*1) Find and/or plot R(x), f(x) and the mean for the following positive random variables:*

* + *Since the hazard function we can invert this relationship to obtain R(x)*
* *Additionally we can find the CDF and PDF in the following way:*
* *The functions are plotted below, verifying their validity as probability functions on*

*Figure 1 - Probability function plot (Problem 1a)*

* *Finally, to find the expected value of this distribution we compute the integral*
* *Therefore we conclude that the mean of this distribution does not exist.*
* *As was done in part (a), the reliability function can be found from:*
* *However,, therefore to ensure R(x) has the correct measure we divide by e. Since we are changing the function by a constant factor there is no impact to the PDF, thus the final expression for the reliability function becomes:*
* *The CDF is thus, the plot below verifies that on, F(x) is a CDF.*

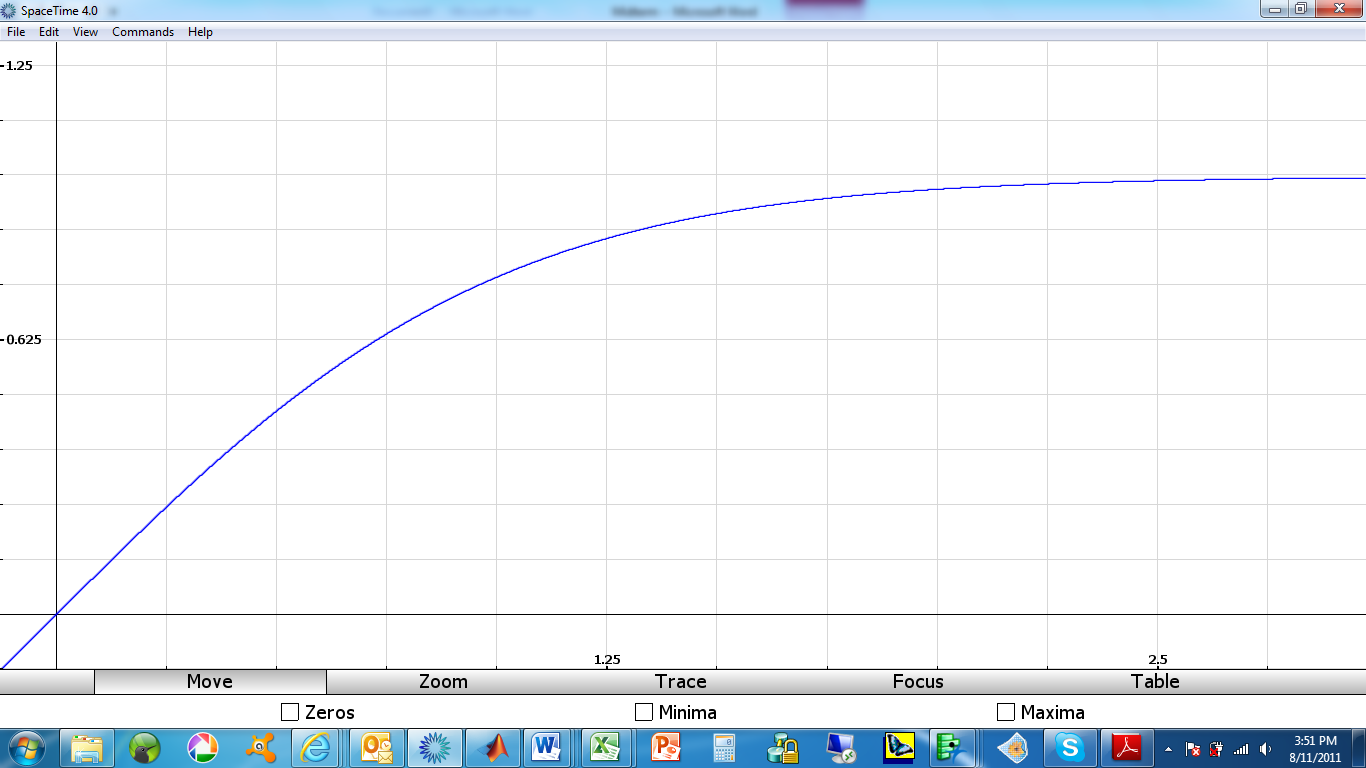
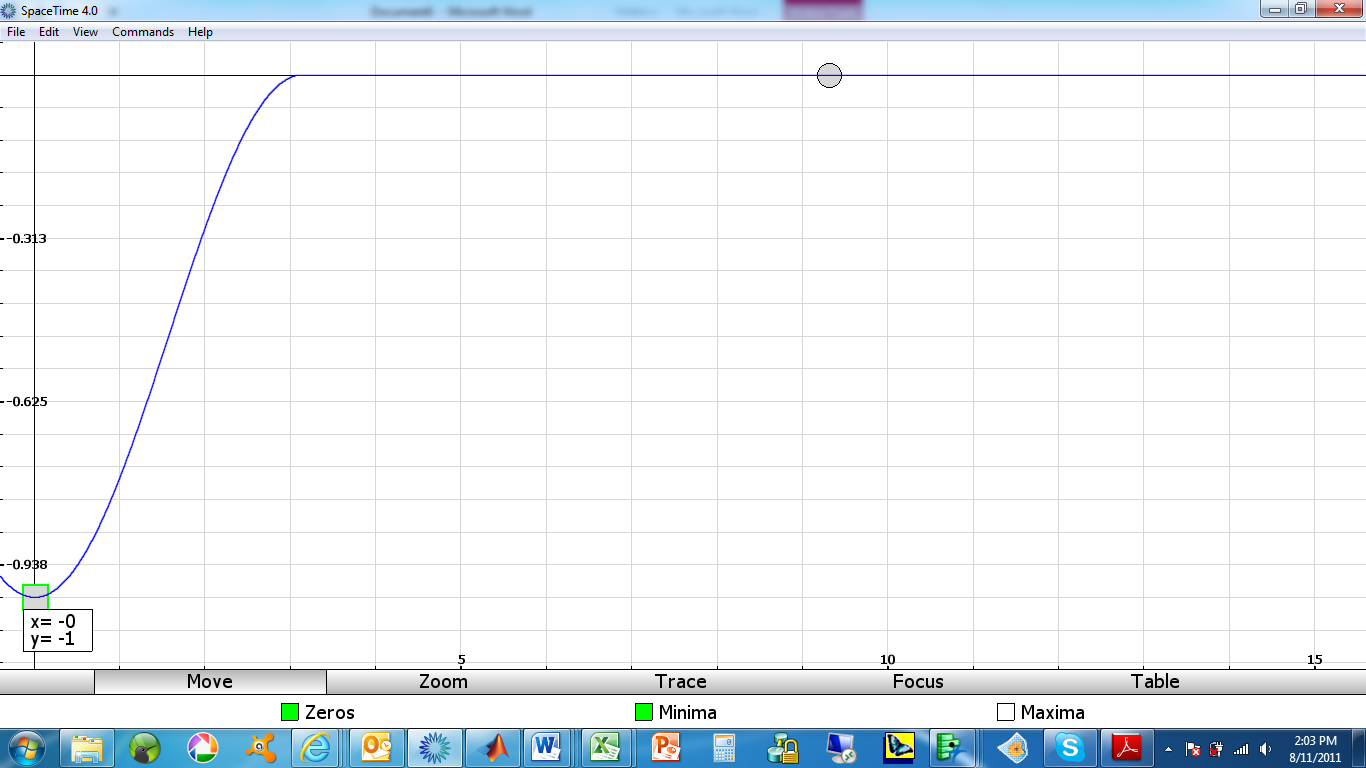
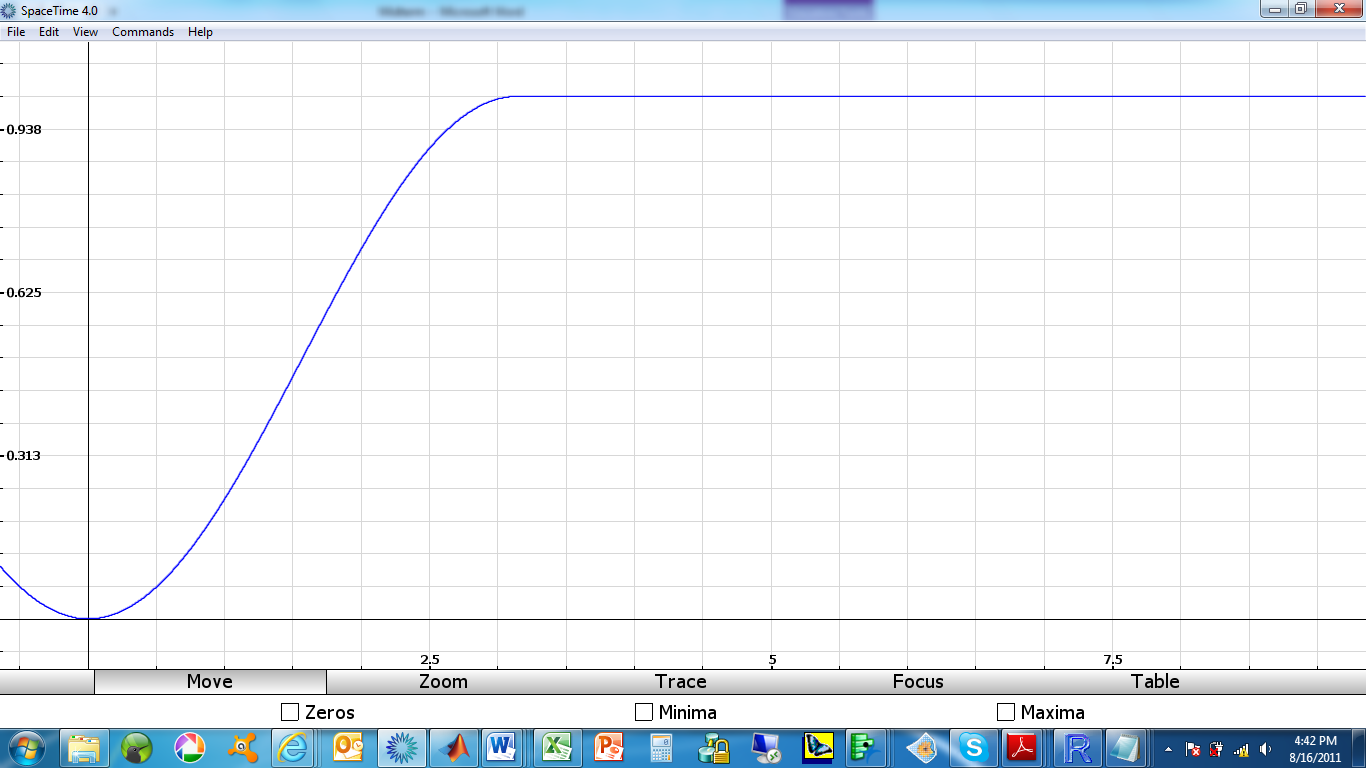
**

