Matlab Assignment 2

by Brent Hickman[1]

1. Write a Matlab program to compute an empirical distribution function (EDF) for a truncated standard normal random variable. An EDF is merely a step function that approximates the true cumulative distribution function of a random variable. The formula for computing an EDF is given by:

$$\hat{F}_V(x) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}(V_t \leq x)$$

where

$$\mathbf{1}(V_t \leq x) = \begin{cases} 1 & \text{if} \quad v_t \leq x; \\ 0 & \text{otherwise,} \end{cases}$$

where T is the number of observations in the sample, V_t is the t^{th} observation, and $\mathbf{1}(\cdot)$ is an indicator function.

   (a) Create a 30×30 matrix whose elements are draws from the standard normal distribution (HINT: see the documentation on the `randn` function in Matlab).

   (b) Convert this matrix into a vector using either the ":" operator or the `reshape` function.

   (c) Separate the data vector into two separate samples of observations above and below zero. For the positive observations, use the `find` function to select entries greater than zero and store them in another vector named "`pos`". For the negative observations, use the `find` function and the "empty-brackets" operator to discard entries greater than zero from the original vector. Then, rename this vector "`neg`".

   (d) Create a grid of points over which you wish to compute the values of the EDF for `neg` using the `linspace` function. Make it with 500 grid points and call it "`negdom`". Do the same for `pos`, but this time using the ":" operator to construct your domain vector. Call it "`posdom`".

   (e) Using at most only ONE "`for`" loop, compute the value of the EDF for each member of your domain grids for both `pos` and `neg`. (HINT: use the `length` and `find` functions, and/or the built-in indicator function in Matlab).

   (f) Plot the results side-by-side, using the `subplot` command in Matlab. Label the axes and provide an explanatory title for the figure.

2. Using Newton's method, write a Matlab program to find the fixed point of the following function:

$$f(x) = \sqrt{x + \pi e}$$

[1] `brent-hickman@uiowa.edu`

accurate to at least 8 decimal places. (HINT: finding the fixed point of $f(x)$ is the same as finding the zero of $g(x) = f(x) - x$.) The output of this program should display in a single table *(i)* the solution for the fixed point, *(ii)* the initial guess, *(iii)* the number of iterations it took to reach an answer, and *(iv)* the stopping rule (*i.e.* the tolerance). Remember that your initial guess should be contained in the interval $[-\pi e, \infty)$, as f is defined only on this interval. Finally, you should also store each successive guess of the final solution in a vector and plot them. give your figure a title and label the horizontal axis "iterations" and the vertical axis "guesses".

3. Write a MATLAB function to numerically compute the following integral by the method of trapezoidal quadrature for inputs of x as supplied by the user:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{n^2}{2}\right) dn.$$

Call the function `trapnormcdf` and make design it so that it can accept vector arguments. Your solution should be accurate to at least 5 decimal places. Next, write a MATLAB script that creates a grid of 1000 domain points on the interval [-1,1] and then inputs the grid into both `trapnormcdf` and MATLAB's built-in `normcdf` with $\mu = 0$ and $\sigma = 1$. Plot the two results on the same axes and include a title, axis labels and a legend in the upper left-hand corner on the inside. Use different line styles and colors for each of the two curves, and increase the line thickness to `2pt`, rather than the default `0.5pt`. (HINT: search the documentation on the plot function.)

The method of trapezoidal quadrature consists of the following steps:

(a) Divide the interval of integration up into $k \geq 2$ subintervals of equal length.

(b) on each of these subintervals, approximate the area under the curve by constructing a trapezoid having a base (b) equal to the length of the interval, left height (lh) equal to the function value at the lower bound, and right height (rh) equal to the functional value at the upper bound.

(c) Add the areas of each of these trapezoids. Recall that the area of a trapezoid is given by

$$b\left(\frac{lh + rh}{2}\right).$$

(d) Increase k by 1 and redo the previous two steps. If the resulting sums are within the desired tolerance of each other then stop and adopt the last approximation as your solution. If not, repeat this and the previous two steps until two consecutive approximations fall within the desired tolerance.