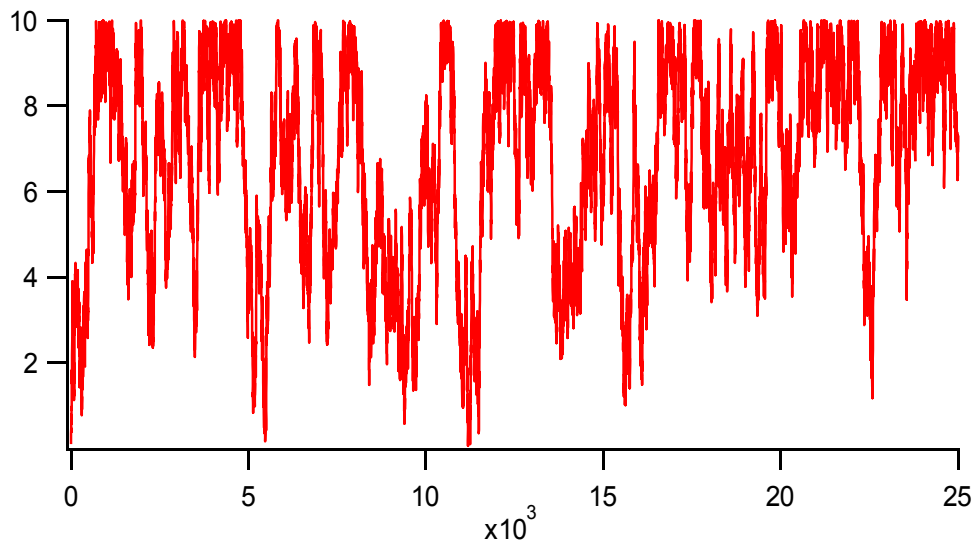


## CE 530

### Assignment #2 Solution

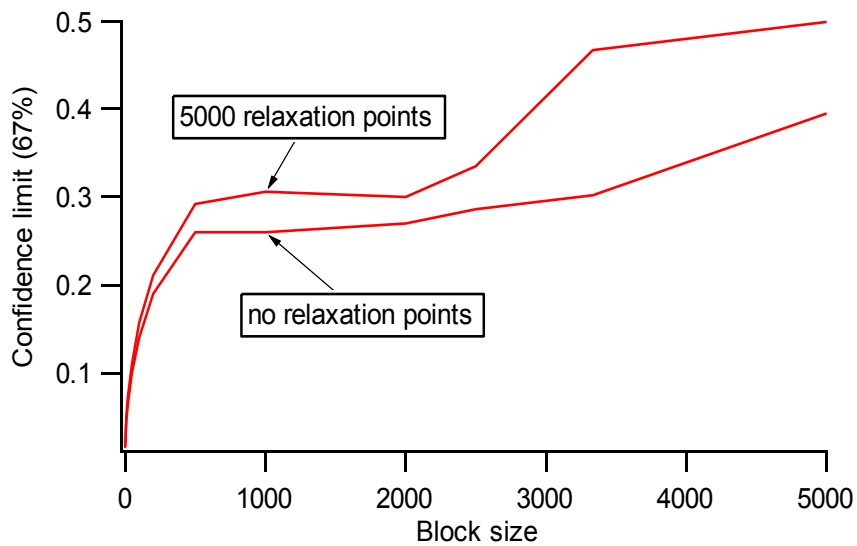
1. Download the data file given at [this link](#) (I suggest right clicking and choosing to save it to disk). The file contains 25,000 data points from a very simple simulation (exclude the first entry in the file, which is just the number of data points). Analyze the data to estimate the average, and determine a proper confidence limit for your result. Note that the data points are not independent. Explain your calculations. Do you need to discard any initial values (does the system need much time to "relax")?

Here is a plot of the data (y-axis) against its position in the sequence (x-axis). The initial state is near  $y = 0$ , and it seems that most of the time is spent near  $y = 10$ . However, there is not much reason to throw away any initial data, as initial datum is not wildly unusual (it is revisited several times in the sequence), and the time to reach the upper range is pretty fast. A simple average of these points yields 6.64 if the first 5000 points are discarded, and 6.72 if all points are used (the true mean of the distribution governing the sampling of the data is 6.667, but of course that fact cannot be known from the data)



The following is a plot of the confidence limits as a function of the block size. Curves are shown for two cases, first in which the initial 5000 points are discarded, and second in which no points are discarded. Both curves exhibit a plateau region, in which the confidence limit is insensitive to the block size. For small block sizes, the confidence limit is underestimated due to the correlation between the points; for larger block sizes the confidence limit becomes inaccurate because there are not enough blocks (only 4 or 5 for the 5000-size block). A reasonable value for the confidence interval is 0.30.

The program used to compute these curves is given on the next page.



Here's the Fortran program that was used for the data analysis

```
dimension data(50000)
  parameter (nSize=13)
  integer iblock(nSize)
  data iblock /1,2,10,20,50,100,200,500,1000,2000,2500,3333,5000/

C   Read in and save all data
  write(6,*) "opening"
  open(20, file="random.dat")
  read(20,*) n
  write(6,*) "read data"
  do i=1,n
    read(20,*) data(i)
  end do

  write(6,*) 'Enter nRelax'
  read(5,*) nRelax

C   Loop over all block sizes
  do k=1,nSize
    nb = iblock(k)

C   Zero all sums
    sum = 0.0
    icount = 0
    blockSum = 0.0
    blockSum2 = 0.0
    nBlock = 0

C   Loop through all points
    do i=nRelax+1,n
      sum = sum + data(i)
      icount = icount + 1
      if(mod(i-nRelax,nb) .eq. 0.0) then
C   Another block is complete
        blockAvg = sum/icount
        blockSum = blockSum + blockAvg
        blockSum2 = blockSum2 + blockAvg*blockAvg
        nBlock = nBlock + 1
        sum = 0.0
        icount = 0
      end if
    end do

C   We're not too careful about making sure all blocks are same size
    avg = blockSum/nBlock
    error = sqrt((blockSum2/nBlock - avg*avg)/(nBlock-1))
    write(6,*) nb,avg, error
  end do

end
```