**Problem Set 3 – Statistics and Polynomial Interpolation**

**Due MIDNIGHT Feb 6, 2019**

Collect all m-files in a single .zip file and upload the .zip file to the course webpage by midnight on Wednesday, February 6, 2019. Please note any collaborations in the header comments of your m-files. Each student must upload their own individual copy of the work.

1. Problem 17.13 from text. Plot the points given and the fit curves on the same plot. Print errors to the screen. Note: to compute the true value, it is besselj(1, 2.1)



1. Superheated steam internal energy (U – given in BTU/lb) varies depending on its temperature and pressure. Use the attached table from the Smith, Van Ness, Abbott thermodynamics text to create a matrix for interpolation (8 pressures vs. 7 temperatures – NOTE: you can do this by hand or use OCR tools). Show a surface plot to illustrate how the internal energy changes with temperature and pressure. Use bilinear interpolation to find the internal energies at the following conditions (table to the right 🡪)
2. An electrocardiogram (EKG or ECG) is a test that checks for problems with the electrical activity of your heart. The spikes and dips in the electrical signal are called waves. The ‘data.csv’ file contains some EKG readings taken every second. The clinician is worried about readouts greater than 880. Create a scatter plot that marks all the worrisome points in red. To help the clinician, plot the data in three equal segments with them vertically stacked (subplot(3,1)).
3. The ‘HorseKickData.xlsx’ file records the number of deaths by horse kick in the Prussian army for 16 different cavalry units from 1875 to 1894. Create an M-file that does the following:
	1. Create one histogram for all the measured counts (all cavalries over all the years)
	2. Fit the data with a Poisson distribution
	3. Overlay the PDF on the normalized histogram as we did in class
	4. Answer – why is the PDF only defined at integer values?
	5. What is the probability of observing >4 deaths in a single unit in a given year (hint: use CDF)? What is the average number of samples you would have to look at to make this observation?
	6. Use a random number generator from your distribution fit (part b) that simulates and validates part e.
4. One reason we care about Taylor series (or Maclaurin series when they are commonly centered at zero) is that they are linear approximations of non-linear functions. This allows us to use them for linear algebra and on matrices. But how many terms are needed for a good approximation? Does this differ depending on the function? Plot the % error (compared to the TRUE value) vs. number of terms for the following common series on a single plot. Evaluate at x = 0.5. Use a log-log plot that goes from n = 0 to 100. Place a marker on the plot to show which function needs the most terms to minimize error at x = 0.5.

 

**Group Credit Points –** Remember to collaborate on SLACK for +2 points.