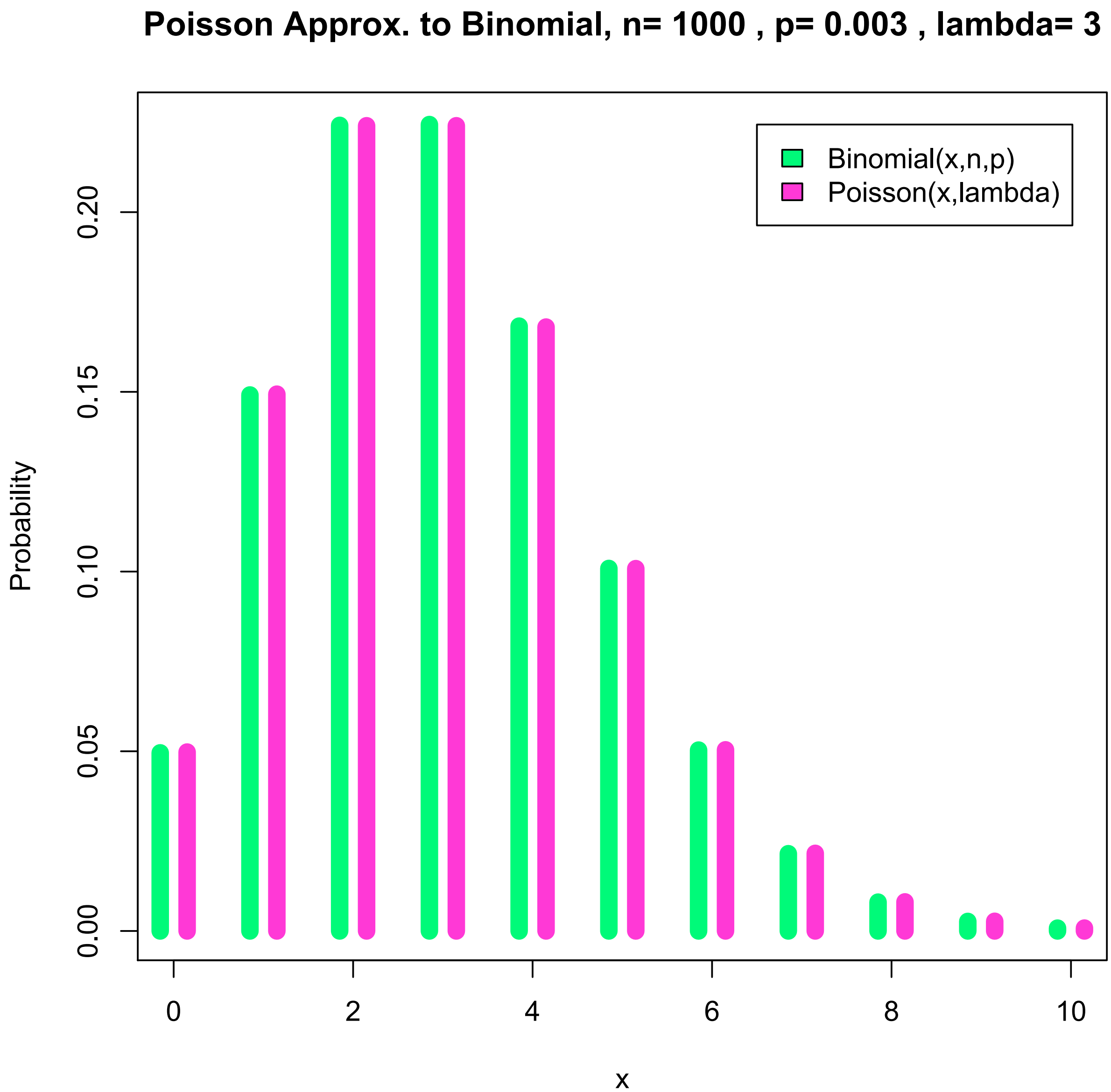
```
> lis(100000,.00003,0,10)
Poisson Approx to Binomial
 n = 1e+05
 p = 3e-05
 lambda = 3
       Binomial      Poisson    Difference
0  0.0497848280 0.0497870684 -2.240412e-06
1  0.1493589646 0.1493612051 -2.240468e-06
2  0.2240429279 0.2240418077  1.120201e-06
3  0.2240451684 0.2240418077  3.360697e-06
4  0.1680338763 0.1680313557  2.520523e-06
5  0.1008193175 0.1008188134  5.040802e-07
6  0.0504086505 0.0504094067 -7.561833e-07
7  0.0216030592 0.0216040315 -9.722092e-07
8  0.0081008232 0.0081015118 -6.886341e-07
9  0.0027001394 0.0027005039 -3.645625e-07
10 0.0008099932 0.0008101512 -1.579727e-07

 Error (Maximum Absolute Difference) =  3.360697e-06


> # Devore's rule of thumb does hold. Error = .000003 is small.
```

```
> ################### ERROR FOR VARIOUS n AND p ############################
> ps <- c(.00001,.0001,.001,.0025,.005,.01,.02,.04,.1,.2,.4)
> ns <- c(10,100,1000,10000)
> ps
 [1] 0.00001 0.00010 0.00100 0.00250 0.00500 0.01000 0.02000 0.04000 0.10000 0.20000 0.40000
> ns
[1]    10   100  1000 10000
> z <- matrix(1:88,ncol=8)
> for (i in 1:4)
+               {
+                for(j in 1:11)
+                              {
+                               z[j,2*i-1]<- round(mx(ns[i],ps[j]),10)
+                               if((ns[i] > 50) &  (ns[i]*ps[j]< 5))
+                                                                     {
+                                                                      z[j,2*i]<-"<-RoT OK   "