Figure 2 - Final CDF plot (Problem 1b)

* *Taking the derivative of F(x) gives the PDF, f(x):*
* *And the expected value is found from:*
* *Through numerical integration using R:*
* *Since this is an MGF we can easily find the mean of the distribution by:*
* *To find the PDF associated with this generating function, we need to convert the MGF to a Laplace transform. Since our RV was defined to be positive we are only interested in the univariate Laplace transform, thus by replacing in the MGF with, gives the univariate Laplace transform for this distribution:*
* *When the Laplace transform is inverted in MATLAB the PDF is found to be:*
* *Where H is the Heaviside step function. Integrating this function gives the following expression:*
* *From the intermediate CDF graph shown on the next page, this function is non-decreasing, but is shifted down by a value of one everywhere along the support, thus by adding one to the above function we ensure that all of the properties of a CDF are met in the interval without affecting the PDF. Therefore the final expression for the CDF is:*
* *And the reliability function is:*

**

*Figure 3 - Intermediate CDF plot (Problem 1c)*

**

*Figure 4 - Final CDF plot (Problem 1c)*

*Problem 2*

***OVERVIEW***

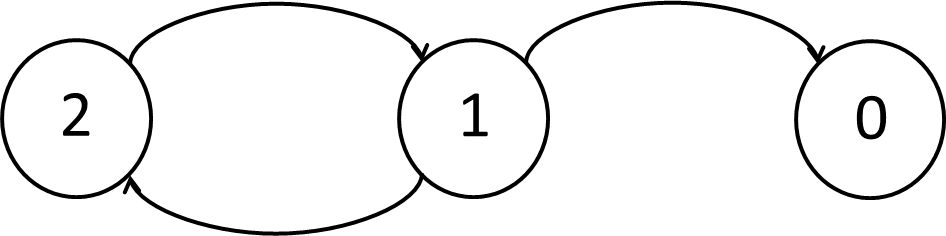
*In creating a stochastic model for this problem the first issue was to identify the state space, S and create the transition diagram. For this model S is comprised of three states that a patient may occupy:*

*State 2: Remission of active symptoms*

*State 1: Recurrence of active symptoms*

*State 0: Death due to the disease under study*

*The transition diagram is shown below:*

**

*Figure 5 - Initial semi-Markov transition diagram*

*To determine the distribution of transition times between the states, the data was entered into both JMP and an AFIT-developed Excel-based distribution calculator called “Bryon’s Tool.” Bryon’s Tool was developed by Capt. Bryon McClain in 2008 as part of his Master’s Thesis and is based on the parameterization and calculations found in the Introduction to Reliability Engineering text by Dr. Charles Ebeling. The results from both tools were compared and found in most cases to be almost exactly the same. However two caveats must be made in using Bryon’s Tool: 1) because the calculations are based on Dr. Ebeling’s book, the parameterization is different than what is typically found in other texts and must be adjusted before entering the distribution into R and 2) Brian’s Tool can only accommodate four distributions (Normal, Log-normal, exponential and Weibull). How these issues were dealt with is addressed below during the analysis of the data and development of these distributions.*

***DATA ANALYSIS & DISTRIBUTION FITTING***

* *Recurrence Time*

*Using the tools described above, the times to recurrence of active symptoms were analyzed. Both tools returned that a Log-normal distribution fit the data best and the JMP probability plot below shows that there is clearly there is a good fit to the data. Additional distributions were analyzed, but based on the Akaike Information Criterion (AIC), which measures the relative goodness of fit of various distributions; the lognormal was confirmed as the best fit.*

**

*Figure 6 - Lognormal probability plot of recurrence time data*

*Coding the parameters of the lognormal distribution into R requires the mean and standard deviation of the associated Normal distribution after taking the logarithm of the lognormal data. By taking the logarithm of the mean produced by Bryon’s Tool, both analyses resulted in the same distribution with parameters: LOGNOR( (however R requires the standard deviation .*

* *Remission & Death Times*

*Before analyzing the remission and death time data, we must first address how best to handle the eight observations that did not move from state 1 to either state 2 or state 0 during the study. Because each observation represents a significant amount of time, their inclusion will shift either transition time distribution significantly to the right, therefore the decision was made to include all eight observations as censored data points for both the remission and death transition time distributions. With this assumption in place we analyze the remaining data.*

*Analyzing the remission and death time data in the same was as was described above for the recurrence data, results in the following distributions:*

*Remission time transition distribution: Log-Logistic*

*Death time transition distribution: Log-Normal*

*Clearly, because Bryon’s tool does not calculate for the log-logistic distribution, there was no agreement between the models. The selection of the log-logistic model was based on a significant difference in AIC values for the log-normal and log-logistic distributions. The JMP plots below illustrate the fit of the data to the respective distributions.*

