**ECE 8527: Introduction to  
Machine Learning and Pattern Recognition**

# HW No. 5: ML vs. Bayesian Estimation

The tasks to be accomplished in this homework assignment are:

1. Generate 10 independent sets of random data consisting of 106 points from a 1D GRV with a mean of (1,1) and a variance of 1.
2. For the first set, estimate the mean value using a maximum likelihood estimate. Plot the error in this estimate as a function of the number of points used in the estimate (e.g., N=1, 10, 100, … 106). Use the square of the Euclidean distance between the true mean and your estimate of the mean as a proxy for the error.
3. Repeat no. 2, but for each value of N, average across the 10 sets. For example, pick the first 100 points of each set. Compute the mean for each set using these 100 points, and then average the means across the 10 sets. Plot the error as a function of the number of points.
4. Next, assume an initial mean guess of 2. Construct the Bayesian estimate of the mean and plot the error as a function of N.

Explain what you observe in each step of this assignment and comment on the significance. What does this suggest about algorithms to build high performance machine learning systems?

In this homework assignment, we are instructed to create a one-dimension array and then apply maximum likelihood estimator and Bayesian estimate for the same data set with different amount of input data with the aim of comparing the performance of each estimator.

The data set we generated is based on random Gaussian distribution with specific mean and variance. For each run of the function, 10 different sets of data are generated and stored before any of the data set is fed into the estimator for the estimation of the mean value.

The maximum likelihood estimator directly takes the data set and calculate the mean value, while the Bayesian estimator takes a guessed mean value and a guessed standard deviation with corresponding weights, which is generated based on the standard deviation from the input data set and the guessed standard deviation, for the estimation.

The following diagram shows the error rates that we obtained by compiling single set of data for maximum likelihood estimator, average of 10 sets of data for maximum likelihood estimator, and single set of data for Bayesian estimator for three different runs.

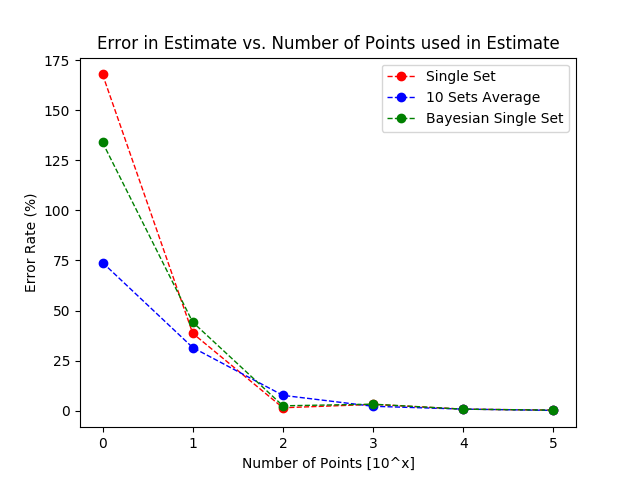


Figure Run No.1

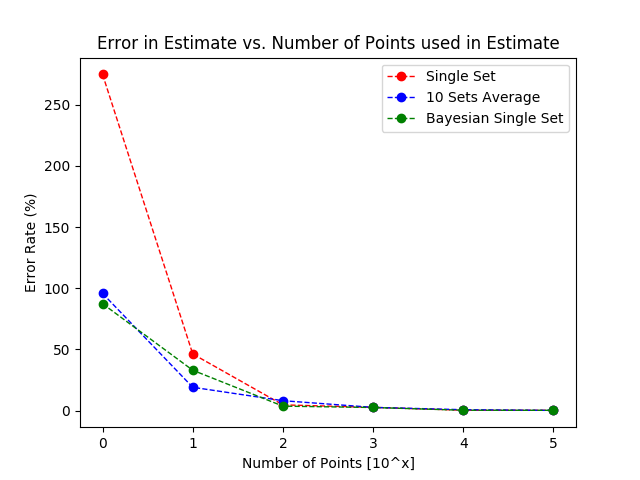


Figure Run No.2

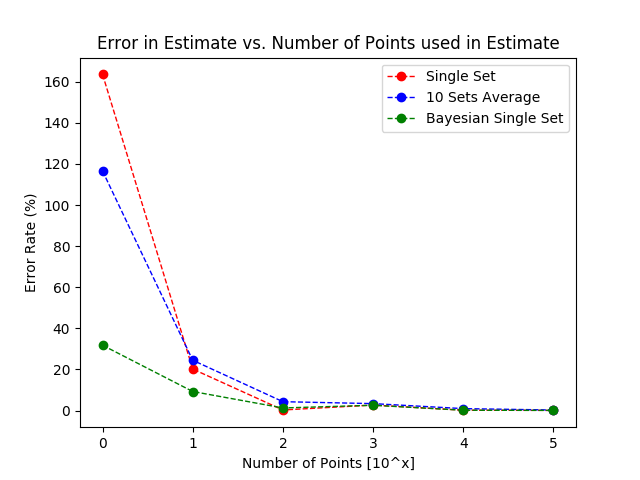


Figure Run No.3

Based on the finding from these three diagrams, the initial error rate is relatively the highest with the one data point for maximum likelihood estimator, but this error rate is case specific, and it could change drastically when different values are inputted into the estimator. However, the trend of all of the estimators will stays relatively the same that when 10000 or more data points are involved, the error rate approaches and stays close to 0. The converge towards an error rate of 0 for all estimator happens with 1000 data points input into the estimator. As for how fast the converge happens for each estimator, with different data set inputted into the estimator, it varies between data sets. For some instances, the Bayesian estimator converges the fastest, while in other cases, the average of 10 sets of maximum likelihood is the fastest, and with some cases the single set of maximum likelihood converges the fastest.

For a data set or data sets with maximum number of points of 10000, the result of both estimators shares the same characteristic result that the error rate approaches 0 at the end of the calculation. With Bayesian estimator requires a guessed input in order to calculating the estimated result, it could be influencing the produced result for this estimator. And the result of the estimation also shows how different the estimation is from the actual value.