Lab Two: Visualizing the Poisson and Normal Approximations to the Binomial

STA 111: Probability & Statistical Inference

Lab Objective

To visualize and understand the Poisson approximation to the binomial, and the normal approximation to both the Poisson and binomial.

Introduction

Today we will verify that the Poisson approximation to the binomial, and the normal approximation to both the Poisson and binomial all work, and also verify the conditions under which they work. First, lets learn some commands to help us draw random samples for all the distributions and plot the results.

1. To draw a random sample of size $N = 50$ from $Bin(n = 10, p = 0.4)$ for example and save the result, type `N = 50; n = 10; p = 0.4` to save the values we need beforehand and type `X = rbinom(N,n,p)` to draw the sample. You can visualize the distribution by plotting a histogram or density plot. Type `plot(density(X),col="green")` or `hist(X,freq=FALSE,col="green")`. *I chose green for the color but you can pick any color you want. R has a lot of color options built in but that's not important right now. You can read about the color options online.*

2. Now to visualize two different samples in the same plot, first draw another random sample not necessarily from the same distribution. In fact, lets try the Poisson. Remember that we can approximate the $Bin(n, p)$ distribution using the $Poi(\lambda = np)$ distribution. Type `Y = rpois(N,lambda=(n*p))` and add a new plot to the previous one by typing `lines(density(Y),col="red")` *I chose a different color to help differentiate the two distributions.*

3. How about one more sample. Let's draw from a normal distribution this time. Remember that we can also approximate the $Bin(n, p)$ distribution using the $No(\mu = np, \sigma^2 = np(1 - p))$ distribution. Type `Z = rnorm(N, mean=(n*p),sd =sqrt(n*p*(1-p)))` and add a new plot to the previous one by typing `lines(density(Z),col="blue")`. *I chose yet another color to help differentiate the three distributions.*

Lab Questions

Set $N = 10000$ to make the distributions fairly smooth. For every possible pair of $n = 10, 50, 100, 500$ and $p = 0.05, 0.2$ and 0.7, compare the $Bin(n, p)$, $Poi(\lambda = np)$ and $N(\mu = np, \sigma^2 = np(1 - p))$ distributions like we did above by looking at an overlay of the three distributions and answer the following questions:

a. For what pairs of n and p is the Poisson approximation to the binomial good?

b. For what pairs of n and p is the normal approximation to the binomial good?

It turns out that we can also approximate the $Poi(\lambda)$ distribution with $N(\mu = \lambda, \sigma^2 = \lambda)$. Compare the Poisson and Normal distributions for the same n, p pairs above by looking at an overlay of the $Poi(\lambda = np)$ and $No(\mu = np, \sigma^2 = np)$ distributions.

c. For what pairs of n and p is the normal approximation to the Poisson good?

Note: You should have 12 different pairs of n and p. Also, "good" here is quite subjective so you should describe what you mean by good.

This ends the lab. Remember to turn in your lab reports on Sakai.