Research Projects for SURIEM 2017

1. **Recursive polynomials.** Consider a Fibonacci-type polynomial sequence given by the recurrence relation

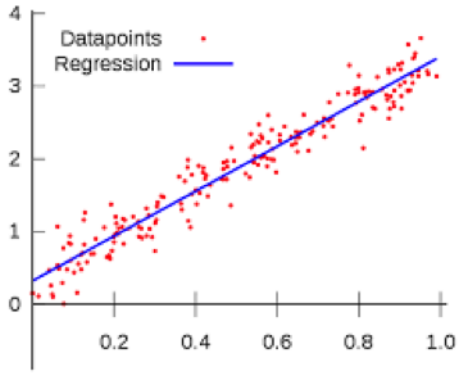
   $$G_0(x) = \alpha,\ G_1(x) = x + \beta,\ G_n(x) = \gamma(x)G_{n-1}(x) + G_{n-2}(x),\ n \geq 2.$$

   Here α and β are integers and γ is some function of x. If $\alpha = 1,\ \beta = 0,$ and $\gamma(x) = x,$ then $G_n(x)$ is the classical Fibonacci polynomial sequence $F_n(x)$. For $\alpha = 2,\ \beta = 0,$ and $\gamma(x) = x$ one gets the Lucas polynomial sequence $L_n(x)$.

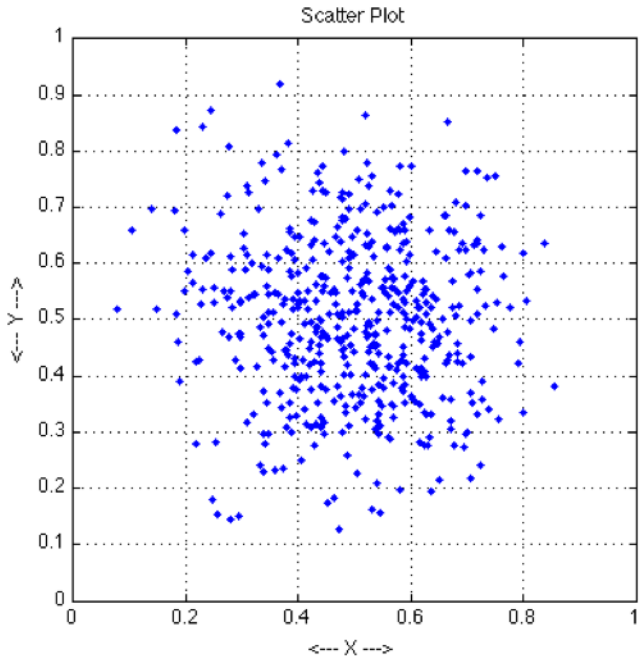
   The following are sample projects for REU participants.

   - Find the asymptotic behavior of the roots of the k-th derivative of the Fibonacci-type polynomial sequence $G_0(x) = \alpha, G_1(x) = \beta,$ and $G_n(x) = x^k G_{n-1}(x) + G_{n-2}(x),\ n \geq 1.$. How does the order of the derivative impact the behavior of the maximum roots?
   - What is the asymptotic behavior of the roots of other Fibonacci-type polynomials? For example a Fibonacci-type polynomial sequence generated using three initial conditions? For what other $\gamma's$ and initial conditions can one get results of the same nature?

2. **Data processing and mathematical modeling for biological applications.** Data driven mathematical modeling has been a major tool for applications in many fields. Before the advanced computer technologies and the internet, the data from experiments and other collecting methods are relatively clean in the sense that the data clearly show the correlation between the related variables. As an example, the following graph strongly suggests that the input and the output have a linear relationship:



However, it is not clear what type of correlations we have for the next graph:



Although statistics suggests that y increases when x increases for the above data set, there is no clear fitting function as for the previous graph. This causes problems especially for biological applications because data collected from biological experiments and digital images are usually very noisy. This project addresses the problem of how one should process biological data obtained from biological experiments so that the data meets mathematical integrity standards.

Students will work on statistical and mathematical methods for data/image processing and come up with solutions for processing the data to be used for mathematical models. These methods include statistical tests, regularization methods and machine learning tools.

Any data processing procedure involves in manipulation of data that could be transformed into loss of information or significant change of information. Our goal is to develop a data processing procedures

based scientific approaches so that biological information will be kept as much as possible.

3. **Biostatistics and Neuroscience.**

   Students will work on extracting the best signal from individual neurons obtained from tiff stacks representing calcium imaging data generated by the BRAIN Initiative. They will use a Bayesian approach incorporating general knowledge of geometry of cells and temporal characteristics of calcium signals.