Contents

0.1   Pseudo-R Squared

Cox and Snell R Square and Nagelkerke R Square - These are pseudo R-squares. Logistic regression does not have an equivalent to the R-squared that is found in OLS regression; however, many people have tried to come up with one.

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.

0.2   Psuedo R Squared Values

0.3   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.

Logistic regression does not have an equivalent to the R-squared that is found in OLS regression; however, many researchrs have tried to come up with one. There are a wide variety of pseudo-R-square statistics. Cox and Snell's 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.

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.

Nagelkerke's R-Square is a further modification of the Cox and Snell coefficient to assure that it can vary from 0 to 1. Nagelkerke's 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.

1   R Squared Diagnostics

- In order to understand how much variation in the dependent variable can be explained by the model (the equivalent of R^2 in multiple regression), you should consult ***Model Summary*** statistics.

- 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 24.0% to 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.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	102.088[a]	.240	.330

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Figure 1: SPSS output

- Although there is no close analogous statistic in logistic regression to the coefficient of determination R^2 the Model Summary Table provides some approximations. 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 55.2% of the variation in the DV is explained by the logistic model.

- Logistic regression does not have an equivalent to the R-squared that is found in OLS regression; however, many people have tried to come up with one. Cox and Snell R Square and Nagelkerke R Square - These are pseudo R-squares.

- Nagelkerke's R-Square is a further modification of the Cox and Snell coefficient to assure that it can vary from 0 to 1. Nagelkerke's 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.

- 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.

- **Cox and Snell's 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.

1.0.1 Nagelkerke's 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 73.7% between the predictors and the prediction.