# Homework No.3 (MSDS 954:567) Spring 2022

Due date: 4/21/2022

**Problem 1.** (a) Generate 200 replicas of uniform $[-\pi, \pi]$ and 200 normal with mean 0 and standard deviation 1/8. Set data $x$ from this uniform, error $\epsilon$ from this normal distribution. The response $y$ is by model:

$$y = \sin(x) + \epsilon$$

Fit the data with two types of smoothing techniques. Plot both the data and your fitted smooth curves.
(b) The same as (a) except changing the standard deviation from 1/8 to 1/2.

[*Remark: Use a computer for you calculation; explain your analysis and results carefully*]

**Problem 2.** (a) Use a linear regression model to analyze the GAG in urine data in data frame GAGurine. Produce a chart to help a pediatrician to assess if a child's GAG concentration is 'normal' or not (hint: plot in one graph the estimated line and confidence bands at different levels)
(b) Consider using a smooth regression to analyze the GAG in urine data

[*Remark: See the data set named "GAGurine.csv" in the assignment. Use a computer for you calculation; explain your analysis and results carefully*]

**Problem 3.** Service times of a queuing system follow Exponential distribution with an unknown parameter $\theta$. A sample of service times $X_1, X_2, \cdots, X_n$ is observed.
(a) Show that the $Gamma(\alpha, \lambda)$ family of prior distributions is conjugate.
(b) Find the posterior parameters, posterior mean and variance. (As functions of $\theta, \alpha, \lambda$ and $X_i$'s)
(c) Suppose that we can allow $\alpha = 0$ and consider a prior density $\pi(\lambda) = 1/\lambda$ for $\lambda > 0$. Of course, this is not a proper density. Nevertheless, find the posterior distribution, its mean and variance.

[*Remark: Solve this problem by paper and pencil. Write down the detailed calculation process.*]

**Problem 4.** Write a computing code to calculate the integration of $\int_{-5}^{5}(x^3 - x^2)e^{-x^2/2}$ using Monte Carlo simulation with $N$ samples from a uniform distribution, for $N = 10, 100, 1000$. For each choice of N, repeat the experiment for 500 times, compute the variance and visualize the relationship between the variance and N.

[*Remark: Use a computer for you calculation; explain your results carefully*]