TIME SERIES AND FORECASTING ASSIGNMENT-7

AIC(Akaike information criterion)

The Akaike information criterion (**AIC**) is an estimator of out-of-sample prediction error and thereby the relative quality of statistical models for a given set of data. AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.

Although the AIC will choose the best model from a set, it won’t say anything about *absolute*quality. In other words, if all of our models are poor, it will choose the best of a bad bunch. Therefore, once we have selected the best model, consider running a [hypothesis test](https://www.statisticshowto.com/probability-and-statistics/hypothesis-testing/) to figure out the relationship between the variables in your model and the outcome of interest.

[Akaike’s Information Criterion is usually calculated with the software. The basic formula is defined as:  
 **AIC = -2(log-likelihood) + 2K**  
Where:](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

* [K is the number of model parameters (the number of variables in the model plus the intercept).](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)
* [Log-likelihood is a measure of model fit. The higher the number, the better the fit. This is usually obtained from statistical output.](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

[For small sample sizes (n/K < ≈ 40), use the second-order AIC:  
 **AICc = -2(log-likelihood) + 2K + (2K(K+1)/(n-K-1))**  
Where:](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

* [n = sample size,](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)
* [K= number of model parameters,](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)
* [Log-likelihood is a measure of model fit.](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

[An alternative formula for](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)[least squares regression](https://www.statisticshowto.com/least-squares-regression-line/)[type analyses for](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)[normally distributed](https://www.statisticshowto.com/probability-and-statistics/normal-distributions/)[errors:  
 **AIC = n log(rss) + 2K**  
Where:](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

* [rss= Residual Sum of Squares/n,](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)
* [n = s](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)[ample size](https://www.statisticshowto.com/probability-and-statistics/find-sample-size/)[,](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)
* [K is the number of model parameters.](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

[Note that with this formula, the estimated](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)[variance](https://www.statisticshowto.com/probability-and-statistics/variance/)[must be included in the parameter count.](https://en.wikipedia.org/wiki/Akaike_information_criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC,a%20given%20set%20of%20data.&text=AIC%20estimates%20the%20relative%20amount,the%20quality%20of%20that%20model.)

BIC( Bayesian information criterion)

The Bayesian information criterion or Schwarz information criterion is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. It is based, in part, on the likelihood function and it is closely related to the Akaike information criterion.

The Bayesian Information Criterion (BIC) is defined as

*k log(n)- 2log(L(θ̂)).*

Here n is the sample size; the number of observations or number of data points you are working with. k is the number of [parameters](https://www.statisticshowto.com/parametrization-parameterize/)which your model estimates, and θ is the set of all parameters.

L(θ̂) represents the likelihood of the model tested, given your data, when evaluated at [maximum likelihood](https://www.statisticshowto.com/maximum-likelihood-estimation/) values of θ. You could call this the likelihood of the model given everything aligned to their most favourable.

Another way of understanding L(θ̂) is that it is the probability of obtaining the data which you have, supposing the model being tested was a given.

Comparing Models

Comparing models with the Bayesian information criterion simply involves **calculating the BIC for each model.**The model with the lowest BIC is considered the best, and can be written BIC\*(or SIC\* if you use that name and abbreviation).

We can also calculate the Δ BIC; the difference between a particular model and the ‘best’ model with the lowest BIC, and use it as an argument against the other model.

**Δ BIC i=BICmodel – BIC\*,**

Where,

BIC\* is the best model.

If Δ BIC is less than 2, it is considered ‘barely worth mentioning’ as an argument either for the best theory or against the alternate one. The edge it gives our best model is too small to be [significant](https://www.statisticshowto.com/what-is-statistical-significance/).

If Δ BIC is between 2 and 6, one can say the evidence against the other model is positive; i.e. we have a good argument in favour of our ‘best model’.

If it’s between 6 and 10, the evidence for the best model and against the weaker model is strong. A Δ BIC of greater than ten means the evidence favouring our best model vs the alternate is very strong indeed.