+                                                                     }
+                               else
+                                                                     {
+                                                                      z[j,2*i]<-"   "
+                                                                     }
+                              }
+                }
> print(z,quote=FALSE)
        10                 100                        1000                       10000
0.00001 0.000000001        0.00000001   <-RoT OK    0.0000000985 <-RoT OK    0.0000008596 <-RoT OK
0.0001  0.0000000999       0.0000009852 <-RoT OK    0.0000085968 <-RoT OK    0.000018395  <-RoT OK
0.001   0.0000098608       0.0000860427 <-RoT OK    0.0001840471 <-RoT OK    0.0000626009
0.0025  0.0000603445       0.0004268818 <-RoT OK    0.0002944837 <-RoT OK    0.0000995888
0.005   0.0002330077       0.0011419247 <-RoT OK    0.0004402467             0.0001413404
0.01    0.0008679829       0.0018501965 <-RoT OK    0.0006301754             0.000200809
0.02    0.0030034018       0.0027433451 <-RoT OK    0.0009017519             0.0002862738
0.04    0.0088855799       0.0040216823 <-RoT OK    0.0012977906             0.0004112312
0.1     0.0195410478       0.0067553111             0.0021557941             0.0006823442
0.2     0.0313193215       0.0104648974             0.0033276335             0.0010528718
0.4     0.0554558412       0.0182721056             0.00580186               0.0018354604
>
> # "RoT OK" = Rule of Thumb holds.
```

