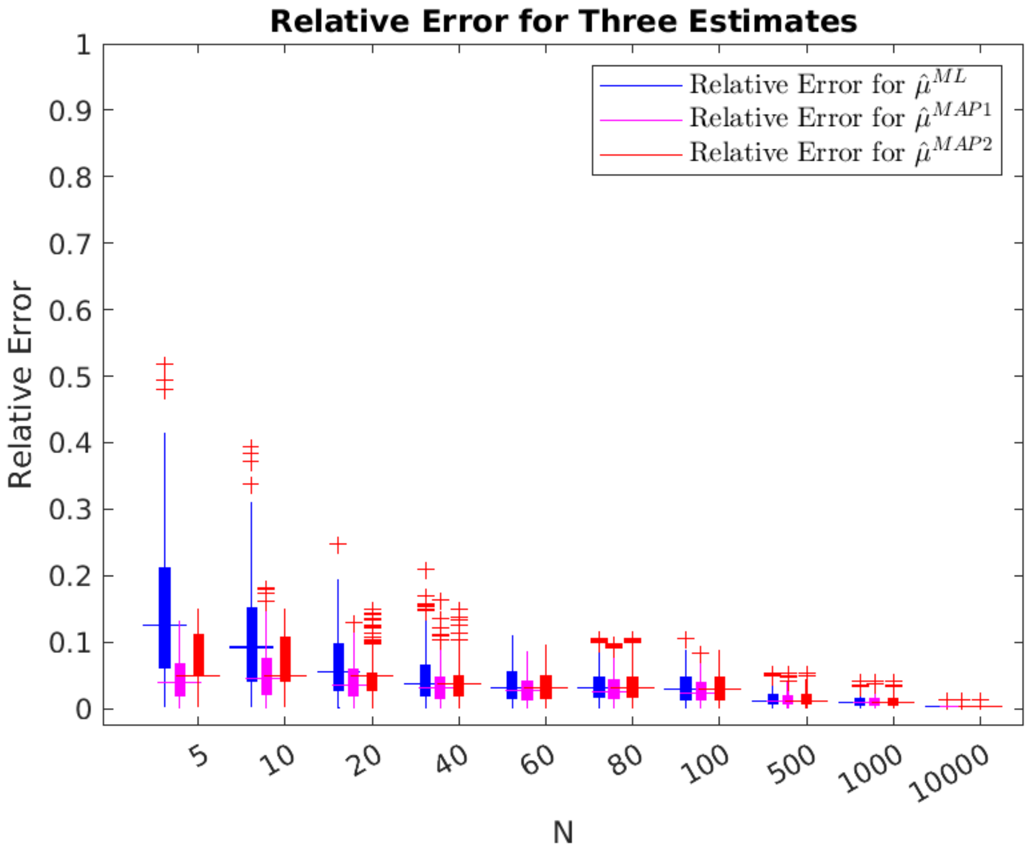
Problem 1

Part 1



Part 2

- **Interpretation** : The range for relative error in $\hat{\mu}^{ML}$ is higher than the range of relative error in $\hat{\mu}^{MAP1}$ and $\hat{\mu}^{MAP2}$. Also the median values are lower for MAP estimates than ML estimate for small values of N. Further more as the value of N is increasing, the box and whisker plot for
  the three estimates are getting identical. Gaussian prior being a better prior than a uniform prior is giving good estimates than the latter.

  (i) As the value of N increases the relative error for the three estimates is reducing and converging to 0 i.e, all the three estimates are approaching μ_{true} which is evident as follows:

$$\hat{\mu}^{ML} = \frac{\sum x_i}{N}$$

By law of large numbers $\hat{\mu}^{ML} \to \mu_{true}$ as N ~~tenty~~ tends to ∞.

$$\hat{\mu}^{MAP1} = \frac{\bar{x}\sigma_0^2 + \mu_0\sigma^2/N}{\sigma_0^2 + \sigma^2/N}$$

As $N \to \infty$, $\sigma^2/N \to 0$ and $\mu_0\sigma^2/N \to 0$

$$\therefore \hat{\mu}^{MAP1} \to \frac{\bar{x}\cancel{\sigma_0^2}}{\cancel{\sigma_0^2}} = \bar{x} \to \mu_{true}$$

$$\hat{\mu}^{MAP2} = \begin{cases} b & \hat{\mu}^{ML} \geq b. \\ \hat{\mu}^{ML} & a \leq \hat{\mu}^{ML} \leq b. \\ a & \hat{\mu}^{ML} \leq a. \end{cases}$$

Here because, $a \leq \mu_{true} \leq b$, $\hat{\mu}^{MAP2}$ converged to $\hat{\mu}^{ML}$ as $N \to \infty$ which converged to μ_{true}.

If our uniform prior on μ had been wrong, it would have converged to either a or b as $N \to \infty$. and would have given constant non zero relative error. So, MAP estimate are good estimates only for accurate prior.

(ii) I would prefer the $\hat{\mu}^{MAP1}$ because even for lower values of N it is more accurate, as it consides a good (More closer to μ_{true}) prior distribution of μ for giving the estimate. ML estimation starts only with the probability of observation given the parameter and tries to find the parameter best accords with the observation. It takes into no consideration the prior knowledge.

While if our prior on parameter is reasonable as it was in this case it is better to prefer MAP as it does take into consideration the prior knowledege and given accurate estimation even with less data on hand.