AML610 Homework #2

For all questions, please submit your R code, and a doc file with a copy of the screen output and plots (where applicable) to [smtowers@asu.edu](mailto:smtowers@asu.edu). Due Thus Feb 14th at 4pm.

The code in your R files should exhibit all of the good coding practices mentioned in <http://sherrytowers.com/2012/12/14/good-programming-practices-in-any-language/>

Please do not just copy and paste code from the examples given in class into your code.

All plots should exhibit all of the good practices in creating figures, mentioned in

<http://sherrytowers.com/2013/01/04/good-practices-in-producing-plots/>

Please submit your files in a format hwk2\_<first name>\_<initial of last name>.R and hwk2\_<first name>\_<initial of last name>.doc

**Question 1**

a) Generate 1000 random numbers from the Poisson distribution with mean lambda=10 and histogram the results.

b) Calculate the mean and the variance of the sample, and the variance of the mean and variance.

Note that the 2nd moment of the Poisson distribution is

M\_2=lambda^2+lambda

and the 4th moment of the Poisson distribution is

M\_4=lambda+7\*lambda^2+6\*lambda^3+lambda^4

Compare the 2nd and 4th Poisson moments with the 2nd and 4th sample moments.

c)Test the null hypothesis that the mean and variance are drawn from mu=10 and sigma^2=10, respectively.

d)Test the null hypothesis that the mean and variance of the sample are drawn from mu=10.05 and sigma^2=9.95, respectively.

e) Repeat c) and d) with a sample of 100,000 random numbers from the Poisson distribution with mean lambda=10. Are the p-values quite different than those you obtained in c) and/or d)? If so, why do you think that is the case?

**Question 2**

a) Generate 1000 random numbers from the Poisson distribution with mean lambda=1. Histogram the results, using the breaks option in hist() with breaks = seq(-1,100,1). Save the histogram to an object called histpois.

Now generate 1000 random numbers from the Normal distribution with mean=lambda and variance=lambda (make sure before you do this that you read the help file for rnorm()... the sigma argument is the standard deviation, not the variance!).

Histogram the results.

Now overlay the histogram from the Poisson generated sample, stored in histpois.

*Note that the x-values of the Poisson sample are histpois$mids+0.5*

*Also note that you will need to scale the y axis to fit both the distributions from the Poisson and Normal samples on the plot.*

The plot should look something like this (with variations because your random seed will have been different from mine):



b) repeat a) with lambda=2, lambda=5, lambda=10, and lambda=100, and put all four plots on the same page. Comment on the results.

**Question 3)**

Prove that the Ehrlang distribution with k=2 is not memoryless. Is there a value of k for which the Ehrlang distribution is memoryless?

While the Exponential distribution is clearly not very realistic for describing the time spent within a compartment of a biological or epidemic model, can you comment on why the memoryless property of the Exponential distribution can make it convenient to use in stochastic modeling like Agent Based models?

**Question 4)**

By now you should have identified data that you wish to use for your project. Write an R script to read in the data, and produce relevant plot(s); for instance, if the data is a time series of epidemic data, plot the time series.

By now you should also have identified a simple compartmental model that you plan to fit to the data. Using the deSolve package, write an R script to solve the ODE’s associated with the model (note; if you need to guess at what some of the model parameters are, that is fine at this point). Using the model results, produce relevant plot(s) that would overlay onto your data distribution(s). Don’t worry if the model distributions don’t look exactly like the data distributions at this point. In fact, if the model and data distributions look quite dis-similar because you have had to guess at some of the model parameters, show the model and data distributions in separate plots on the same page.