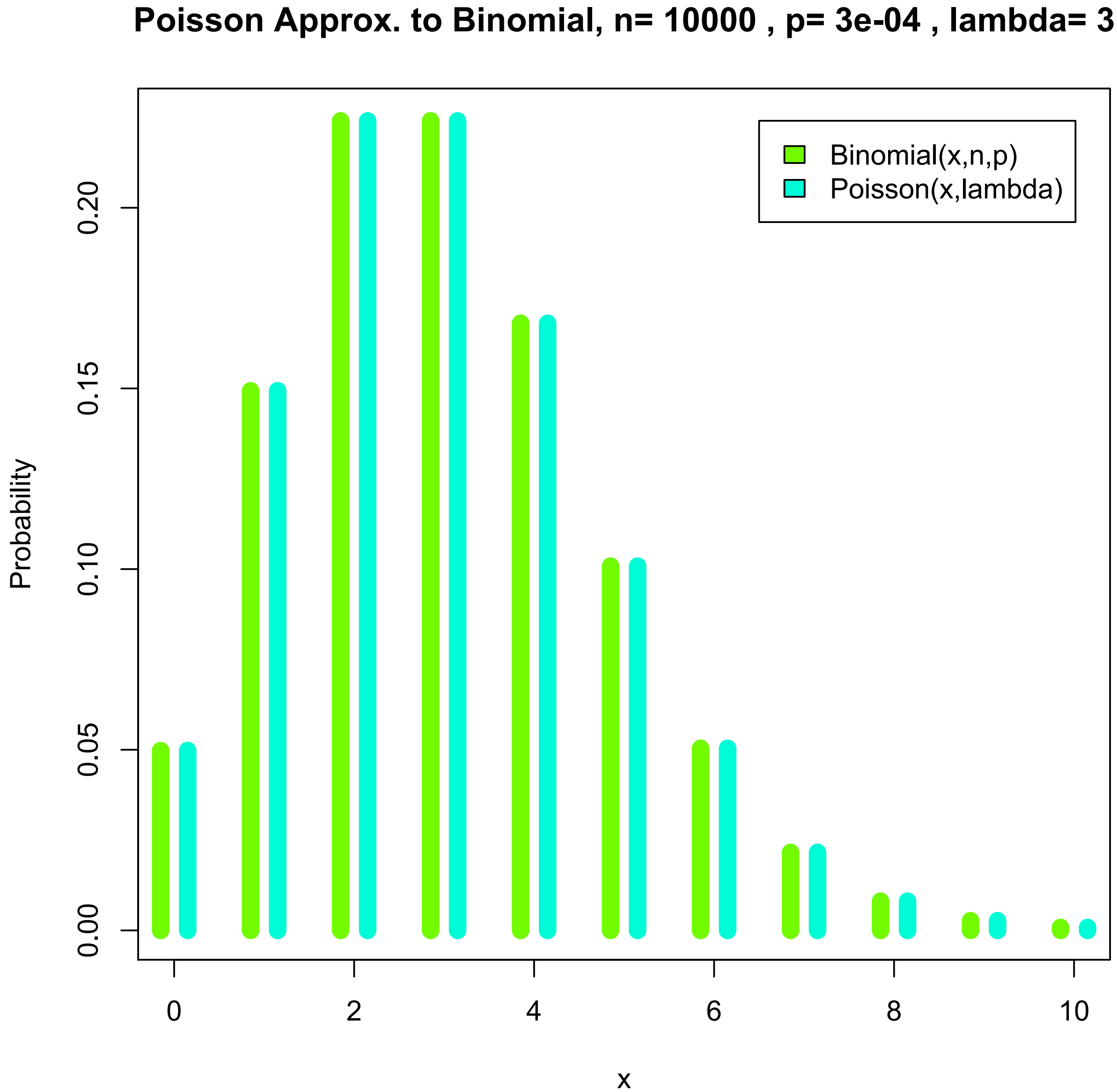
```
> ################## PLOT POISSON AND BINOMIAL PMF'S ##############
> # Write a subprogram to do plot of Binomial vs. Poisson
> pl <- function(n,p,a,b)
+                {
+                 clr<-rainbow(15)[ceiling(runif(2,4,15))]
+                 lambda <- n*p
+                 mx <- max(dbinom(a:b,n,p))
+                 plot(c(a:b,a:b),c(dbinom(a:b,n,p), dpois(a:b,lambda)),
+                 type="n", main = paste("Poisson Approx. to Binomial, n=", n,
+                 ", p=", p, ", lambda=",lambda), ylab = "Probability", xlab="x")
+                 points((a:b)-.15,dbinom(a:b,n,p), type = "h",
+                 col = clr[1], lwd = 10)
+                 points((a:b)+.15,dpois(a:b,lambda), type="h",
+                 col = clr[2], lwd=10)
+                 legend(b-3.5, mx, legend=c("Binomial(x,n,p)",
+                 "Poisson(x,lambda)"), fill = clr, bg="white")
+                }
> pl(10,.3,0,10)
> # M3074PoisApprox1.pdf
>
> pl(100,.03,0,10)
> # M3074PoisApprox2.pdf
>
> pl(1000,.003,0,10)
> # M3074PoisApprox3.pdf
>
> pl(10000,.0003,0,10)
> # M3074PoisApprox4.pdf
```


Poisson Approx. to Binomial, n= 10 , p= 0.3 , lambda= 3
Binomial(x,n,p)
Poisson(x,lambda)
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0
2
4
6
8
10
x


Poisson Approx. to Binomial, n= 100 , p= 0.03 , lambda= 3
Binomial(x,n,p)
Poisson(x,lambda)
Probability
0.00
0.05
0.10
0.15
0.20
0
2
4
6
8
10
x



Poisson Approx. to Binomial, n= 1000 , p= 0.003 , lambda= 3
Binomial(x,n,p)
Poisson(x,lambda)
Probability
0.00
0.05
0.10
0.15
0.20
0
2
4
6
8
10
x


Poisson Approx. to Binomial, n= 10000 , p= 3e-04 , lambda= 3
Binomial(x,n,p)
Poisson(x,lambda)
Probability
0.00
0.05
0.10
0.15
0.20
0
2
4
6
8
10
x