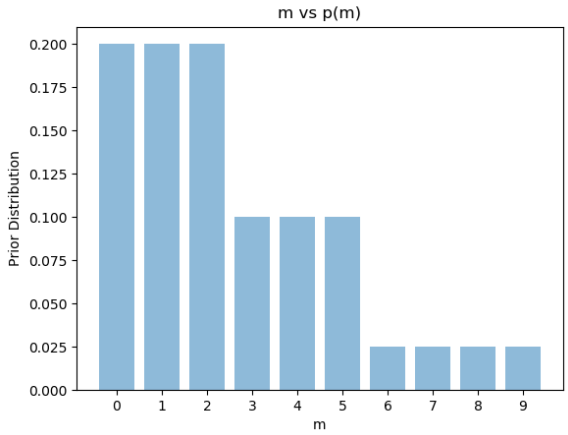
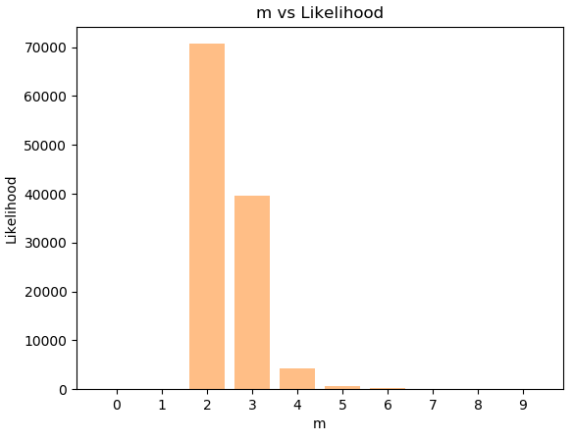
1 Prior Distribution

:



2 Likelihood Function

:



If log-scale it, then the bar chart is not as illustrative as the above because when $m = 0$ the likelihood is 0 which blows up the rest in the picture. As here, the highest is when $m = 2$, with a value around 10^{-2}, thus I chose to show this bar chart, even though the y-axis labels may be confusing.

3 Formula for Posterior Distribution

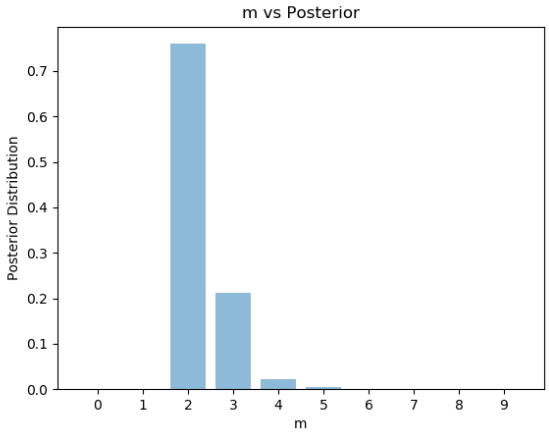
The posterior distibution is obtained by applying the Bayes rule to the likelihood function. We have that $p(m)$ is the prior distribution, $p(Data|m)$ is the likelihood and $p(Data)$ is the marginal distribution aka the normalization term. Applying Bayes rule, we get

$$p(m|Data) = \frac{p(Data|m) \cdot p(m)}{p(Data)} = \frac{p(Data|m) \cdot p(m)}{\sum_{m=0}^{9} p(Data|m) \cdot p(m)},$$

where $p(m|Data)$ is the posterior distribution and the denominator of the second equality results from conditioning.

4 Posterior Distribution

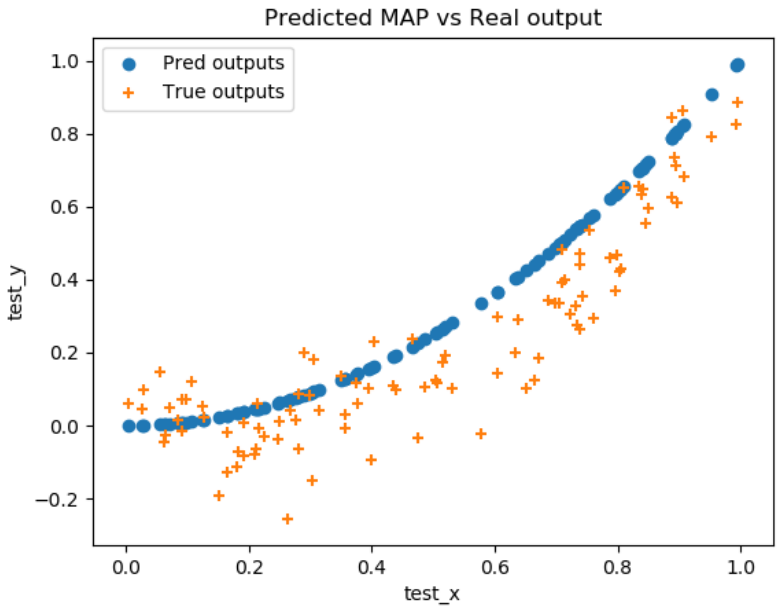
:



5 MAP estimate

The best m is when $m = 2$. The value of the posterior probability for when $m = 2$ is $p(2|Data) = 0.7598037833515152$.

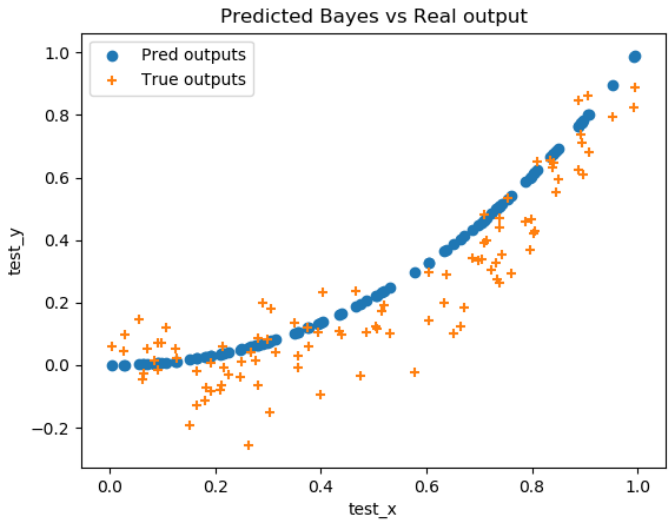
6 MAP predictions



7 MAP error

The mean-squared error of the MAP estimate is 0.02146290417007275.

8 Bayes estimate



9 Bayes error

The mean-squared error of the Bayes estimate is 0.01630936546872594.

10 Bayes vs MAP

Yes, the Bayes error is lower than the MAP error. The bar chart of the posterior distribution shows us that, while for $m = 2$ we get the best predicted values, for $m = 3$ we notice that the posterior is around 0.2 which is significant. Hence, while the MAP error considers only the predicted values for $m = 2$, the Bayes error considers all the values of m, since the Bayes estimate is the average over m, which results in a lower error. In general, if the posterior $p(\hat{m}_{MAP}|Data)$ is close to 1, then MAP error $\sim$ Bayes error, but if the posterior is spreaded across m, MAP error $>$ Bayes error (Bayes approach tends to outperforms MAP).