AIC and BIC

Vaibhav Sherkar (BS2041)

While comparing two different models, we look at maximum value attained by the log likelihood function and the number of free parameters estimated while fitting the model. Now, we present two commonly used methods to balance the maximum value attained by the log likelihood and the number of free parameters estimated while fitting the model in order to select the best model among different models.

Akaike Information Criterion

Originally proposed for 'An Information Criterion' by Akaike, the acronym AIC is used for 'Akaike Information Criterion'. AIC is defined as

$$AIC = -2l(\hat{\theta}) + 2p$$

Here, p denotes the number of free parameters in the model and $l(\hat{\theta})$ denotes the maximum value attained by the log likelihood function(l is the log likelihood function and $\hat{\theta}$ is the MLE.). If the maximum value attained by log likelihood increases, then AIC decreases. On the other hand, if number of free parameters estimated in the model increases then AIC increases. Lesser number of free parameters estimated and higher maximum value attained by log likelihood function are favourable conditions for good fit. Hence, the model with lower AIC is preferred over model with higher AIC.

Bayes Information Criterion

The acronym 'BIC' or 'SBC'(Schwarz Bayes Criterion) is used for this information criterion. BIC is defined as

$$BIC = -2l(\hat{\theta}) + plogn$$

Here, p denotes the number of free parameters in the model, $l(\hat{\theta})$ denotes the maximum value attained by the log likelihood function(l is the log likelihood function and $\hat{\theta}$ is the MLE.) and n is the sample size. Here, the weight of p depends on the sample size. Model with lower BIC is preferred over model with higher BIC.

Note that values of AIC and BIC do not have any interpretation. AIC and BIC values need to be used solely for the purpose of comparison.