Carl Cortez

CIS 735 – Assignment 4

This assignment will be due by midnight on the day of live session 11

1. For the Bayesian Information Criterion (BIC) and the Akaiki Information Criterion (AIC), you measure the fitness of a model based on the accuracy and complexity of the model. Where the complexity is viewed as a negative and accuracy as a positive. For each of the information criterions, lower score value is considered a better fit for the data given. For a particular data set using the same model but with different numbers of parameters i.e. complexity, K, the BIC and AIC return values are shown in the table below. Looking for optimal K value for both. Minimum values of each.

K	BIC(K) down then up values	AIC(K) down then up values
1	931.8	540.1
2	925.5	539.7
3	920.4	534.3
4	922.2	533.8
5	929.7	530.2
6	931.8	529.3
7	940.1	532.1

A. Using BIC what K value gives the best fitness for the data? AIC?

For BIC, K=3 would give the best fitness.
For AIC, K=6 would give the best fitness.

B. If the K values differ for the best fitting model, why would that happen? What could that mean?

The K value may be assessing a certain region of the data instead of the entire data set.

This may mean some parts of the data set are not being assessed correctly. This may also mean there's new data being filtered in that will sway the K value accordingly.

C. If the two K values differ when you are trying to select a model, what would be the steps you would walk through to determine the number parameters you would use to select a model? In other words, how would YOU pick a model using AIC and BIC if they don't line up?

Given that I like convenience, I would go for the lease complex model.