**MAT 331 Fall 2017, Project 6**
**Last digits of consecutive primes**

This project is based on the 2016 preprint “Unexpected biases in the distribution of primes” by Robert J.K. Oliver and Kannan Soundararajan. A prime (larger than 2) must end in 1, 3, 7 or 9, and a versiobn of the prime number theorem says that there are asymptotically the same number of primes that end in each digit. It has long been thought that the occurence of digits should look fairly random, e.g., if one prime ends with a 1 next prime is equally likely to end with any of the four possibilities. However, this paper presents evidence that this is not correct; the next prime is more likely to end with certain digits than with others. This project is to numerically verify this.

(1) Define and explain the notation $\pi(x, q, a)$ and $\pi(x, q, \mathbf{a})$ from the first page of the paper.

(2) Replicate the table from page2 of the paper. That is, compute a table showing the probability that if a prime ends with digit $k$ then the next prime ends in digit $j$ Do ths by checking primesp to $N = 10^5, 10^6, 10^7$. Time your program and estimate how many primes you can test in a reasonable time. Do that experiment.

(3) Read about the chi-squared test to see if a given number of items in $N$ bins is uniformly distributed or not. There is a chi squared test built into `MATLAB`. Does it accept or reject the hypothesis that the distribution about is uniform? For example, the following code checks take `obsCounts` as the number of balls in each of four bins and estimates whether it is uniform or not (in this case it accepts the uniform hypotheis).

```
obsCounts=[12,8,12,8];   % number of objects per bin
L=length(obsCounts);     % number of bins
bins=[1:L];              % list of bins
expCounts=mean(obsCounts)*ones(1,L);  % expected number of objects per bin
[h,p,st] = chi2gof(bins,'Ctrs',bins,...
                   'Frequency',obsCounts, ...
                   'Expected',expCounts)
```

(4) Choose a small prime $p$ and look at the distribution of conseqtive pairs modulo $p$. Do you observe irregularities?

(5) What about the second to last digit? This can be anything from 0 to 9, so there are 100 possibilities to check. Does the distribution between consecutive primes look random, or are certain pairs more likely than others?