

Project 5 - extra credit

Task5

- I used the Multinomial HMM over 1000 data points.
- I generate the values of P,B,O the same way as in Task1.
- I calculate the likelihood, AIC, BIC over number of states ranging from 2 -
- The AIC and BIC are calculated as
 - $AIC = -2\log L + 2p$
 - $BIC = -2\log L + p\log T$
 - $p = M*M + NM - 1$,
 - L is the likelihood, T is the number of observations, and p is
 - the total number of independent parameters, where N is the number of states, and M is the number of objects =3.
 - logL is calculated directly from the model function “score” which returns the log probability of the sequence, under the parameters of the model under consideration.
- AIC is an estimate of the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model. So that a lower AIC means a model is considered to be closer to the truth.
- BIC is an estimate of a function of the posterior probability of a model being true. So that a lower BIC means that a model is considered to be more likely to be the true model.
- The log-likelihood, AIC and BIC are plotted and we that there is a variation in the AIC and BIC values. We see that the AIC and BIC values almost overlap. We see that the AIC, BIC value slowly increase over time. But there is variation. The log likelihood also has more variation and increases slightly as the number of states increases. The model takes a very long time to train. (contd below)

- We see that we get the best model with the least AIC, BIC at num_states=63