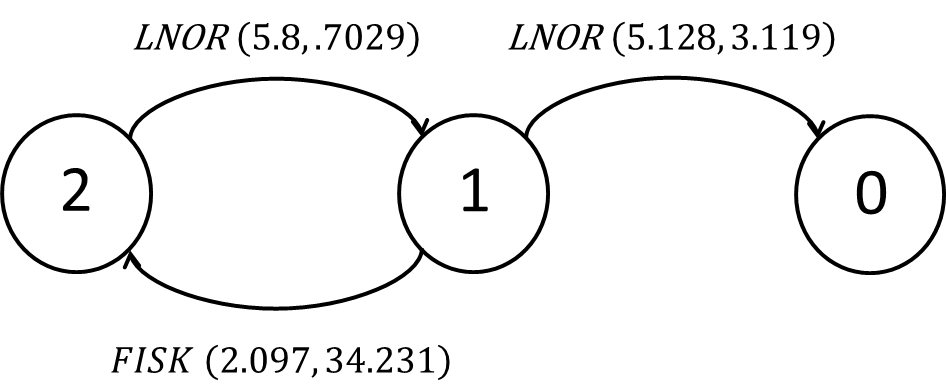
**

*Figure 7 – Log-logistic distribution fit to remission time data*

**

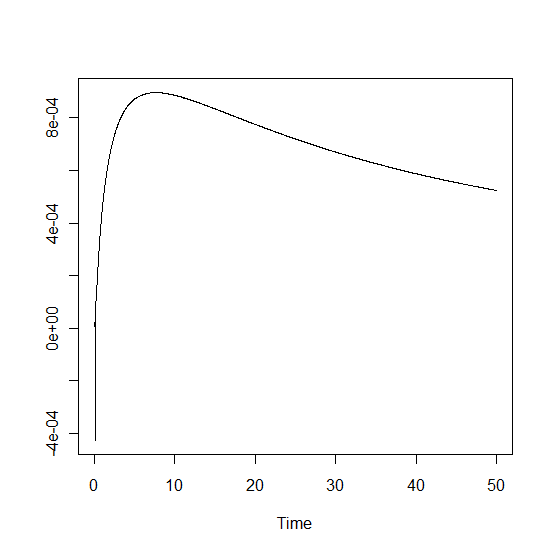
*Figure 8 – Log-normal distribution fit to death time data*

*Before continuing, it should be noted that the log-logistic distribution is also called that Fisk distribution and because the parameterization of the Fisk distribution in R is much simpler than that of the log-logistic distribution, the remission time transition distribution will be labeled as FISK(34.231, 2.097) for the remainder of this report. With the transition distributions established, the final transition diagram, infinitesimal generator matrix and probability matrix are shown below. With these values calculated, we are now ready to code this model into R and produce the quantities of interest.*

**

*Figure 9 - Final semi-Markov Transition diagram*

*In this analysis we are asked to model the first passage distribution from the time of a person who has just been diagnosed with active symptoms (state 2) to the time that this person dies from the disease (state 1). Since state 1 is clearly an absorbing state we’re interested in modeling the first passage hazard rate of this disease,. The graphs below shows a plot of the first passage hazard rate, as computed using R. The reader will notice that the plot extends below zero at, this error has been verified using Excel and found to be a numerical anomaly over the interval, and thus the actual plot does stop at zero.*

**

*Figure 10 - First passage hazard plot*

*The hazard plot was created over a 2000 day interval with using a 40 day time steps, so each time unit along the x axis must be multiplied accordingly. Using this plot along with the R code attached, the average time until death was found to be 3698.364, or 10.135 years. Additionally, the mode time (most typical/most likely) can be found from the highest point on the plot and is approximately 1 year. Finally, the remaining quantities of interest are found below:*

* *Probability of being in remission 4 yrs. after diagnosis of active symptoms?* ***(0.55332, 55.33%)***
* *Expected # visits into remission within 4 yrs. of diagnosis of active symptoms?* ***(2.728 visits)***
* *Long run probability a patient with active symptoms remains in remission?* 
  + *So there is a chance this disease could make you live forever ☺*