0.1 Akaike Information Criterion

Akaike's information criterionis a measure of the goodness of fit of an estimated statistical model. The AIC was developed by Hirotsugu Akaike under the name of "an information criterion" in 1971. The AIC is a ***model selection*** tool i.e. a method of comparing two or more candidate regression models. The AIC methodology attempts to find the model that best explains the data with a minimum of parameters. (i.e. in keeping with the law of parsimony)

The AIC is calculated using the "likelihood function" and the number of parameters (Likelihood function : not on course). The likelihood value is generally given in code output, as a complement to the AIC. Given a data set, several competing models may be ranked according to their AIC, with the one having the lowest AIC being the best. (Although, a difference in AIC values of less than two is considered negligible).

The Akaike information criterion is a measure of the relative goodness of fit of a statistical model. It was developed by Hirotsugu Akaike, under the name of "an information criterion" (AIC), and was first published by Akaike in 1974.

$$\text{AIC} = 2p - 2\ln(L)$$

- p is the number of free model parameters.
- L is the value of the Likelihood function for the model in question.
- For AIC to be optimal, n must be large compared to p.

0.1.1 Schwarz's Bayesian Information Criterion

An alternative to the AIC is the Schwarz BIC, which additionally takes into account the sample size n.

$$\text{BIC} = p\ln n - 2\ln(L)$$

1 Information Criterions

We define two types of information criterion: the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). In AIC and BIC, we choose the model that has the minimum value of:

$$AIC = 2log(L) + 2m,$$

$$BIC = 2log(L) + mlogn$$

where

- L is the likelihood of the data with a certain model,
- n is the number of observations and
- m is the number of parameters in the model.

1.1 AIC

The Akaike information criterion is a measure of the relative **goodness of fit** of a statistical model.

When using the AIC for selecting the parametric model class, choose the model for which the AIC value is lowest.

1.2 Akaike Information Criterion

Akaike's information criterionis a measure of the goodness of fit of an estimated statistical model. The AIC was developed by Hirotsugu Akaike under the name of "an information criterion" in 1971. The AIC is a ***model selection*** tool i.e. a method of comparing two or more candidate regression models. The AIC methodology attempts to find the model that best explains the data with a minimum of parameters. (i.e. in keeping with the law of parsimony)

The AIC is calculated using the "likelihood function" and the number of parameters (Likelihood function : not on course). The likelihood value is generally given in code output, as a complement to the AIC. Given a data set, several competing models may be ranked according to their AIC, with the one having the lowest AIC being the best. (Although, a difference in AIC values of less than two is considered negligible).

The Akaike information criterion is a measure of the relative goodness of fit of a statistical model. It was developed by Hirotsugu Akaike, under the name of "an information criterion" (AIC), and was first published by Akaike in 1974.

$$\text{AIC} = 2p - 2\ln(L)$$

- p is the number of free model parameters.
- L is the value of the Likelihood function for the model in question.
- For AIC to be optimal, n must be large compared to p.

1.2.1 Schwarz's Bayesian Information Criterion

An alternative to the AIC is the Schwarz BIC, which additionally takes into account the sample size n.

$$\text{BIC} = p\ln n - 2\ln(L)$$

1.3 Model Metrics for Logistic Regression Models

- In order to understand how much variation in the dependent variable can be explained by a logistic regression model (the equivalent of R 2 in multiple regression), you should consult Model Summary statistics.
- Although there is no close analogous statistic in logistic regression to the coefficient of determination R 2 the Model Summary Table provides some approximations.

- Logistic regression does not have an equivalent to the R-squared that is found in OLS regression; however, many researchers have tried to come up with one.
- The SPSS output table below contains the Cox & Snell R Square and Nagelkerke R Square values, which are both methods of calculating the explained variation. These values are sometimes referred to as pseudo R 2 values (and will have lower values than in multiple regression).
- However, they are interpreted in the same manner, but with more caution. Therefore, the explained variation in the dependent variable based on our model ranges from 24.0to 33.0%, depending on whether you reference the Cox & Snell R 2 or Nagelkerke R 2 methods, respectively.
- Nagelkerke R 2 is a modification of Cox & Snell R 2 , the latter of which cannot achieve a value of 1. For this reason, it is preferable to report the Nagelkerke R 2 value.
- The Nagelkerke modification that does range from 0 to 1 is a more reliable measure of the relationship.
- Nagelkerkes R 2 will normally be higher than the Cox and Snell measure. Figure 1: SPSS output
- Cox and Snells R-Square attempts to imitate multiple R-Square based on likelihood, but its maximum can be (and usually is) less than 1.0, making it difficult to interpret. Here it is indicating that 55.2logistic model.

1.4 Pseudo R-squares

- Cox & Snell R Square and Nagelkerke R Square are two measures from the pseudo R-squares family of measures.
- There are a wide variety of pseudo-R-square statistics (these are only two of them). Because this statistic does not mean what R-squared means in OLS regression (the proportion of variance explained by the predictors), we suggest interpreting this statistic with great caution.

1.5 Cox & Snell R Square

Cox and Snells R-Square is an attempt to imitate the interpretation of multiple R-Square based on the likelihood, but its maximum can be (and usually is) less than 1.0, making it difficult to interpret. It is part of SPSS output.

1.6 Nagelkerkes R-Square

- Nagelkerkes R 2 is part of SPSS output in the Model Summary table and is the most-reported of the R-squared estimates.
- In our case it is 0.737, indicating a moderately strong relationship of 73.7predictors and the prediction.

- Nagelkerkes R-Square is a further modification of the Cox and Snell coefficient to assure that it can vary from 0 to 1. Nagelkerkes R-Square will normally be higher than the Cox and Snell measure. It is part of SPSS output and is the most-reported of the R-squared estimates.